Social Psychological and Personality Science
1-11

DOI: 10.1177/1948550617734616
journals.sagepub.com/home/spp
SAGE

Faces Tell Everything in a Just and Biologically Determined World: Lay Theories Behind Face Reading

Atsunobu Suzuki[1], Saori Tsukamoto[2], and Yusuke Takahashi[3]

Abstract

The present research investigated an often presumed but rarely assessed construct named the *physiognomic belief*—a generic belief that various traits can be inferred from faces. Studies in Japan and the United States have demonstrated that this belief can be measured reliably and invariantly across cultures and that those having stronger beliefs make more extreme trait inferences from faces. Of note, in both countries, the physiognomic belief is positively associated with a biologically deterministic view of personality traits and a belief in a just world. These findings suggest two types of naive justifications for the physiognomic belief: faces and traits should be related because they are both manifestations of biological essences and because the world is an orderly place wherein people get faces they deserve. This highlights an understudied role of folk concepts involving faces and traits in the popularity of face-based trait inference among laypeople.

Keywords

physiognomy, psychological essentialism, just world belief, trustworthiness, first impressions

Face-based trait inference (FBTI) is very popular among laypeople (Todorov, 2017; Zebrowitz, 1997), such that they spontaneously judge basic social attributes, like trustworthiness, from faces (Klapper, Dotsch, van Rooij, & Wigboldus, 2016). Consequently, facial appearance affects high-stakesreal-world decisions, like voting and criminal sentencing (Olivola, Funk, & Todorov, 2014; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). Some scientists warn against the situations and call FBTI "face-ism" to emphasize its basis in inaccurate stereotypes (Olivola, Funk, et al., 2014), while others focus on the "kernels of truth" in it (Bonnefon, Hopfensitz, & De Neys, 2015).

The most actively studied aspect of the FBTI is "face–trait mapping," or a set of rules people commonly (and probably implicitly) use to relate specific facial features to specific traits (Todorov, Dotsch, Porter, Oosterhof, & Falvello, 2013; Todorov et al., 2015; Todorov & Oosterhof, 2011). In contrast, relatively little research has been done on lay beliefs about the validity of FBTI. It has so far been indicated that people's introspections into the accuracy of the trait inferences they make from others' faces tend to be poor (Ames, Kammrath, Suppes, & Bolger, 2010; Hassin & Trope, 2000; Olivola, Eubanks, & Lovelace, 2014; Olivola & Todorov, 2010a; but see also Biesanz et al., 2011). The previous studies, however, are concerned with specific kinds of face judgments and are somewhat mute on a generic belief that various traits can be inferred from faces. This belief is termed here as the *physiognomic belief*, after the name "physiognomy," which refers to the pseudoscience of face reading (Hassin & Trope, 2000; Todorov, 2017; Zebrowitz, 1997). Although plenty of anecdotal cases imply this belief, like Cicero's saying that "the countenance is the reflection of the soul" (as cited in Zebrowitz, 1997), it has rarely been assessed in empirical studies. The paucity of research is unfortunate, given that the physiognomic belief has a close connection with FBTI. Namely, this belief could be shared among people who trust in FBTI, even if they adhere to different ways of face–trait mapping (Mattarozzi, Todorov, Marzocchi, Vicari, & Russo, 2015). Therefore, identifying the origins and consequences of the physiognomic belief is important to understand the popularity of FBTI in human society.

A pioneering theory by Secord (1958) proposed FBTI as a type of temporal extension; that is, an inference of enduring

[1] Department of Cognitive and Psychological Sciences, Graduate School of Informatics, Nagoya University, Nagoya, Aichi, Japan
[2] Division of Liberal Arts and Sciences, Aichi Gakuin University, Nisshin, Aichi, Japan
[3] Hakubi Center for Advanced Research, Kyoto University, Kyoto, Japan

Corresponding Author:
Atsunobu Suzuki, Department of Psychology, Graduate School of Humanities and Sociology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.
Email: atsuzuki@l.u-tokyo.ac.jp

traits from momentary states. For example, people's tendency to attribute trustworthiness to happy-looking faces is explained as an overgeneralization of a friendly intent at the moment (Oosterhof & Todorov, 2008; Sutherland, Oldmeadow, & Young, 2016). It has been shown that dispositional inference from momentary behaviors, which is another type of temporal extension, is enhanced by *the entity theory of personality*, or the lay belief in the fixedness of personal attributes (Chiu, Hong, & Dweck, 1997; Dweck, Chiu, & Hong, 1995). Thus, the physiognomic belief may have a positive association with the entity theory as well.

The entity theory can be placed in a broader folk conceptual framework of psychological essentialism (Bastian & Haslam, 2006; Medin, 1989), which postulates that individuals have their own "essences" that make them who they are. People often reify the abstract concept of essences into biological entities like genes (Haslam, Bastian, & Bissett, 2004; Keller, 2005). Intriguingly, such *biological determinism* often underlies scientific arguments in favor of FBTI. For example, a large facial width to height ratio (fWHR) may be an honest signal of aggression because the high level of testosterone may increase both fWHR and aggression (Lefevre, Lewis, Perrett, & Penke, 2013). Although this particular claim has received mixed support (Bird et al., 2016; Haselhuhn, Ormiston, & Wong, 2015), similar biologically deterministic views of human traits are likely to foster the physiognomic belief.

Furthermore, like the words of Oscar Wilde that "a man's face is his autobiography" (as cited in Zebrowitz, Voinescu, & Collins, 1996), there may be a common belief that people get faces that they deserve. This may remind one of a well-studied construct of the *belief in a just world* (BJW), stating that the world is a just place where people get what they deserve (Lerner & Miller, 1978). While BJW has been typically related to the judgment of deservingness of extrinsic outcomes, such as crime victimization (Hafer & Sutton, 2016), Rubin and Peplau (1975) have suggested that it is associated with the attractiveness halo effect, a type of FBTI, whereby positive personality traits are attributed to physically attractive individuals (Zebrowitz, 1997). That is, according to BJW, attractive-looking people must have desirable characteristics that make them deserve "the reward of beauty" (Rubin & Peplau, 1975). Indeed, Dion and Dion (1987) have demonstrated that believers in a just world exhibit a stronger halo effect than nonbelievers do (see also Callan, Powell, & Ellard, 2007). Likewise, BJW may generally contribute to the physiognomic belief.

Finally, the physiognomic belief is naturally expected to bolster FBTI. That is, compared to nonbelievers who should view FBTI as unfeasible, believers should make more extreme trait judgments from faces (Chiu et al., 1997; Dweck et al., 1995; Livingston, 2001). However, people having faith in the physiognomic belief can possess their own ideas about face–trait relations (Mattarozzi et al., 2015). Therefore, although stronger physiognomic belief should be associated with more extreme face judgments, the direction of extremization might be idiosyncratic. Namely, different individuals who commonly hold the physiognomic belief may judge the same face as being highly trustworthy and highly untrustworthy. Nonetheless, it would also be interesting to see if believers in physiognomy tend to make similar face judgments, given the socially shared stereotypes that connect specific facial appearances with specific traits (Todorov et al., 2015). In short, it is important to distinguish between *simple extremity* and *consensual extremity* of FBTI. Here, simple extremity is defined to be high when any extreme judgments of faces are made, whereas consensual extremity is defined to be high when not only extreme but also stereotype-consistent judgments are made; for instance, judging a consensually trustworthy face (e.g., happy-looking face) as highly untrustworthy indicates high simple extremity but low consensual extremity. As discussed above, simple extremity of FBTI is predicted to be related to the physiognomic belief, whereas it is somewhat an open question whether consensual extremity has the same relation.

Study 1

In Study 1A, the relationships between the physiognomic belief and other constructs described above, which are theoretically predicted but not yet demonstrated, were tested in Japan and the United States. Since FBTI is a ubiquitous practice across cultures (Olivola & Todorov, 2010b; Rule et al., 2010; Walker, Jiang, Vetter, & Sczesny, 2011), we expected to obtain consistent results across the two countries, which would provide stronger evidence than data from one nation. As part of the project, a cross-culturally applicable measure of the physiognomic belief was developed. While the Perceptual Reliance Index (Livingston, 2001)—which quantifies the propensity to make person judgments from perceptual cues in general—might seem to suffice for the present purpose, it includes items containing racial stereotypes specific to the United States, which precludes its use in other countries. The test–retest reliability of our original measure of the physiognomic belief was examined in Study 1B.

Study 1A

Method

Participants

Participants were recruited through an online research company. Japanese people speaking Japanese and U.S. residents speaking English as their primary language (20–49 years old) were eligible for participation. Responses from 896 Japanese (444 males and 452 females; 35.21 $\pm$ 8.25 years old) and 500 Americans (250 males and 250 females; 34.76 $\pm$ 8.25 years old; 333 Caucasians, 47 African Americans, 39 Hispanics, 30 Asians, and 51 other races) were collected. The sample sizes were determined considering budgetary limitations. The size of the Japanese sample was larger than we had requested ($n = 800$) due to an accident.

Table 1. Loadings on the Single Factor by the Target Items of the PBS.

Items	Japan	United States
An aggressive personality is something that shows in a person's face.	.59	.66
I know an emotionally cold person when I see their face.	.72	.68
I know a dominant person when I see their face.	.65	.67
An egoistic personality is something that shows in a person's face.	.64	.58
I know an immoral person when I see their face.	.69	.59
I know a mean person when I see their face.	.70	.61
I know an emotionally unstable person when I see their face.	.55	.69
I know a caring person when I see their face.	.76	.70
Competence is something that shows in a person's face.	.65	.62
A cooperative personality is something that shows in a person's face.	.70	.69
I know a forgiving person when I see their face.	.74	.67
Intelligence is something that shows in a person's face.	.59	.63
I know a kind person when I see their face.	.72	.66
I know a trustworthy person when I see their face.	.66	.70
Proportion of variance explained	.45	.43

Note. Table S1 in the Online Supplemental Material provides fuller descriptions of the PBS, including the Japanese version. PBS = Physiognomic Belief Scale.

Procedure and Tasks

Participants completed a series of tasks (questionnaires and face judgments) in a fixed order on a research website. Only the tasks of interest in this article are described in detail below (see the Online Supplemental Material for the full set of tasks). The order of the items was randomized within each task.

Physiognomic belief. A questionnaire named the Physiognomic Belief Scale (PBS) was developed for this study, which contains 14 "target items" that assess the physiognomic belief (Table 1) as well as 4 "control items" used for screening purposes (see Table S1 in the Online Supplemental Material). Participants rated how closely each item matched their own thoughts using a scale ranging from 1 (*disagree*) to 4 (*agree*).

Target items were written in one of the following formats: "I know a ____ person when I see their face" or "____ is something that shows in a person's face," wherein the blanks were substituted with specific trait words. Those traits that were spontaneously inferred from faces (e.g., trustworthiness, dominance) or were related to basic dimensions of person perception (e.g., warmth, competence) were used (Oosterhof & Todorov, 2008; Sutherland et al., 2016). A total of 26 items were generated, among which inappropriate ones (e.g., those showing ceiling or floor effects) were screened out through pilot studies in Japan, resulting in the 14 final items.

Control items included 2 "YES" items and 2 "NO" items that should normally be endorsed (e.g., "I can tell babies and the elderly apart by their faces") or rejected (e.g., "I can guess a person's address correctly when I see their face"). These items were used to detect and discard "inattentional" (i.e., giving the same response to all items) and "unusual" responses (i.e., agreeing and disagreeing with all NO and YES items, respectively).

The PBS was first created in Japanese by the authors, and was then translated into English by a native speaker. The authors examined the translation and made minor corrections as needed to minimize discrepancies between the Japanese and English versions.

Entity theory of personality. The Implicit Person Theory Measure (Levy, Stroessner, & Dweck, 1998) was used, which was translated into Japanese by the authors. The scale consists of 8 items assessing a belief in the fixedness of personal attributes (e.g., "The kind of person someone is, is something basic about them, and it can't be changed very much"). Participants indicated how much they agreed with each statement using a scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*).

Belief in a just world. The Global Belief in a Just World Scale (Lipkus, 1991) was used, which was translated into Japanese by Shirai (2011). The scale consists of 7 items that assess people's belief that others get what they deserve and deserve what they get (e.g., "I feel that people get what they are entitled to have"). Participants indicated how much they agreed with each statement using a scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*).

Biological determinism. A questionnaire named the Belief in Biological Basis Scale was devised by the authors, using a similar method to that of Haslam, Bastian, and Bissett (2004). Specifically, the measure listed the same 14 trait words included in the PBS. Participants rated how strongly they thought each trait was based on a person's biological nature (genes, brain structure, hormone levels, etc.) using a scale ranging from 1 (*not based on biological nature*) to 4 (*based on biological nature*).

Perceptual reliance. The Perceptual Reliance Index (Livingston, 2001) consists of 9 items that measure one's propensity to make social judgments from perceptual cues (e.g., "In general, light-skinned Blacks tend to be more attractive than dark-skinned Blacks"). Participants indicated how much they agreed or disagreed with each statement using a scale ranging from 1 (*disagree*) to 5 (*agree*). The scale was administered only in the U.S. study since several items mention racial stereotypes with which Japanese people would not be familiar.

Face judgment. In this task, pairs of faces were presented side by side on the screen. The faces were paired so that one was consensually more trustworthy (i.e., likely to be rated as more trustworthy) than the other. Two and four pairs of faces were presented in a random order in the studies conducted in Japan and the United States, respectively. See the Online Supplemental Material for more details about the face stimuli.

Participants judged which person in each pair was more trustworthy based on their impressions of the faces. The judgments were made on a scale ranging from 1 (*The person on the right is far more trustworthy*) to 6 (*The person on the left is far more trustworthy*). When computing simple extremity, ratings were recoded so that higher values reflected a greater tendency to make extreme face judgments, regardless of directionality (i.e., ratings of 6, 5, 4, 3, 2, and 1 were recoded into 3, 2, 1, 1, 2, and 3, respectively, for all pairs). When computing consensual extremity, ratings were recoded so that higher values reflected a greater tendency to rate consensually more trustworthy faces as being more trustworthy; that is, the ratings were reversed for those pairs wherein the consensually more trustworthy face was placed on the right.

Preference for extreme options. Correlations between the physiognomic belief and simple extremity of face judgments, even if observed, could be spurious ones that are mediated by an individual's tendency to respond extremely to any question. To address this problem, each participant's preference for extreme options was estimated from their responses to a personality questionnaire (the short 15-item Big Five Inventory; Lang, John, Lüdtke, Schupp, & Wagner, 2011). Participants indicated how much they agreed with each of the 15 statements describing one of the Big Five personality traits (three per trait) on a 5-point scale. The ratings of 5, 4, 3, 2, and 1 were recoded into 3, 2, 1, 2, and 3, respectively, for all items, and the average of the recoded ratings was computed as an index of preference for extreme options.

Results

Participants who met either of the following criteria were discarded: (1) giving the same response to all items in any questionnaire with reverse items (including the PBS) or (2) agreeing and disagreeing with all NO and YES items in the PBS, respectively. This screening procedure resulted in the inclusion of 770 Japanese (372 males and 398 females; 35.33 $\pm$ 8.23 years old) and 443 U.S. participants (221 males and 222 females; 34.80 $\pm$ 8.22 years old; 300 Caucasians, 40 African Americans, 35 Hispanics, 22 Asians, and 46 other races).

First, exploratory factor analyses of the PBS (target items) were performed separately for the Japanese and U.S. data. In both samples, parallel analysis and Velicer's minimum average partial test (O'Connor, 2000) indicated a one-factor model, and the single factor was heavily loaded by all target items (Table 1; for the item statistics, see Table S2 in the Online Supplemental Material). Multigroup confirmatory factor analysis was then conducted on the pooled data using the "lavaan" package for R (Hirschfeld & von Brachel, 2014) to examine measurement invariance between Japanese and English versions of the PBS (Raju, Laffitte, & Byrne, 2002). Specifically, three models were compared for their fit to the PBS data pooled across Japan and the United States. (See Figure 1 for a detailed and graphical description of the models.) In short, Model 1 (configural invariance) was the baseline model, wherein a single-factor model was separately applied to the PBS data of each country. Model 2 (metric invariance) was a constrained version of Model 1, wherein the vectors of factor loadings were set to be equal across countries. Model 3 (scalar invariance) was the most constrained model, positing that the vectors of factor loadings and item intercepts were equal across countries. Table 2 shows goodness-of-fit indices for the three models. Model 2 was selected according to the "ΔCFI < .01" rule (i.e., if the difference in the comparative fit index [CFI] between two models is smaller than .01, the more constrained model is adopted; Hirschfeld & von Brachel, 2014). This confirmed that the two language versions exhibited metric invariance; that is, they measured the same latent factor (i.e., physiognomic belief) in that the strength of the relationships between the items and the factor was equivalent across versions (Raju et al., 2002).

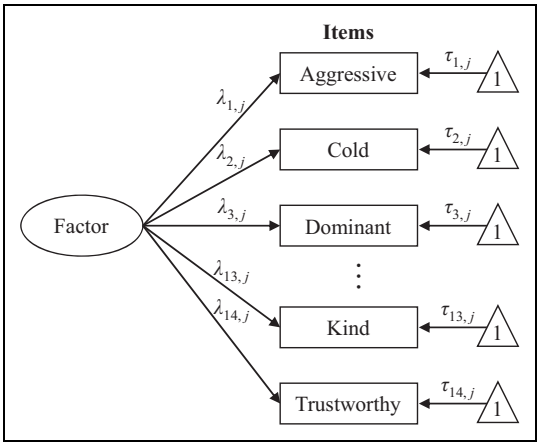

Figure 1. Graphical representation of the multigroup confirmatory factor analysis model applied to the Physiognomic Belief Scale data pooled across Japan and the United States. The oval, rectangles, and triangles indicate a latent variable (factor), observed variables (item ratings), and constants, respectively. $\lambda_{i,\,j}$s and $\tau_{i,\,j}$s are the parameters to be estimated, which represent factor loadings and item intercepts, respectively; the subscripts of *i* and *j* denote the item (1, . . . , 14) and the group (1 [Japan] or 2 [United States]), respectively. $\lambda_{1,\,1}$ and $\lambda_{1,\,2}$ were fixed to one, and factor means were fixed to zero for the model to be identified. In Model 1, no further constraints were imposed on the parameters. In Model 2, the factor loadings were constrained to be equal across groups (i.e., $\lambda_{i,\,1} = \lambda_{i,\,2}$ for any *i*). In Model 3, both the factor loadings and item intercepts were constrained to be equal (i.e., $\lambda_{i,\,1} = \lambda_{i,\,2}$ and $\tau_{i,\,1} = \tau_{i,\,2}$ for any *i*). See Gregorich (2006) for more details of the notational conventions.

Table 2. Goodness-of-Fit Indices of the Models With Different Types of Measurement Invariance Between Japanese and English Versions of the PBS.

Model No.	Type of Invariance	*df*	CFI	RMSEA	ΔCFI
1	Configural	154	.934	.073	—
2	Metric	167	.929	.073	.005
3	Scalar	180	.874	.093	.054

Note. See the main text and Figure 1 for the detailed descriptions of the models. PBS = Physiognomic Belief Scale; CFI = comparative fit index; RMSEA = root mean square error of approximation.

For each measure, ratings were averaged to obtain an aggregate score per participant. Table 3 shows descriptive statistics of those scores and their correlations. First, PBS scores had a positive and moderate correlation with the scores on the Perceptual Reliance Index in the United States. This result, together with the abovementioned metric equivalence, indicates that the PBS is a cross-culturally valid measure of a construct related to a type of perceptual reliance, namely, reliance on face perception in social inference. Moreover, as predicted, across Japan and the United States, PBS scores significantly correlated with scores on measures of biological determinism, BJW, and simple extremity in face judgments. Correlations between the PBS score and simple extremity remained significant even after adjusting for the preference for extreme options (Japan: $r = .219$, 95% confidence interval [CI] [.151, .285]; the United States: $r = .217$, 95% CI [.126, .304]). On the other hand, two of the significant correlations were specific to either country: the physiognomic belief significantly correlated with the entity theory of personality only in the United States and with the consensual extremity of face judgments only in Japan.

Study 1B

To confirm the test–retest reliability of the PBS, Japanese and the U.S. participants were asked to complete the PBS twice separated by an interval of approximately a month (see the Online Supplemental Material for details about the participants and procedure). The results showed that PBS scores were highly correlated between the two occasions both in Japan, $r = .771$, 95% CI [.700, .826] and in the United States, $r = .799$, 95% CI [.716, .860].

Study 2

Study 1 showed that the PBS is a cross-culturally valid and reliable measure and supports most of the predictions regarding the physiognomic belief. However, we found almost no correlation between the PBS and consensual extremity of face judgments in the United States. The result was not totally unexpected but may challenge the existence of a natural association between the physiognomic belief and FBTI. Moreover, it may be insufficient to only demonstrate the PBS's association with the judgments of facial trustworthiness. Therefore, additional studies were conducted in the United States to corroborate the relationships between the physiognomic belief and FBTI. In Study 2A, face-based judgments of dominance were examined in addition to trustworthiness, as these are the two basic dimensions accounting for a variety of FBTI (Oosterhof & Todorov, 2008; Sutherland et al., 2016). Study 2B focused on competence judgments, which are known to show high agreement across raters and affect people's real-world choices such as political elections (Olivola, Funk, et al., 2014; Olivola & Todorov, 2010b; Todorov et al., 2015). The means of consensual extremity in Study 1A (see Table 3) indicate that the trustworthiness impressions of paired faces presented to the U.S. participants differed to a lesser extent than those presented to Japanese participants. We speculated that this might have obscured the relationship between the physiognomic belief and consensual extremity in the U.S. data. Thus, in Studies 2A and 2B, face stimuli were prepared so that differences in the relevant trait impression between paired faces would be greater than those in Study 1A. Moreover, participants in Study 2B completed a series of questionnaires as in Study 1A. Importantly, the task order was randomized this time to eliminate a concern that our findings in Study 1A might have been affected by the fixed order.

Study 2A

Method

Participants

U.S. residents were recruited for participation via an online research company. Eligibility requirements were the same as in Study 1A. Assuming an effective response rate of 90%, power analyses determined the necessary sample size to be 515 to detect a correlation of .15 with 90% power using a 5% significance level. The obtained sample included 288 males and 227 females (29.99 ± 6.79 years old; 368 Caucasians, 29 African Americans, 20 Hispanics, 42 Asians, and 56 other races).

Procedure and Tasks

Participants completed the PBS, the short 15-item Big Five Inventory, and face judgments in the order herein on a research website. Face judgments were further divided into trustworthiness and dominance ratings, the order of which was randomized across participants.

Like in Study 1A, for both face judgment tasks, pairs of faces were presented side by side on the screen. More specifically, pairs of consensually trustworthy and untrustworthy faces were used for trustworthiness ratings, and pairs of consensually dominant and nondominant faces were used for dominance ratings. See the Online Supplemental Material for more details about the face stimuli. Four pairs of faces were presented in a random order for each type of ratings.

Trustworthiness and dominance ratings required participants to judge which person in each pair was more trustworthy and more dominant, respectively, according to their impressions of the faces. As in Study 1A, the ratings were made on a 6-point scale, which were used to compute simple and consensual extremities.

Results

A screening procedure similar to that of Study 1A was used, resulting in the inclusion of 501 participants (281 males and 220 females; 30.07 ± 6.80 years old; 359 Caucasians, 29 African Americans, 20 Hispanics, 39 Asians, and 54 other races). Table 4 shows descriptive statistics regarding the average scores from the PBS and face judgment tasks as well as their correlations. The physiognomic belief was significantly and

Table 3. Study 1A: Average Scores of Each Measure and Their Correlations.

	Japan			United States			Correlation						
							Upper Diagonal: Japan; Lower Dagonal: United States						
Measure	*M*	*SD*	α	*M*	*SD*	α	1	2	3	4	5	6	7
1. PBS	2.49	0.57	.92	2.50	0.66	.91	—	—	.014 [−.057, .084]	**.211** [.142, .278]	**.381** [.319, .440]	**.235** [.167, .301]	**.176** [.107, .244]
2. Perceptual Reliance Index	—	—	—	3.13	0.69	.74	**.445** [.367, .517]	—	—	—	—	—	—
3. Implicit Person Theory Measure	3.71	0.73	.82	3.12	0.79	.75	**.185** [.093, .274]	**.224** [.134, .311]	—	**.091** [.020, .161]	**−.110** [−.180, −.040]	**.079** [.008, .149]	.023 [−.048, .093]
4. Belief in Biological Basis Scale	2.39	0.63	.94	2.52	0.67	.92	**.374** [.291, .452]	**.402** [.321, .478]	**.219** [.128, .306]	—	**.168** [.098, .236]	**.071** [.001, .141]	.023 [−.047, .094]
5. Global Belief in a Just World Scale	3.36	0.74	.77	3.49	1.01	.87	**.394** [.312, .470]	**.441** [.363, .513]	**.121** [.028, .212]	**.371** [.287, .449]	—	**.155** [.086, .224]	**.073** [.002, .143]
6. Simple extremity in face judgment	1.59	0.66	.74[a]	1.48	0.59	.87	**.235** [.145, .321]	**.281** [.192, .364]	−.037 [−.130, .057]	**.175** [.083, .264]	**.196** [.104, .284]	—	**.650** [.607, .689]
7. Consensual extremity in face judgment	4.29	0.92	.59[a]	3.93	0.77	.61	.089 [−.005, .181]	**.110** [.016, .201]	.008 [−.085, .101]	.060 [−.033, .153]	.054 [−.040, .147]	**.445** [.367, .517]	—

Note. All correlations are corrected for participant's age and gender, and significant correlations at the 5% level are in boldface. Values in square brackets indicate 95% confidence intervals. α = Cronbach's α; PBS = Physiognomic Belief Scale.

[a]Correlations between the ratings of the two pairs of faces.

Table 4. Study 2A: Average Scores of Each Measure and Their Correlations.

Measure	*M*	*SD*	α	Correlation 1	2	3	4
1. PBS	2.45	0.58	.91	—	—	—	—
2. Simple extremity in face judgment (trustworthiness)	1.72	0.49	.79	**.325** [.244, .401]	—	—	—
3. Simple extremity in face judgment (dominance)	1.96	0.47	.73	**.266** [.182, .345]	**.571** [.509, .627]	—	—
4. Consensual extremity in face judgment (trustworthiness)	4.53	0.65	.79	**.238** [.154, .319]	**.768** [.730, .802]	**.395** [.318, .466]	—
5. Consensual extremity in face judgment (dominance)	4.73	0.58	.56	**.168** [.081, .252]	**.364** [.285, .438]	**.708** [.661, .749]	**.365** [.287, .439]

Note. All correlations are corrected for participant's age and gender, and significant correlations at the 5% level are in boldface. Values in square brackets indicate 95% confidence intervals. α = Cronbach's α; PBS = Physiognomic Belief Scale.

positively correlated with both simple and consensual extremities in both trustworthiness and dominance judgments of faces. Correlations between the PBS score and simple extremity remained significant even after adjusting for the preference for extreme options (trustworthiness: $r = .300$, 95% CI [.218, .378]; dominance: $r = .237$, 95% CI [.153, .319]).

Study 2B

Method

Participants

Five hundred and fifteen U.S. residents speaking English as their primary language (20–49 years old) were recruited for participation via an online research company as in Study 2A (298 males and 217 females; 30.07 ± 7.08 years old; 380 Caucasians, 30 African Americans, 13 Hispanics, 36 Asians, and 56 other races).

Procedure and Tasks

Participants completed the same questionnaires as described in the Method section of Study 1A and a face judgment task in a random order on a research website.

The face judgment task again involved displaying pairs of faces side by side on the screen. Twenty pairs of consensually competent and incompetent faces were used in total, which were presented in a random order. See the Online Supplemental Material for more details about the face stimuli.

Participants judged which person in each pair was more competent based on their impressions of the faces using a 6-point scale, which were used to compute simple and consensual extremities.

Results

The same screening procedure as before was used, resulting in the inclusion of 484 participants (279 males and 205 females; 30.30 ± 7.11 years old; 361 Caucasians, 29 African Americans, 10 Hispanics, 32 Asians, and 52 other races). Table 5 shows descriptive statistics regarding the average scores from all the measures as well as their correlations. The physiognomic belief significantly and positively correlated with perceptual reliance, biological determinism, BJW, and simple extremity of face judgments, replicating the results of Study 1A. The correlation between the PBS score and simple extremity remained significant even after the preference for extreme options was adjusted, $r = .257$, 95% CI [.171, .338]. On the other hand, different from Study 1A, the physiognomic belief did not correlate significantly with the entity theory of personality, while it did correlate significantly and positively with consensual extremity of face judgments.

Discussion

Across Japan and the United States, the physiognomic belief was positively associated with biologically deterministic views of personality and BJW. This implies that people may be seduced by the physiognomic belief for at least two reasons. First, it may be due to an essentialist logic that faces and traits are manifestations of common essences (or biological entities) and are thus related to each other (Lefevre et al., 2013). Second, since the world is a just place, people are endowed with faces they deserve as reward or punishment (Rubin & Peplau, 1975). In brief, the results provide unique evidence for an understudied role of intuitive theories about faces and traits in underpinning laypeople's adherence to FBTI (Hugenberg & Wilson, 2013).

The relationships of the physiognomic belief with biological determinism and BJW suggest a couple of interesting predictions. First, learning scientific evidence against biological determinism is shown to mitigate its related attitudes and thoughts, such as racial prejudice (Tsukamoto, Enright, & Karasawa, 2013; Williams & Eberhardt, 2008) and pessimism about mental illness (Lebowitz, Ahn, & Nolen-Hoeksema, 2013). Similar interventions may thus reduce the physiognomic belief and, in turn, the practice of FBTI. Second, according to

Table 5. Study 2B: Average Scores of Each Measure and Their Correlations.

Measure	*M*	*SD*	α	Correlation 1	2	3	4	5	6
1. PBS	2.35	0.63	.93	—	—	—	—	—	—
2. Perceptual Reliance Index	3.09	0.66	.74	**.484** [.413, .550]	—	—	—	—	—
3. Implicit Person Theory Measure	3.02	1.01	.93	.088 [−.002, .176]	**.214** [.128, .298]	—	—	—	—
4. Belief in Biological Basis Scale	2.32	0.60	.92	**.253** [.168, .335]	**.264** [.179, .345]	**.328** [.245, .405]	—	—	—
5. Global Belief in a Just World Scale	3.11	1.00	.91	**.293** [.209, .372]	**.387** [.308, .460]	**.115** [.026, .202]	**.136** [.047, .222]	—	—
6. Simple extremity in face judgment	1.56	0.40	.92	**.266** [.181, .347]	**.318** [.235, .396]	.054 [−.035, .143]	.019 [−.070, .109]	**.250** [.165, .332]	—
7. Consensual extremity in face judgment	4.28	0.42	.79	**.253** [.167, .335]	**.324** [.242, .402]	.068 [−.021, .157]	.086 [−.004, .173]	**.217** [.130, .301]	**.772** [.733, .806]

Note. All correlations are corrected for participant's age and gender, and significant correlations at the 5% level are in boldface. Values in square brackets indicate 95% confidence intervals. α = Cronbach's α; PBS = Physiognomic Belief Scale.

the BJW literature, people who witness an injustice attempt to restore justice using various strategies (Ellard, Harvey, & Callan, 2016; Lerner & Miller, 1978). The affirmation of the physiognomic belief could be one such compensatory tactic (Callan et al., 2007) because it allows one to view the world as just and orderly to the extent that "a book can be judged by its cover." Consequently, for instance, court cases that elicit a sense of injustice may intensify a judicial bias to convict and give harsh sentences to untrustworthy-looking defendants (Porter, ten Brinke, & Gustaw, 2010; Wilson & Rule, 2015; see also Olivola, Funk, et al., 2014) due to a motivated belief in the "regularity" with which faces are meaningfully related to inner traits or due to a desire to see "just deserts" delivered to those who appear to deserve them. These real-world possibilities are nontrivial and merit empirical investigation.

As expected, the positive association between the physiognomic belief and simple extremity of FBTI was also observed consistently across studies. That is, those who had stronger physiognomic beliefs tended to make more extreme trait judgments based on faces. On the other hand, whether the correlation between the physiognomic belief and consensual extremity of FBTI was significant was dependent on the quality and number of face stimuli. Understandably, a general trend was that the more reliable the score of consensual extremity, the larger the correlation. Thus, future studies in which consensual extremity is reliably assessed using a sufficient number of distinct stimuli (as in Study 2B) are important to establish its link with the physiognomic belief. Furthermore, although the present research examined the basic and representative trait dimensions (i.e., trustworthiness, dominance, and competence) to increase the generalizability of the findings, it remains to be determined if the physiognomic belief actually has a broader relationship with FBTI that goes beyond these traits.

The expected correlation between the physiognomic belief and the entity theory of personality was observed in the United States but not in Japan (Study 1A). However, the correlation failed to reach significance also in the United States in Study 2B, suggesting that the association between these two scales may be weak and fragile. Their weak association is somewhat surprising, considering that traditional explanations of FBTI describe it in terms of dispositional reasoning on the basis of facial cues (Secord, 1958; Todorov et al., 2015). That is, the physiognomic belief and FBTI are expected to be enhanced by the entity theory of personality just as dispositional inference from behavioral cues is (Chiu et al., 1997; Dweck et al., 1995). Do our results challenge this theoretical prediction? Future studies using different measures of the entity theory are necessary to answer this question. The scale used to assess the entity theory in the current study consists of items worded in highly abstract terms, such as "kind of person" (Levy et al., 1998), in contrast to the PBS, which contains items that refer to everyday trait words. Thus, the relationship between the entity theory and the physiognomic belief may be more clearly observed if the former is assessed using more tangible items; measurement of implicit associations between the concept of fixedness and personal attributes (Fujii & Uebuchi, 2010; Mascret, Roussel, & Cury, 2015) may be useful for this purpose.

In addition to the limitations mentioned above, a major disadvantage of the present research is that it is correlational in nature. Although we suggest that biologically deterministic views of personality and BJW may foster the physiognomic belief, verification of such causal hypotheses awaits further experimental work. Another limitation is that the present research did not address the relationships between individuals' physiognomic beliefs and their actual abilities of FBTI (Olivola, Eubanks, et al., 2014; Rule, Krendl, Ivcevic, & Ambady, 2013). That is, it remains to be seen whether people with stronger physiognomic beliefs are indeed more accurate in FBTI. Additionally, studies have shown that adult-like FBTI emerges in children as young as 3 years of age (Cogsdill, Todorov, Spelke, & Banaji, 2014) and that older adults adhere to FBTI even when being confronted with its uselessness (Suzuki,

2016). Thus, determining the life span trajectory of the physiognomic belief is also an interesting future direction.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by JSPS KAKENHI grants (no. 15K17326 and 15H03446) awarded to Atsunobu Suzuki.

Supplemental Material

The supplemental material is available in the online version of the article.

References

Ames, D. R., Kammrath, L. K., Suppes, A., & Bolger, N. (2010). Not so fast: The (not-quite-complete) dissociation between accuracy and confidence in thin-slice impressions. *Personality and Social Psychology Bulletin*, *36*, 264–277. doi:10.1177/0146167209354519

Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology*, *42*, 228–235. doi:10.1016/j.jesp.2005.03.003

Biesanz, J. C., Human, L. J., Paquin, A.-C., Chan, M., Parisotto, K. L., Sarracino, J., & Gillis, R. L. (2011). Do we know when our impressions of others are valid? Evidence for realistic accuracy awareness in first impressions of personality. *Social Psychological and Personality Science*, *2*, 452–459. doi:10.1177/1948550610397211

Bird, B. M., Jofré, V. S. C., Geniole, S. N., Welker, K. M., Zilioli, S., Maestripieri, D., . . . Carré, J. M. (2016). Does the facial width-to-height ratio map onto variability in men's testosterone concentrations? *Evolution and Human Behavior*, *37*, 392–398. doi:10.1016/j.evolhumbehav.2016.03.004

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2015). Face-ism and kernels of truth in facial inferences. *Trends in Cognitive Sciences*, *19*, 421–422. doi:10.1016/j.tics.2015.05.002

Callan, M. J., Powell, N. G., & Ellard, J. H. (2007). The consequences of victim physical attractiveness on reactions to injustice: The role of observers' belief in a just world. *Social Justice Research*, *20*, 433–456. doi:10.1007/s11211-007-0053-9

Chiu, C. Y., Hong, Y. Y., & Dweck, C. S. (1997). Lay dispositionism and implicit theories of personality. *Journal of Personality and Social Psychology*, *73*, 19–30. doi:10.1037/0022-3514.73.1.19

Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from faces: A developmental study. *Psychological Science*, *25*, 1132–1139. doi:10.1177/0956797614523297

Dion, K. L., & Dion, K. K. (1987). Belief in a just world and physical attractiveness stereotyping. *Journal of Personality and Social Psychology*, *52*, 775–780. doi:10.1037/0022-3514.52.4.775

Dweck, C. S., Chiu, C. Y., & Hong, Y. Y. (1995). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry*, *6*, 267–285. doi:10.1207/s15327965pli0604_1

Ellard, J. H., Harvey, A., & Callan, M. J. (2016). The justice motive: History, theory, and research. In C. Sabbagh & M. Schmitt (Eds.), *Handbook of social justice theory and research* (pp. 127–143). New York, NY: Springer.

Fujii, T., & Uebuchi, H. (2010). Assessment of participants' theories of intelligence: Reliability and validity of the Implicit Association Test. *Japanese Journal of Educational Psychology*, *58*, 263–274. doi:10.5926/jjep.58.263

Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, *44*, S78–S94. doi:10.1097/01.mlr.0000245454.12228.8f

Hafer, C. L., & Sutton, R. (2016). Belief in a just world. In C. Sabbagh & M. Schmitt (Eds.), *Handbook of social justice theory and research* (pp. 145–160). New York, NY: Springer.

Haselhuhn, M. P., Ormiston, M. E., & Wong, E. M. (2015). Men's facial width-to-height ratio predicts aggression: A meta-analysis. *PLoS One*, *10*, e0122637. doi:10.1371/journal.pone.0122637

Haslam, N., Bastian, B., & Bissett, M. (2004). Essentialist beliefs about personality and their implications. *Personality and Social Psychology Bulletin*, *30*, 1661–1673. doi:10.1177/0146167204271182

Hassin, R., & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology*, *78*, 837–852. doi:10.1037/0022-3514.78.5.837

Hirschfeld, G., & von Brachel, R. (2014). Improving multiple-group confirmatory factor analysis in R: A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research & Evaluation*, *19*, Article 7. Retrieved from http://pareonline.net/getvn.asp?v=19&n=7

Hugenberg, K., & Wilson, J. P. (2013). Faces are central to social cognition. In D. E. Carlston (Ed.), *The Oxford handbook of social cognition* (pp. 167–193). New York, NY: Oxford University Press.

Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology*, *88*, 686–702. doi:10.1037/0022-3514.88.4.686

Klapper, A., Dotsch, R., van Rooij, I., & Wigboldus, D. H. J. (2016). Do we spontaneously form stable trustworthiness impressions from facial appearance? *Journal of Personality and Social Psychology*, *111*, 655–664. doi:10.1037/pspa0000062

Lang, F. R., John, D., Lüdtke, O., Schupp, J., & Wagner, G. G. (2011). Short assessment of the Big Five: Robust across survey methods except telephone interviewing. *Behavior Research Methods*, *43*, 548–567. doi:10.3758/s13428-011-0066-z

Lebowitz, M. S., Ahn, W. K., & Nolen-Hoeksema, S. (2013). Fixable or fate? Perceptions of the biology of depression. *Journal of Consulting and Clinical Psychology*, *81*, 518–527. doi:10.1037/a0031730

Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. (2013). Telling facial metrics: Facial width is associated with testosterone levels in men. *Evolution and Human Behavior*, *34*, 273–279. doi:10.1016/j.evolhumbehav.2013.03.005

Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin*, *85*, 1030–1051. doi:10.1037/0033-2909.85.5.1030

Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology*, *74*, 1421–1436. doi:10.1037/0022-3514.74.6.1421

Lipkus, I. (1991). The construction and preliminary validation of a global belief in a just world scale and the exploratory analysis of the multidimensional belief in a just world scale. *Personality and Individual Differences*, *12*, 1171–1178. doi:10.1016/0191-8869(91)90081-L

Livingston, R. W. (2001). What you see is what you get: Systematic variability in perceptual-based social judgment. *Personality and Social Psychology Bulletin*, *27*, 1086–1096. doi:10.1177/0146167201279002

Mascret, N., Roussel, P., & Cury, F. (2015). Using implicit measures to highlight science teachers' implicit theories of intelligence. *European Journal of Psychology of Education*, *30*, 269–280. doi:10.1007/s10212-015-0249-6

Mattarozzi, K., Todorov, A., Marzocchi, M., Vicari, A., & Russo, P. M. (2015). Effects of gender and personality on first impression. *PLoS One*, *10*, *Article* e0135529. doi:10.1371/journal.pone.0135529

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist*, *44*, 1469–1481. doi:10.1037/0003-066X.44.12.1469

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, *32*, 396–402. doi:10.3758/Bf03200807

Olivola, C. Y., Eubanks, D. L., & Lovelace, J. B. (2014). The many (distinctive) faces of leadership: Inferring leadership domain from facial appearance. *The Leadership Quarterly*, *25*, 817–834. doi:10.1016/j.leaqua.2014.06.002

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences*, *18*, 566–570. doi:10.1016/j.tics.2014.09.007

Olivola, C. Y., & Todorov, A. (2010a). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, *46*, 315–324. doi:10.1016/j.jesp.2009.12.002

Olivola, C. Y., & Todorov, A. (2010b). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*, *34*, 83–110. doi:10.1007/s10919-009-0082-1

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. doi:10.1073/pnas.0805664105

Porter, S., ten Brinke, L., & Gustaw, C. (2010). Dangerous decisions: The impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology Crime & Law*, *16*, 477–491. doi:10.1080/10683160902926141

Raju, N. S., Laffitte, L. J., & Byrne, B. M. (2002). Measurement equivalence: A comparison of methods based on confirmatory factor analysis and item response theory. *Journal of Applied Psychology*, *87*, 517–529. doi:10.1037/0021-9010.87.3.517

Rubin, Z., & Peplau, L. A. (1975). Who believes in a just world. *Journal of Social Issues*, *31*, 65–89. doi:10.1111/j.1540-4560.1975.tb00997.x

Rule, N. O., Ambady, N., Adams, R. B. Jr, Ozono, H., Nakashima, S., Yoshikawa, S., & Watabe, M. (2010). Polling the face: Prediction and consensus across cultures. *Journal of Personality and Social Psychology*, *98*, 1–15. doi:10.1037/a0017673

Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology*, *104*, 409–426. doi:10.1037/A0031050

Secord, P. F. (1958). Facial features and inference processes in interpersonal perception. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior* (pp. 300–315). Stanford, CA: Stanford University Press.

Shirai, M. (2011). For conceptual understanding in "personal severity to crime" and "just world beliefs" (2). *Bulletin of the Graduate School, Toyo University*, *47*, 151–166. Retrieved from https://www.toyo.ac.jp/uploaded/attachment/8065.pdf

Sutherland, C. A. M., Oldmeadow, J. A., & Young, A. W. (2016). Integrating social and facial models of person perception: Converging and diverging dimensions. *Cognition*, *157*, 257–267. doi:10.1016/j.cognition.2016.09.006

Suzuki, A. (2016). Persistent reliance on facial appearance among older adults when judging someone's trustworthiness. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. Advance online publication. doi:10.1093/geronb/gbw034

Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton, NJ: Princeton University Press.

Todorov, A., Dotsch, R., Porter, J. M., Oosterhof, N. N., & Falvello, V. B. (2013). Validation of data-driven computational models of social perception of faces. *Emotion*, *13*, 724–738. doi:10.1037/a0032335

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*, 519–545. doi:10.1146/annurev-psych-113011-143831

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

Tsukamoto, S., Enright, J., & Karasawa, M. (2013). Psychological essentialism and nationalism as determinants of interethnic bias. *Journal of Social Psychology*, *153*, 515–519. doi:10.1080/00224545.2013.795926

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

Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology*, *94*, 1033–1047. doi:10.1037/0022-3514.94.6.1033

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, *26*, 1325–1331. doi:10.1177/0956797615590992

Zebrowitz, L. A. (1997). *Reading faces: Window to the soul?*Boulder, CO: Westview Press.

Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-eyed" and "crooked-faced": Determinants of perceived and real honesty

across the life span. *Personality and Social Psychology Bulletin*, *22*, 1258–1269. doi:10.1177/01461672962212006

Author Biographies

Atsunobu Suzuki is a cognitive psychologist and an associate professor at Department of Psychology at The University of Tokyo. The present work was completed when he was at Department of Cognitive and Psychological Sciences at Nagoya University. He is broadly interested in person perception, focusing in particular on trait inference and emotion recognition from faces.

Saori Tsukamoto is a social psychologist and a lecturer at Division of Liberal Arts and Sciences at Aichi Gakuin University. Her research centers on perception of social categories, with special interest in psychological essentialism and the role of culture.

Yusuke Takahashi is a personality psychologist and an associate professor at Hakubi Center for Advanced Research at Kyoto University. He studies genetics and the development of personality traits through statistical modeling of twin, longitudinal, and cross-national data.

Handling Editor: Dominique Muller