Received: 20 November 2023 | Revised: 7 February 2024 | Accepted: 10 March 2024

DOI: 10.1002/ajp.23623

RESEARCH ARTICLE

AMERICAN JOURNAL OF PRIMATOLOGY WILEY

Capuchin monkeys' (*Sapajus [Cebus] apella*) categorization of photos of unknown male conspecifics suggests attention to fWHR and a dominance bias

Ashley M. Meacham[1] | Meghan J. Sosnowski[1,2] | Heather M. Kleider-Offutt[1,3] | Sarah F. Brosnan[1,2,3,4]

[1]Department of Psychology, Georgia State University, Atlanta, Georgia, USA

[2]Language Research Center, Georgia State University, Decatur, Georgia, USA

[3]Neuroscience Institute, Georgia State University, Atlanta, Georgia, USA

[4]Center for Behavioral Neuroscience, Georgia State University, Atlanta, Georgia, USA

Correspondence
Ashley M. Meacham, Department of Psychology, Georgia State University, P.O. Box 5010, Atlanta, GA 30302-5010, USA.
Email: ameacham1@student.gsu.edu

Present address
Meghan J. Sosnowski, California National Primate Research Center, University of California, Davis, California, USA.

Funding information
Georgia State University,
Grant/Award Numbers: 2CI PSCEB Doctoral Fellowship, Brains & Behavior Doctoral Fellowship; National Science Foundation,
Grant/Award Numbers: IBSS 2135621,
SES 1919305

Abstract

The ability to quickly perceive others' rank minimizes costs by helping individuals behave appropriately when interacting with strangers. Indeed, humans and at least some other species can quickly determine strangers' rank or dominance based only on physical features without observing others' interactions or behavior. Nonhuman primates can determine strangers' ranks by observing their interactions, and some evidence suggests that at least some cues to dominance, such as facial width-to-height ratio (fWHR), are also present in other primates. However, it is unknown whether they can determine strangers' rank simply by looking at their faces, rather than observing their interactions. If so, this would suggest selective pressure across the primates on both cues to dominance and the ability to detect those cues accurately. To address this, we examined the ability of male and female tufted capuchin monkeys (*Sapajus [Cebus] apella*) to categorize images of the faces of unknown conspecifics (*Sapajus* from different colonies) and humans (computer-generated and real) as dominant or nondominant based only on still images. Capuchins' categorization of unknown conspecific faces was consistent with fWHR, a cue to dominance, although there was a strong tendency to categorize strangers as dominant, particularly for males. This was true despite the continued correct categorization of known individuals. In addition, capuchins did not categorize human strangers in accordance with external pre-ratings of dominance by independent human raters, despite the availability of the same cue, fWHR. We consider these results in the context of capuchin socio-ecology and what they mean for the evolution of rapid decision-making in social contexts.

KEYWORDS

categorization, cognition, dominance, face processing, judgment

---

Abbreviations: fWHR, facial width-to-height ratio.

1 | INTRODUCTION

Humans use first impressions as important cues of how a stranger might behave (e.g., aggression and violence) so that they can determine how to respond appropriately. These first impression cues come primarily from face judgments, which can occur in as few as 100 ms (Willis & Todorov, 2006), suggesting that these responses are automatic. The automaticity of these judgments suggests that they could be based on cues that have evolved to be reliable and easy to perceive, such that they can be assessed rapidly (Mealey et al., 1996; Todorov, 2017). Indeed, one would anticipate that this might be a particularly important ability for any highly social species, not just humans, in which individuals must routinely—and rapidly—decide how to best interact with both familiar and unfamiliar others, and for which the ability to make judgments quickly and accurately would allow an individual to predict how another might behave and respond accordingly (Lefevre et al., 2014; Pineda et al., 1994; Todorov, 2017). In particular, individuals should be especially proficient at recognizing and categorizing others who may be relevant to their well-being, for instance, as a potential threat (Mealey et al., 1996).

One ever-present threat in the social arena is a more dominant individual, who may best one in competition for resources. As might be expected, humans are particularly sensitive to any facial cues (e.g., wide nose, thin lips, and broad jaw; Kleider-Offutt et al., 2021; Windhager et al., 2011) correlated with dominance, a trait that increases with perceived masculinity (Oosterhof & Todorov, 2008), body size (Han et al., 2017), and physical strength (Windhager et al., 2011). When faced with an unknown individual, dominance cues may be particularly informative when deciding how to respond, especially among individuals who may be less dominant. Among humans, less dominant men show an increased attentiveness toward such cues (e.g., masculinity and height; Watkins et al., 2010) as compared to more dominant men, and women, who are often shorter and smaller than men (Batres et al., 2015; Gray & Wolfe, 1980; Zebrowitz, 2017), and therefore, presumably, less dominant, are more likely to rate male faces as more dominant overall (e.g., Kleider-Offutt et al., 2021). Some have even argued that this suggests a potential cognitive bias toward dominance-related threat potentials (Mealey et al., 1996). Humans are not alone in this, either; birds identify dominance from the size of “badges,” such as head tufts, that correlate with strength and rank (Rohwer, 1975, 1977, 1982), giving us reason to believe that visual cues may be an important signal across the animal kingdom to the assessment of strangers.

Indeed, dominance is a key feature in structuring both human (Keating et al., 1981; Kleider-Offutt et al., 2021) and animal (Boehm, 1999; de Waal, 1982) social groups. Dominant individuals gain significant benefits (such as longevity, infant survival, and success in mating; Murray et al., 2007; Palombit et al., 2001; Rhodes et al., 2005; Silk et al., 2003; Silk et al., 2010; Valentine et al., 2014), and individuals know both their place in the hierarchy and, at least in some animals, others' places in it as well (for instance, capuchin monkeys preferentially recruit allies who outrank their adversaries, suggesting that they monitor relative rank of those around them; *Cebus capuchinus*: Perry et al., 2004).

Research Highlights

- Tufted capuchins categorize unknown male faces as dominant or not consistent with facial width-to-height ratios.
- Categorizations are biased toward dominance, suggesting that all strangers are initially treated as a threat.
- This suggests that capuchins do best by assuming dominance in unknown males.

Perhaps not surprisingly, then, dominance cues are important in primate species. Many primates give standardized submissive signals (Preuschoft & van Hooff, 1995), such as the pant grunt in chimpanzees (de Waal, 1982), which are unidirectional, unambiguous standardized signals given from the subordinate to the dominant. Experimental work shows that nonhuman primates can learn to identify and discriminate dominance cues through observing social interactions (Bovet & Washburn, 2003; Paxton et al., 2010), suggesting that monkeys can also utilize abstract social concepts to assist in navigating these social settings. An important question that remains, however, is whether this dominance recognition in nonhuman primates extends to information derived primarily from the face, as is true in humans. If nonhuman primates, too, use facial structures to make rapid judgments about social features, such as dominance, it would further suggest that facial cues to dominance have been conserved, at minimum, across primate species.

Parallel lines of research suggest that this could be the case. For instance, facial width-to-height ratio (fWHR), a widely debated indicator of dominance and dominance-related behavior (see Durkee & Ayers, 2021), is a cue common to both human and nonhuman primates. The ratio of facial width (i.e., the distance between cheekbones) to upper facial height (i.e., the distance between the upper lip and mid-brow) increases, along with testosterone levels, after puberty (Dixson, 2017; Valentine et al., 2014), indicating that this structural change might confer some fitness benefit as individuals enter sexual maturity and mating competition (e.g., Lefevre et al., 2014). From a physiological standpoint, increased facial width (i.e., broader jaw) may be linked to greater bite strength and teeth being utilized as a weapon (Lefevre et al., 2014; McArthur & Apatow, 1984), signifying increased fighting ability and an overall capacity to inflict harm. Supporting the possibility of an evolutionary link, fWHR was positively related to alpha status and assertiveness in adult brown capuchin monkeys (*Sapajus [Cebus] apella*; including some of the same subjects in the current study), along with significant sexual dimorphism (i.e., visible difference in appearance) in adults compared to juveniles (Lefevre et al., 2014). These results suggest that there are structural indicators of dominance in the faces of other species, an essential step if animals are to accurately assess dominance from visual cues alone. Moreover, if fWHR is used by capuchins as a structural cue to dominance because it is a feature that is shared by at least some other primates (i.e., humans show a similar pattern of fWHR variation), in principle, it could be used to assess dominance in non-conspecific faces as well.

To this end, we tested the ability of tufted capuchin monkeys (*Sapajus [Cebus] apella*) to categorize faces of both other capuchin monkeys and humans—the other primate with whom they have the most direct experience—as dominant or nondominant using only photographs. Capuchins are a particularly good nonhuman primate to use for this question. They are an unusually long-lived, highly social primate species (Benitez et al., 2021; Fragaszy et al., 2004) in which females live their entire lives in the same social group and develop close-knit, reciprocal relationships that include mutual grooming (di Bitetti, 1997), food sharing (de Waal, 2000), and even allo-nursing of one another's infants (Baldovino and Di Bitetti, 2008), suggesting that relationships are critical. Recognizing dominance is important in capuchins; males must constantly assess whether they are in a position to take over a group or whether an unfamiliar male may be in a position to take over from them, and females must recognize when a new male will take over, as incoming males can be highly infanticidal (Benitez et al., 2021; Janson & van Schaik, 2000). As mentioned above, adult capuchins' fWHR is positively related to alpha status (Lefevre et al., 2014), so we can be reasonably confident that there is at least one cue for the primates to use if they are so inclined. Finally, our particular group of capuchins is ideal as our monkeys live in stable, mixed-sex social groups, allowing them to have species-typical relationships and have visual and vocal access to multiple other groups in their outdoor enclosures, allowing them a complex social repertoire that facilitates the development of species-typical social behavior. Capuchins also recognize both familiar groupmates and familiar out-group members in photographs (although not unfamiliar out-group members; Talbot et al., 2016), indicating that capuchins are able to extract identifying information from two-dimensional images of conspecific faces.

In the present study, we first trained tufted capuchin monkeys to categorize photographs of familiar male conspecifics (i.e., within visual access at the same facility; Talbot et al., 2016) as dominant (alpha male) or nondominant. We then assessed whether they could generalize these trained dominance categories to images of unfamiliar male conspecifics. To do so, we solicited photos of male capuchins (both dominant and nondominant) from other facilities housing captive group-housed monkeys that were unknown to and unrelated to our monkeys, so that we could determine whether our monkeys could successfully identify the real-world dominance of unknown individuals by their image alone, based upon whatever cue they were using. As monkeys make accurate dominance assessments based on behavior (Bovet & Washburn, 2003; Paxton et al., 2010), we predicted that they would correctly categorize photos of unfamiliar male faces as “dominant” or “nondominant” at above chance levels. Given the differences in costs and benefits for males and females, we also predicted that categorizations might differ between subjects of different sexes, although we did not have a directional prediction.

To explore whether these facial cues of dominance were common across primate species (and if so, if capuchins could generalize these dominance categories to non-conspecific faces), we then tested whether capuchins would categorize images of computer-generated and real male human faces, all unfamiliar to our subjects, in the same way as human raters. Humans were an appropriate choice for two reasons. First, our capuchins interact with a variety of humans daily, including both males and females, so this is the most familiar non-conspecific primate species with which to test them. Second, fWHR varies with dominance in similar ways in humans as capuchins (see Geniole et al., 2015). We did not have a prediction with respect to the human faces, and indeed, success would not tell us whether they were using the same cues to categorize humans and monkeys. However, an ability to do so would at least suggest that they could generalize across primates.

Finally, we did not initially include a measure of each stimulus' fWHR, but given its potential importance to identifying dominance, we added it post hoc to the model to determine if this may have been related to their choices (because this was post hoc, we did not have an a priori prediction).

2 | METHODS

2.1 | Transparency and openness

We report how we determined our sample size, all data exclusions (there were none), all manipulations, and all measures in the study.

3 | STUDY 1 – TRAINING

3.1 | Subjects

We tested 25 adult tufted capuchin monkeys (7 males and 18 females, age range: 9–34 years) housed at the Language Research Center (LRC) at Georgia State University. Capuchin monkeys at the LRC live in one of five mixed-sex social groups, with the exception of one bachelor pair of males who live together with one another and adjacent to another social group with whom they always have visual and vocal access. Every capuchin group, including the bachelor pair, has its own large indoor–outdoor enclosure that includes a large outdoor play yard to which they have access except during inclement weather or if they choose to separate for voluntary testing. Each day, monkeys are given the opportunity to voluntarily separate from their groupmates to participate in cognitive and behavioral testing. Monkeys are never deprived of food, water, or access to the outdoors or their groupmates to encourage participation in testing, and the only consequence of choosing not to separate from their group is not being able to participate in the day's cognitive testing. All monkeys are fed a species-appropriate diet of vegetables, fruit, and supplemental monkey chow in several meals each day (testing rewards are in addition to this diet) and water is available ad libitum, including during testing.

All monkeys at the LRC have extensive experience with the LRC Computerized Testing System (Evans et al., 2008), consisting of a joystick-controlled computer with a monitor and attached pellet dispenser, which automatically dispenses 45 mg Bioserv reward pellets for correct responses. Computers are either a desktop computer running Windows XP or a laptop computer running Windows 7; the program, which was coded in Python 2.7, was consistent across all computers regardless of Windows version.

All procedures for the present study were approved by the Georgia State University Institutional Animal Care and Use Committee (IACUC, #A16031, #A19027) and complied with all relevant legal requirements that govern research involving animal subjects in the United States of America. All procedures with respect to the treatment of our sample complied with the American Psychological Association's ethical standards for the treatment of animals in research and the American Society of Primatologists's statement on the Principles for the Ethical Treatment of nonhuman Primates.

3.2 | Categorization task

3.2.1 | General trial procedure

We tested subjects on a computerized two-choice categorization task in which monkeys determined if a presented face image represented a dominant or nondominant individual. Stimuli consisted of head-on images of conspecific (or, in later phases, human) faces with neutral expressions looking directly ahead that were cropped to include only the face, ears, and tufts (two patches of longer dark hair on top of the head, a distinguishing feature in this species; see Figure 2 below) with whitespace surrounding; we standardized images to appear at 200 × 200 pixels and controlled for image size as much as possible given the irregular shape by ensuring that the face comprised at least 75% of the image presentation.

Unfortunately, we had a very small number of images due to limitations in our ability to source photos that (1) were of males that were unknown and unrelated to our monkeys, (2) were of sufficient quality (i.e., high enough pixel count and in-focus image), (3) that had the proper pose and facial expression (facing the camera with a neutral expression), and (4) for which we had information about dominance, all of which were essential for the current study. There are very few facilities housing tufted capuchins, and we relied on their generosity in providing photos (we could not use photos available on open-source photography websites because we did not have dominance information on these animals). This meant that we had an unbalanced number of photos per individual, and that the number of "dominant" and "nondominant" monkeys was dictated by the number of those males housed at each facility, including our own. Captive capuchin groups often have only one adult male (the dominant male), both to minimize conflict and because this group structure reflects a common robust capuchin demographic in the wild (Benitez et al., 2021). While we and some of our colleagues have groups with multiple males (and, thus, nondominant males), this tendency toward groups with a single male resulted in more dominant than nondominant individuals in our training stimuli set, although our test stimuli set was better balanced. Thus, we overall were limited in our ability to standardize the number of individuals in each dominance category and the number of images per individual, which we had to account for by repeating photos to avoid an unbalanced design.

To deal with our small sample set while maintaining the same number of exposures to dominant and nondominant stimuli during training and non-probe trials during testing (when differential reinforcement was used), stimuli were used more than once, using random sampling with replacement. As a result, in training and in non-probe trials during testing, the presentation of *categories* (dominant/nondominant) was always balanced, but subjects saw the same *images* more than once and saw the same images of nondominant individuals (our smaller sample set for familiar stimuli) repeated more frequently than those of dominants (our larger sample set of familiar stimuli). However, the selection of images was fully randomized within each phase, so the number of times each subject saw a particular image varied across subjects, ensuring no systematic difference in which image or face was seen more frequently. More importantly, for probe trials during testing, we had a much more balanced sample of dominant versus nondominant males and a more balanced sample of images; in fact, it was reversed from that of training and non-probe trials during testing, with slightly more images of nondominants than dominants (there were more images than individuals because we had multiple images of most males). Moreover, subjects were not differentially reinforced for probe trials, which should have avoided issues with frequency effect biases during testing. The number of dominant and nondominant individuals and stimuli per category utilized in each phase of the study is described in greater detail in the following sections and in Table 1. Although this unbalanced number of images was clearly not ideal, it allowed us to run the study given the limited sample of images met our criteria that we were able to obtain.

All trials began with a start screen with a randomly selected face image displayed in the upper-center and the cursor in the lower-center of the screen (Figure 1). To begin a trial, monkeys moved the cursor up to contact with the stimulus. Once the trial began, the face stimulus remained on the screen for 2 s before two response symbols were presented on either side of the bottom of the screen. One symbol, a blue triangle, represented the "dominant" category, and the other, a yellow cross, represented the "nondominant" category. To emphasize that the response categories remained the same over different types of stimuli and to provide an additional cue, the dominant response symbol was always presented on the left, while the nondominant response symbol was always presented on the right. If the monkey correctly categorized the face in the training phases, they received positive auditory feedback (a chime) and a food pellet. After a brief intertrial interval (2 s), monkeys were allowed to begin the next trial. However, following incorrect responses, monkeys received negative auditory feedback (a buzz) and a longer "time out" period (5-s intertrial interval). In testing probe trials (see

TABLE 1 Number of individuals and stimuli per category utilized during probe trials in each phase of the study.

Phase	Dominant individuals *N*	Dominant images *N*	Nondominant individuals *N*	Nondominant images *N*
Training 1	5	19	2	10
Training 2	6	35	3	14
Unfamiliar conspecifics	5	25	6	41
Computer-generated humans	25	25	25	25
Real humans	17	17	18	18

Note: *N* = sample size.

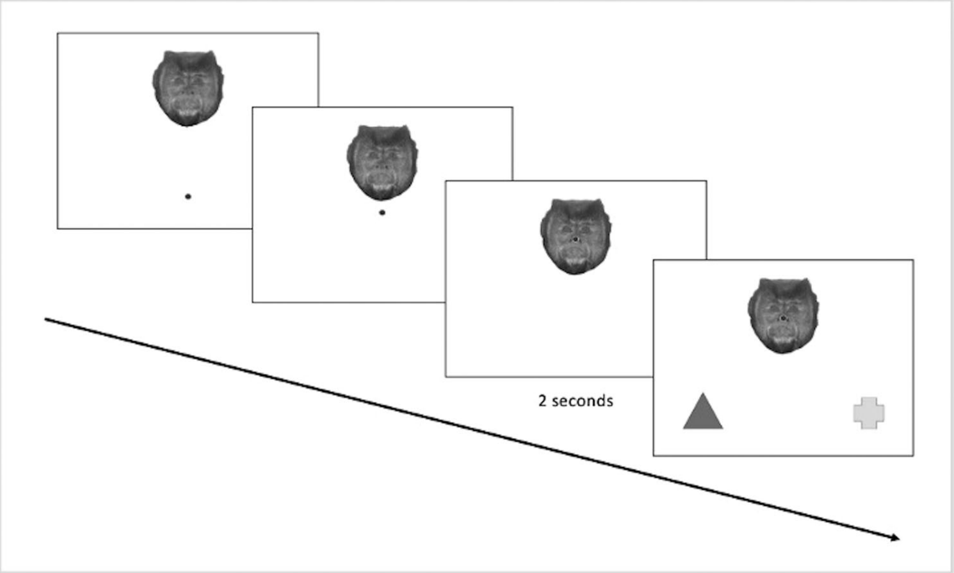

FIGURE 1 Sample trial shown to subjects. *Note*: While images are recreated in grayscale here, monkeys were presented color images for conspecific face stimuli, cursor, and choice symbols.

below), monkeys were rewarded for either choice to avoid influencing their categorizations.

3.2.2 | Training phases

We first trained the monkeys to associate the response symbols with the dominance categories. To do so, we used positive reinforcement to train the monkeys to associate one symbol with dominance (i.e., alpha male) and one with nondominance using photos of familiar male conspecifics that they regularly see at the LRC. Capuchins recognize photos of familiar conspecifics (Talbot et al., 2016), so we expected that, based on their knowledge of each male's dominance status in their respective groups, they would extrapolate to categorizing the face images as they would the actual monkeys. Which males were alpha males (i.e., the dominant) was assessed through a combination of researcher and caregiver report and group scan behavioral observations.

Because our training was based on familiar individuals, it was limited by the number of individuals housed at our facility. Only the alpha male of each group was categorized as "dominant," and other males (there were never more than two males in any group) were categorized as "nondominant." In Training Phase 1, for which the goal was to train the categories, we had five alpha males and two nondominant males. In Training Phase 2, for which the goal was to see if the categories generalized, we used novel images of the same seven individuals to see if the monkeys extrapolated to these males' new photos. In addition, to test whether their categorization generalized to photos of other familiar monkeys whose photos they had not seen in Training Phase 1, we included two additional individuals, one dominant and one nondominant male, for whom we had been able to acquire appropriate photos since testing commenced (getting photos of males in the correct pose and facial expression without another individual or element of their cage obstructing can be difficult). Of course, two monkeys are not a sufficiently large sample to test generalization conclusively, but the goal was not to test generalization but to teach the categories, and additional exemplars gave them a greater chance of doing so. Thus, Training Phase 2 consisted of a total of nine (six alpha males and three nondominant) familiar conspecific faces. Finally, as noted previously, although we were limited in the number of alpha and nondominant males, we ensured that dominant and nondominant images were presented equally often within each phase of training by using sampling with replacement.

In Phase 1, monkeys saw a single set of full-color images of familiar male conspecific faces (Figure 2; 7 males, $N_{\text{Dominant}} = 19$ images, $N_{\text{Nondominant}} = 10$ images), presented in a randomized order, such that of every eight trials, four were "dominant" and four were "nondominant" images, and were rewarded for correctly categorizing those faces as dominant or nondominant. We set the training criterion at 80% accuracy in two consecutive sessions, at which point they moved to the next phase.

The goal of Phase 2 was to ensure that the monkeys would generalize this learned categorization rule to a novel set of stimuli utilizing the same procedure from Phase 1. Thus, we presented novel photos of the seven males in Phase 1 and two novel males (9 males, $N_{\text{Dominant}} = 35$ images, $N_{\text{Nondominant}} = 14$ images); none of these images had been shown in Training Phase 1. The criterion was again set at 80% accuracy on two consecutive sessions. Two female capuchins failed to complete Training Phase 1, and one male capuchin failed to complete Training Phase 2 and were therefore dropped from the study.

3.2.3 | Testing phases

Testing phases consisted of fully randomized probe and non-probe trials. One of every seven trials was a non-differentially reinforced probe trial that contained an unfamiliar stimulus—depending on testing phase, this was either an image of an unknown conspecific face (Testing Phase 1; see details below), a computer-generated human face (Testing Phase 2), or a real human face (Testing Phase 3)—and six non-probe trials, containing three "dominant" and three "nondominant" trials, consisting of familiar conspecific face images from Training Phase 2. For non-probe trials, the feedback and reward/consequence for correct and incorrect categorizations were consistent with training. For probe trials, monkeys received positive auditory feedback and a pellet regardless of their categorization to avoid biasing their overall

FIGURE 2 Sample dominant (top row) and nondominant (bottom row) male capuchin faces. *Note*: While images are reproduced in grayscale here, monkeys were presented color images for conspecific face stimuli.

categorizations of these stimuli. Monkeys completed 1000 total trials in each testing phase, divided among as many sessions as they required ($Mdn_{Test1} = 2$ sessions, IQR: 2–2 sessions, range: 1–5 sessions; $Mdn_{Test2} = 2$ sessions, IQR: 2–3 sessions, range: 1–5 sessions; $Mdn_{Test3} = 4.5$ sessions, IQR: 3.25–7 sessions, range: 2–15 sessions).

Testing Phase 1 assessed monkeys' ability to categorize images of the faces of unfamiliar conspecifics. We used full-color photographs ($N = 66$) of 11 unfamiliar male conspecifics (five dominant, six nondominant) from four other captive capuchin colonies that were unrelated to and unfamiliar to our monkeys to create face stimuli using the same procedure that we described for creating the training stimuli. We used only photos in which the unfamiliar conspecifics had a neutral expression; although we did our best to use similar quality photographs as to our training stimuli (and chose to exclude some that were clearly of lower quality), as all of these monkeys were group-living (i.e., unrestrained) and different photographers took the photos at each facility, there was some variation in image quality. For these photos, the males were rated as “dominant” ($N = 25$ images) or “nondominant” ($N = 41$ images) by caretakers familiar with them at their respective facilities (not the authors of the current study). Again, as with the training stimuli, the number of “dominant” and “nondominant” monkeys was dependent upon the number of those individuals housed at each facility, and the number of images of each unfamiliar individual was determined by the images that colleagues had available.

Testing Phase 2 assessed whether monkeys could categorize computer generated images of human male faces. The computer-generated human faces (FaceGen Modeller; Singular Inversions, Toronto, Canada) consisted of a subset of images taken from a larger database of 25 unique face identities that varied on seven levels of dominance (for examples, see Oosterhof & Todorov, 2008; Todorov & Oh, 2021; Todorov & Oosterhof, 2011). To remove extraneous variation and match our available sample for Testing Phase 3, we chose to use faces that were generated to be bald, White males with neutral expressions. We also converted all images to grayscale because humans are more variable in coloration (eyes, hair, and skin) than capuchins, which vary very little. For the purposes of the current study, we used the dominant “extremes” to alleviate concerns that a face would be perceived as ambiguously dominant; thus, we used the pre-determined least (i.e., “nondominant”; $N = 25$) and most (i.e., “dominant”; $N = 25$) dominant images for each distinct face identity.

Finally, in Testing Phase 3, we assessed whether monkeys could categorize images of real, but unknown human faces. To test this, we used photographs of real male humans from the Chicago Face Database (for examples, see Ma et al., 2015). All faces were White males, again converted to grayscale, with neutral expressions. Faces were pre-rated on a variety of physical and subjective attributes (e.g., dominance and attractiveness); thus, using the dominance ratings provided in the Chicago Face Database norming data and codebook (see Ma et al., 2015 for further details), we performed a median split to determine our “nondominant” ($N = 18$) and “dominant” ($N = 17$) face-type categories. Additionally, the photos varied in the amount and length of hair; thus, the faces were cropped to include only a thin line of hair to preserve the natural shape of the skull while also limiting the effects of hairstyle on categorization.

Because all cognitive testing at the LRC is voluntary, monkeys can choose not to participate or can choose to stop participating at any time. One female subject failed to complete Testing Phase 2, and two females and one male failed to complete Testing Phase 3. Overall, 18 monkeys completed all phases of the study.

3.2.4 | fWHR measurement

In previous literature, fWHR was positively associated with hierarchical rank in adult brown capuchin monkeys (*Sapajus [Cebus] apella*), suggesting that fWHR might be a key physical cue related to dominance judgments (Lefevre et al., 2014). Therefore, as a post hoc measure, we assessed if our subjects in the present study might have used fWHR to help inform their decision-making when categorizing conspecifics. To do so, we developed a methodology to calculate fWHR for each individual depicted in the stimuli, based on that described in previous literature that measured fWHR in nonhuman primates (Lefevre et al., 2014; Wilson et al., 2013).

Two fully frontal images of each individual were randomly selected from our stimulus set, deidentified, and placed into a new, random order; we horizontally aligned and scaled each of these photos according to interpupillary distance using WebMorph software (DeBruine, 2018). Two independent coders blind to the experimental conditions used the WebMorph software to identify facial points using these aligned and scaled stimulus photos. Coders identified nine points on each face that allowed us to calculate the horizontal distances between the left to right facial boundaries in two places: first, the horizontal distance across the face at the height of the monkey's eyelids, and second, the horizontal distance across the face at the height of the center of the monkey's upper lip. Then, they identified the vertical distance from the midpoint of the monkey's upper lip to the highest point of the eyelids. One coder measured all photos (a total of 40 photos from 20 individuals), and a second coder independently measured a randomly selected 25% of these (10 photos total, no more than one photo per stimulus monkey; neither coder had previously participated in the study and so did not know the monkeys or our hypotheses). Inter-rater reliability for these measurements was high ($r(8) = 0.92$, $p < 0.001$, 95% CI [0.70, 0.98]). Using these measurements, we calculated fWHR for each photo by dividing the maximum horizontal distance by the height of the face; the measures for the two photos were then averaged, inspected for high variability between measurements, and included in our model for Testing Phase 1.

As the calculation of fWHR is almost identical for human faces (Carré & McCormick, 2008), we used the same methodology to align, scale, and identify the relevant facial points among the computer-generated human faces (Testing Phase 2) in WebMorph; however, instead of measuring the vertical distance using the highest point of the eyelids, we opted for the mid-brow (Hehman et al., 2015). One coder (AMM) measured all photos (a total of 50 photos, each of a unique identity), and a second coder blind to the hypotheses independently measured a randomly selected 20% of these (10 photos total). Inter-rater reliability for these measurements was high ($r(8) = 0.91$, $p < 0.001$, 95% CI [0.67, 0.98]). For the real human faces (Testing Phase 3), we used the fWHR measurements provided in the norming data and codebook for the Chicago Face Database (Ma et al., 2015).

3.2.5 | Data analysis

Data analyses were conducted in R v4.0.3 (R Core Team, 2020) using the "glmer" function available in the *lme4* package (Bates et al., 2015) and the "ANOVA" function in the *stats* package (R Core Team, 2020). We analyzed the likelihood of categorizing a face as dominant on testing trials that included unfamiliar conspecific (Testing Phase 1), computer-generated human (Testing Phase 2), and real human (Testing Phase 3) faces. Within each phase, we isolated the probe trials from non-probe trials (i.e., familiar conspecific faces) to assess the trials of interest. For each testing phase, we built separate mixed-effect binary logistic regression models to assess the fixed effects of actual rated dominance (0 = nondominant, 1 = dominant) and the standardized continuous variable of fWHR on the outcome variable of dominance categorization (0 = nondominant, 1 = dominant) for both male and female subjects (Bonferroni corrections were conducted to account for multiple comparisons). The form of the model is as follows:

$$\text{Dominance Categorization} \sim \text{Rated Dominance} + \text{fWHR} + (1|\ \text{Subject}).$$

Furthermore, within each model, we included subject identity as a random effects variable (i.e., grouping variable) to account for the different baseline rates of the outcome. We also compared the fit of each model that we built to that of a null model consisting of only the intercept and random effects terms:

$$\text{Dominance Categorization} \sim 1 + (1|\ \text{Subject}).$$

Aggregated models controlling for the age of the subjects are presented in Supplement 1 (Supporting Information S1: Tables S1–S3).

Additionally, to determine whether monkeys were simply trained to categorize unfamiliar faces as dominant, we calculated kappa coefficients (k) to assess whether they were choosing one category over the other at greater than expected values given how often a particular type of face (i.e., dominant and nondominant) appeared for both probe (i.e., unfamiliar conspecific, computer-generated human, and real human) and non-probe (i.e., familiar conspecific) trials. In this way, we could evaluate and compare the distribution of erroneous responses when categorizing unfamiliar and known/trained faces to determine if a systematic bias was present across one or both types of trials for males and/or females. The aggregated and individual summaries of the categorization pattern and error distribution for each testing phase are presented in Supplement 2 (Supporting Information S1: Tables S4–S26).

4 | RESULTS

4.1 | Training

Twenty-two capuchins (6 males and 16 females) completed training and moved onto testing. Due to the logistical constraints of obtaining stimuli, the study lacks a proper transfer set and has fewer exemplars than is recommended for capuchins to show full transfer (~128 stimuli; Wright et al., 2021); nonetheless, the median number of trials to 80% criterion across all 78 training stimuli dropped from 7535 trials in Training Phase 1 ($\mathrm{Mdn}_{\mathrm{Train1}} = 7535$ trials, IQR: 5315–14,596.75 trials, range: 1555–26,198 trials) to 2215 trials in Training Phase 2 ($\mathrm{Mdn}_{\mathrm{Train2}} = 2215$ trials, IQR: 1474.25–3794 trials, range: 804–16,043 trials); suggesting that capuchins likely abstracted the categorization rule to a reasonable degree.

To better determine the degree to which the monkeys generalized the categories that they learned to the new set of stimuli rather than re-learning the task in Training Phase 2, we compared overall accuracy between the first 50 trials of Training Phase 2 and the last 50 trials of Training Phase 1. Accordingly, monkeys were performing at 72% accuracy (above chance) for the first 50 trials of Training Phase 2 relative to 95% accuracy for the last 50 trials of Training Phase 1, indicating that they only partially abstracted (training performance > transfer performance > chance (50%)) the categorization rule to the novel set of stimuli (see Katz et al., 2007). Additional supplemental analyses (see Supplement 2, Supporting Information S1: Table S27 and Figure S1) show that overall performance increased across sessions of Training Phase 2—with all monkeys reaching criterion by session 15. Finally, overall categorization

TABLE 2 Summary of the mixed effects binary logistic regression analyses using unfamiliar conspecific faces for variables predicting probe trial dominance categorization among (A) male subjects and (B) female subjects.

A: Male subjects ($N = 6$; 143 probe trials per subject)					
Predictor	**b**	**SE**	**e^b**	**$1/e^b$**	**95% CI**
Intercept	2.01^{***}	–	–	–	–
Rated dominance	0.20	0.31	1.22	–	[−0.41, 0.83]
fWHR	0.57^{***}	0.16	1.77	–	[0.26, 0.89]
Marginal R^2/ conditional R^2	0.075/0.395				
B: Female Subjects ($N = 16$; 142–144 probe trials per subject)					
Predictor	**b**	**SE**	**e^b**	**$1/e^b$**	**95% CI**
Intercept	1.00^{***}	–	–	–	–
Rated dominance	0.28	0.15	1.32	–	[−0.01, 0.56]
fWHR	0.93^{***}	0.08	2.54	–	[0.77, 1.09]
Marginal R^2/ conditional R^2	0.207/0.358				

Note: Rated dominance: nondominant target is the reference category. e^b: odds ratio; $1/e^b$: inverse odds ratio.

(A) Full versus null model: $\chi^2(2) = 35.43$, $p < 0.001$.

(B) Full versus null model: $\chi^2(2) = 287.59$, $p < 0.001$.

$^{***}p < 0.001$.

accuracy across the first 50 trials of Training Phase 2 was influenced by the fWHR of the familiar conspecific faces, such that the likelihood of a dominant category response being accurate was significantly greater as fWHR *increased*, while the likelihood of a nondominant category response being accurate was greater as fWHR *decreased* (see Supplement 3, Supporting Information S1: Table S28 and Figure S2).

4.2 | Testing Phase 1 (unfamiliar conspecific faces)

4.2.1 | Male subjects

Six male capuchins completed Testing Phase 1. fWHR was a significant predictor of dominance categorization ($b = 0.57$, SE = 0.16, $e^b = 1.77$, $p < 0.001$, 95% CI [0.26, 0.89]; Table 2A), such that as fWHR increased, males were more likely to categorize a face as dominant. Externally rated dominance (i.e., real-world rank) was not a significant predictor of dominance categorization (lines are close together in Figure 3a). Instead, male subjects had a propensity to categorize any face as dominant, regardless of externally rated dominance (83% of trials), despite being accurate for known/trained faces, indicating that this dominance bias is specific to the unknown/untrained images (kappa coefficients for probe ($k = 0.10$; $N_{\text{Trials}} = 858$) and non-probe ($k = 0.87$; $N_{\text{Trials}} = 5142$) trials, along with the error rates, are presented in Table 3A).

4.2.2 | Female subjects

Sixteen female capuchins completed Testing Phase 1. As with males, fWHR was a significant predictor of dominance categorization ($b = 0.93$, SE = 0.08, $e^b = 2.54$, $p < 0.001$, 95% CI [0.77, 1.09]; Table 2B), such that females were more likely to categorize a face as dominant as fWHR increased (Figure 3b). Again, like males, externally rated dominance was not a significant predictor of dominance categorization, with females categorizing faces as dominant at a high rate (71% of trials) in probe trials despite being accurate for known/trained faces (kappa coefficients for probe ($k = 0.19$; $N_{\text{Trials}} = 2286$) and non-probe ($k = 0.85$; $N_{\text{Trials}} = 13{,}714$) trials, along with the error rates, are presented in Table 3B).

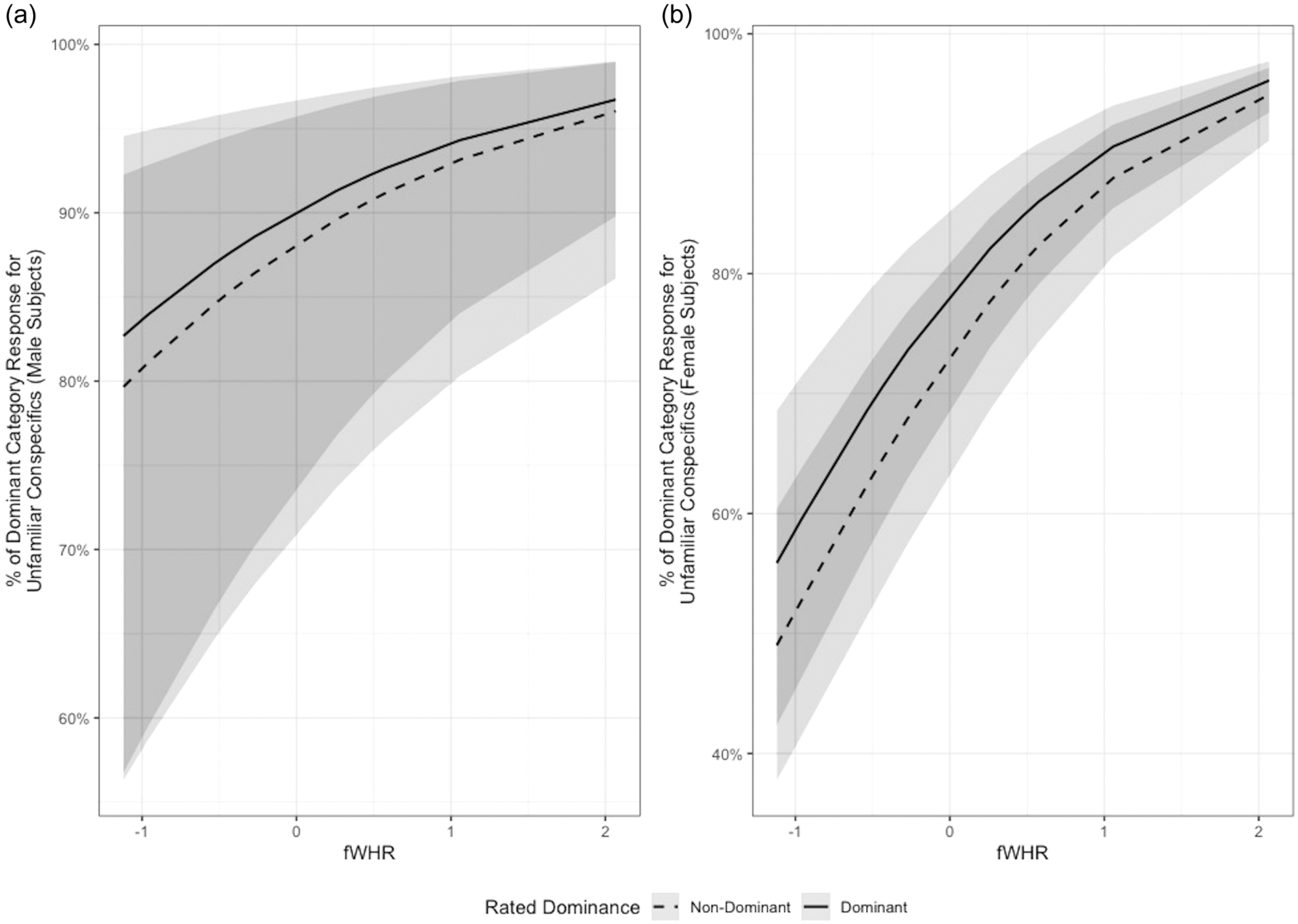

FIGURE 3 Percentage of dominant category responses made by (a) male subjects and (b) female subjects according to the rated dominance (dominant = solid line, nondominant = dashed line) and standardized fWHR measurements of the unfamiliar conspecific target faces. Gray-shaded areas represent 95% CI.

TABLE 3 Summary of the categorization pattern and error distribution for probe (i.e., unfamiliar conspecific) and non-probe (i.e., familiar conspecific) trials presented in Testing Phase 1 among (A) male subjects and (B) female subjects.

A: Male subjects (*N* = 6)								
	Probe trials				*Non-probe trials*			
	Dominance categorization				Dominance categorization			
External Rating	Dominant	Nondominant	Total rating	% Errors	Dominant	Nondominant	Total rating	% Errors
Dominant	293	33	326	0.10	2401	171	2572	0.07
Nondominant	415	117	532	0.78	169	2401	2570	0.07
Total cats.	708	150	**858**		2570	2572	**5142**	
% Errors	0.59	0.22			0.07	0.07		
B: Female subjects (*N* = 16)								
	Probe trials				*Non-probe trials*			
	Dominance categorization				Dominance categorization			
External rating	Dominant	Nondominant	Total rating	% Errors	Dominant	Nondominant	Total rating	% Errors
Dominant	733	129	862	0.15	6258	600	6858	0.09
Nondominant	897	527	1424	0.63	429	6427	6856	0.06
Total cats.	1630	656	**2286**		6687	7027	**13,714**	
% Errors	0.55	0.20			0.06	0.09		

Note: Total cats.: total number of categorizations made according to dominance category; % Errors (row): percentage of errors made according to the dominance category chosen relative to the external rating of the stimulus (e.g., proportion of dominant responses made for nondominant stimuli); Total rating: total number of trials containing each type of externally rated stimulus; % Errors (column): percentage of errors made according to the type of externally rated stimulus being categorized relative to the dominance category chosen (e.g., proportion of dominant stimuli categorized as nondominant). Total number of probe and non-probe trials presented in Testing Phase 1 are bolded.

TABLE 4 Summary of the mixed effects binary logistic regression analysis using computer-generated human faces for variables predicting probe trial dominance categorization among (A) male subjects and (B) female subjects.

A: Male subjects (*N* = 6; 142–143 probe trials per subject)					
Predictor	b	SE	e^b	$1/e^b$	95% CI
Intercept	0.91	–	–	–	–
Rated dominance	−0.07	0.17	0.93	1.07	[−0.40, 0.26]
fWHR	0.18*	0.08	1.20	–	[0.02, 0.34]
Marginal R^2/ conditional R^2	0.006/0.340				
B: Female subjects (*N* = 15; 142–144 probe trials per subject)					
Predictor	b	SE	e^b	$1/e^b$	95% CI
Intercept	−1.42*	–	–	–	–
Rated dominance	0.18	0.14	1.19	–	[−0.10, 0.46]
fWHR	−0.12	0.07	0.89	1.13	[−0.26, 0.02]
Marginal R^2/ conditional R^2	0.002/0.607				

Note: Rated dominance: nondominant target is the reference category. e^b: odds ratio; $1/e^b$: inverse odds ratio.

(A) Full versus null model: $\chi^2(2) = 4.90$, $p = 0.09$.

(B) Full versus null model: $\chi^2(2) = 3.32$, $p = 0.19$.

*$p < 0.05$.

4.3 | Testing Phase 2 (computer-generated human faces)

4.3.1 | Male subjects

Six male capuchins completed Testing Phase 2. The overall model predicting dominance categorization was not significant, suggesting that our model predictors do not reliably influence dominance categorization (Table 4A). However, consistent with Testing Phase 1, males were slightly more likely to categorize a face as dominant as fWHR increased (Figure 4a). Furthermore, the kappa coefficients for probe ($k = 0.01$; $N_{\text{Trials}} = 856$) and non-probe ($k = 0.85$; $N_{\text{Trials}} = 5144$) trials suggest that males continued to show a dominance bias (64% of trials) when categorizing computer-generated human faces despite maintaining high accuracy rates for known/trained faces (Table 5A).

4.3.2 | Female subjects

Fifteen female capuchins completed Testing Phase 2. The overall model predicting dominance categorization was not significant, suggesting that our model predictors do not reliably influence dominance categorization (Table 4B). Inconsistent with previous

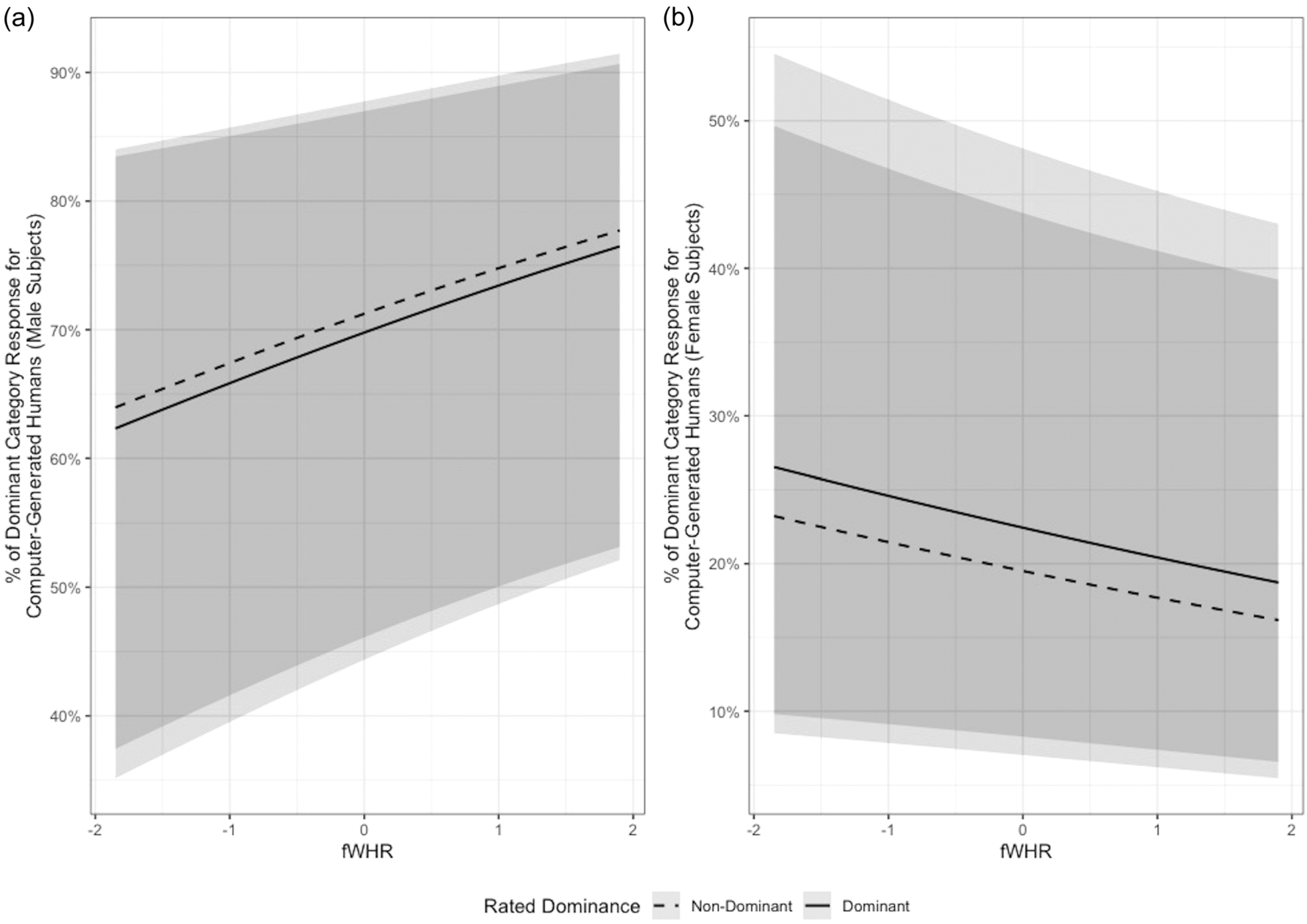

FIGURE 4 Percentage of dominant category responses made by (a) male subjects and (b) female subjects according to the rated dominance (dominant = solid line, nondominant = dashed line) and standardized fWHR measurements of the computer-generated human target faces. Gray-shaded areas represent 95% CI.

results, however, the likelihood of categorizing a face as dominant decreases slightly as the fWHR of the computer-generated human faces increases (Figure 4b). Accordingly, the kappa coefficients for probe ($k = 0.00$; $N_{\text{Trials}} = 2147$) and non-probe ($k = 0.88$; $N_{\text{Trials}} = 12{,}853$) trials indicate that females shifted their bias toward non-dominance (69% of trials) when categorizing computer-generated human faces, again despite maintaining great accuracy for known/trained faces (Table 5B).

4.4 | Testing Phase 3 (real human faces)

4.4.1 | Male subjects

Five male capuchins completed Testing Phase 3. The overall model predicting dominance categorization was not significant, suggesting that our model predictors do not reliably influence dominance categorization (Table 6A). Interestingly, and in line with the results for females from Testing Phase 2, the likelihood of categorizing a face as dominant now, if anything, decreased as the fWHR of the real human faces increased (Figure 5a). Accordingly, the kappa coefficients for probe ($k = 0.08$; $N_{\text{Trials}} = 715$) and non-probe ($k = 0.81$; $N_{\text{Trials}} = 4285$) trials suggest that males shifted their bias towards nondominance (68% of trials) when categorizing real human faces while sustaining their accuracy for known/trained faces (Table 7A).

4.4.2 | Female subjects

Thirteen female capuchins completed Testing Phase 3. The results suggest that fWHR was a significant predictor of dominance categorization ($b = -0.22$, SE = 0.07, $e^b = 0.81$, $1/e^b = 1.24$, $p < 0.01$, 95% CI [−0.36, −0.07]; Table 6B). Consistent with the findings from Testing Phase 2, as fWHR increased, females were less likely to categorize a face as dominant (Figure 5b). Moreover, the kappa coefficients for probe ($k = -0.04$; $N_{\text{Trials}} = 1852$) and non-probe ($k = 0.84$; $N_{\text{Trials}} = 11{,}148$) trials suggest that females continued to show a bias towards nondominance (69% of trials) when categorizing real human faces despite remaining accurate for known/trained faces (Table 7B).

TABLE 5 Summary of the categorization pattern and error distribution for probe (i.e., computer-generated human) and non-probe (i.e., familiar conspecific) trials presented in Testing Phase 2 among (A) male subjects and (B) female subjects.

A: Male subjects ($N = 6$)

	Probe trials				*Non-probe trials*			
	Dominance categorization				Dominance categorization			
External rating	Dominant	Nondominant	Total rating	% Errors	Dominant	Nondominant	Total rating	% Errors
Dominant	272	147	419	0.35	2466	105	2571	0.04
Nondominant	278	159	437	0.64	285	2288	2573	0.11
Total cats.	550	306	**856**		2751	2393	**5144**	
% Errors	0.51	0.48			0.10	0.04		

B: Female subjects ($N = 15$)

	Probe trials				*Non-probe trials*			
	Dominance categorization				Dominance categorization			
External rating	Dominant	Nondominant	Total rating	% Errors	Dominant	Nondominant	Total rating	% Errors
Dominant	324	710	1034	0.69	6055	374	6429	0.06
Nondominant	344	769	1113	0.31	393	6031	6424	0.06
Total cats.	668	1479	**2147**		6448	6405	**12,853**	
% Errors	0.51	0.48			0.06	0.06		

Note: Total cats.: total number of categorizations made according to dominance category; % Errors (row): percentage of errors made according to the dominance category chosen relative to the external rating of the stimulus (e.g., proportion of dominant responses made for nondominant stimuli); Total rating: total number of trials containing each type of externally rated stimulus; % Errors (column): percentage of errors made according to the type of externally rated stimulus being categorized relative to the dominance category chosen (e.g., proportion of dominant stimuli categorized as nondominant). Total number of probe and non-probe trials presented in Testing Phase 2 are bolded.

TABLE 6 Summary of the mixed effects binary logistic regression analysis using real human faces for variables predicting probe trial dominance categorization among (A) male subjects and (B) female subjects.

A: Male subjects ($N = 5$; 143 probe trials per subject)

Predictor	b	SE	e^b	$1/e^b$	95% CI
Intercept	−1.13	–	–	–	–
Rated dominance	0.21	0.21	1.23	–	[−0.20, 0.61]
fWHR	−0.08	0.11	0.93	1.08	[−0.30, 0.14]
Marginal R^2/ conditional R^2	0.002/0.362				

B: Female subjects ($N = 13$; 140–143 probe trials per subject)

Predictor	b	SE	e^b	$1/e^b$	95% CI
Intercept	−1.28*	–	–	–	–
Rated dominance	−0.20	0.14	0.82	1.22	[−0.48, 0.08]
fWHR	−0.22**	0.07	0.81	1.24	[−0.36, −0.07]
Marginal R^2/ conditional R^2	0.010/0.528				

Note: Rated dominance: nondominant target is the reference category. e^b: odds ratio; $1/e^b$: inverse odds ratio.

(A) Full versus null model: $\chi^2(2) = 1.12$, $p = 0.57$.

(B) Full versus null model: $\chi^2(2) = 14.49$, $p < 0.001$.

$^*p < 0.05$; $^{**}p < 0.01$.

5 | DISCUSSION

In this study, we found that monkeys trained to categorize familiar conspecific faces as dominant or nondominant partially generalized these categories to unfamiliar capuchin faces, despite the small training set, suggesting that subjects were able to extract information about dominance from unknown conspecifics. Their choices also mirrored the fWHR of the sample faces, suggesting that this cue, which is known to correlate with dominance, may be the (or one of the) mechanisms the monkeys were using to make this discrimination. Notably, however, capuchins showed a strong bias towards categorizing unknown individuals as dominant while maintaining high accuracy on known faces, suggesting that there is an overall bias towards assuming any stranger is a potential threat (i.e., dominant). Finally, monkeys did not show the same categorization for either unfamiliar computer-generated or unfamiliar real human faces, suggesting that this ability does not generalize to the evaluation of other species' faces, despite these monkeys' extensive familiarity with humans. Finally, males and females showed varying systematic biases during categorization of probe, but not non-probe, trials across each testing phase. Below, we discuss each of these in turn.

Capuchin monkeys were able to learn to categorize based on dominance (dominant vs. nondominant) and then generalize it to unfamiliar conspecifics based entirely on static images (i.e., photographs). This extends previous research on dominance categorization in nonhuman primates, which found that monkeys could make similar

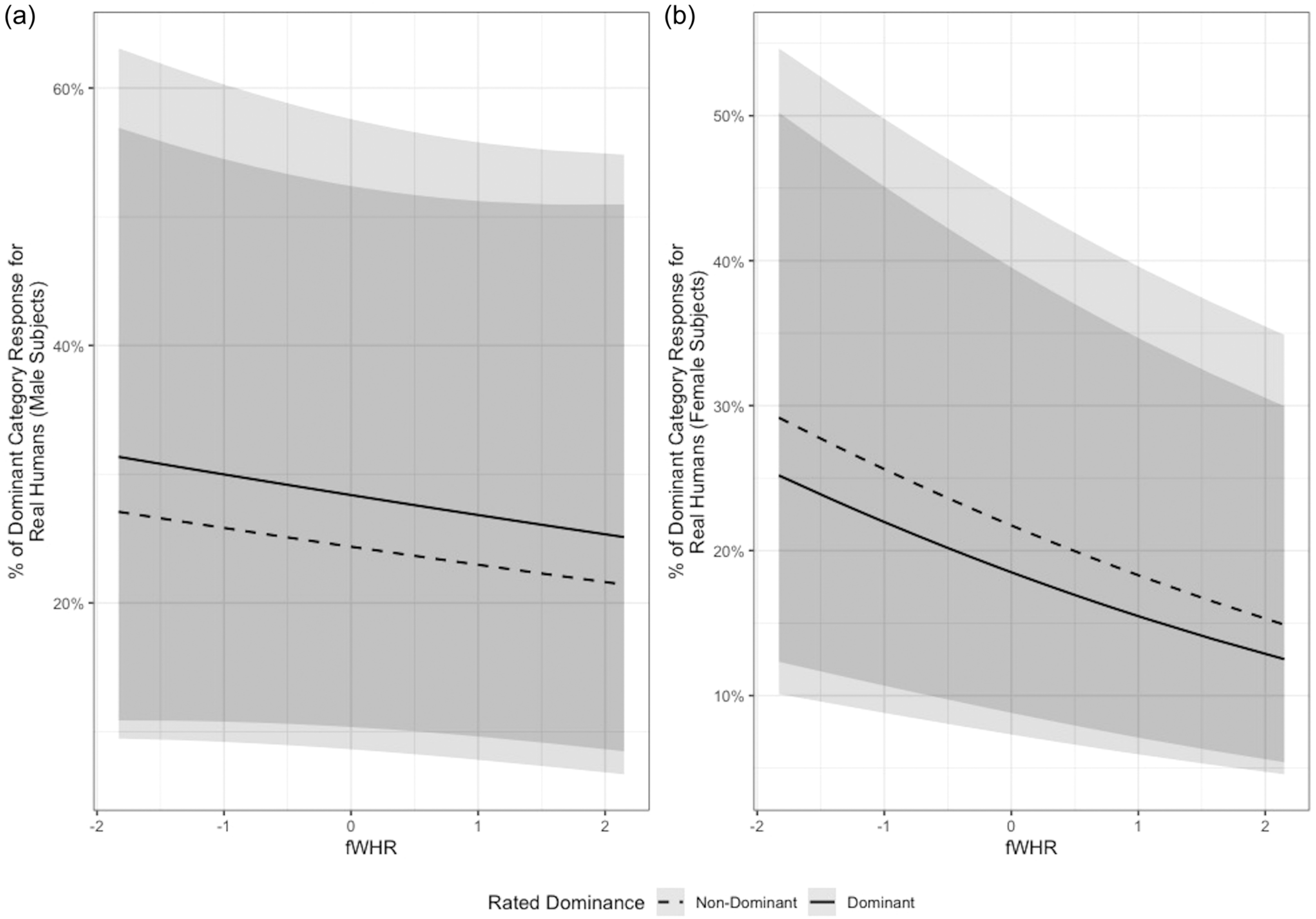

FIGURE 5 Percentage of dominant category responses made by (a) male subjects and (b) female subjects according to the rated dominance (dominant = solid line, nondominant = dashed line) and standardized fWHR measurements of the real human target faces. Gray-shaded areas represent 95% CI.

categorizations based on social interactions (Bovet & Washburn, 2003; Paxton et al., 2010), and suggests that a static image, perhaps analogous to a glance, is sufficient to convey at least some dominance information. Indeed, the ability to rapidly assess stranger monkeys to determine whether they pose a threat to the current alpha, which is a risk of overthrow for the male and infanticide for the female (Benitez et al., 2021), would benefit capuchins.

An obvious question is how they did so. fWHR in capuchins correlates with dominance (Lefevre et al., 2014), or possibly with the capacity to harm, which would indicate threat (i.e., the wider jaw may be associated with greater bite strength; McArthur & Apatow, 1984). Thus, we explored the correlation between our monkeys' choices and fWHR and found that both males and females appeared to be using it as a cue, at least for unfamiliar conspecifics, in the current study. This fits with previous work finding that both fWHR and other cues are associated with perceived formidability in humans (Todorov, 2017; Toscano et al., 2014; Zilioli et al., 2015) and work in other species on cues (i.e., plumage patch size in golden-crowned sparrows; *Zonotrichia atricapilla*; Chaine et al., 2018) that are used to guide behavior during social encounters (Rohwer, 1975, 1977, 1982). Of course, there are potentially other factors, for instance, those associated with secondary sexual characteristics (which include fWHR, but also enhanced tufts and changes in facial shape), that could have been used as cues as well. It is also possible that our subjects may have relied on other aspects of the stimuli, such as the maximum width and/or height of the image, or the ratio between the two. However, monkeys consistently categorized images of the same individual with respect to dominance for 10 of 11 unfamiliar conspecifics, suggesting that other variations across photos were not influencing responses in most cases.

One of our more interesting results is that in the non-differentially reinforced probe trials, capuchins displayed a systematic bias towards judging the faces of strangers as dominant, regardless of the external dominance rating of the target faces, while maintaining high accuracy rates for familiar faces. Their continued accuracy in the baseline trials suggests that this was not an artifact of training (i.e., we did not simply train them to always choose the dominant icon). From an ecological standpoint, this makes sense. Male capuchins should be sensitive to other males as potential threats to their dominant

TABLE 7 Summary of the categorization pattern and error distribution for probe (i.e., real human) and non-probe (i.e., familiar conspecific) trials presented in Testing Phase 3 among (A) male subjects and (B) female subjects.

A: Male subjects (*N* = 5)

	Probe trials				*Non-probe trials*			
	Dominance categorization				Dominance categorization			
External rating	Dominant	Nondominant	Total rating	% Errors	Dominant	Nondominant	Total rating	% Errors
Dominant	120	210	330	0.64	1994	156	2150	0.07
Nondominant	111	274	385	0.29	261	1874	2135	0.12
Total cats.	231	484	**715**		2255	2030	**4285**	
% Errors	0.48	0.43			0.12	0.08		

B: Female subjects (*N* = 13)

	Probe trials				*Non-probe trials*			
	Dominance categorization				Dominance categorization			
External rating	Dominant	Nondominant	Total rating	% Errors	Dominant	Nondominant	Total rating	% Errors
Dominant	259	647	906	0.71	5021	555	5576	0.10
Nondominant	306	640	946	0.32	350	5222	5572	0.06
Total cats.	565	1287	**1852**		5371	5777	**11,148**	
% Errors	0.54	0.50			0.07	0.10		

Note: Total cats.: total number of categorizations made according to dominance category; % Errors (row): percentage of errors made according to the dominance category chosen relative to the external rating of the stimulus (e.g., proportion of dominant responses made for nondominant stimuli); Total rating: total number of trials containing each type of externally rated stimulus; % Errors (column): Percentage of errors made according to the type of externally rated stimulus being categorized relative to the dominance category chosen (e.g., proportion of dominant stimuli categorized as nondominant). Total number of probe and non-probe trials presented in Testing Phase 3 are bolded.

position or individuals who must be overcome to move up the hierarchy, thus the safest error is to over-estimate dominance. Relatedly, females risk infanticide if a new male takes over the group, so they should be cautious of strangers. None of our females were pregnant or had young offspring, however; thus, future work could further explore whether this finding reflects a generalized response, rather than one based on immediate circumstances, for instance, by testing whether female capuchins who had infants (and thus were at risk of infanticide) responded differently than the females in our study. These results are also similar to human females' choices, who tend to rate unknown male faces as more dominant overall (e.g., Kleider-Offutt et al., 2021). Humans are also more sensitive to faces that are more similar to their own (Watkins & Jones, 2012; Watkins et al., 2010), so it would be interesting to see how capuchin females respond to female faces in future work.

Unfortunately, due to the challenges of acquiring images, we lacked sufficient data for a full transfer set (Wright et al., 2021), and the lack of full transfer in the first 50 trials between training phases indicates that the categories were only partially generalized by the rule. In addition, of course, we do not know what rule they actually learned. While we chose the photos based on dominance (i.e., who was alpha male or not), it is possible that we trained them to discriminate based on fWHR (or some other cue) and that this, to the monkeys, had nothing to do with dominance. Nonetheless, the fact that fWHR correlates with dominance in other contexts suggests that this may be a way that capuchins, and presumably other primates, extrapolate information about dominance from unknown individuals, likely by learning to associate individuals with certain facial ratios with dominant behaviors or relative hierarchical positions. Nevertheless, again, our small sample size for both the number of subjects and the number of stimuli limits our ability to fully extrapolate our findings.

Despite their success with conspecifics, monkeys did not reliably categorize human faces. On the one hand, this is not surprising; after all, humans are not conspecifics, and there is no obvious, ecologically valid reason that capuchin monkeys should have evolved to respond to human dominance cues as they do to conspecific ones. On the other hand, if our specific monkeys were to generalize to another species, one might expect that it would be humans, because our faces are structurally similar, share at least some of the same characteristics of dominance (e.g., fWHR), and, perhaps most importantly, our population of capuchins is extremely familiar with humans. As a captive group, they interact closely with multiple humans on a daily basis and have interacted with dozens of different humans over the course of their lives. Indeed, there is evidence that some captive primates are equally good at recognizing human faces as conspecific faces (Myowa-Yamakoshi & Tomonaga, 2001; Yamaguchi et al., 2003).

So, why did they not reliably use facial cues? We see several possible explanations that we cannot disentangle with these data (and that are not mutually exclusive). First, perhaps they are only attuned to conspecific faces and so did not pay sufficient attention to the human faces to discriminate them. Second, perhaps the facial features (e.g., nose width, lip fullness, and skin properties/reflectance; Kleider-Offutt et al., 2021) that discriminate dominant from nondominant humans are different from those that discriminate dominant from nondominant capuchins, and so the monkeys did not know what cue to use or were using a cue that was appropriate with the capuchins but was not with the humans. However, fWHR, which they appeared to use for conspecifics, was also relevant for humans, so their failure to do so suggests that they did not generalize this specific cue across species. Third, perhaps the capuchins have not had the opportunity to see sufficient dominance interactions to learn what a dominant human looks like, so even with the shared features, they failed to understand how to pick a dominant human. This, however, implicitly presupposes that they understood their task to choose the dominant, rather than to choose based on fWHR, which may not be valid. Fourth, and perhaps most likely, capuchins *underestimated* dominance in humans, despite overestimating it in stranger capuchins. It may be that humans are rarely a threat to them, except in a few cases that come with specific other cues (i.e., the veterinarian, who may wear different clothes or bring specific equipment); thus, their failure may have been a combination of failure to pay as much attention to the human faces and, when they did, a failure to see them as potential threats (which would, presumably, also have reduced attention). Of course, given our subjects' familiarity with humans, they may have categorized based on their own personal experiences with a limited subset of humans. Indeed, the same could also be true for capuchin faces, as subjects have (even more) limited exposure to different capuchins. For future research, it would be interesting to see how they respond to unfamiliar allospecifics. We did not do so to avoid novelty confounds, but this could help to address this question.

Our results suggest that capuchins may have the ability to rapidly assess dominance from a static image, possibly based on fWHR, which previous work suggests is a reliable cue to dominance in other primates. That being said, capuchins also showed a strong bias toward assuming strangers were dominant. While there is an ecological explanation for this, future work is needed to determine the degree to which this holds in other contexts, for other faces (i.e., females, juveniles), and for other species. We suspect that additional investigations would uncover similar mechanisms in other highly social species and encourage further research to determine whether the cues to dominance, too, are conserved and the degree to which the contexts that influence dominance assessments may differ across species. Better understanding this will help clarify how individuals are able to rapidly make decisions to successfully navigate their social worlds.

AUTHOR CONTRIBUTIONS

Ashley M. Meacham: Data curation (equal); formal analysis (lead); investigation (supporting); methodology (supporting); resources (equal); visualization (lead); writing—original draft (lead); writing—review and editing (equal). **Meghan J. Sosnowski**: Data curation (equal); formal analysis (supporting); investigation (equal); methodology (equal); project administration (lead); resources (equal); software (lead); visualization (supporting); writing—original draft (equal); writing—review and editing (equal). **Heather M. Kleider-Offutt**: Conceptualization (equal); methodology (equal); supervision (lead); writing—original draft (supporting); writing—review and editing (equal). **Sarah F. Brosnan**: Conceptualization (equal); methodology (equal); resources (equal); supervision (lead); writing—original draft (equal); writing—review and editing (equal).

ACKNOWLEDGMENTS

We thank Kate Talbot, Kelly Leverett, and Jhonatan Saldaña Santisteban for allowing us to use additional photos of the LRC monkeys as stimuli, as well as colleagues at the National Institutes of Health, Bucknell University, and Yerkes National Primate Research Center for photos of their capuchin monkeys that we used as unfamiliar conspecifics. We thank Dr. Alexander Todorov and Drs. Debbie Ma, Joshua Correll, and Bernd Wittenbrink for granting permission to use their computer-generated and real human face stimuli, respectively, in our study. We also thank members of the CEBUS lab for aid in data collection, along with Sara Sims, Hadley Mueller-Hill, and Rachel Heiter for their part in obtaining facial measurements. Finally, we thank the Division of Animal Resources staff for their dedicated and compassionate care of our nonhuman primate subjects. Ashley M. Meacham was supported by the Georgia State University 2CI Primate Social Cognition, Evolution, and Behavior Doctoral Fellowship during this project. Meghan J. Sosnowski was supported, in part, by the Georgia State University Brains & Behavior Doctoral Fellowship. Sarah F. Brosnan was supported, in part, by NSF grants SES 1919305 and IBSS 2135621.

CONFLICTS OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Ashley M. Meacham http://orcid.org/0000-0001-5556-993X
Meghan J. Sosnowski http://orcid.org/0000-0002-6690-5086
Sarah F. Brosnan http://orcid.org/0000-0002-5117-6706

REFERENCES

Baldovino, M. C., & Di Bitetti, M. S. (2007). Allonursing in tufted capuchin monkeys (*Cebus nigritus*): Milk or pacifier? *Folia Primatologica*, *79*(2), 79–92. https://doi.org/10.1159/000108780

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Batres, C., Re, D. E., & Perrett, D. I. (2015). Influence of perceived height, masculinity, and age on each other and on perceptions of dominance in male faces. *Perception*, *44*(11), 1293–1309. https://doi.org/10.1177/0301006615596898

Benitez, M., Brosnan, S. F., & Fragaszy, D. M. (2021). Behavioral biology of capuchin monkeys (genera *Cebus* and *Sapajus*). In K. Coleman & S. J. Schapiro (Eds.), *The behavioral biology of laboratory animals*. CRC Press.

di Bitetti, M. S. (1997). Evidence for an important social role of allogrooming in a platyrrhine primate. *Animal Behaviour*, *54*(1), 199–211. https://doi.org/10.1006/anbe.1996.0416

Boehm, C. (1999). *Hierarchy in the forest: The evolution of egalitarian behavior*. Harvard University Press.

Bovet, D., & Washburn, D. A. (2003). Rhesus macaques (*Macaca mulatta*) categorize unknown conspecifics according to their dominance relations. *Journal of Comparative Psychology*, *117*(4), 400–405. https://doi.org/10.1037/0735-7036.117.4.400

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. https://doi.org/10.1098/rspb.2008.0873

Chaine, A. S., Shizuka, D., Block, T. A., Zhang, L., & Lyon, B. E. (2018). Manipulating badges of status only fools strangers. *Ecology Letters*, *21*, 1477–1485. https://doi.org/10.1111/ele.13128

Core Team, R. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

DeBruine, L. M. (2018). *debruine/webmorph: Beta release 2 (version v0.0.0.9001)*. Zenodo. https://doi.org/10.5281/zenodo.1162670

Dixson, B. J. W. (2017). Facial width to height ratio and dominance. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 1–4). Springer International. https://doi.org/10.1007/978-3-319-19650-3_1419

Durkee, P. K., & Ayers, J. D. (2021). Is facial width-to-height ratio reliably associated with social inferences. *Evolution and Human Behavior*, *42*(6), 583–592. https://doi.org/10.1016/j.evolhumbehav.2021.06.003

Evans, T. A., Beran, M. J., Chan, B., Klein, E. D., & Menzel, C. R. (2008). An efficient computerized testing method for the capuchin monkey (*Cebus apella*): Adaptation of the LRC-CTS to a socially housed non-human primate species. *Behavior Research Methods*, *40*(2), 590–596. https://doi.org/10.3758/BRM.40.2.590

Fragaszy, D. M., Visalberghi, E., & Fedigan, L. M. (2004). *The complete capuchin: The biology of the genus Cebus*. Cambridge University Press.

Geniole, S. N., Denson, T. F., Dixson, B. J., Carré, J. M., & McCormick, C. M. (2015). Evidence from meta-analyses of the facial width-to-height ratio as an evolved cue of threat. *PLoS One*, *10*(7), e0132726. https://doi.org/10.1371/journal.pone.0132726

Gray, J. P., & Wolfe, L. D. (1980). Height and sexual dimorphism of stature among human societies. *American Journal of Physical Anthropology*, *53*(3), 441–456. https://doi.org/10.1002/ajpa.1330530314

Han, C., Kandrik, M., Hahn, A. C., Fisher, C. I., Feinberg, D. R., Holzleitner, I. J., DeBruine, L. M., & Jones, B. C. (2017). Interrelationships among men's threat potential, facial dominance, and vocal dominance. *Evolutionary Psychology*, *15*(1), 1–4. https://doi.org/10.1177/1474704917697332

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*(8), 1123–1134. https://doi.org/10.1177/0146167215591495

Janson, C. H., & van Schaik, C. P. (2000). The behavioural ecology of infanticide by males. In C. P. van Schaik & C. H. Janson (Eds.), *Infanticide by males and its implications* (pp. 469–494). Cambridge University Press.

Katz, J. S., Wright, A. A., & Bodily, K. D. (2007). Issues in the comparative cognition of abstract-concept learning. *Comparative cognition & behavior reviews*, *2*, 79–92. https://doi.org/10.3819/ccbr.2008.20005

Keating, C. F., Mazur, A., & Segall, M. H. (1981). A cross-cultural exploration of physiognomic traits of dominance and happiness. *Ethology and Sociobiology*, *2*(1), 41–48. https://doi.org/10.1016/0162-3095(81)90021-2

Kleider-Offutt, H., Meacham, A. M., Branum-Martin, L., & Capodanno, M. (2021). What's in a face? The role of facial features in ratings of dominance, threat, and stereotypicality. *Cognitive Research: Principles and Implications*, *6*, 53. https://doi.org/10.1186/s41235-021-00319-9

Lefevre, C. E., Wilson, V. A. D., Morton, F. B., Brosnan, S. F., Paukner, A., & Bates, T. C. (2014). Facial width-to-height ratio relates to alpha status and assertive personality in capuchin monkeys. *PLoS One*, *9*(4), e93369. https://doi.org/10.1371/journal.pone.0093369

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*, 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

McArthur, L. Z., & Apatow, K. (1984). Impressions of baby-faced adults. *Social Cognition*, *2*(4), 315–342. https://doi.org/10.1521/soco.1984.2.4.315

Mealey, L., Daood, C., & Krage, M. (1996). Enhanced memory for faces of cheaters. *Ethology and Sociobiology*, *17*, 119–128. https://doi.org/10.1016/0162-3095(95)00131-X

Murray, C. M., Mane, S. V., & Pusey, A. E. (2007). Dominance rank influences female space use in wild chimpanzees, *Pan troglodytes*: Towards an ideal despotic distribution. *Animal Behaviour*, *74*(6), 1795–1804. https://doi.org/10.1016/j.anbehav.2007.03.024

Myowa-Yamakoshi, M., & Tomonaga, M. (2001). Development of face recognition in an infant gibbon (*Hylobates agilis*). *Infant Behavior and Development*, *24*(2), 215–227. https://doi.org/10.1016/S0163-6383(01)00076-5

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Palombit, R. A., Cheney, D. L., & Seyfarth, R. M. (2001). Female-female competition for male “friends” in wild chacma baboons (*Papio cynocephalus ursinus*). *Animal Behaviour*, *61*(6), 1159–1171. https://doi.org/10.1006/anbe.2000.1690

Paxton, R., Basile, B. M., Adachi, I., Suzuki, W. A., Wilson, M. E., & Hampton, R. R. (2010). Rhesus monkeys (*Macaca mulatta*) rapidly learn to select dominant individuals in videos of artificial social interactions between unfamiliar conspecifics. *Journal of Comparative Psychology*, *124*(4), 395–401. https://doi.org/10.1037/a0019751

Perry, S., Barrett, H. C., & Manson, J. H. (2004). White-faced capuchin monkeys show triadic awareness in their choice of allies. *Animal Behaviour*, *67*(1), 165–170. https://doi.org/10.1016/j.anbehav.2003.04.005

Pineda, J. A., Sebestyen, G., & Nava, C. (1994). Face recognition as a function of social attention in non-human primates: An ERP study. *Cognitive Brain Research*, *2*, 1–12. https://doi.org/10.1016/0926-6410(94)90015-9

Preuschoft, S., & van Hooff, J. A. R. A. M. (1995). Homologizing primate facial displays: A critical review of methods. *Folia Primatologica*, *65*(3), 121–137. https://doi.org/10.1159/000156878

Rhodes, G., Simmons, L. W., & Peters, M. (2005). Attractiveness and sexual behavior: Does attractiveness enhance mating success? *Evolution and Human Behavior*, *26*(2), 186–201. https://doi.org/10.1016/j.evolhumbehav.2004.08.014

Rohwer, S. (1975). The social significance of avian winter plumage variability. *Evolution*, *29*(4), 593–610. https://doi.org/10.2307/2407071

Rohwer, S. (1977). Status signaling in Harris Sparrows: Some experiments in deception. *Behaviour*, *61*(1/2), 107–129. https://www.jstor.org/stable/4533813

Rohwer, S. (1982). The evolution of reliable and unreliable badges of fighting ability. *American Zoologist*, *22*(3), 531–546. https://doi.org/10.1093/icb/22.3.531

Silk, J. B., Alberts, S. C., & Altmann, J. (2003). Social bonds of female baboons enhance infant survival. *Science*, *302*(5648), 1231–1234. https://doi.org/10.1126/science.1088580

Silk, J. B., Beehner, J. C., Bergman, T. J., Crockford, C., Engh, A. L., Moscovice, L. R., Wittig, R. M., Seyfarth, R. M., & Cheney, D. L. (2010). Strong and consistent social bonds enhance the longevity of female baboons. *Current Biology*, *20*(15), 1359–1361. https://doi.org/10.1016/j.cub.2010.05.067

Talbot, C. F., Leverett, K. L., & Brosnan, S. F. (2016). Capuchins recognize familiar faces. *Animal Behaviour*, *122*, 37–45. https://doi.org/10.1016/j.anbehav.2016.09.017

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

Todorov, A., & Oh, D. (2021). The structure and perceptual basis of social judgments from faces. *Advances in Experimental Social Psychology*, *63*, 189–245. https://doi.org/10.1016/bs.aesp.2020.11.004

Todorov, A., & Oosterhof, N. (2011). Modeling social perception of faces. *IEEE Signal Processing Magazine*, *28*(2), 117–122. https://doi.org/10.1109/MSP.2010.940006

Toscano, H., Schubert, T. W., & Sell, A. N. (2014). Judgments of dominance from the face track physical strength. *Evolutionary Psychology*, *12*(1), 1–18. https://doi.org/10.1177/147470491401200101

Valentine, K. A., Li, N. P., Penke, L., & Perrett, D. I. (2014). Judging a man by the width of his face: The role of facial ratios and dominance in mate choice at speed-dating events. *Psychological Science*, *25*(3), 806–811. https://doi.org/10.1177/0956797613511823

de Waal, F. B. M. (1982). *Chimpanzee Politics: Power and sex among apes*. The Johns Hopkins University Press.

de Waal, F. B. M. (2000). Attitudinal reciprocity in food sharing among brown capuchin monkeys. *Animal Behaviour*, *60*(2), 253–261. https://doi.org/10.1006/anbe.2000.1471

Watkins, C. D., & Jones, B. C. (2012). Priming men with different contest outcomes modulates their dominance perceptions. *Behavioral Ecology*, *23*(3), 539–543. https://doi.org/10.1093/beheco/arr221

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*(8), 967–971. https://doi.org/10.1016/j.paid.2010.08.006

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100ms exposure to a face. *Psychological Science*, *17*(7), 592–598. https://doi.org/10.1111/j.1467-9280.2006.01750.x

Wilson, V., Lefevre, C. E., Morton, F. B., Brosnan, S. F., Paukner, A., & Bates, T. C. (2014). Personality and facial morphology: Links to assertiveness and neuroticism in capuchins (*Sapajus [Cebus] apella*). *Personality and Individual Differences*, *58*, 89–94. https://doi.org/10.1016/j.paid.2013.10.008

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*(6), 805–814. https://doi.org/10.1002/ajhb.21219

Wright, A. A., Kelly, D. M., & Katz, J. S. (2021). Same/different concept learning by primates and birds. *Learning & Behavior*, *49*, 76–84. https://doi.org/10.3758/s13420-020-00456-z

Yamaguchi, M. K., Kanazawa, S., Tomonaga, M., & Murai, C. (2003). Development of face recognition in humans and Japanese monkeys. In M. Tomonaga, M. Tonaka, & T. Matsuzawa (Eds.), *Cognitive and behavioral development in chimpanzees* (pp. 347–352). Kyoto University Press.

Zebrowitz, L. A. (2017). First impressions from faces. *Current Directions in Psychological Science*, *26*(3), 237–242. https://doi.org/10.1177/0963721416683996

Zilioli, S., Sell, A. N., Stirrat, M., Jagore, J., Vickerman, W., & Watson, N. V. (2015). Face of a fighter: Bizygomatic width as a cue of formidability. *Aggressive Behavior*, *41*(4), 322–330. https://doi.org/10.1002/ab.21544

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Meacham, A. M., Sosnowski, M. J., Kleider-Offutt, H. M., & Brosnan, S. F. (2024). Capuchin monkeys' (*Sapajus [Cebus] apella*) categorization of photos of unknown male conspecifics suggests attention to fWHR and a dominance bias. *American Journal of Primatology*, e23623. https://doi.org/10.1002/ajp.23623