Journal of Experimental Social Psychology 74 (2018) 1–7

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

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

Case Report

Configural processing and social judgments: Face inversion particularly disrupts inferences of human-relevant traits☆

John Paul Wilson[a,*], Steven G. Young[b], Nicholas O. Rule[c], Kurt Hugenberg[d]

[a] Montclair State University, Department of Psychology, 1 Normal Ave, Montclair, NJ 07043, United States
[b] Baruch College & The Graduate Center, CUNY, Box B8-215, 55 Lexington Ave, New York, NY 10010, United States
[c] University of Toronto, Department of Psychology, 100 St. George Street, Toronto, ON M5S 3G3, Canada
[d] Miami University, 90 N Patterson Ave, Oxford, OH 45056, United States

ARTICLE INFO

Keywords:
Trait inferences
Person perception
Configural processing

ABSTRACT

Perceivers tend to strongly agree about the basic trait information that they encode from faces. Although some research has found significant consistency for social inferences from faces viewed at multiple angles, disrupting configural processing can substantially alter the traits attributed to faces. Here, we reconciled these findings by examining how disruptions to configural processing (via face inversion) selectively impairs trait inferences from faces. Across four studies (including a pre-registered replication), we found that inverting faces disrupted inferences about particularly human-relevant traits (trustworthiness and humanness) more than it did for a trait relevant to both human and non-human animals (dominance). These findings contribute to emerging research linking configural processing to the humanization of social targets, helping to provide a clearer understanding of how visual cognition may moderate perceptions of humanness.

Perceivers evaluate faces on basic social traits like trustworthiness and dominance quickly, efficiently, and with great consistency (e.g., Bar, Neta, & Linz, 2006; Willis & Todorov, 2006). Such evaluations can also have a strong impact: predicting the outcome of political elections (Hehman, Carpinella, Johnson, Leitner, & Freeman, 2014), economic decisions (van't Wout & Sanfey, 2008), and even life-and-death criminal sentences (Wilson & Rule, 2015, 2016). The cognitive and perceptual processes that people use to extract traits from faces remain a matter of some debate, however.

1. Featural versus configural processing

Central to the literature on face perception, *featural processing* and *configural processing* represent two fundamentally different ways of encoding faces. Faces have both features (e.g., eyes, nose, mouth) and a configuration (the eyes-over-nose-over-mouth arrangement typical of faces). When perceivers process faces featurally, they encode specific facial features to identify someone (such as a prominent nose or notable birthmark) without integrating features into a gestalt (Madera & Hebl, 2012; Tanaka & Farah, 1993). When perceivers process faces configurally, however, they extract and integrate information about multiple features in parallel and integrate them into a single representation that includes how the features relate to each other (see Maurer, Le Grand, & Mondloch, 2002, for a review).

These two processing modes can be distinguished using manipulations that interfere with the perception of configural versus featural information. For example, inverting a face (i.e., turning it upside-down) disrupts its eyes-over-nose-over-mouth configuration but not the perception of its constituent features. Because face inversion distinctly undermines configural processing, it has been used to explore what aspects of face perception specifically depend on configural information. Indeed, researchers commonly employ face inversion to demonstrate how configural processing influences a variety of outcomes, ranging from face memory (Yin, 1969) to recognizing facial expressions (Young & Hugenberg, 2010).

Recent evidence also suggests that face inversion can actually encourage dehumanization (e.g., Fincher & Tetlock, 2016). Hugenberg et al. (2016) found that face inversion slowed the processing of human-related concepts, disrupted categorizations of human faces as human, and reduced the ascription of human-like traits (e.g., the capacity for

☆ This research was supported by NSF grant BCS-1423765 and a grant from the Social Sciences and Humanities Research Council of Canada.
* Corresponding author.
E-mail addresses: wilsonjoh@montclair.edu (J.P. Wilson), Steven.Young@baruch.cuny.edu (S.G. Young), rule@psych.utoronto.ca (N.O. Rule), hugenbk@miamioh.edu (K. Hugenberg).

http://dx.doi.org/10.1016/j.jesp.2017.07.007
Received 29 April 2016; Received in revised form 2 May 2017; Accepted 20 July 2017
0022-1031/

emotional and cognitive sophistication). Put simply, the typical eyes-over-nose-over-mouth configuration of faces appears to serve as a bottom-up signal of humanness. This configural-humanness link runs in the opposite direction as well: Fincher and Tetlock (2016) found that the faces of dehumanized people were processed less configurally. Moreover, people tend to rely on featural processing when perceiving nonhuman faces (Dahl, Rasch, & Chen, 2014; Mondloch, Maurer, & Ahola, 2006). Thus, we propose that inferences of uniquely human traits might rely more on configural processing than do inferences of traits that humans share with other animals. To inform this hypothesis, we turn next to a brief description of how people infer traits from others' faces.

2. Inferring traits from faces

Perceivers extract trait information from faces both easily and with surprising consensus. For instance, Zebrowitz, Montepare, and Lee (1993) observed strong inter-rater consensus for a number of personality traits (warmth, dominance, strength, honesty, shrewdness) across faces of multiple ethnicities, and even young children make reliable inferences from faces following very brief exposures (e.g., 39 ms; Bar et al., 2006; Cogsdill, Todorov, Spelke, & Banaji, 2014). Data-driven models of social inferences have found that two orthogonal dimensions (facial *trustworthiness* and facial *dominance*) capture much of the variance in these consensual evaluations, and overlap strongly with facial expressions of emotion (Oosterhof & Todorov, 2008; Sutherland et al., 2013).

Consistent with our hypothesis, past studies have indirectly suggested that facial trustworthiness and dominance may rely on configural and featural information differently. In one example, participants judged facial halves as more trustworthy when paired with a trustworthy versus untrustworthy complementary half that they were instructed to ignore (Todorov, Loehr, & Oosterhof, 2010). Their seemingly involuntary integration of these “irrelevant” facial features with focal face characteristics into a unified percept suggests the importance of configural processing for extracting facial trustworthiness. Similarly, Hehman, Flake, and Freeman (2015) found that inferences related to intentions (i.e., trustworthiness) varied more across multiple presentations of the same face than did inferences related to ability (i.e., competence). Ability-related inferences typically relate to static and structural facial features whereas perceptions of trustworthiness typically depend on dynamic facial characteristics, such as facial affect (Carré, McCormick, & Mondloch, 2009; Hehman, Leitner, Deegan, & Gaertner, 2015; Oosterhof & Todorov, 2008). Thus, dynamic traits like trustworthiness seem to rely on constellations of features but structural traits like dominance may be gleaned from single cues (e.g., brow prominence, jaw size; Burton & Rule, 2013).

Given their foundational nature and potential processing distinction, we therefore focused on the “big two” traits of trustworthiness and dominance to examine the link between dehumanization and face perception. Specifically, we hypothesized that disrupting configural processing would affect inferences of especially human traits (e.g., trustworthiness) more than inferences of traits shared with animals (e.g., dominance). Indeed, although it may be difficult to imagine a cat or frog as trustworthy, animals' displays of dominant behavior tend to be quite clear. In fact, even *facial* dominance can be accurately observed in other mammals. Kramer, King, and Ward (2011) found that humans could accurately categorize chimpanzees' dominance by looking at their faces, but could not categorize them as sociable or sympathetic (i.e., traits often associated with humanness). We believe this may be due to different trait signals in faces: If features signal dominance but configurations signal sophisticated human-like traits (e.g., trustworthiness, empathy), then perceivers should be able to reliably extract dominance (but not sociability) from animal faces, as Kramer et al. (2011) found. No one has yet tested this, however.

3. Current research

Here, we therefore examined the role of configural processing in inferences of dominance and trustworthiness from people's faces. We predicted that inverting faces would disrupt the perception of traits considered uniquely human (e.g., trustworthiness) but not the perception of traits believed to be shared by humans and animals (e.g., dominance). We tested this in four studies.

In Study 1, we assessed the correspondence between ratings of the same faces presented upright and inverted, finding that inversion reduced the consistency of trustworthiness ratings more than the consistency of dominance ratings. In Study 2, participants categorized the dominance or trustworthiness of upright and inverted faces in a speeded categorization task, showing less ability to identify inverted faces as trustworthy than as dominant. We replicated these results in Study 3, which we pre-registered using the Open Science Framework (see public registration at https://osf.io/zjevk/). Finally, in Study 4, we demonstrated that trustworthiness is considered a more uniquely human trait than dominance (Study 4a), and that face inversion disrupts inferences of trustworthiness and humanness more than dominance (Study 4b). Together, these studies provide evidence that face inversion selectively disrupts humanity-related trait inferences.

4. Study 1

In Study 1, we randomly assigned participants to rate the trustworthiness or dominance of individual male faces. Participants rated each face both upright and inverted. This allowed us to calculate correlations for each participant's ratings of upright and inverted presentations of the same identity, and thus test how much judgments of upright faces corresponded to judgments of inverted faces. Consistent with our hypothesis that configural processing plays a greater role in inferences of human-specific traits, we predicted that the correlation between upright and inverted faces would be stronger for the dominance than trustworthiness ratings.

4.1. Method

4.1.1. Materials

We downloaded all 37 Black and 36 White male faces from the Chicago Face Database (version 1.0; Ma, Correll, & Wittenbrink, 2015) and eliminated the last Black target so that we would have equal numbers of each race. We resized the images to 450×316 pixels (72 pixels/in.). We selected male faces for the current studies because we did not wish to introduce additional social group factors that might strongly influence trait ratings, and other recent work looking at consistency across multiple presentations also focused on male faces (Hehman, Flake, et al., 2015; Hehman, Leitner, et al., 2015). Moreover, recent work indicates that dominance may be represented differently in male and female faces (Sutherland, Oldmeadow, & Young, 2016)—an issue to which we return in the General Discussion.

4.1.2. Participants and procedure

We recruited 150 American Mechanical Turk (MTurk) Workers so that we could achieve approximately equal Type-I ($\alpha = 0.05$) and Type-II ($\beta = 0.05$) error rates when assuming a medium effect size (i.e., Cohen's $d = 0.6$) in an independent-samples t-test. Half of the participants ($n = 75$) rated each target face on trustworthiness from 1 (*Not trustworthy*) to 7 (*Very trustworthy*), whereas the other half ($n = 75$) rated each face on dominance from 1 (*Not dominant*) to 7 (*Very dominant*). Participants saw each of the 72 targets presented upright in one block and inverted in another block for a total of 144 trials. We randomized the presentation of the faces within these blocks and counterbalanced the order of the blocks. We did not collect any additional measures and report all manipulations and exclusions herein.

4.2. Results

We eliminated one participant in the trustworthiness condition for responding the same to every item. The final sample thus consisted of 75 participants in the dominance condition (38 male, 37 female; $M_{age} = 33.0$ years, $SD = 10.9$), and 74 participants in the trustworthiness condition (47 male, 27 female; $M_{age} = 33.7$ years, $SD = 11.1$).

4.2.1. Sensitivity correlations

We were primarily interested in the consistency between ratings of upright and inverted faces, and whether that consistency differed for dominance versus trustworthiness. As such, we calculated the Pearson correlation between each participant's ratings of the upright and inverted faces for his or her randomly assigned trait, and subsequently converted the correlation coefficients to Fisher's *z*-scores for analysis (see Judd, Ryan, & Park, 1991). We predicted that the average upright-inverted correlation would be higher in the dominance condition than in the trustworthiness condition, indicating that inversion disrupted judgments of trustworthiness more than dominance. Although participants' correlations between the upright and inverted faces significantly exceeded chance (i.e., $r_z = 0$) for judgments of both traits ($ts \geq 8.40$, $ps < 0.001$, Cohen's $ds \geq 0.97$), we conducted an independent-samples *t*-test to test our primary hypothesis that the correlations for dominance would be greater than those for trustworthiness. Results showed just that: The ratings between the upright and inverted faces correlated significantly more strongly among participants rating dominance ($M = 0.35$, $SD = 0.24$) than among participants rating trustworthiness ($M = 0.24$, $SD = 0.25$), $t(147) = 2.84$, $p = 0.005$, Cohen's $d = 0.47$, 95% CI of difference [0.03, 0.19]. Additional information regarding descriptive statistics and supplementary analyses can be found in the Supplemental Materials.

4.3. Discussion

Face inversion disrupted judgments of trustworthiness significantly more than dominance in male faces. Although participants' ratings of inverted faces significantly correlated with their ratings of upright faces for both traits, they correlated less for trustworthiness than for dominance. These findings accord with recent work showing that perceptions of human-relevant traits (e.g., trustworthiness) may depend more on constellations of facial features than do perceptions of traits conveying abilities shared by humans and other animals (e.g., dominance; Hehman, Flake, et al., 2015; Hugenberg et al., 2016; Todorov et al., 2010).[1] Study 1 provided initial evidence that the consistency between upright and inverted face ratings would be lower for trustworthiness than dominance, but without direct evidence that inversion disrupts these perceptions. We conducted Study 2 to test our hypothesis using a different methodology.

5. Study 2

In Study 1, we interpreted the difference in correspondence between ratings for upright and inverted faces as evidence that inversion affects distinct trait inferences differently. In Study 2, we addressed whether disrupting configural processing by inverting male faces would reduce perceivers' ability to accurately differentiate between high- and low-trustworthiness and high- and low-dominance faces (as pre-rated by a separate sample of participants; see Santos & Young, 2008, for a similar procedure). Based on our findings in Study 1, we predicted that perceivers would perform better on this task for upright than inverted faces. Critically, however, we expected that the trait judged would qualify the magnitude of this difference. Specifically, we predicted that inversion would impair the categorization of high- and low-trustworthiness faces more than the categorization of high- and low-dominance faces, consistent with the greater reliance on configural cues when judging trustworthiness.

5.1. Method

5.1.1. Materials and design

We selected the 12 White faces rated highest and the 12 White faces rated lowest on each of dominance and trustworthiness based on ratings of upright faces by a separate sample of 80 participants (40 rated targets for trustworthiness and 40 for dominance).[2] Interrater agreement was high for both trustworthiness ($\alpha = 0.92$) and dominance ($\alpha = 0.93$). We only used White faces to preclude participants from using race as a cue for trustworthiness or dominance (e.g., Stanley, Sokol-Hessner, Banaji, & Phelps, 2011). Some of the faces ranked at the extremes of both traits, causing overlap between the trait conditions (i.e., they appeared in both stimulus sets), but this should not bias the results because we used a between-subjects design to manipulate whether participants attempted to discriminate between high and low trustworthiness or high and low dominance faces. We manipulated Face Orientation (upright vs. inverted) within subjects.

5.1.2. Participants and procedure

Although we recruited 150 US residents through MTurk, two additional participants completed the study without collecting compensation for a total of 152 participants (81 male, 69 female, 2 unknown; $M_{age} = 34.2$ years, $SD = 11.8$). We excluded two participants from analysis because they provided the same response on every trial, resulting in 150 participants (75 in each between-subjects trait condition). We determined the sample size by following the same parameters as in Study 1, but note that this sample size would provide approximately 85% power to find an effect the size of that observed in Study 1 ($d = 0.47$).

We first informed the participants that the faces that they would see had been pre-rated on various traits, and that their task was to determine whether each face was high or low on a particular trait. We then randomly assigned the participants to judge each face on either trustworthiness or dominance, for which they completed 48 trials (24 upright, 24 inverted). Participants saw each face individually at a self-paced rate and were asked to provide a binary *Trustworthy*-versus-*Untrustworthy*, or *Dominant*-versus-*Not Dominant* response via mouse click, respective to their condition. As in Study 1, the faces appeared in random order within counterbalanced blocks organized by face orientation (i.e., upright, inverted).

5.1.3. Analytic strategy

To measure the participants' ability to discriminate the consensually pre-determined dominance and trustworthiness of the faces, we computed measures of perceptual sensitivity using signal detection analysis separately for the upright and inverted faces (MacMillan & Creelman, 2005). We therefore counted trials on which participants categorized consensus-dominant (consensus-trustworthy) faces as dominant (trustworthy) as hits, and trials on which participants categorized consensus-not dominant (consensus-untrustworthy) faces as dominant (trustworthy) as false alarms. We then used these values to calculate individual discriminability scores (A' and d') for every participant for the upright and inverted faces, which we submitted to a 2 (Face Orientation: upright, inverted) × 2 (Trait Condition: dominance, trustworthiness) ANOVA with repeated-measures on the first factor (see Table 1 for descriptive statistics). By using the pretest consensus ratings as the judgment criterion, we effectively measured the degree of other-other

[1] See the SOM for analysis of the effect of inversion on mean trait ratings.

[2] Incidentally, using the ratings of either the upright or inverted faces from Study 1 would return virtually the same collection of targets (spare 5 targets in total).

Journal of Experimental Social Psychology 74 (2018) 1–7

Table 1
Means and standard deviations for the signal detection data in Study 2.

Trait condition	Upright				Inverted			
	Hits	FAs	A'	d'	Hits	FAs	A'	d'
Trustworthiness	0.67 (0.23)	0.28 (0.21)	0.78 (0.14)	1.20 (0.78)	0.69 (0.24)	0.38 (0.25)	0.72 (0.16)	0.91 (0.72)
Dominance	0.61 (0.23)	0.24 (0.19)	0.75 (0.18)	1.13 (0.87)	0.59 (0.23)	0.24 (0.18)	0.75 (0.15)	1.05 (0.78)

Note. Standard deviations in parentheses. FAs = False Alarms.

agreement between the perceivers in this task and the perceivers in the previous rating task (Funder, 2012).

5.2. Results

Preliminary analyses indicated that participants' discriminated the faces' dominance and trustworthiness significantly better than chance for both the upright ($M_{A'} = 0.76$, $SD = 0.16$), $t(149) = 20.02$, $p < 0.001$, Cohen's $d = 1.64$, and inverted faces ($M_{A'} = 0.74$, $SD = 0.15$), $t(149) = 18.69$, $p < 0.001$, Cohen's $d = 1.53$. In other words, participants' categorizations matched the consensus judgments from the pretest: they identified the consensus-trustworthy and consensus-untrustworthy faces as trustworthy and untrustworthy, respectively; and identified the consensus-dominant and consensus-not dominant faces as dominant and not dominant, respectively.

More important, our primary analysis showed that the two-way ANOVA yielded no main effect of Trait Condition, $F(1, 148) = 0.01$, $p = 0.91$, $\eta^2_{\text{partial}} < 0.001$, but did show a main effect of Face Orientation, $F(1, 148) = 5.28$, $p = 0.02$, $\eta^2_{\text{partial}} = 0.034$, such that participants discriminated the dominance and trustworthiness of the faces better when upright ($M_{A'} = 0.76$, $SD = 0.16$) than when inverted ($M_{A'} = 0.74$, $SD = 0.15$).

The predicted Trait Condition × Face Orientation interaction qualified this effect, $F(1, 148) = 4.21$, $p = 0.04$, $\eta^2_{\text{partial}} = 0.028$ (see Fig. 1). Planned comparisons showed that inversion ($M_A' = 0.72$, $SD = 0.16$) significantly reduced participants' sensitivity to differences in the faces' trustworthiness relative to upright presentation ($M_A' = 0.78$, $SD = 0.14$), $t(74) = 2.96$, $p = 0.004$, 95% CI of difference [0.02, 0.09], Cohen's $d = 0.34$. Inversion ($M_A' = 0.75$, $SD = 0.15$) did not significantly reduce sensitivity to differences in dominance relative to upright presentation ($M_A' = 0.75$, $SD = 0.18$), $t(74) = 0.18$, $p = 0.86$, 95% CI of difference [−0.03, 0.04], Cohen's $d = 0.02$.

5.2.1. Additional tests

Although the primary signal detection measure was *A*', we analyzed the data using *d*' as well. In this analysis, the interaction between Trait Condition and Face Orientation was only marginally significant, $F(1, 148) = 3.23$, $p = .07$, $\eta^2_{\text{partial}} = 0.021$. The planned contrasts, however, still showed that inversion ($M_d' = 0.91$, $SD = 0.72$) significantly reduced agreement for trustworthiness categorizations relative to upright presentation ($M_d' = 1.20$, $SD = 0.78$), $t(74) = 3.75$, $p < 0.001$, 95% CI of difference [0.14, 0.45], Cohen's $d = 0.43$. Inversion ($M_d' = 1.05$, $SD = 0.78$) did not significantly reduce sensitivity to differences in dominance relative to upright presentation ($M_d' = 1.13$, $SD = 0.87$), $t(74) = 2.02$, $p = 0.29$, 95% CI of difference [−0.08, 0.25], Cohen's $d = 0.12$.

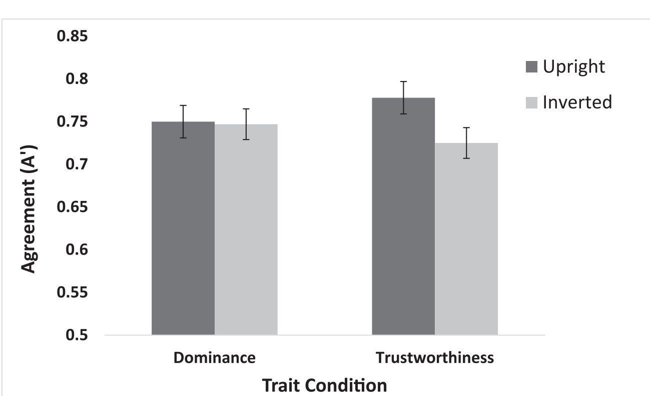

Fig. 1. Means and standard errors for agreement with consensus trait categorizations in Study 2.

5.3. Discussion

Participants in Study 2 classified faces' dominance about the same whether they were upright or inverted. Their discrimination of the faces' trustworthiness suffered significantly when inverted, however. These data support our hypothesis that configural processing plays a greater role in inferences of human-relevant traits (e.g., trustworthiness) than in inferences of traits shared by humans and animals (e.g., dominance). They also complement and conceptually replicate the within-subject correlations in Study 1 by showing that the consensus dominance and trustworthiness ratings of separate participants in a pretest corresponded to the current participants' judgments of the traits. Although the results of the main analysis were conventionally significant, an analysis using *d*' as the signal detection sensitivity measure yielded only marginally significant results. As such, these results should be treated with caution and call for replication. We undertook this effort in Study 3.

6. Study 3

The results of Studies 1 and 2 suggested that face inversion affects inferences of trustworthiness and dominance differently. To confirm this, we conducted a pre-registered replication of Study 2 via the Open Science Framework (https://osf.io/zjevk/) for two reasons. First, we acknowledge the need for increased reproducibility in psychology research (Open Science Collaboration, 2015). This may be especially important when existing evidence relies on *p*-values very near the criterion for significance (Simonsohn, Nelson, & Simmons, 2014), as in our Study 2. Second, we wanted to improve the methodology of Study 2 by doubling the number of target faces seen by each participant (see Westfall, Judd, & Kenny, 2015, for a convincing argument advocating the need to increase stimulus sample size) and by cropping all of the faces to (a) exclude the neck and shoulders, and (b) eliminate excess hair. The latter ensured that participants inferred trustworthiness and dominance based on the *faces* versus extrafacial features.

We pre-registered the hypothesis that trait condition would interact with face orientation, such that inversion would disrupt the categorization of trustworthiness more than the categorization of dominance, measured using signal detection analyses. Although we predicted that categorizations of upright and inverted faces would significantly differ for participants judging trustworthiness, we registered no specific hypothesis regarding the dominance judgments. However, given that previous research (e.g., Santos & Young, 2008) has shown that perceptions of multiple traits are at least somewhat affected by inversion, we did plan to test for such a difference with this higher-powered study.

6.1. Method

6.1.1. Materials and design

We selected faces from the updated Chicago Face Database (version 2.0; Ma et al., 2015). We resized the images to 450 × 316 pixels (72 pixels/in.), and cropped each one to remove the shoulders, neck, and stray hair. We asked 30 participants to rate all of the 93 White male faces on trustworthiness and 30 separate participants to rate them on dominance. We then selected the 24 faces with the highest and lowest average scores on each trait. The study design was thus identical to Study 2, except that participants viewed twice as many faces in each block (48 upright and 48 inverted). We once again manipulated whether participants judged trustworthiness or dominance.

6.1.2. Participants and procedure

We recruited 300 participants from MTurk, as specified in our preregistration plan. We determined that this sample size would provide approximately 99% power to detect an effect size as large as the interaction observed in Study 2 in a design with twice as many target faces. As in Study 2, two extra participants completed the study without claiming compensation. Per our analysis plan, we included these participants. Of the 302 participants, seven provided the same response for every trial in at least one experimental block, rendering their data unusable. This left us with a final sample of 295 participants (131 male, 162 female, 2 unknown; $M_{\text{age}} = 35.1$ years, $SD = 11.9$), with 132 randomly assigned to the trustworthiness condition and 163 randomly assigned to the dominance condition.

6.2. Results

6.2.1. Pre-registered hypothesis test

Once again, signal detection analyses showed that participants' discrimination of the faces' dominance and trustworthiness corresponded to the prior raters' consensus judgments for both the upright ($M_{A'} = 0.81$, $SD = 0.10$), $t(295) = 50.96$, $p < 0.001$, Cohen's $d = 2.97$, and inverted faces ($M_{A'} = 0.76$, $SD = 0.12$), $t(295) = 37.54$, $p < 0.001$, Cohen's $d = 2.19$; see Table 2 for descriptive statistics. Having confirmed this, we proceeded to implement our primary analytic strategy, as outlined in Study 2.

In keeping with our plan, we subjected the A' scores to a 2 (Face Orientation: upright, inverted) × 2 (Trait Condition: trustworthiness, dominance) ANOVA with repeated-measures on the first factor. As predicted, we observed a main effect of Face Orientation, $F(1, 293) = 76.28$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.21$, showing higher sensitivity for trait differences in the upright faces ($M_A' = .81$, $SD = .10$) than in the inverted faces ($M_A' = 0.76$, $SD = 0.12$), 95% CI of difference [0.04, 0.06]. Next, confirming the focal pre-registered hypothesis, we observed an interaction between Face Orientation and Trait Condition, $F(1, 293) = 12.49$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.041$. A planned comparison showed that inversion significantly reduced sensitivity for trustworthiness, $t(131) = 7.28$, $p < 0.001$, 95% of difference [0.05, 0.09], Cohen's $d = 0.63$, such that participants discriminated the faces better when upright ($M_A' = 0.78$, $SD = 0.11$) than inverted ($M_A' = 0.71$, $SD = 0.12$). Based on the lack of difference between upright and inverted targets for dominance categorizations in Study 2, we made no prediction about whether such a difference would emerge in Study 3. Yet, a planned comparison showed that inversion did significantly reduce the participants' sensitivity to dominance differences, $t(162) = 4.43$, $p < 0.001$, 95% of difference [0.02, 0.04], Cohen's $d = 0.35$. Although this result is not entirely consistent with Study 2, it clearly shows that inversion disrupted perceptions of trustworthiness more than perceptions of dominance.

6.2.2. Additional tests

We also observed an unhypothesized main effect of Trait Condition, $F(1, 293) = 38.62$, $p < 0.001$, $\eta^2_{\text{partial}} = 0.12$, such that the participants' judgments matched the prior raters' consensus more for dominance ($M_{A'} = 0.81$, $SD = 0.08$) than for trustworthiness ($M_{A'} = 0.75$, $SD = 0.10$), 95% CI [0.05, 0.09]. Although we did not specifically predict this difference, it is not surprising, given the tendency for inversion to impair judgments much more strongly for trustworthiness than for dominance.

We once again conducted an additional analysis to determine whether the primary results would be confirmed using d' as our signal detection measure of sensitivity. Whereas the focal interaction was rendered marginally significant in Study 2, here it remained significant, $F(1, 293) = 8.88$, $p = 0.003$, $\eta^2_{\text{partial}} = 0.029$. Planned comparisons showed that inversion significantly reduced participants' ability to discriminate both trustworthiness, $t(131) = 8.33$, $p < 0.001$, 95% of difference [0.33, 0.53], Cohen's $d = 0.74$, and dominance, $t(162) = 4.91$, $p < 0.001$, 95% of difference [0.13, 0.31], Cohen's $d = 0.39$.

6.2.3. Meta-analysis

To summarize our findings across the first three studies, we meta-analytically aggregated our effect sizes to derive a reliable estimate of the overall size of the effects observed. We first converted the focal effect size in each study to r, then Fisher-transformed these values (see Table 3). The mean effect size was significant and of small-to-medium size based on Cohen's (1988) criteria.

7. Study 4

The results of the first three studies showed consistent support for our hypothesis that inversion disrupts perceptions of trustworthiness more than it does perceptions of dominance. Thus far, however, we lack direct evidence implicating humanness. We therefore returned to the method used in Study 1 with a more direct focus on humanness. First, we asked participants to rate the extent to which trustworthiness, dominance, and a number of other traits are uniquely human (vs. shared with other animals) in Study 4A. Then, we asked participants to rate upright and inverted male faces for the extent to which they appeared to be human in Study 4B. As in Study 1, we calculated a sensitivity correlation between each participant's ratings of the faces when upright and inverted. This allowed us to additionally test whether inversion disrupted humanness judgments even more than it did traits only indirectly linked to humanness (i.e., dominance and trustworthiness), by comparing these sensitivity correlations to those from Study 1. We predicted that participants would rate trustworthiness as more

Table 2
Means and standard deviations for the signal detection data in Study 3.

Trait condition	Upright				Inverted			
	Hits	FAs	A'	d'	Hits	FAs	A'	d'
Trustworthiness	0.65 (0.22)	0.28 (0.21)	0.78 (0.11)	1.22 (0.60)	0.63 (0.22)	0.37 (0.21)	0.71 (0.12)	0.79 (0.54)
Dominance	0.64 (0.17)	0.16 (0.15)	0.83 (0.09)	1.55 (0.63)	0.59 (0.18)	0.18 (0.15)	0.80 (0.10)	1.33 (0.63)

Note. Standard deviations in parentheses. FAs = False Alarms.

Table 3
Effect sizes and overall meta-analytic estimate for Studies 1–3.

	r	Z_r	*SE*	*LL*	*UL*	Z	*p*
Study 1	0.23	0.23	0.08	0.07	0.38	2.80	0.005
Study 2	0.17	0.17	0.08	0.01	0.32	2.04	0.042
Study 3	0.20	0.21	0.06	0.09	0.31	3.50	< 0.001
Overall		0.20	0.04	0.12	0.28	4.90	< 0.001

Note. Z_r = Fisher-transformed *r* effect size, *SE* = standard error of effect size, *LL* = lower limit of 95% confidence interval around the effect size, *UL* = upper limit of 95% confidence interval around the effect size, *Z* = one-sample test statistic for the effect size.

human than dominance and that inversion would disrupt humanness judgments more than either dominance or trustworthiness judgments.

7.1. Study 4A

We recruited 36 participants (10 female, 26 male; $M_{age} = 29.5$, $SD = 7.8$) to rate 10 traits on the extent to which they apply exclusively to humans, from 1 (*Applies Equally to Humans and Non-Humans*) to 9 (*Applies Uniquely to Humans*). This sample size provided > 90% power to detect a medium-large difference (Cohen's $d = 0.6$) between the critical traits of interest. Critically, dominance and trustworthiness appeared among the list of traits, with the other traits serving as fillers (see Supplemental Materials for full list).

Because each participant rated both dominance and trustworthiness, we subjected their ratings to a paired-samples *t*-test. Consistent with our hypothesis, this showed that they considered trustworthiness ($M = 6.06$, $SD = 1.66$) as significantly more unique to humans than dominance ($M = 3.58$, $SD = 2.18$), $t(35) = 5.80$, $p < 0.001$, Cohen's $d = 0.98$.

7.2. Study 4B

7.2.1. Method

Emulating the design and sample size of Study 1, we recruited 75 American MTurk Workers to rate the humanness of each of the same 144 images of 72 upright and inverted faces in random order within blocks counterbalanced by orientation. To promote variability in their ratings, we instructed the participants that the faces had been very subtly morphed with an artificial face such that none were fully human, and that they were to assess each face on how “humanlike” it was, from 1 (*Not at all humanlike*) to 7 (*Very humanlike*). In fact, the targets they viewed were exactly the same as those viewed by the participants in Study 1. Although the participants responded to this manipulation by giving varied responses, we did eliminate data from seven participants who gave the maximum value of 7 to every target (final $N = 68$; 42 male, 26 female; $M_{age} = 33.1$ years, $SD = 9.9$).

7.2.2. Results and discussion

As in Study 1, we calculated the Pearson correlation between each participant's ratings of the upright and inverted faces, converting it to a Fisher's *z* for analysis (Judd et al., 1991). The mean score again exceeded 0, $M_z = 0.11$, $SD = 0.21$, $p < 0.001$, 95% CI [0.06, 0.16], $d = 0.53$, indicating that the humanness ratings of the upright and inverted faces significantly correlated across participants. Critically, however, the results of a one-way ANOVA including these ratings and the two additional trait conditions from Study 1 showed that the relation between upright and inverted faces for humanness ratings significantly differed from the relations between upright and inverted faces for dominance and trustworthiness ratings in Study 1, $F(2, 214) = 18.90$, $p < 0.001$, $\eta^2_{partial} = 0.15$.

Planned contrasts (Bonferroni-corrected $\alpha = 0.025$) showed that the upright-inverted correlations were much stronger for perceptions of dominance than humanness, $t(214) = 6.15$, $p < 0.001$, Cohen's $d = 1.06$, and for perceptions of trustworthiness than humanness, $t(214) = 3.25$, $p = 0.001$, Cohen's $d = 0.56$. Although we interpret these differences cautiously because the data come from separate studies collected at different times without random assignment across conditions, they do lend further support to our hypothesis that perceptions of humanity rely on configural processing and thus suffer when configural processing is undermined.

8. General discussion

Across four studies, we have presented novel evidence that configural processing distinctly affects inferences of dominance and trustworthiness from male faces. Inverting faces interferes with configural processing and, thus, interfered with inferences of trustworthiness more than inferences of dominance. This supports the hypothesis that trustworthiness judgments rely more heavily on configural facial cues than do dominance judgments for male faces. Judgments of trustworthiness also appear to be supported by some featural cues as well, however. Whereas inversion did disrupt trustworthiness judgments, it did not do so strongly enough to fully ablate the consistency between upright and inverted face judgments. Thus, perceivers appear to rely on both featural and configural processing in the extraction of facial trustworthiness, but extracting dominance seems to rely exclusively (or at least to a much greater extent) on featural cues.

Notably, we observed this pattern despite using the same photos for the upright and inverted images, thereby holding constant everything but face orientation (cf. Santos & Young, 2008). Unlike past work in which manipulations of viewing angle resulted in the display of different parts of the face and head (Rule, Ambady, & Adams, 2009) or that displayed different poses of the same identity (Hehman, Flake, et al., 2015), we only manipulated vertical rotation (i.e., inversion). Doing so preserves all of the face's features, only affecting perceivers' ability to process it configurally (Rossion & Gauthier, 2002; Valentine, 1988). The present work therefore provides a key methodological advantage not present in some of the past research on related topics.

Beyond the specific demonstration that trustworthiness inferences rely more heavily on configural processing than do dominance inferences, the current research provides valuable links to other literatures as well. First, and perhaps most obvious, our findings helps to further wed social psychological research on person perception to cognitive science research on face perception. Although a number of studies have made similar links between person perception and face processing in recent years (e.g., Derks, Stedehouder, & Ito, 2015; Dotsch, Wigboldus, & van Knippenberg, 2011; Hackel, Looser, & Van Bavel, 2014; Ratner & Amodio, 2013; Van Bavel, Packer, & Cunningham, 2011; see also Adams, Ambady, Nakayama, & Shimojo, 2010), understanding the perceptual processes that undergird social judgments remains an important goal. Second, and more specifically, the current research also connects recent findings on configural face processing to motivated social perception. Indeed, configural face processing both facilitates the perception that others are human (Hugenberg et al., 2016) and supports perceptions and cognitions related to ingroup members and other targets who follow ingroup norms (Fincher & Tetlock, 2016; Hugenberg & Corneille, 2009). It therefore seems logical that extracting cues to trustworthiness relies heavily on configural processing because trustworthiness uniquely applies to humans and signals the willingness to adhere to consensual norms. Future research more directly testing this mechanism would thus be an interesting, and likely fruitful, avenue.

One notable limitation of this research is that we only tested male faces. We did this for methodological and theoretical reasons. Methodologically, we not only wanted to limit the number of ratings required of each participant but were concerned that adding female faces might bias responses, such that gender's salience might outweigh the subtler variability in dominance and trustworthiness between targets (e.g., Biernat & Manis, 1994). Theoretically, recent research has suggested that dominance may be represented differently for male and female faces (Sutherland et al., 2016), potentially leading to different processes best examined in separate work. Indeed, we report a pre-registered attempt to replicate Study 1 using female faces in the Supplemental Materials that did not confirm the central theoretical hypothesis of this paper, lending credence to this concern. There, the upright-inverted sensitivity correlation for dominance ($M = 0.28$, $SD = 0.29$) was similar to that for trustworthiness ($M = 0.32$, $SD = 0.27$), $t(196) = 1.12$, $p = 0.26$, $d = -0.16$. Thus, more research will certainly be needed to determine the causes and impact of this sex difference in how face perception relates to perceptions of humanness with the present studies merely supplying a first step on this path.

In conclusion, the current data offer novel evidence that face inversion, known to hamper configural encoding (Yin, 1969), selectively disrupts judgments of trustworthiness more than dominance in male faces. Whereas dominance and trustworthiness are among the most fundamental cues rapidly extracted from faces, they are not equally sensitive to visuospatial orientation. Rather, facial trustworthiness in male faces seems to rely more on configural processing than does facial dominance.

Open practices

This article earned Open Data and Preregistration badges for transparent practices. Data and registrations for the experiments are available at https://osf.io/t26xq.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jesp.2017.07.007.

References

Adams, R. B., Ambady, N., Nakayama, K., & Shimojo, S. (Eds.), (2010). *The science of social vision*. New York, NY: Oxford University Press.

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, *6*, 269–278.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, *66*, 5–20.

Burton, C. M., & Rule, N. O. (2013). Judgments of height from faces are informed by dominance and facial maturity. *Social Cognition*, *31*(6), 672–685.

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science*, *20*, 1194–1198.

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.

Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum and Associates.

Dahl, C. D., Rasch, M. J., & Chen, C. C. (2014). The other-race and other-species effects in face perception – A subordinate-level analysis. *Frontiers in Psychology*, *5*, 10–3389.

Derks, B., Stedehouder, J., & Ito, T. A. (2015). Social identity modifies face perception: An ERP study of social categorization. *Social Cognitive and Affective Neuroscience*, *10*, 672–679.

Dotsch, R., Wigboldus, D. H., & van Knippenberg, A. (2011). Biased allocation of faces to social categories. *Journal of Personality and Social Psychology*, *100*, 999–1014.

Fincher, K. M., & Tetlock, P. E. (2016). Perceptual dehumanization of faces is activated by norm violations and facilitates norm enforcement. *Journal of Experimental Psychology: General*, *145*, 131–146.

Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, *21*, 177–182.

Hackel, L. M., Looser, C. E., & Van Bavel, J. J. (2014). Group membership alters the threshold for mind perception: The role of social identity, collective identification, and intergroup threat. *Journal of Experimental Social Psychology*, *52*, 15–23.

Hehman, E., Carpinella, C. M., Johnson, K. L., Leitner, J. B., & Freeman, J. B. (2014). Early processing of gendered facial cues predicts the electoral success of female politicians. *Social Psychological and Personality Science*, *5*, 815–824.

Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and dynamic facial cues differentially affect the consistency of social evaluations. *Personality and Social Psychology Bulletin*, *41*, 1123–1134.

Hehman, E., Leitner, J. B., Deegan, M. P., & Gaertner, S. L. (2015). Picking teams: When dominant facial structure is preferred. *Journal of Experimental Social Psychology*, *59*, 51–59.

Hugenberg, K., & Corneille, O. (2009). Holistic processing is tuned for in-group faces. *Cognitive Science*, *33*, 1173–1181.

Hugenberg, K., Young, S. G., Rydell, R. J., Almaraz, S., Stanko, K. A., See, P. E., & Wilson, J. P. (2016). The face of humanity: Configural face processing influences ascriptions of humanness. *Social Psychological and Personality Science*, *7*, 167–175.

Judd, C. M., Ryan, C. S., & Park, B. (1991). Accuracy in the judgment of in-group and out-group variability. *Journal of Personality and Social Psychology*, *61*, 366–379.

Kramer, R. S., King, J. E., & Ward, R. (2011). Identifying personality from the static, nonexpressive face in humans and chimpanzees: Evidence of a shared system for signaling personality. *Evolution and Human Behavior*, *32*, 179–185.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*, 1–14.

MacMillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.

Madera, J. M., & Hebl, M. R. (2012). Discrimination against facially stigmatized applicants in interviews: An eye-tracking and face-to-face investigation. *Journal of Applied Psychology*, *97*, 317–330.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*, 255–260.

Mondloch, C. J., Maurer, D., & Ahola, S. (2006). Becoming a face expert. *Psychological Science*, *17*, 930–934.

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.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*, aac4716.

Ratner, K. G., & Amodio, D. M. (2013). Seeing "us vs. them": Minimal group effects on the neural encoding of faces. *Journal of Experimental Social Psychology*, *49*, 298–301.

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, *1*, 63–75.

Rule, N. O., Ambady, N., & Adams, R. B., Jr. (2009). Personality in perspective: Judgmental consistency across orientations of the face. *Perception*, *38*, 1688–1699.

Santos, I. M., & Young, A. W. (2008). Effects of inversion and negation on social inferences from faces. *Perception*, *37*, 1061–1078.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*, 534–547.

Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences*, *108*(19), 7710–7715.

Sutherland, C. A., Oldmeadow, J. A., Santos, I. M., Towler, J., Burt, D. M., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*, 105–118.

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

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, *46*, 225–245.

Todorov, A. T., Loehr, V. G., & Oosterhof, N. N. (2010). The obligatory nature of holistic processing of faces in social judgments. *Perception*, *39*, 514–532.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, *79*, 471–491.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: Evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience*, *23*, 3343–3354.

van't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition*, *108*, 796–803.

Westfall, J., Judd, C. M., & Kenny, D. A. (2015). Replicating studies in which samples of participants respond to samples of stimuli. *Perspectives on Psychological Science*, *10*, 390–399.

Willis, J., & Todorov, A. (2006). First impressions making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*, 592–598.

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

Wilson, J. P., & Rule, N. O. (2016). Hypothetical sentencing decisions are associated with actual capital punishment outcomes: The role of facial trustworthiness. *Social Psychological and Personality Science*, *7*, 331–338.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145.

Young, S., & Hugenberg, K. (2010). Mere social categorization modulates identification of facial expressions of emotion. *Journal of Personality and Social Psychology*, *99*, 964–977.

Zebrowitz, L. A., Montepare, J. M., & Lee, H. K. (1993). They don't all look alike: Individuated impressions of other racial groups. *Journal of Personality and Social Psychology*, *65*, 85–101.