Personality and Social Psychology Bulletin
2016, Vol. 42(12) 1603–1616

DOI: 10.1177/0146167216666266
pspb.sagepub.com
SAGE

Physical Strength as a Cue to Dominance: A Data-Driven Approach

Hugo Toscano[1], Thomas W. Schubert[2], Ron Dotsch[3], Virginia Falvello[4], and Alexander Todorov[4]

Abstract

We investigate both similarities and differences between dominance and strength judgments using a data-driven approach. First, we created statistical face shape models of judgments of both dominance and physical strength. The resulting faces representing dominance and strength were highly similar, and participants were at chance in discriminating faces generated by the two models. Second, although the models are highly correlated, it is possible to create a model that captures their differences. This model generates faces that vary from dominant-yet-physically weak to nondominant-yet-physically strong. Participants were able to identify the difference in strength between the physically strong-yet-nondominant faces and the physically weak-yet-dominant faces. However, this was not the case for identifying dominance. These results suggest that representations of social dominance and physical strength are highly similar, and that strength is used as a cue for dominance more than dominance is used as a cue for strength.

Keywords

face perception, dominance, physical strength, data-driven methods

Received August 19, 2015; revision accepted July 29, 2016

People rapidly form impressions about others based on seeing their faces (e.g., Bar, Neta, & Linz, 2006; Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006). In addition to cues of interpersonal warmth and trustworthiness, people are particularly tuned to features related to dominance (Jones et al., 2010; Oosterhof & Todorov, 2008; Watkins et al., 2010; Watkins, Jones, & DeBruine, 2010). In an analysis of the personality attributes most frequently inferred from faces, dominance emerges as one of the central dimensions of face evaluation (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Todorov, Said, Engell, & Oosterhof, 2008).

Judgments of dominance affect social behavior in predictable ways: In times of war, people are more likely to vote for dominant looking politicians (Little, Burriss, Jones, & Roberts, 2007), cadets with dominant facial appearance are more likely to reach higher military ranks (Mueller & Mazur, 1996), and dominant looking males have higher reproductive success (Rhodes, Simmons, & Peters, 2005). Looking dominant can also have its pitfalls in contexts involving threat: In the courtroom, more dominant looking defendants receive higher penalties than nondominant looking defendants if plaintiffs are baby-faced (Zebrowitz & McDonald, 1991).

Some research suggests that one likely determinant of being judged as dominant is appearing to be physically strong, because strength is important for prevailing in antagonistic physical contests as well as in securing resources (Lukaszewski, Simmons, Anderson, & Roney, 2016; Sell, Cosmides, et al., 2009; Toscano, Schubert, & Sell, 2014; Windhager, Schaefer, & Fink, 2011). Muscularity is particularly important for attaining high rank in hierarchies based on dominance rather than prestige (Blaker & van Vugt, 2014). As a result, diagnostic cues to what a dominant face looks like may be similar to diagnostic cues to what a strong face looks like. Here, we tested whether the representations of facial dominance and facial strength are indeed *similar* by using a data-driven approach (see Todorov, Dotsch, Wigboldus, & Said, 2011). However, because other attributes, such as abilities to build alliances, have long been important for attaining high social dominance (Cheng, Tracy, Foulsham, Kingstone, & Henrich, 2013), we also intend to capture *differences* between representations of facial dominance and facial strength. In sum, the goal of this research is

[1]Instituto Universitário de Lisboa (ISCTE-IUL), Centro de Investigação e Intervenção Social, Portugal
[2]University of Oslo, Norway
[3]Utrecht University, The Netherlands
[4]Princeton University, NJ, USA

Corresponding Author:
Hugo Toscano, Centro de Investigação e Intervenção Social, Instituto Universitário de Lisboa (ISCTE-IUL), Av. Forças Armadas, Ed. ISCTE, Lisboa, 1649-026, Portugal.
Email: toscano.hugo@gmail.com

to understand similarities and differences between representations of faces of dominant and strong people.

Dominance and Bodily Strength

In ancestral humans, bodily strength was certainly associated with greater resource-holding potential (Parker, 1974; Sell, Cosmides, et al., 2009). Stronger persons would have had more ability to inflict harm, to withhold or secure resources. In contexts of repeated intergroup fighting during human evolution, physical strength may have played an important role in conflict resolution and building social hierarchies together with other strategies such as alliances (Van Vugt, De Cremer, & Janssen, 2007). Puts (2010) argues that fighting contests between men were the central process of sexual selection in the evolutionary history of men. Indeed, greater physical strength increases the chances of reproductive success (Von Rueden, Gurven, & Kaplan, 2011).

Given the importance of bodily strength to inflict costs on others and to increase the resource-holding potential, it is likely that humans have evolved the capacity to assess bodily strength from visual cues to decide whether to enter or avoid a fight, and thus persevere or defer in conflicts. Such judgments can obviously be made best from direct inspection of the upper body (its strength is particularly important for both hand-to-hand combat and body-powered weapons—Lassek & Gaulin, 2009), but Sell, Cosmides, et al. (2009) argued and empirically confirmed that they can also be reliably made based on the face alone, especially for male targets. In line with this work, we assume that humans have evolved specialized cognitive capacities to judge formidability from the face, and spontaneously engage in such judgments.

As an inherently social species, humans also evolved the capacity to form various types of social relations, including social hierarchies (Fiske, 1991, 2000). Given its importance for resource holding potential and formidability, bodily strength was likely crucial for setting up social hierarchies, which are characterized by some individuals having more access to resources than others. Thus, social status may have been in part derived from physical strength over a considerable time of our evolutionary past (von Rueden, Gurven, & Kaplan, 2008). Even today, boys are ranked by their peers as having higher status when they are involved in more play fighting (e.g., Pellegrini, 1995).

Of course, in our current social reality, status does not exclusively depend on physical strength, but is, and probably always has been, determined by many other factors (e.g., cognitive abilities, expert knowledge, social networks and support, material capital; Cheng et al., 2013). Hierarchies based on dominance exist alongside hierarchies based on prestige (the ability to share relevant skills and knowledge), and often they overlap (Blaker & van Vugt, 2014). Nevertheless, initial adaptations that *Homo sapiens* and its ancestors evolved to judge bodily strength may still contribute to judgments of dominance (see Smuts, Cheney, Seyfarth, Struhsaker, & Wrangham, 1987). Supporting evidence comes from Fink, Neave, and Seydel (2007), who showed that judgments of men's dominance are correlated with these men's actual handgrip strength, indicating that physical strength is recognized from perceiving the body and incorporated into dominance judgments.

Inferring dominance from bodily strength seems to have some validity: Gallup, O'Brien, White, and Wilson (2010) found high correlations between actual handgrip strength and socially dominant behaviors. Likewise, stronger individuals are more likely to be aggressive (but see Isen, McGue, & Iacono, 2015) and feel entitled (Sell, Tooby, & Cosmides, 2009). Some studies also show that high levels of circulating testosterone are correlated with both dominance and strength (Penton-Voak & Chen, 2004). Together, this suggests that inferences of dominance from physical strength may have a kernel of truth.

In line with this association of physical strength and both dominant behavior and perceptions of dominance, past research has identified similarities between facial cues indicating strength and facial cues indicating dominance. For instance, the facial width-to-height ratio (fWHR) tends to be associated with physical aggression (Carré & McCormick, 2008; Geniole, Keyes, Carré, & McCormick, 2014; Goetz et al., 2013; Haselhuhn, Ormiston, & Wong, 2015; but see Deaner, Goetz, Shattuck, & Schnotala, 2012) and physical strength (Hehman, Flake, & Freeman, 2015), as well as with self-reported dominance (Lefevre, Etchells, Howell, Clark, & Penton-Voak, 2014).

Using handgrip strength as a proxy of actual physical strength (Rantanen et al., 1999; Wind, Takken, Helders, & Engelbert, 2010), Windhager and colleagues (2011) derived the facial shape associated with actual strength in men. The derived shape was found to be similar to a shape created from dominance judgments of the faces of the same men. Consistent with this, Toscano et al. (2014) found that several facial features predicted both physical strength and dominance judgments, in particular eyebrow height, eye and chin length, and the widths of the nose and the mouth. As a result, individuals who were judged as socially dominant were also judged as physically strong, even if those judgments were made by different perceivers.

In sum, due to an evolutionary history in which strength and social dominance have been associated and due to a social environment where physical strength continues to be associated with actual social dominance, diagnostic cues to what a dominant face looks like should be similar to diagnostic cues to what a strong face looks like. Here, we build data-driven computational models of impressions of dominance and physical strength to identify the diagnostic cues for these impressions.

Modeling Faces

When investigating how traits are represented in the human face, it is tempting to focus on specific features, such as small eyes, large chins, or low eyebrows. Changes in those features would lead in most cases to different perceptions of dominance and strength. However, the feature-by-feature approach has limitations because it is unable to identify the complete constellations of features that underlie the perceptual representation of face traits (Todorov et al., 2011). Features are perceived in the context of the whole face, and face perception is inherently holistic.

An alternative approach is to identify the changes in holistic face shape critical for trait judgments without explicitly manipulating any features in the face (Dotsch & Todorov, 2012; Jack, Caldara, & Schyns, 2012; Oosterhof & Todorov, 2008; Todorov, Dotsch, Porter, Oosterhof, & Falvello, 2013; Todorov & Oosterhof, 2011; Walker, Jiang, Vetter, & Sczesny, 2011; Walker & Vetter, 2009). This approach is based on the notion of the face space (Valentine, 1991; see also Todorov et al., 2011), where each face is represented as a point in a multidimensional face space (Blanz & Vetter, 1999, 2003). Typically, to obtain the dimensions of the face space, differences between multiple faces are analyzed and reduced to orthogonal dimensions. For example, Blanz and Vetter (1999, 2003) laser scanned real faces in three dimensions (3D) and quantified the shape differences between faces. Because these differences are correlated, data reduction techniques such as principal components analysis can be used to represent these differences in a multidimensional space. In this space, each face is represented as a set of coordinates on the dimensions. Moreover, the statistical face space allows for the generation of novel faces in the space. Given that each face is fully determined by a set of coordinates in this statistical space, it is possible to model any trait judgment (see Todorov et al., 2013). Specifically, a random sample of faces is generated, and then participants are asked to rate the faces on the trait of interest (Oosterhof & Todorov, 2008; Todorov & Oosterhof, 2011). As long as the judgments are reliable, they can be modeled in the statistical face space. Akin to a regression approach, the mean trait judgment is modeled as a function of the coordinates of the randomly sampled faces. The resulting model is a new vector in the face space that captures the maximal variation of the judgments (Oosterhof & Todorov, 2008). Moreover, this model allows for the visualization of all changes in the face that are important for the respective trait judgment.

With this approach, one can visualize how the representations of extremely nondominant faces differ from those of extremely dominant faces. Using this approach, Oosterhof and Todorov (2008) built facial shape models of judgments of trustworthiness, threat, and dominance. The set of judgments has been extended recently and reflectance models of changes in skin and texture have been added (Todorov & Oosterhof, 2011). These models of trait judgments have been validated and shown to capture unique variance specific to each judgment (Todorov et al., 2013).

Modeling Physical Strength and Dominance in the Human Face

Trait judgments from faces are often highly correlated (Oosterhof & Todorov, 2008). For instance, a person who is considered attractive will also be considered likable. Important for the current purposes, the modeling approach allows for the statistical separation of even highly correlated dimensions. Shared variance of trait dimensions can be removed through the subtraction of one dimension from the other. For instance, Todorov et al. (2013) were able to create faces that were trustworthy, but not attractive, and faces that were attractive, but not trustworthy.

Following this approach, in the current article, we created a model that investigates the differences between physical strength and dominance. This works in both directions: We can maximize what differentiates dominance from physical strength and, consequently, visualize identities that are dominant but not strong. Alternatively, we can maximize what differentiates strength from dominance, and visualize identities that are strong, but not dominant.

Despite the introductory arguments about the associations between strength and status, as well as the correlations between physical strength and perceptions of dominance, it is reasonable to assume that perceptions of dominance cannot be fully explained by perceptions of physical strength. For instance, Toscano et al. (2014) investigated the correlation between a man's actual physical strength and how strong and how dominant he was perceived to be based on the face only. They found that the actual strength and judgments of strength were correlated after controlling for dominance judgments, but dominance judgments and actual strength were not correlated after controlling for physical strength judgments. Moreover, social dominance may have always been and is surely now, in modern democratic societies, determined by factors other than physical strength, such as cognitive and emotional capabilities (see Henrich & Gil-White, 2001).

With this in mind, we used face-space-based data-driven methods to disentangle the highly correlated dimensions of social dominance and physical strength. The use of these methods allowed us to maximize the visual differences between representations of dominance and representations of physical strength. In other words, we investigate which faces are judged to belong to very dominant but physically weak individuals, and which faces are judged to belong to very nondominant but physically strong individuals. Furthermore, the resulting differences in dimensions allow us to test whether physical strength is used as a cue for judgments of social dominance but not vice versa.

In sum, our main objective is to create models of judgments of physical strength and dominance, and then to create a model that differentiates these two models, using a data-driven approach (Dotsch & Todorov, 2012; Todorov et al., 2013; Todorov & Oosterhof, 2011). To the best of our knowledge, strength, in contrast to dominance, has not been modeled in this way before. This will enable us to visualize both the similarities and differences between physical strength and dominance.

In addition, one of the advantages of this approach is the ability to generate new identities for each model (see Todorov et al., 2013). We use both male and female identities, because

Figure 1. Examples of the stimuli used in Experiment 1.

facial masculinity is associated with facial dominance for both men and women (e.g., Jones et al., 2010), and because physical strength and dominance may be judged differently for men and women. According to Sell, Tooby, and colleagues (2009), physical strength is associated with entitlement in men, but not in women.

In Study 1, we create new identities for both models of strength and dominance judgments. In addition, we create a model of the differences between these models. In Study 2, using the identities from Study 1, we ask participants to select from pairs of identities of both models those that look more socially dominant or physically stronger. This permits us to test whether participants can discriminate socially dominant faces from physically strong ones and vice versa.

Study 1

We asked participants to judge faces on physical strength and dominance. The judged faces were computer-generated from a face space. First, the judgments were used to model the representations of dominance and physical strength in the human face. Second, we created a model of the difference between the models of physical strength and dominance, that is, the difference dimension. Thus, we created three dimensions: dominance, physical strength, and their difference. To denote the poles of the difference dimension, in the following text, we will name the subtracted dimension in brackets after the maximized dimension, for example, *strong (non-dominant)* visualizes faces that appear very strong but lack appearance of dominance. Third, we randomly created new identities. For each identity, we generated versions that varied on dominance, physical strength, and their difference. This permitted us to visualize how a face varies across different levels (−3 to +3 *SD*) of the three models of judgments.

Method

Participants

In total, 194 participants (99 male, M_{age} = 36.05 years, *SD* = 13.15) from the United States and Western Europe were recruited and paid US$2 through Amazon Mechanical Turk (MTurk; Buhrmester, Kwang, & Gosling, 2011) for participating in the study. Ten additional cases were deleted because of more than 20% missing cases.[1]

Materials

The stimuli consisted of 300 computer-generated faces originally developed by Oosterhof and Todorov (2008; see Figure 1). These faces were previously used in the development of computational models of trait judgments of faces (see Oosterhof & Todorov, 2008, supporting information). These faces were created randomly purely on the basis of a general statistical face space with FaceGen Software (Singular Inversions, Toronto, Canada). This statistical face space was derived from statistical analysis of 3D laser scans of about 300 real people. The generation of these faces was not based on any previous trait ratings (trustworthiness, dominance, or other). The stimuli are mostly White faces of both genders without hair, facial hair, or clothing (for some stimuli, gender and ethnicity are hard to judge).

Procedure

The data were collected online using Qualtrics (www.qualtrics.com). Participants were asked to judge faces and instructed to rely on their intuition. On a single trial, one face was presented at the center of the screen with a question below: "How physically strong this person is compared to others of the same age?" or "How dominant this person is compared to others of the same age?" Answers were given on a 9-point scale, anchored with *very weak* and *very strong*, or *not dominant* and *very dominant*, respectively.

Each participant judged the faces on only one dimension (strength or dominance). The order of faces was randomized. Participants who judged dominance were instructed that by dominance, we meant "how much this person wants to influence other people and how much she or he is able to do so" (see Toscano et al., 2014). For physical strength, no specific definition was given.

Presenting all 300 faces to the participants of online experiments is likely to result in fatigue and low reliability. Therefore, we randomly divided the set of faces into four groups of 75 faces for each participant to answer. Each face

Table 1. Number of Participants in Study 2 by Substudy, Target, Participant Gender, and Judgment.

	Judgment			
	Dominance		Strength	
	Participant gender			
	Male	Female	Male	Female
Study 2a				
Target gender				
Male	28	41	32	49
Male	52	29	58	24
Study 2b				
Target gender				
Male	47	53	40	64
Female	44	34	45	36

was rated twice by every participant to increase reliability (150 ratings in total per participant). Dominance was judged by 87 participants (23, 21, 23, and 20 participants for each group of faces, respectively), and physical strength by 107 participants (24, 26, 29, and 28 participants for each group of faces, respectively)—see Table 1.

Results and Discussion

To assess inter-rater reliabilities, we computed intraclass correlations (ICC), asking whether the raters in each of the total eight samples were consistent with each other (four groups of raters drawn for each judgment). ICCs were estimated with SPSS's RELIABILITY module using the two-way random model and determining consistency for average measures (Landers, 2015; Shrout & Fleiss, 1979). For dominance judgments, we observed ICC(2, 23) = .93, ICC(2, 21) = .91, ICC(2, 23) = .92, and ICC(2, 20) = .93 for groups 1 to 4, respectively. For strength judgments, we observed ICC(2, 24) = .96, ICC(2, 26) = .94, ICC(2, 29) = .96, and ICC(2, 28) = .97 for groups 1 to 4, respectively. These values indicate that ratings from different raters shared considerable variance, allowing us to average them.

For each face, we aggregated the judgments of physical strength and dominance of all participants. Using the face as the unit of analysis, we computed Pearson correlations between the judgments of dominance and physical strength. Our analysis showed that the correlation was very high, $r = .84$, $p < .001$, $df = 298$, 95% confidence interval (CI) = [0.82, 0.95]. Replicating previous findings (Toscano et al., 2014), faces that are perceived as physically strong are also perceived as dominant, even when the two dimensions are judged by different observers.

For each judged face, the judgments of dominance and physical strength were averaged across participants. Following the approach outlined in Oosterhof and Todorov (2008), we used the mean judgments of the two traits to find the vectors in the face space that changed maximally dominance and physical strength. Therefore, we created new shape dimensions[2] visualizing impressions of physical strength and impressions of dominance (see Figure 2). The rendered images show that as dominance and physical strength increase, the face and the nose become wider, the eyes smaller, the eyebrows lower, and the chin larger. Moreover, as the dominance and physical strength increase, the faces become more masculine and mature. The models of both judgments are almost indistinguishable. In fact, the correlation between the two models is .93, which confirms the similarity of physical strength and dominance judgments.

In addition, we subtracted the vector of dominance from the vector of physical strength to create a model of the differences of these two vectors—the Difference Dimension model. As faces move along the difference dimension, they change from very dominant and simultaneously weak (−3 *SD*) to very strong but nondominant (+3 *SD*)—see Figure 2. The renderings resulting from the difference dimension suggest that strong (nondominant) faces are less threatening than simply strong faces from the model based on only the physical strength judgments. The dominant (weak) faces also appear to be less threatening than the pure dominant faces. In both types of faces, strong (nondominant) and dominant (weak) faces, the expression of negative emotions seems less present, although it is more present in the latter than in the former. In addition, we can see that from the dominant (weak) faces to the strong (nondominant) faces, the shape becomes rounder, and there is an increase of facial fat. The eyes also become closer and smaller, the lips become thinner, and the eyebrows slightly lower. In men's faces, the nose also becomes wider.

We then randomly generated 20 new identities (10 men and 10 women) and applied to each identity the vectors of the three models—Physical Strength, Dominance, and Difference Dimension (for details, Todorov & Oosterhof, 2011; Todorov et al., 2013). Thus, for each facial identity and all three dimensions, we created 7 renderings that show this face after applying a positive extreme of the respective dimension (+3 *SD*), a negative extreme of that dimension (−3 *SD*), and every 1 *SD* step in between.

Study 2

In Study 1, we replicated the finding that physically strong individuals are also judged as socially dominant, created computational models of dominance and strength in a face space, and generated renderings resulting from these dimensions as well as their difference. We discussed that the resulting faces look very much alike and that this appears to be true for both men and women, but, of course, an empirical test of this interpretation of the faces is missing. To provide such a test of the outcome of Study 1 is the first goal of Study 2.

First, we want to validate that the models of physical strength and dominance are hard to tell apart. To empirically test the images created in Study 1, we first show participants a face maximizing just strength and a face maximizing just dominance, and

Figure 2. Models of trait judgments created in Experiment 1.
Note. Example rendering using one male identity and one female identity, with model applied from −3 *SD* to +3 *SD*. (A) Model of Physical Strength Judgments; (B) Model of Dominance Judgments; (C) Difference Model, where −3 *SD* creates dominant (weak) and +3 *SD* creates strong (nondominant).

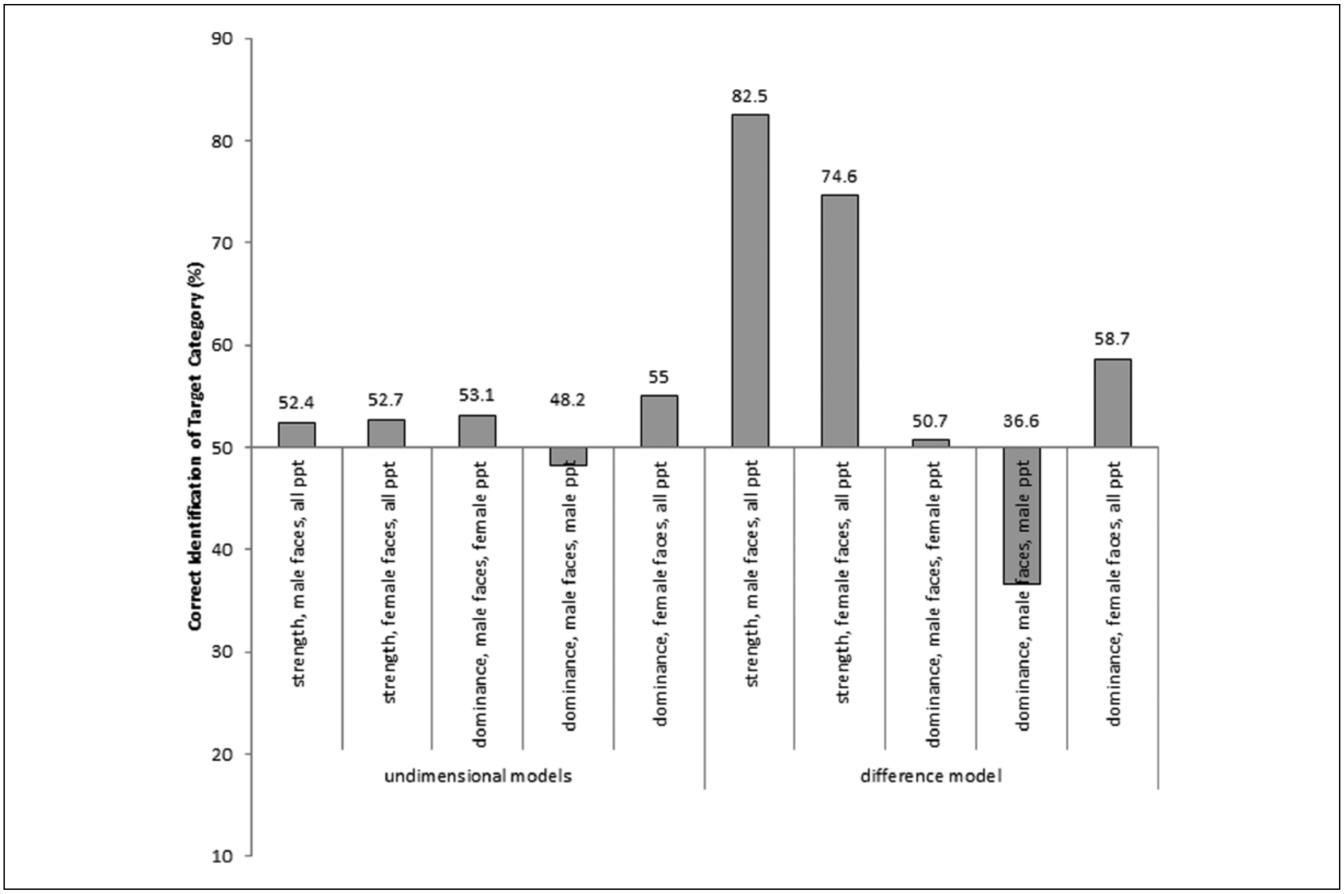

Figure 3. Correct identification (%) of target category in Studies 2a and 2b. In Study 2a, participants selected the stronger (bars a, b) or more dominant (bars c-d) face from a pair of faces where one maximized strength and one maximised dominance based on unidimensional models. In Study 2b, participants selected the stronger (bars f, g) or more dominant (bars h-j) face from a pair of faces that maximised strength or dominance based on the difference model. Separate bars for male and female participants are shown if their performance differed significantly. High values indicate correct identification of face maximising the trait over alternative trait. Chance performance equals 50%.

ask participants to select the one face they considered as physically stronger, or, in another condition, socially more dominant. Our hypothesis is that participants will perform at chance level.

Second, we repeat this procedure with the faces generated from the difference dimension to test whether physical strength and dominance can be discriminated here. More specifically, we investigate whether faces maximizing the difference in one direction can be successfully distinguished from faces maximizing the difference in the other direction—for example, whether strong (nondominant) faces indeed appear stronger than dominant (weak) faces. Finally, we investigate whether the pattern is the same for male and female faces.

Method

Participants

We recruited 676 participants (346 men; see Footnote 1), M_{age} = 34.74 years, *SD* = 11.45, from the United States and Western Europe via Amazon MTurk (paid: US$1.50). Different participants judged faces from the dominance or strength dimension (Study 2a) or faces from the difference dimension (Study 2b). Table 1 provides details of the number of participants by study.

Materials

The stimuli consisted of the 20 identities created in Experiment 1. We used 20 extremely strong faces and 20 extremely dominant faces (manipulated to 3 *SD* on their respective dimensions). In addition, 20 faces were maximized on the difference dimension and thus were extremely strong but not at all dominant. In the following, we will refer to such faces from the difference dimension as strong (nondominant). Finally, 20 faces were created to be extremely dominant but physically weak (always +3 *SD*; see Figure 3). Thus, we used the exemplars from Experiment 1 that were maximizing each specific dimension (i.e., physical strength, dominance, the subtraction of dominance from physical strength, and the subtraction of physical strength from dominance).

Procedure

Data were collected online, using Qualtrics (www.qualtrics.com). Participants were asked to choose from 10 pairs of faces the person that they considered as physically stronger or, in another condition, more dominant. We also used 10 filler pairs where one of the targets was much more dominant than the other (taken from the Dominance Model Database from Oosterhof & Todorov, 2008). Participants were told there were no right or wrong responses and to respond intuitively.

Each participant only judged faces on either physical strength or dominance, selecting which person appeared to be physically stronger or more dominant, respectively. Each pair appeared twice, and the position (left or right) of the faces was counterbalanced. The pairs of faces were presented at the center of the screen with a question below. The questions were "Which person do you consider as physically stronger?" or "Which person do you consider as more dominant?" As in Experiment 1, we defined dominance as "how much this person wants to influence other people and how much she or he is able to do so." In Study 2a, each pair consisted of one face that was extremely strong (+3 *SD* physical strength) and one that was extremely dominant (+3 *SD* dominant). In Study 2b, each pair consisted of one face that maximized what differentiated dominance from strength (i.e., +3 *SD* dominant weak faces), and another face maximized what differentiated strength from dominance (i.e., +3 *SD* strong nondominant faces). The design of both Studies 2a and 2b was 2 (Picture: high strength vs. high dominance, within) × 2 (judgment: stronger vs. more dominant, between) × 2 (position: left vs. right, within) × 2 (Target Gender: Men vs. Women, between).

Results

Answers were coded as 1 if the normatively correct picture was chosen (e.g., if the very strong instead of the very dominant face was chosen when the task was to select the stronger face), and 0 otherwise. Separate scores were created for the two tasks (identify physically stronger vs. more dominant), and male or female target faces. We tested the averaged scores against .5 (i.e., chance). For these comparisons, we also computed effect size $r = (t^2 / (t^2 + df))^{½}$. In addition, we also explored whether gender of participant had an influence, by submitting the scores to independent samples *t* tests. We report the outcomes of these *t* tests if they were significant.

Study 2a

In Study 2a, on each trial, two faces were presented, each maximizing one dimension (strength or dominance). When instructed to choose the stronger face from a pair of male faces, participants chose the very strong face instead of the very dominant face in 52.4% of the cases. One-sample *t* tests against chance (0.5) showed that this choice was not significantly different from chance, $t(80) = 1.51$, $p = .134$, $r = .17$, 95% CI of difference [−0.01, 0.05].

When instructed to choose the more dominant face from a pair of male faces, choices were also not better than chance ($p = .362$), but there were significant differences due to participants' gender. This difference emerged because male and female participants erred in different ways. Female participants chose the dominant face more often (53.1%) than male participants did (48.2%), $t(67) = 2.05$, $p = .044$, 95% CI of difference [−0.10, −0.001]. Men chose (wrongly) the strong face slightly more often (51.8%) instead of the dominant one (48.2%), but performed still at chance, $t(27) = −1.03$, $p = .31$, $r = .19$, 95% CI of difference [−0.05, 0.02]. Women chose (correctly) the dominant face slightly more often (53.0%), but this differed from chance only marginally, $t(40) = 1.97$, $p = .06$, $r = .30$, 95% CI of difference [−0.001, 0.06].

When choosing the stronger face from a pair of female faces, participants selected the strong faces (+3 *SD* physical strength) rather than the dominant face (+3 *SD* dominance) in 52.7% of the times. This selection was not better than chance, $t(81) = 1.41$, $p = .163$, $r = .15$, 95% CI of difference [−0.01, 0.06].

When choosing the more dominant face from a pair of female faces, the highly dominant female face was chosen in 55% of the cases over the very strong female face. This was significantly different from chance, $t(80) = 2.56$, $p = .012$, $r = .28$, 95% CI of difference [0.01, 0.09].

In sum, performance was largely at chance when participants had to select the more dominant or the stronger face from two faces that were maximizing either the dominance or the strength dimension. In none of the tests did performance differ more than 5% from chance level. Thus, as expected, participants performed at chance level.

Study 2b

Study 2b showed two faces, each pitting one dimension against the other using the difference dimension. Again, participants were instructed to select either the more dominant face, or the stronger face.

When instructed to choose the stronger face from a pair of male faces, participants selected the strong (nondominant) instead of the dominant (weak) face in 82.5% of the cases. This preference was significantly different from chance, $t(103) = 14.87$, $p < .001$, $r = .83$, 95% CI of difference [0.28, 0.37].

When instructed to choose the more dominant from a pair of male faces, participants wrongly selected the strong (nondominant) face in 56.6% of the cases instead of the dominant (weak) face (43.4%), which differed from chance performance, $t(99) = 2.21$, $p = .029$, $r = .42$, 95% CI of difference [−0.13, −0.01]. However, we found a gender difference, $t(98) = 2.15$, $p = .034$, $r = .21$, 95% CI of difference [−0.24, −0.01]. The unexpected difference was completely due to the male participants: They chose the strong (nondominant) face more often (63.4%) than the dominant (weak) face (36.6%), which was different from chance, $t(46) = -3.12$, $p = .003$, $r = .42$, 95% CI of difference [−0.22, −0.05].[3] Women, in contrast, selected the strong (nondominant) face slightly more often (50.7%), but were at chance, $t(52) = -0.16$, $p = .872$, $r = .02$, 95% CI of difference [−0.09, 0.08].

When instructed to choose the stronger face from a pair of female faces, the strong (nondominant) face was selected more often (74.6%). This preference was significantly different from chance, $t(80) = 7.62$, $p < .001$, $r = .65$, 95% CI of difference [0.18, 0.31].

Finally, instructing to choose the more dominant face from a pair of female faces, participants selected the dominant (weak) faces more often (58.7%) than the strong (nondominant) face (41.3%). This preference was significantly different from chance, $t(77) = 3.03$, $p = .003$, $r = .33$, 95% CI of difference [0.03, 0.14].

In sum, when having to pick the more dominant or the stronger face from two faces where one was generated from the dominance dimension and the other was generated from the strength dimension, participants could not distinguish between the models; they basically performed simply at chance, within ±5% around the 50% chance level.

However, we observed a double asymmetry when the same task was done with pictures generated from the difference dimension: When looking for the stronger face, participants correctly and significantly chose more often the face that maximized strength while minimizing dominance, and this was true for both male and female faces. This performance differed substantially from chance (>74%). When looking for the more dominant among two *female* faces of which one maximized strength over dominance and the other maximized dominance over strength, participants still chose (correctly) more often the face that maximized dominance, although less clearly so (59%).

However, when participants looked for the more dominant among two *male* faces, they actually mistakenly chose more often the face maximizing strength over dominance, committing this error above chance levels. At first, this appeared to happen especially for male participants, but the replication reported in Footnote 1 suggests this may happen for both genders.[4]

What is remarkable is that the task of choosing the stronger face and the task of choosing the more dominant face both use the *same faces*. Nevertheless, participants largely succeed in choosing the face that maximizes strength over dominance as stronger but fail to choose the face that maximizes dominance over strength as the more dominant one.

General Discussion

In the current studies, we created models of physical strength and dominance judgments to explore whether these two judgments differ. These models were built on the notion of statistical face space (Valentine, 1991). The models resulted from judgments of 300 computer-generated faces (Oosterhof & Todorov, 2008) along the two dimensions, where each dimension represents the maximum amount of explained variance from the respective judgment. Importantly, these models are data-driven and are not biased by the manipulation of single features. More specifically, these models enable us to investigate all the information people use when evaluating physical strength and dominance. Consequently, we can visualize the representations of physical strength and dominance.

We found several indications that the judgments of physical strength and dominance are highly aligned. First, replicating earlier findings, faces that were rated as very dominant were also rated as physically strong, even though these ratings came from different participants. Second, as a result, we show for the first time that the model that maximizes physical strength and the model that maximizes dominance based on the participants' ratings led to faces that look very much alike. Study 2a empirically confirmed that participants were at chance at discriminating faces generated from these two models. Thus, increases in dominance and physical strength are expressed similarly in changes in facial shape.

The facial properties that appear to change with increases in both dimensions are lengths of eyes and chin, eyebrows height, and width of nose and face, confirming earlier results (Toscano et al., 2014). In addition, as faces become more dominant or stronger, they seem to show more masculinity and anger (see also Sell, Tooby, & Cosmides, 2009). Note that this description is not exhaustive—more subtle changes are happening along the dimensions as a result of the data-driven method.

Some gender differences emerge from the visualizations: In men, the increase in dominance and physical strength seems to be linked primarily to muscularity. Dominance and physical strength in women, however, appear to be more associated with bodily weight. Note, however, that the face models we created based on the judgments of the initial 300 stimuli did not discriminate between genders. These apparent differences may be simply due to the application of the models to different head geometries. The initial 300 faces judged by the participants in Study 1 are more male than female, and the resulting model may thus be male-biased. In addition, the generated female faces might have some limitations. The baldness of all heads might have biased the judgments of the participants because bald heads tend to be judged as males (Dotsch & Todorov, 2012). As reported in Footnote 2, female faces that resulted from Dominance and Physical strength models may have been misidentified as males in some cases. However, female faces from the Difference model were identified in most cases as females.

Even though physically strong and dominant faces look very much alike, our face models make it possible to compute what differentiates strength from dominance, and to create faces that maximize this difference in one or the other direction. Thus, we can see how a face goes from extremely dominant (weak) to extremely strong (nondominant). This permits us to visualize the perceptual cues that discriminate perceptions of physical strength from perceptions of dominance.

The visualization of this difference model—see Figure 2—suggests that the strong (nondominant) faces have a rounder shape characterized by more facial fat and a less threatening look than the strong faces from the models of physical strength judgments. Moreover, dominant (weak) faces seem to have less muscularity and also a less angry look than the dominant faces from the models of dominance judgments. Both strong (nondominant) and dominant (weak) faces seem to show less angry expressions than the faces from the respective unidimensional models, although these expressions are more present in dominant (weak) faces than in strong (nondominant) faces. In addition, facial fat seems to increase from the dominant (weak) to the strong(nondominant) extreme. Apparently, strong (nondominant) persons are thought to have more body fat, which suggest an above average weight (Donofrio, 2000). Previous studies also show that physical strength and body fat shared some resemblances (Windhager et al., 2011). Other changes also seem to occur, as eyes become smaller, the eyebrows lower, and the noses flatter. (Note again that these observations are based on our inspection of the visualizations because the approach does not identify particular individual features but the constellations of all features changing as a function of judgments.)

When we presented the faces that maximized one difference over the other to participants in Study 2b, we found two interesting asymmetries. When participants had to choose the stronger face, they correctly selected the face maximizing physical strength over dominance more often than chance, and they did this for both male and female faces. Thus, participants were able to correctly differentiate strong (nondominant) faces from dominant (weak) faces when deciding about physical strength. However, when participants had to choose the more dominant face, the picture got complicated, even though only the judgment changed while the pictures stayed the same. When participants did this task for female faces, they had some success doing so and selected the dominant (weak) face more often than the strong (nondominant) face. That was not the case for male faces. When selecting the more dominant male face, participants tended to choose, in fact, the strong (nondominant) face more often than the dominant (weak) face.

In sum, pure strength can be distinguished from pure dominance in both male and female faces, but pure dominance can only be distinguished from pure strength in female faces, and not very well. Pure strength, however, tends to be mistaken for dominance in male faces. These asymmetries suggest that the degree to which dominance is equated with strength differs between male and female faces. Note that this result only became obvious when using the difference model, because under normal circumstances, dominance and strength are too conflated.

This double asymmetry cannot be purely explained by different morphologies of male and female faces because the two tasks (selecting the most dominant vs. selecting the stronger one) used exactly the same pictures. It is more likely to have its root in the conflation of the concepts of strength and dominance, especially for male targets. According to the social role perspective, men are more encouraged by our culture to apply physical force and use more physical aggression than women (Bettencourt & Miller, 1996; Eagly & Steffen, 1986). Furthermore, the use of bodily strength seems to play a more instrumental role in men. Men use their physical strength to gain influence, whereas for women, the use of physical strength tends to be seen as a loss of power (Alexander, Allen, Brooks, Cole, & Campbell, 2004; Campbell & Muncer, 1994; Campbell, Muncer, & Coyle, 1992; Schubert, 2004). Therefore, it is likely that the association between physical strength and social dominance will be more present for male targets.

From an evolutionary viewpoint, another explanation of this asymmetry is that the difficulties in dissociating dominance from physical strength in men are linked to the role that physical strength played in the bargaining position of an individual in our ancestors. Physically strong individuals would acquire a better bargaining position given their higher ability to inflict costs on others. As a result, these individuals would have more resource-holding potential or formidability. Thus, characteristics related to physical strength evolved to enhance the power of negotiation of an individual, which influenced their social dominance (Parker, 1974; Sell, Tooby, & Cosmides, 2009). A related evolutionary explanation is linked to mechanisms of fighting competition within males to exclude same-sex competitors (Darwin, 1871). Thus, physical characteristics that could increase the fighting ability (i.e., physical strength) would increase the mating opportunities for males and, consequently, their social influence. The mechanisms of sexual selection favored strong and larger bodies in males (see Puts, 2010), but not in females. For instance, men have on average 61% more total muscle mass than women, especially in upper body muscles (Lassek & Gaulin, 2009). Therefore, females could not monopolize men or mate with them through the use of sheer physical force. Moreover, typical female traits such as larger breasts and hips are not associated with physical strength, but with physical attractiveness (see Puts, 2010). As a result, their social dominance seems to be less associated with physical strength. For instance, Sell, Tooby, and Cosmides (2009) found that women's physical strength does not increase their feelings of entitlement, but attractiveness does.

Another related reason might be the role of weight cues for attractiveness judgments. As Coetzee, Chen, Perrett, and Stephen (2010) have demonstrated, weight can be accurately judged from the human face. Moreover, overweight women are seen as less attractive than women with normal or lower weight (see also Hume & Montgomerie, 2001). We can see that the dominant (weak) faces have less facial fat than the strong (nondominant) faces. This might increase the unattractiveness of the latter and, consequently, diminish their social dominance. In addition, more facial fat is also correlated with more health problems (Tinlin et al., 2013). Therefore, more facial fat might increase the perception of less fitness.

Caveats and Limitations

We want to point out several limitations of the current studies.

First, we have to recognize that the absence of a social context might have played a role. It is possible that participants would be more prone to identify dominance with the dominant (weak) faces than with strong (nondominant) faces in a context where skills such as emotional intelligence or technical expertise are more important. Instead, in the absence of an external judgment context, the stimuli themselves constitute the context. Because these stimuli are for the greater part seemingly young individuals with bald heads, the definition of dominance may have shifted toward focusing on bodily cues. This may have increased the salience of strength. Note, however, that it could alternatively have brought about a salience of other aspects (e.g., various facial expressions of emotions, such as smiling, anger, or contempt). Our evidence for the link between dominance and strength should nevertheless be interpreted in the light of the absence of other nonbodily cues.

A similar point applies to how participants interpret the concept of dominance itself during ratings in both Studies 1 and 2. We provided them with a definition of social dominance, but the context may have fed back to the definition of dominance itself; that is, the interpretation of dominance as the ability and desire to affect others might have been influenced by a context where participants saw faces only. Future studies might profit from contextualizing the judgments more strongly. In addition, just as our original rating stimuli were sampled from an empirically derived space of possible faces, situations of social dominance could be sampled from a theoretically or empirically derived space of dominance contests (see Brunswick, 1956). This would obviously broaden the generalizability of the evidence.

Another limitation of our data-driven approach is due to the fact that our models of dominance and physical strength derive from a specific set of faces where some important factors, such as age, do not differ substantially. If a feature does not vary substantially in the original face space, it cannot be picked up reliably by the data-driven approach (see Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). For instance, Sutherland and colleagues (2013) showed that age influences dominance perception. In particular, older men tend to be judged as more dominant. Thus, it seems that with stimuli with larger age differences, the relationship and similarity between dominance and physical strength could have been attenuated (elderly adults also have reduced muscle mass and strength compared with young adults). Therefore, using a face space that better encodes variance in age-related facial features could have led to older people being considered more dominant, but not necessarily physically stronger than younger people. This could result in larger differences between the models of physical strength and dominance than the ones seen in our studies. Future research may show how models of dominance and physical strength change as a function of age. When interpreting our data, it is important to keep in mind that they apply to a certain age group, young adults between around 20 and 40.

Taken together, our data are consistent with evolutionary models that posit a strong relationship between physical strength and social dominance primarily for men. In addition, the way that our society expects men and women to use physical strength might also have contributed (see Scott et al., 2014).

Conclusion

In this article, we investigated whether mental representations of faces of socially dominant individuals differ from representations of faces of physically strong individuals. Our data show that facial representations of social dominance and physical strength, when visualized from a computational model of both judgments, are highly similar. In addition, participants in an empirical study were not able to distinguish exemplars of both dimensions. However, it is possible to combine the model of strength with the model of dominance, and to generate faces that maximize one over the other. When participants were presented with such faces, they could identify the faces maximized for strength when searching for strength, as well in the case of female faces maximized for dominance when searching for dominance. However, they had trouble recognizing male faces maximizing dominance when looking for dominance. In sum, our data suggest that physical strength and social dominance judgments are strongly correlated, and representations of both look alike, but strength tended to be mistaken for dominance more than dominance was mistaken for strength, especially in male faces. Our data show how a data-driven approach can lead to interesting insights into mental representations, but we also identify several limitations of generalizability that derive from the necessarily large amount of judgments involved in building the models. We hope that this inspires future extensions of this approach.

Acknowledgments

We thank one anonymous reviewer and Frank Bernieri for helpful comments on an earlier version.

Author Contributions

H.T. and T.S. designed the studies and planned the data analysis. H.T. collected the data and conducted the analysis. A.T., R.D., and V.F. planned the face space analysis. V.F. conducted the face space

analysis and generated the classification images. H.T. wrote the first draft of the manuscript, and all authors contributed revisions.

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 funding provided by the University Institute of Lisbon (ISCTE-IUL) and the Center for Social Research and Intervention (CIS-IUL) to H.T. and T.S.

Supplemental Material

The online supplemental material is available at http://pspb.sagepub.com/supplemental.

Notes

1. We did not base the number of participants on a power analysis. However, in both studies, we used at least 20 participants for each group of faces (see Simmons, Nelson, & Simonsohn, 2011).
2. FaceGen Software (Singular Inversions, Toronto, Canada) uses 50 dimensions to represent face shape and 50 dimensions to represent face reflectance (i.e., texture, color, brightness). We wanted to have male and female identities of the same race and, with this goal in mind, we only manipulated facial shape, not skin texture (reflectance). The manipulation of reflectance would have created, at some levels of the respective dimension, faces with different skin color and gender.
3. Because of the unexpected direction of the effect, we conducted another experiment to replicate the results of Study 2b for the task of selecting the more dominant face from a pair of male faces generated from the difference dimension. We recruited 90 (47 female) participants from Amazon Mechanical Turk (MTurk). The mean age was 35 years (SD = 10.17). Again, we found that the strong (nondominant) face was chosen more often (66.3%) than the dominant (weak) face (33.7%), $t(89) = -5.33$, $p < .001$, $r = .49$, 95% confidence interval (CI) of difference [−0.22, −0.10]. However, we did not find gender differences as in the previous study, $p = .39$.
4. Because our models may be somewhat male-biased (see section "General Discussion"), we ran an additional study to check whether female faces were correctly identified as females. We asked participants (N = 98) recruited on MTurk to select the gender for each one of the faces used in Study 2. As we expected, male faces from dominance and physical strength models were identified in most cases as males (99%), $p < .001$. However, female faces were considered as males more than half of the time (67%), $p < .001$. Regarding faces generated from the difference model, male faces were still identified as males in almost every case (99%), $p < .001$. Female faces from this model were correctly identified as females in the majority of cases (89%), $p < .001$. Thus, the fact that we did not find strong effects of target gender in Study 2a could be due to miscategorization of the face gender. However, the face gender effects in Study 2b are likely grounded in largely correct gender categorizations.

References

Alexander, F., Allen, C., Brooks, J., Cole, C., & Campbell, A. (2004). Reason to believe: Representations of aggression as phenomenological read-out. *Sex Roles*, *51*, 647-659.

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

Bettencourt, B., & Miller, N. (1996). Gender differences in aggression as a function of provocation: A meta-analysis. *Psychological Bulletin*, *119*, 422-447.

Blaker, N. M., & van Vugt, M. (2014). The status-size hypothesis: How cues of physical size and social status influence each other. In J. T. Cheng, J. L. Tracy, & C. Anderson (Eds.), *The psychology of social status* (pp. 119-137). New York, NY: Springer Science+Business Media.

Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 187-194). Los Angeles: Addison Wesley Longman.

Blanz, V., & Vetter, T. (2003). Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*, 1063-1074.

Brunswick, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: University of California Press.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*, 3-5.

Campbell, A., & Muncer, S. (1994). Sex differences in aggression: Social representation and social roles. *British Journal of Social Psychology*, *33*, 233-240.

Campbell, A., Muncer, S., & Coyle, E. (1992). Social representation of aggression as an explanation of gender differences: A preliminary study. *Aggressive Behavior*, *18*, 95-108.

Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B: Biological Sciences*, *275*, 2651-2656.

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*, 103-125.

Coetzee, V., Chen, J., Perrett, D. I., & Stephen, I. D. (2010). Deciphering faces: Quantifiable visual cues to weight. *Perception*, *39*, 51-61.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London, England: John Murray.

Deaner, R. O., Goetz, S. M. M., Shattuck, K., & Schnotala, T. (2012). Body weight, not facial width-to-height ratio, predicts aggression in pro hockey players. *Journal of Research in Personality*, *46*, 235-238.

Donofrio, L. M. (2000). Fat distribution: A morphologic study of the aging face. *Dermatologic Surgery*, *26*, 1107-1112.

Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological & Personality Science*, *3*, 562-571.

Eagly, A. H., & Steffen, V. J. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, *100*, 309-330.

Facegen Main Software Development Kit 3.1. (2006). Vancouver, British Columbia, Canada: Singular Inversions.

Fink, B., Neave, N., & Seydel, H. (2007). Male facial appearance signals physical strength to women. *American Journal of Human Biology*, *19*, 82-87.

Fiske, A. P. (1991). *Structures of social life: The four elementary forms of human relations*. New York, NY: Free Press.

Fiske, A. P. (2000). Complementarity theory: Why human social capacities evolved to require cultural complements. *Personality and Social Psychology Review*, *4*, 76-94. doi:10.1207/S15327957PSPR0401_7

Gallup, A. W., O'Brien, D. T., White, D. D., & Wilson, D. S. (2010). Handgrip strength and socially dominant behavior in male adolescents. *Evolutionary Psychology*, *8*, 229-243.

Geniole, S. N., Keyes, A. E., Carré, J. M., & McCormick, C. M. (2014). Fearless dominance mediates the relationship between the facial width-to-height ratio and willingness to cheat. *Personality and Individual Differences*, *57*, 59-64.

Goetz, S. M., Shattuck, K. S., Miller, R. M., Campbell, J. A., Lozoya, E., Weisfeld, G. E., & Carré, J. M. (2013). Social status moderates the relationship between facial structure and aggression. *Psychological Science*, *24*, 2329-2334.

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.

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. doi:10.1177/0146167215591495

Henrich, J., & Gil-White, F. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution & Human Behavior*, *22*, 165-196.

Hume, D. K., & Montgomerie, R. (2001). Facial attractiveness signals different aspects of "quality" in women and men. *Evolution & Human Behavior*, *22*, 93-112.

Isen, J. D., McGue, M. K., & Iacono, W. G. (2015). Aggressive-antisocial boys develop into physically strong young men. *Psychological Science*, *26*, 444-455.

Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General*, *14*, 19-25.

Jones, B. C., DeBruine, L. M., Main, J. C., Little, A. C., Welling, L. L., Feinberg, D. R., & Tiddman, B. P. (2010). Facial cues of dominance modulate the short-term gaze-cuing effect in human observers. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 617-624.

Landers, R. N. (2015). Computing intraclass correlations (ICC) as estimates of interrater reliability in SPSS. *The Winnower*, *2*, Article e143518.81744. doi:10.15200/winn.143518.81744

Lassek, W. D., & Gaulin, S. J. C. (2009). Costs and benefits of fat-free muscle mass in men: Relationship to mating success, dietary requirements, and natural immunity. *Evolution & Human Behavior*, *30*, 322-328.

Lefevre, C. E., Etchells, P. J., Howell, E. C., Clark, A. P., & Penton-Voak, I. S. (2014). Facial width-to-height ratio predicts self-reported dominance and aggression in males and females, but a measure of masculinity does not. *Biology Letters*, *10*, Article 20140729. doi:10.1098/rsbl.2014.0729

Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007). Facial appearance affects voting decisions. *Evolution & Human Behavior*, *28*, 18-27.

Lukaszewski, A. W., Simmons, Z. L., Anderson, C., & Roney, J. R. (2016). The role of physical formidability in human social status allocation. *Journal of Personality and Social Psychology*, *110*, 385-406.

Mueller, U., & Mazur, A. (1996). Facial dominance of West Point cadets as a predictor of later military rank. *Social Forces*, *74*, 823-850.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*, 11087-11092.

Parker, G. A. (1974). Assessment strategy and the evolution of fighting behaviour. *Journal of Theoretical Biology*, *47*, 223-243.

Pellegrini, A. D. (1995). A longitudinal study of boys' rough-and-tumble play and dominance during early adolescence. *Journal of Applied Developmental Psychology*, *16*, 77-93.

Penton-Voak, I. S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution & Human Behavior*, *25*, 229-241.

Puts, D. A. (2010). Beauty and the beast: Mechanisms of sexual selection in humans. *Evolution & Human Behavior*, *31*, 157-175.

Rantanen, T., Guralnik, J. M., Foley, D., Masaki, K., Leveille, S., Curb, J. D., & White, L. (1999). Midlife hand grip strength as a predictor of old age disability. *Journal of the American Medical Association*, *28*, 558-560.

Rhodes, G., Simmons, L. W., & Peters, M. (2005). Attractiveness and sexual behavior: Does attractiveness enhance mating success? *Evolution & Human Behavior*, *26*, 186-201.

Schubert, T. W. (2004). The power in your hand: Gender differences in bodily feedback from making a fist. *Personality and Social Psychology Bulletin*, *30*, 757-769.

Scott, I. M., Clark, A. P., Josephson, S. C., Boyette, A. H., Cuthill, I. C., Fried, R. L., . . . Penton-Voak, I. S. (2014). Human preferences for sexually dimorphic faces may be evolutionarily novel. *Proceedings of the National Academy of Sciences*, *111*, 14388-14393.

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., Von Rueden, C., & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society B: Biological Sciences*, *276*, 575-584.

Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences*, *106*, 15073-15078.

Shrout, P., & Fleiss, J. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420-428.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359-1366.

Smuts, B., Cheney, D., Seyfarth, R., Struhsaker, T., & Wrangham, R. (1987). *Primate societies*. Chicago, IL: University of Chicago Press.

Sutherland, C. A. M., 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.

Tinlin, R. M., Watkins, C. D., Welling, L. L., DeBruine, L. M., Al-Dujaili, E. A., & Jones, B. C. (2013). Perceived facial adiposity conveys information about women's health. *British Journal of Psychology*, *104*, 235-248.

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.

Todorov, A., Dotsch, R., Wigboldus, D., & Said, C. P. (2011). Data-driven methods for modeling social perception. *Social & Personality Psychology Compass*, *5*, 775-791.

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.

Todorov, A., & Oosterhof, N. N. (2011, March). Modeling social perception of faces. *IEEE Signal Processing Magazine*, *28*, 117-122.

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, *27*, 813-833.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*, 455-460.

Toscano, H., Schubert, T. W., & Sell, A. N. (2014). Judgments of dominance from the face track physical strength. *Evolutionary Psychology*, *12*, 1-18.

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

Van Vugt, M., De Cremer, D., & Janssen, D. P. (2007). Gender differences in cooperation and competition: The male-warrior hypothesis. *Psychological Science*, *18*, 19-23.

Von Rueden, C., Gurven, M., & Kaplan, H. (2008). The multiple dimensions of male social status in an Amazonian society. *Evolution & Human Behavior*, *29*, 402-415.

Von Rueden, C., Gurven, M., & Kaplan, H. (2011). Why do men seek status? Fitness payoffs to dominance and prestige. *Proceedings of the Royal Society B: Biological Sciences*, *278*, 2223-2232.

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

Walker, M., & Vetter, T. (2009). Portraits made to measure: Manipulating social judgments about individuals with a statistical face model. *Journal of Vision*, *9*, 1-13.

Watkins, C. D., Fraccaro, P. J., Smith, F. G., Vukovic, J., Feinberg, D. R., DeBruine, L. M., & Jones, B. C. (2010). Taller men are less sensitive to cues of dominance in other men. *Behavioral Ecology*, *21*, 943-947.

Watkins, C. D., Jones, B. C., & DeBruine, L. M. (2010). Individual differences in dominance perception: Dominant men are less sensitive to facial cues of male dominance. *Personality and Individual Differences*, *49*, 967-971.

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

Wind, A. E., Takken, T., Helders, P. J., & Engelbert, R. H. (2010). Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? *European Journal of Pediatrics*, *30*, 477-488.

Windhager, S., Schaefer, K., & Fink, B. (2011). Geometric morphometrics of male facial shape in relation to physical strength and perceived attractiveness, dominance, and masculinity. *American Journal of Human Biology*, *23*, 805-814.

Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*, *15*, 603-623.