Journal of Nonverbal Behavior (2023) 47:545–568
https://doi.org/10.1007/s10919-023-00443-8

ORIGINAL PAPER

People Watching: Social Perception and the Ensemble Coding of Bodies

Flora Oswald[1,2] · Jason W. Griffin[3] · Max Weisbuch[4] · Reginald B. Adams Jr.[1]

Accepted: 10 September 2023 / Published online: 22 September 2023

Abstract
Bodies are rich and important social stimuli, which we often encounter in the context of social groups. Yet, little attention has been paid to how we process these groups, and what information perceivers might extract from groups of bodies. Drawing from work on the perception of individual bodies, we conducted two studies to test the ability of human observers (college students; N_{total} = 375) to ensemble code (i.e., rapidly extract summary statistics about attributes of stimulus groups) human bodies. Specifically, we examined whether participants extracted summary statistics of lower-level (body mass index, waist-to-chest ratio, and waist-to-hip ratio) and higher-level (emotion, gender) properties from groups of bodies. Participants were relatively accurate in extracting summary statistics for both lower-level and higher-level characteristics from groups of bodies, consistent with the view that visual processes rapidly summarize group characteristics from bodily information.

Keywords body perception · summary perception · people perception · human body · social groups

Introduction

Perceptions of human bodies underlie some of the most pervasive forms of discrimination, including anti-fat prejudice (see Rubino et al., 2020), police harassment of Black men (which varies by height and weight; Hester & Gray, 2018), and skeptical responses to women's sexual assault claims (Paganini et al., 2023). Body-related discrimination is culturally

✉ Flora Oswald
feo5020@psu.edu

1 Department of Psychology, The Pennsylvania State University, University Park, PA, USA

2 Department of Women's, Gender, & Sexuality Studies, Pennsylvania State University, University Park, PA, USA

3 Yale Child Study Center, Yale University School of Medicine, New Haven, CT, USA

4 Department of Psychology, University of Denver, Denver, CO, USA

pervasive, but some group contexts are more likely to foster discrimination than others. Identifying which group contexts are “threatening” versus “safe” is thus an important task for people who may be targets of discrimination, as a means of preserving their well-being (e.g., Alt et al., 2019). Among the many cues that may be relevant to body-discrimination in a group context, the most obvious may be the body size and shape of people in that group—when a group is slim, a fat perceiver may assume that the group members will devalue their non-normative body size (see Oswald et al., 2022).

Such evaluations are likely to be especially useful to targets of discrimination but can aide anyone—not just people who identify as fat or as a sexual assault survivor —in determining whether their body will be valued or devalued by a relevant social group. This sort of protective process would be rooted in rapid identification of group norms, yet little is known about if or how *anyone* rapidly perceives the average size of bodies in a group. Thus, it is unclear if group perceptions of body size and shape are precise enough to aide perceivers in navigating a biased social world. In the current work, we examine the precision of body-size perceptions and how bodily perceptions, more generally, give rise to conceptual judgments of a group.

Perceiving Bodies

It seems obvious that body-related discrimination requires perception of others’ bodily characteristics, so it may be unsurprising that such discrimination relies on visual processing operations such as configural processing (e.g., Griffin & Oswald, 2022; Reed et al., 2003), size-based threat detection (e.g., McElvaney et al., 2021; Wilson et al., 2017), and early neural processes attuned to body perception (e.g., the n190 event-related potential; Thierry et al., 2006). For example, in a study of body size perception, Wilson et al. (2017) found that – even when controlling for upper body strength – Black male bodies were perceived as bigger and more threatening than White male bodies, leading to increased justification of use of physical force against Black relative to White men. This work demonstrates how low-level processes including size perception are influential for downstream outcomes of body-related bias.

Perceiving Groups

Despite advances in scientific understanding of how individual bodies are perceived and evaluated (de Gelder, 2009), little is known about how people perceive and evaluate groups of bodies. Though much has been made of the impact of individual body perception on well-being – for example, in considering how body perception is related to body image disturbances (e.g., Hartmann et al., 2020) – perceptions of *groups* of bodies have yet to be examined in depth or integrated with the rich literature on body-related discrimination. As noted earlier, group body perception is likely to play a role in perceivers’ evaluations of the threat afforded by groups or social contexts. Indeed, groups attract more visual attention than individuals (Woolhouse & Lai, 2014) and have an outsized influence on well-being through processes sensitive to group belonging and affiliation (e.g., Baumeister & Leary, 1995; Oyserman et al., 2006). That is, group perceptions offer different affordances than do individual perceptions, and may be particularly sensitive to certain visual cues.

If perceivers form precise and accurate representations of group body size, it is likely that those representations strongly inform perceivers' social judgment and behavior. Indeed, rapid group perceptions (1s or less exposure to four or more people) provide some of the earliest and most influential information about social contexts (Phillips et al., 2014), informing perceivers about opportunities for affiliation with a group (Goodale et al., 2018), group cohesion (Dasgupta et al., 1999; Ip et al., 2006), group boundaries (Lamer et al., 2018) and group threat (Mihalache et al., 2021). These findings and others have extended social perception research from *person* perception to *people* perception, or the simultaneous perception of multiple individuals (Alt & Phillips, 2021; Phillips et al., 2018). Given that the events which characterize our social lives often entail interacting with *groups of people* – from classes and meetings to concerts, stores, parties, and funerals – the shift toward understanding people perception reflects a shift toward understanding in greater depth the inherently group-based human experience (e.g., Alt & Phillips, 2021). In terms of the current work, perceptions of the typical (normative) body size would constitute a perceptual means for people to rapidly ascertain information important to their well-being in group-living.

People perception research primarily draws from theories of ensemble coding, and we draw our hypotheses from this literature. *Ensemble coding* refers to the capacity of the human visual system to rapidly extract summary statistics (e.g., averages) from a group of stimuli prior to detailed processing of each group member (Alvarez, 2011; Whitney & Yamanashi Leib, 2018). Ensemble coding is robust for concrete visual features such as orientation (Miller & Sheldon, 1969), color (Webster et al., 2014), size (Ariely, 2001; Chong & Treisman, 2003), and motion direction (Sweeny et al., 2013). This work has also been extended to high-level social vision (Phillips et al., 2014), which focuses on more abstract judgments. For example, when encountering a group of faces, study participants accurately extracted the average facial identity (de Fockert & Wolfenstein, 2009; Neumann et al., 2013), emotional expression (Haberman & Whitney, 2007, 2009) and gender (Alt et al., 2019; Goodale et al., 2018) of the facial group—even when the entire group is seen for less than ½ second (too brief to allow for close inspection of each group member).

Ensemble coding – both of concrete features (e.g. orientation angle) and of abstract features (e.g., facial emotion) – is understood as "an adaptive mechanism which allows for the efficient representation of a large amount of information" (Haberman & Whitney, 2009, p. 719); with social stimuli like faces, ensemble perception allows for rapid extraction of functionally-relevant social information (e.g., emotion, group identity; Goldenberg et al., 2020; Haberman & Whitney, 2007, 2009). In virtually every domain examined to date, study participants have proven capable of accurately and rapidly detecting the average signal value for any perceptible characteristic in a stimulus group (e.g., circle size, facial emotion, orientation direction; Whitney & Yamanashi Leib, 2018). Thus, it may seem trivial to examine if ensemble coding applies to bodily-characteristics. As detailed below, it is not.

Ensemble Perceptions of Body Shape and Size

We know of no research that has examined how people perceive socially-meaningful *bodily* characteristics in groups—whether via ensemble coding or any other process. Yet for many people, perceptions of bodily groups are uniquely informative to navigating social life. First, for millions of fat people, ensemble perceptions of bodies may be the basis for their evaluations of social identity threat (see Oswald & Adams, 2022). Indeed, contextual evaluations of social identity threat represent a critical task for people with stigmatized identities, as such threats impair well-being (via discrimination) and identifying them may help people avoid groups and contexts likely to be harmful (see Schmader & Sedikides, 2018). This perspective has motivated burgeoning research on ecological cues to social identity threat (SIT) and safety (see Oswald et al., 2022) yet there is empirical ambiguity regarding the processes that explain the relationship between social ecology and evaluations of social-identity threat.

In the socially complex and perceptually rich ecologies that characterize group settings, evaluations of SIT would seem to require processes capable of rapidly summarizing cues (across group members) relevant to perceivers' social identities. With respect to body-related SIT, ensemble coding of body shape and size *could* rapidly provide perceivers with information about the body norms in a group, and in this way learn information critical to their well-being. Moreover, self-perceptions of body size/shape strongly influence self-esteem even among non-fat perceivers (e.g., Bratovcic et al., 2015; Thompson & Thompson, 1986) and ensemble perceptions of bodies may provide these perceivers with critical information about the extent to which they will be valued or devalued by a group. Thus, even if ensemble coding *processes* do not differ for bodies versus other stimuli, the outcome of that processing is highly relevant to many populations. Yet absent evidence that people are capable of accurately and rapidly summarizing the bodily features they perceive of a group, little attention has been paid the potential role of group body perceptions in marginalized people's experiences of groups. In the current work, we sought to empirically confirm that ensemble coding processes apply to perceptions of body characteristics.

One way of understanding this contribution is to consider an analogy to social categorization research. Ensemble coding of bodies may reflect a broader, domain-general process of ensemble coding, much as social categorization processes are thought to reflect more general categorization processes that can be applied of *any* object (human or not). Because social categorization has salient consequences for stereotyping and prejudice (e.g., Brewer, 2007; Hugenberg & Sacco, 2008; Johnson et al., 2015), however, scientific understanding of the *social* categorization process has been important to examine in its own right (e.g., Freeman et al., 2013; Kramer et al., 2017; Volpert-Esmond & Bartholow, 2021)— even if social categorizations are simply byproducts of domain-general categorization abilities. In the same way, ensemble perceptions are likely to shape evaluations of group threat and thus group entry and exit (Phillips et al., 2014) even if those perceptions are simply byproducts of domain-general processes. By examining *if* (for the first time) and *how* people rapidly perceive the body shape/size typical of a group, we can begin to understand the processes that underlie evaluations of group threat, much as studies of social categorization supports understanding of processes such as stereotyping and prejudice. With respect to the processes involved in ensemble coding of bodies, scientific understanding can shape theories of body-related stigma and how perceivers respond to it.

The current work also makes a potentially important contribution to the broader ensemble coding literature. Specifically, ensemble coding is thought to be a domain general mechanism (Chang & Gauthier, 2022) but there remains debate about the nature of the mental representation produced in ensemble coding (e.g., Corbett et al., 2023). In the current work, we examine the extent to which ensemble perceptions of body size are represented in a more concrete or abstract manner. Indeed, human bodies have limits in terms of both height and weight, and there are certain size ratios between individual bodily parts that create particular bodily shapes to which humans are visually attuned. For example, waist-to-hip ratio (WHR) and waist-to-chest ratio (WCR) are often implicated in studies of body perception. These ratios are concrete in the sense that they describe the physical features of a body, and yet perceptions of such ratios are related to social judgements including evaluations of another person's attractiveness (e.g., Fan et al., 2005; Singh, 1994), self-evaluations (e.g., Joiner et al., 1994; Pazhoohi et al., 2012), and victim-blaming for sexual assault (Paganini et al., 2023). These ratios inform distinct social judgements; For example, WCR informs perceptions of dominance and physical fitness, particularly among men (Coy et al., 2014), while WHR informs perceptions of fertility and health, particularly among women (e.g., Fink et al., 2003). We thus suspected that ensemble perceptions of WHR, WCR, and BMI would be related but separable, indicating that perceivers encoded these concrete features with dimensional precision.

Importantly, there is an alternative possibility—perceivers may simply use a size heuristic to evaluate all three characteristics. Indeed, BMI is moderately correlated with WHR (e.g., Staiano et al., 2012) and perceivers may have difficulty distinguishing between them. Failures to distinguish among WHR, WCR, and BMI in ensemble coding would support the view that ensemble representations of body size are more abstract ("fat/slim body") than concrete (e.g., WHR). This approach is uncommon in ensemble perception research, which focuses less on judgment processes associated with construal-level than on the visual processes associated with ensemble perceptions. However, research on individual differences does suggest that ensemble perceptions of concrete characteristics (e.g., stimulus orientation angle) differ from those of abstract characteristics (e.g., facial emotion), whereas ensemble perceptions of different concrete characteristics share considerable variance with each other, as do ensemble perceptions of abstract characteristics (Haberman et al., 2015). Such findings lead to the possibility people do not distinguish among (concrete) ensemble perceptions of WCR, WHR, and BMI (owing to shared variance). In the current work, we tested whether ensemble percepts of body characteristics were separable and additionally examined the degree of accuracy for each judgment. To our knowledge, Study 1 is one of the first experiments to examine the degree which ensemble perceptions of concrete features reflect higher-level construals (more abstract) or lower-level construals (more concrete; Trope et al., 2007). Accordingly, the current work may contribute to scientific understanding of ensemble perceptions, especially the representational quality of those percepts.

Ensemble Perceptions of Gender and Emotion

We have argued that ensemble coding of body size and shape—to the extent it operates like other forms of ensemble coding—is likely to play an important role in how people evaluate groups. However, the *impact* of such ensemble coding is likely to depend on the social (abstract) inferences drawn from those bodies. In Study 2, we thus focus on ensemble perceptions of abstract characteristics. We do so with three goals in mind.

First, body norms are heavily sex-typed. For example, a low WHR is more socially desirable for women than for men, whereas a low WCR is more socially desirable for men than women (Coy et al., 2014; Fink et al., 2003). For this reason, ensemble coding of body size and shape might be more functional for perceivers to the extent that they can also ensemble code gender. Indeed, one recent study observed that study participants were unable to ignore gender when perceiving body weight in a garner paradigm (but were able to ignore weight in judgments of gender; Johnstone & Downing, 2017), implying that processing of gender is fundamentally involved in perceiving weight. It thus seemed important to examine if people were capable of ensemble coding the abstract characteristic of gender from a crowd of bodies. That is, even though people seem to be capable of ensemble coding gender from facial crowds (Alt et al., 2019; Goodale et al., 2018), the gender-signal may be noisier (than faces) with respect to body perception. The absence of ensemble coding of gender from bodies would limit the functional advantage of ensemble coding WHR, WCR, and BMI, as the body norms relevant to a specific perceiver vary by gender. Thus, in Study 2, we examined whether people were capable of ensemble coding the abstract characteristics of gender from bodies.

The second reason we focused on abstract ensemble perceptions in Study 2 was to examine differences between ensemble coding of structural and dynamic cues. Whereas WHR, WCR, BMI, and gender are relatively stable characteristics perceived via the structure of the body, other characteristics (e.g., emotion) are more dynamic and perceived from bodily expressions. Indeed, some have argued that there are different processing routes for perceiving structural versus dynamic cues of faces (e.g., Berstein & Yovel, 2015; see also Haxby et al., 2000) but different processing routes are associated with the ensemble coding of faces (compared to individual faces), even though people appear to exhibit similar processing of structural (e.g., gender) and dynamic (e.g., emotion) cues (Im et al., 2017). We speculated that such similarities emerge because ensemble codes of abstract characteristics are two- to five-fold noisier than ensemble codes of abstract characteristics (see Haberman et al., 2015, Table 1). That is, when the relevant "signal" is noisier people may be especially likely to conflate structural and dynamic cues, as in Garner paradigm studies (Weisbuch & Ambady, 2011). In Study 2, we thus examine shared variance in the ability to ensemble code gender and emotion. That is, we examine whether participants who produced less precise ensemble codes would be more likely to exhibit shared variance in their ensemble codes of gender and emotion. This hypothesis test has implications for the ensemble coding literature, especially as it applies to the human form. That is, if covariance among ensemble codes increases with reductions in the signal-fidelity of the relevant cues, the implication is that differences between "low-level" (concrete) and "high-level" (abstract) ensemble codes may be explained by the relatively higher signal-value of concrete characteristics, rather than by a qualitative difference in the ensemble coding of concrete or abstract characteristics. Study 2 may thus offer important insight into ensemble coding processes.

The third reason we focused on abstract ensemble perceptions in Study 2 was to examine ensemble perceptions of emotion. Among ensemble coding studies that examine perceptions of people, facial emotion is commonly examined (e.g., Haberman & Whitney, 2007; Goldenberg et al., 2020). Yet over the last decade, many questions have arisen regarding perceptions of facial versus bodily emotion. For example, extant evidence suggests that people are sensitive to both faces and bodies in their perceptions of another *person's* emotions (e.g., Aviezer et al., 2008). With this in mind, the domain-generality of ensemble coding suggests that people should be capable of ensemble coding emotion from bodies. However, there are also reasons that we may *not* observe ensemble coding of emotion from bodies. Most importantly, existing evidence suggests that perceivers attend more to bodies than faces in perceptions of secondary emotions (e.g., shame, pride) but the reverse is true for primary emotions (e.g., App et al., 2011). If bodies provide a lower-fidelity signal of primary emotion than faces it is possible that people are unable to ensemble code emotion from bodies, or do so with limited accuracy. Accordingly, even though perceivers exhibit substantial accuracy in ensemble perceptions of *primary* facial emotions it is possible that such accuracy does not extend to *primary* bodily emotions. We thus examine whether people are capable of ensemble coding of primary emotions from bodily expressions.

The Current Studies

In the current work we examine if and how study participants ensemble code human bodies. We first test whether people are capable of ensemble coding precise physical characteristics of bodies, and we examine whether the representations produced by such ensemble coding reflect specific physical metrics (e.g., WCR) or a more abstract representation of "body size". We then turn our attention to whether study participants are capable of ensemble coding two characteristics defined more abstractly: gender and emotion. Thus, in two studies, we tested the ability of human observers to ensemble code human bodies. Specifically, we examined whether human observers were able to extract summary statistics of both concrete (BMI, WHR, WCR) and abstract (emotion, gender) properties from groups of bodies.

Transparency and Openness

Data were collected between February and March 2022. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All novel stimuli, data, and analysis code are available on the Open Science Framework (https://osf.io/whqgk/?view_only=65ea0a2edb654727bbb290e6965ed064). All analyses were conducted in R version 4.1.3 (R Core Team, 2022). The study design and analysis were not pre-registered. Procedures were deemed minimal risk and thus exempt by an Institutional Review Board prior to data collection.

Study 1

Study 1 examined the degree to which participants were capable of ensemble coding WCR, WHR, and BMI, and whether the resulting ensemble codes were separable or reflective of a single "size" estimate.

Methods

Participants

Our sample size for Studies 1 and 2 was determined a priori and was based on available resources and expected magnitude of findings. Based on previous body ensemble coding literature (Alt et al., 2017; Goodale et al., 2018), we anticipated the relationship between actual and perceived ensemble representations to be moderate (β = 0.41). By using this expected estimate in conjunction with conversative power analysis parameters (i.e., alpha = 0.01, power [1- β] = 0.90), 90 participants were required to detect the effect of interest.

However, two evaluate interaction models, improve our estimates, and be robust to attrition common in online studies, we aimed to recruit three times as many participants.

We sought to collect 300 participants for Study 1 (and for Study 2), a sample larger than those represented in similar, previous work (e.g., Haberman & Whitney, 2007, 2009; Sweeny et al., 2013) and indeed much larger than most within-subject designs in visual perception research. We projected these large sample sizes in anticipation of potential data quality issues given the online nature of the studies, which were conducted during the ongoing COVID-19 pandemic.

Participants were recruited from the subject pool of a large university; this college student sample was recruited due to convenience. Of the 292 participants who consented, only 276 started the task paradigm. We excluded participants who did not fully complete the task ($n = 26$), did not follow the stimulus screening adjustment ($n = 57$), or who demonstrated evidence of near identical responses to each stimulus item ($n = 23$). Our final sample after these exclusions ($N = 179$) ranged in age from 18 to 25 years ($M = 19.04$; $SD = 1.07$). Participants included 133 women, 44 men, and two who elected not to provide gender information.

Stimuli

In Study 1, we developed a set of body stimuli using the Max Planck Institute Perceiving Systems Body Visualizer tool (Max Planck Institute, 2011), a program which allows for the generation of three-dimensional body stimuli with variable values on several parameters including height, and weight, as well as chest, waist, and hip size. We developed three sets of stimuli varying on one set parameter each (BMI, WHR, or WCR); each set consisted of six bodies for males and six for females (12 bodies total on each parameter). Faces were included on all bodies but were blurred using Adobe Photoshop, and all stimuli were rendered in a consistent blue-grey shade to avoid cuing potential biases associated with skin color. Body mass index (BMI) stimuli varied in weight to height ratio; the height for all bodies was set to the United States population average for the appropriate gender category (Centers for Disease Control and Prevention, 2021) and weight was varied to alter BMI. Each bodily metric was manipulated to have 6 levels; we created stimulus variability on each dimension by dividing the stimulus range by 5. Stimulus BMI ranged from 15 to 40 (rounded) in intervals of five. WHR stimuli varied from ratios of 0.7 to 1.2 for men and from 0.5 to 1.0 for women in intervals of 0.1; these ranges were centered on reported averages for each gender (Molarius et al., 1999). WCR stimuli varied from 0.6 to 1 for both men and women. All bodies were visualized from a frontal view, standing in a half t-pose (arms to the side and slightly raised, feet hip width apart).

Procedure

Participants in Study 1 (and Study 2) were recruited via an institutional research subject pool and earned partial course credit for their participation. After completing informed consent, participants completed a brief demographic questionnaire and were then randomly assigned to complete either Study 1 or Study 2 via a randomization process embedded in Qualtrics. Participants were automatically redirected to Pavlovia (pavlovia.org), where they completed the assigned study (either Study 1 or Study 2). Participants were unaware of this random assignment procedure.

Before they started, participants were asked to use the arrow keys on their keyboard to adjust a rectangle on their screen to be the size of a credit card. Since participants were likely to be using a variety of display resolutions and screen sizes, we recorded the numerical change (from a fixed arbitrary value) required to make the virtual credit card match the known physical dimensions of an actual credit card (5.39×8.56 cm). This scaling factor for the x and y axis, when multiplied by 7, ensured that all body images in the arrays were presented in a standardized size (7×7 cm) for all participants.

Participants in Study 1 completed three randomized blocks where they viewed arrays of human bodies that varied in BMI, WHR, or WCR, and were instructed to provide a judgement about the average body type of the group (see Fig. 1a). Participant responses were measured via a continuous sliding scale that used visual body depictions as intervals across the scale.

For each block, we used stratified random sampling with replacement to select six bodies that were arranged in a 2×3 array. To ensure that participants viewed a range of body arrays with varying levels of mean BMI, WHR, and WCR, we randomly selected body images to form 100 body arrays for each bodily dimension. For each array, the average of the target characteristic was within five equally spaced intervals (BMI: 15–20, 20–25, 25–30, 30–35, 35–40; WHR: 0.5-0.64, 0.64–0.78, 0.78–0.93,0.93–1.06,1.06–1.21; WCR: 0.60-0.68, 0.68–0.76, 0.76–0.84, 0.84–0.92, 0.92-1.00) across the range of our body images. Since the body arrays were generated by randomly selecting body images with replacement, the distribution within body arrays was variable. Specifically, the average standard deviation within body arrays was 7.05 (range: 2.03–12.47) for BMI, 0.15 (range: 0.05–0.24) for WHR, and 0.11 (0.04–0.19) for WCR.

For each trial, participants completed the following sequence: (1) view fixation cross for 500ms, (2) view body array for 500ms, and (3) select along a slider scale (with body

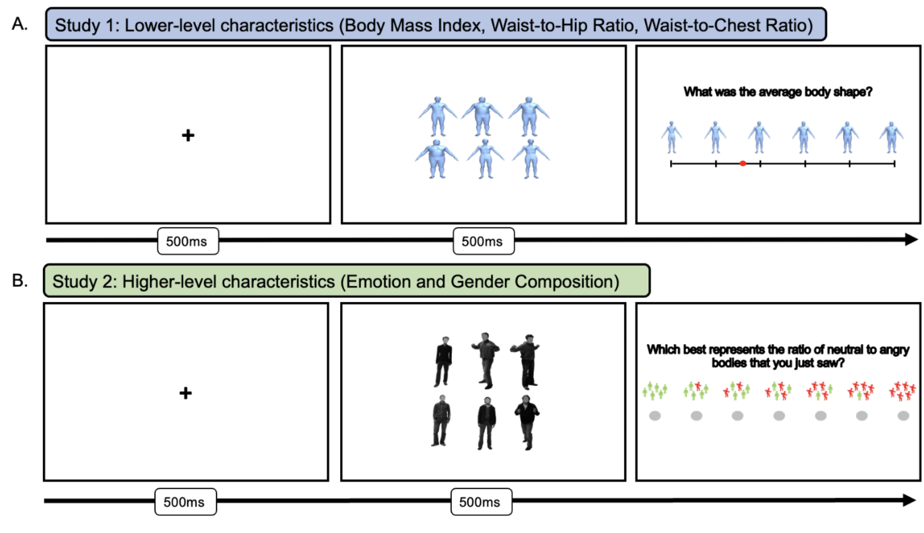

Fig. 1 Depiction of task for Study 1 (A) and Study 2 (B)

images) the average body type for the previously viewed body array. Each block contained 100 trials randomly presented for a total of 300 trials. Participants completed this task in approximately 30 minutes.

Results

We used linear mixed effects modeling to account for repeated observations in participants using *lme4* (Bates et al., 2015) within the statistical software R (R Core Team, 2022). To address our primary aims, we regressed participants summary estimates and degree of error onto the actual average body ensemble scores and included a random intercept for participant. In initial analyses we fit separate models for BMI, WHR, and WCR. In subsequent analyses, we fit the same models but with the other two metrics as covariates (e.g., BMI estimates controlling for WCR and WHR estimates). For all models, we reported unstandardized regression coefficients (i.e., *b*) and used Satterhwaite's approximation for *p*-value estimation for linear mixed-effects models (Kuznetsova et al., 2017).

How Precise are Concrete Representations of Body Shape and Size Ensembles?

We evaluated the capability of human observers to extract summary statistics about human body ensembles that varied on BMI, WHR, and WCR. As expected, participants' estimates of the average body in the group was positively associated with the actual average body, in terms of BMI ($b=0.53$, $se=0.006$, 95% CI [0.52, 0.55], $p<.001$), WHR ($b=0.53$, $se=0.006$, 95% CI [0.51, 0.54], $p<.001$), and WCR ($b=0.46$, $se=0.005$, 95% CI [0.45, 0.47], $p<.001$; see Fig. 2). To test the sensitivity of these results, we also evaluated these associations without excluding participants who did not adhere to the screen scaling procedure or who showed potentially abnormal responding. The pattern of results was similar for BMI ($b=0.47$, $se=0.005$, 95% CI [0.46, 0.48], $p<.001$), WHR ($b=0.47$, $se=0.005$, 95% CI [0.46, 0.48], $p<.001$), and WCR ($b=0.41$, $se=0.005$, 95% CI [0.40, 0.42], $p<.001$). These findings are consistent with ensemble coding of body shape/size—after seeing a crowd of bodies for only 500ms, participants were able to accurately estimate the BMI, WHR, and WCR of those bodies[1].

We next examined whether ensemble perceptions were specific to BMI, WHR, and WCR or were all derived from more abstract "size" perception. We found that accuracy of ensemble perceptions for arrays varying in BMI remained statistically significant even after accounting for WHR ($b=0.57$, $se=0.007$, 95% CI [0.56, 0.58], $p<.001$) and WCR ($b=0.57$, $se=0.007$, 95% CI [0.56, 0.59], $p<.001$). Similarly, accuracy of ensemble perceptions for arrays varying in WHR remained statistically significant even after accounting for BMI ($b=0.51$, $se=0.03$, 95% CI [0.45, 0.56], $p<.001$) and WCR ($b=1.20$, $se=0.07$, 95% CI [1.05, 1.35], $p<.001$). Finally, accuracy of ensemble perceptions for arrays varying in WCR remained statistically significant even after accounting for and BMI ($b=0.53$, $se=0.008$, 95% CI [0.51, 0.54], $p<.001$) and WHR ($b=0.42$, $se=0.007$, 95% CI [0.41, 0.44], $p<.001$).

[1] There was a small and statistically significant effect of learning (i.e., time) on this association for WHR ($b=0.0004$, $se=0.0002$, 95%CI [-0.0005, 0.0002], $p=.02$), but this finding did not generalize to BMI ($b=-0.0001$, $se=0.0002$, 95%CI [-0.0005, 0.0002], $p=.45$) or WCR ($b=-0.0001$, $se=0.0002$, 95%CI [-0.0005, 0.0002], $p=.60$).

These findings suggest that human perceivers are capable of forming extremely precise summary percepts of human bodies that are not purely based on size.

Finally, we also evaluated the effect of participant gender on ensemble perceptions of WHR, WCR, and BMI. We did not find that accuracy differed as a function of participant gender for BMI ($b=0.01$, $se=0.01$, 95% CI [-0.01, 0.04], $p=.27$) or WCR ($b=0.002$, $se=0.01$, 95% CI [-0.02, 0.3], $p=.84$); however, we found a significant effect of gender on the relationship between actual and perceived WHR (WHR x Gender interaction: b = -0.08, $se=0.01$, 95% CI [-0.10, -0.05], $p<.001$). Men were less accurate at perceiving WHR from groups than women, but only for ensembles with lower WHR.

Beyond Correlations: Absolute Accuracy in Ensemble Perceptions of Bodies

As described in the preceding, our main analytic approach was to correlate estimated and actual group body size/shape. However, there are limits to this approach. For example, participants' estimates of group average BMI could be highly correlated with the actual group average BMI, yet systematically underestimate (or overestimate) group BMI. Such findings would indicate that participants are accurate in a relative sense but not in an absolute sense. To illustrate absolute accuracy, we calculated the distance between the estimated and actual crowd average for each dimension (BMI, WCR, WHR)—that is, we calculated the margin of error. Most (64.25%) participant estimates were within 5 BMI points with 91.44% of responses being within 10 BMI points. We observed that 64.33% and 86.74% of participant estimates of WHR and WCR respectively, had a degree of error (i.e., |perceived average ratio – actual average ratio|) less than or equal to 0.15 (on a 6-point scale) (symmetrical), and 92.87% and 99.87% of participant estimates of WHR and WCR respectively, had a degree of error less than or equal to 0.30 (see Fig. 3).

Study 2

Study 1 provides novel evidence consistent with the view that humans have the capacity to rapidly perceive the average body shape/size in a group. As described earlier, however, ensemble perceptions of bodies may differ according to whether those judgments are concrete (Study 1) or abstract (Study 2). In Study 2, we examine whether ensemble coding of bodies extends to abstract characteristics, including characteristics (gender) likely to modulate body perception.

Methods

Participants

Our initial sample consisted of 305 participants from a large research university's subject pool. We excluded participants who did not start ($n=13$) or fully complete the task ($n=9$), did not follow the stimulus screening adjustment ($n=83$), or who demonstrated evidence

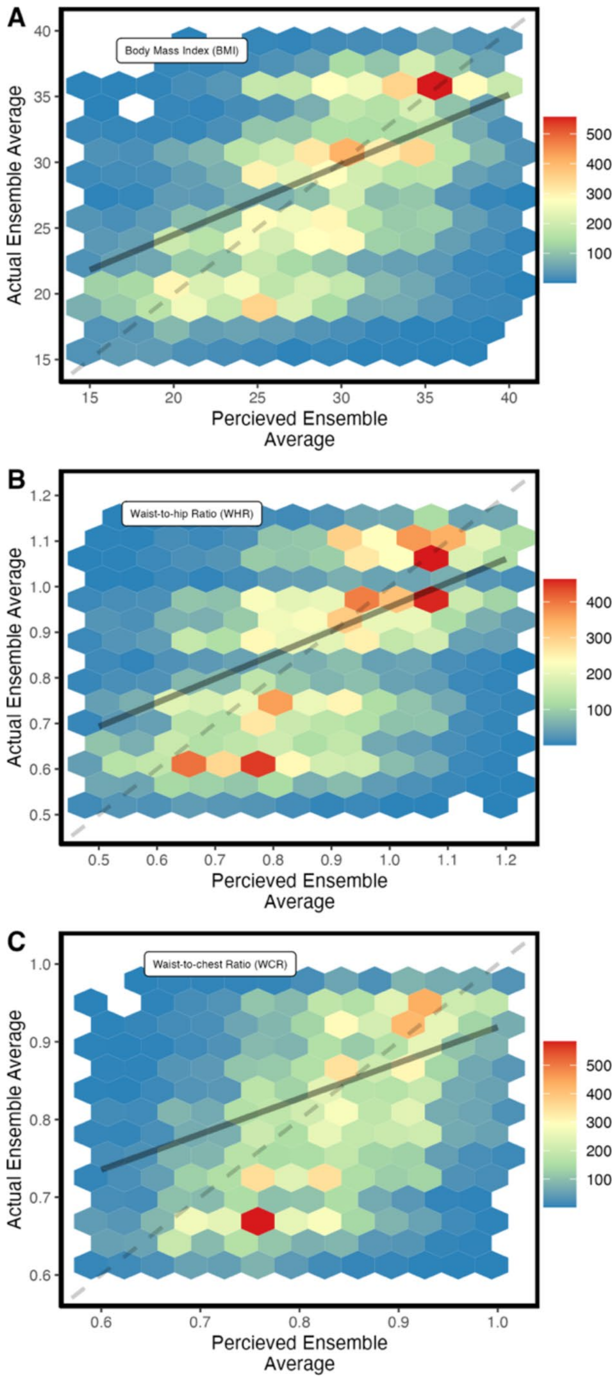

Fig. 2 (A-C) Heatmap distribution plots of individual trials for each participant show perceived ensemble average as a function of the actual ensemble average. Warmer colors reflect more responses for BMI, WHR, and WCR. Dashed line is a reference line reflecting perfect responding; Solid line is predicted regression line based on participant data

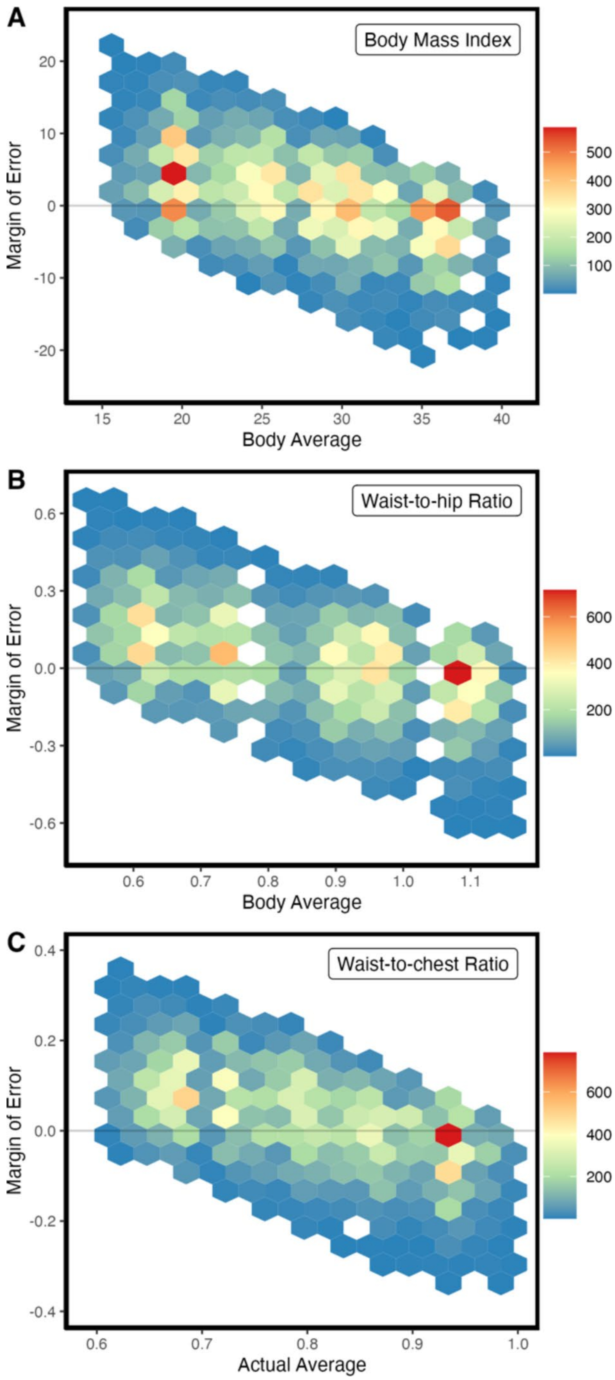

Fig. 3 *(A-C) Heatmap distribution plots of individual trials for each participant show actual body ensemble averages as a function of participants' margin of error. Warmer colors reflect more responses for BMI, WHR, and WCR.* Grey horizontal line at 0 reflects no error

of responding near identically to each stimulus item ($n=4$). Our final sample after these exclusions ($N=196$) ranged in age from 18 to 41 years ($M=19.20$; $SD=2.01$). Participants were mostly women (67.9%, $n=133$); a smaller portion were men (31.6%, $n=62$) and one participant was nonbinary.

Stimuli

We used full-body stimuli from the Bodily Expressive Action Stimulus Test (BEAST; de Gelder & Van den Stock, 2011). Unlike the bodies we *created* for Study 1, these bodily images are of real people expressing emotion in their bodies. We removed the background of all images and resized some images given variable photographing distances in the original stimulus set. Faces of all stimuli were blurred. Images were standardized to 150×300 pixels. We included both male and female bodies, and only used bodies for which both neutral and angry postures were available. We used neutral bodies to examine gender perception and used neutral and angry bodies to examine emotion perception. We opted to use angry bodies, as opposed to other emotional expressions, to make our manipulation of emotion salient; prior research demonstrates highly efficient processing of threat-related information, including angry facial and bodily expressions (e.g., Bannerman et al., 2009; Fox et al., 2005).

Procedure

See Study 1 Procedure for randomization into the studies and screen standardization. Participants assigned to Study 2 completed two randomized blocks (i.e., Gender, Emotion) where they viewed arrays of human bodies and were instructed to provide a judgement about the ratio of women to men and angry to neutral bodies respectively (see Fig. 1b).

For each block, we used stratified random sampling without replacement to select six individual bodies that were arranged in a 2×3 array. To ensure that participants viewed a range of emotional (angry:neutral bodies) and gender (women:men) compositions, we randomly selected body images to form 98 body arrays that contained 14 trials of each possible seven compositions for gender and emotion (i.e., 0:6, 1:5, 2:4, 3:3, 4:2, 5:1, 6:0) arrays. Gender arrays contained only neutral bodies; emotion arrays were gender-constant within, but not between trials, such that participants saw both all-male and all-female arrays.

For each trial, participants completed the following sequence: (1) view fixation cross for 500ms, (2) view body array for 500ms, and (3) select an icon reflecting the ratio of women to men or angry to neutral bodies for the previously viewed array (similar to Alt et al., 2019; Goodale et al., 2018). Each block contained 98 trials randomly presented for a total of 196 trials. Participants completed this task in approximately 30 minutes.

Results

We used linear mixed effects modeling to account for repeated observations in participants using *lme4* (Bates et al., 2015) within the statistical software R (R Core Team). To address our primary aims, we regressed participants summary estimates (estimated number of angry people or women in an ensemble) and summary estimates degree of error onto the actual average body ensemble scores (i.e., actual number of angry people or women in an ensemble) and included a random intercept for participant. For all models with continuous dependent variables, we reported unstandardized regressions coefficients (i.e., *b*) and used a Satterhwaite correction for *p*-value estimation. We used linear regression to examine the interacting effect of accuracy and ensemble perception across emotion and gender conditions.

Given our categorical conceptualization and measurement of gender and emotion, the gender and emotional composition of these body arrays could be exactly estimated by participants; therefore, we also derived accuracy scores (1 = correct, 0 = incorrect) for each body array. To evaluate if participant performed above chance (1/7), we used a binomial test. Finally, we evaluated participants' absolute accuracy at estimating the number of targets using binary logistic regression with a binomial link function. We reported unstandardized regression coefficients in log odds.

How Precise are Abstract Representations of Human Body Ensembles?

We evaluated the ability of human observers to extract summary statistics of emotional expression and gender composition in groups of human bodies (i.e., the number of angry people or females in a group). For clarity, we use the angry bodies and female bodies as reference "targets". As expected, participants' estimates of group composition was positively associated with the actual group composition, in terms of emotion ($b = 0.65$, $se = 0.004$, 95% CI [0.64, 0.65], $p < .001$) and gender ($b = 0.49$, $se = 0.004$, 95% CI [0.48, 0.50], $p < .001$) (see Fig. 4). To test the sensitivity of these results, we also evaluated these associations without excluding participants who did not adhere to the screen scaling procedure or who showed potentially abnormal responding. The pattern of results was similar for emotion ($b = 0.61$, $se = 0.003$, 95% CI [0.60, 0.61], $p < .001$) and gender ($b = 0.45$, $se = 0.004$, 95% CI [0.45, 0.46], $p < .001$). Similar to Study 1, these findings show that participants can ensemble code higher-order representations from groups of bodies.

Finally, since Study 2 used a within-person design, we evaluated how individual differences influenced the relationships between actual and perceived body ensembles. We found a significant effect of participant gender on ensemble perception accuracy for both gender ensembles ($b = -0.04$, $se = 0.01$, 95% CI [-0.06, -0.02], $p < .001$) and emotion ensembles ($b = -0.05$, $se = 0.01$, 95% CI [-0.07, -0.04], $p < .001$). Specifically, men were less accurate at perceiving gender and emotion composition from ensembles than women, but only for ensembles with lower numbers of women and angry bodies respectively.

We also evaluated how ensemble perception of emotion was influenced by the interaction of accuracy and ensemble perception of gender, and how ensemble perception of gender was influenced by the interaction accuracy and ensemble perception of emotion. We did not find a significant emotion accuracy x emotion ensemble perception interaction on ensemble perception of gender ($b = -1.07$, $se = 0.57$, 95% CI [-2.19, 0.05], $p = .06$), nor did we find a significant gender accuracy x gender ensemble perception interaction on ensemble perception of emotion ($b = -0.49$, $se = 0.78$, 95% CI [-2.03, 1.06], $p = .54$)

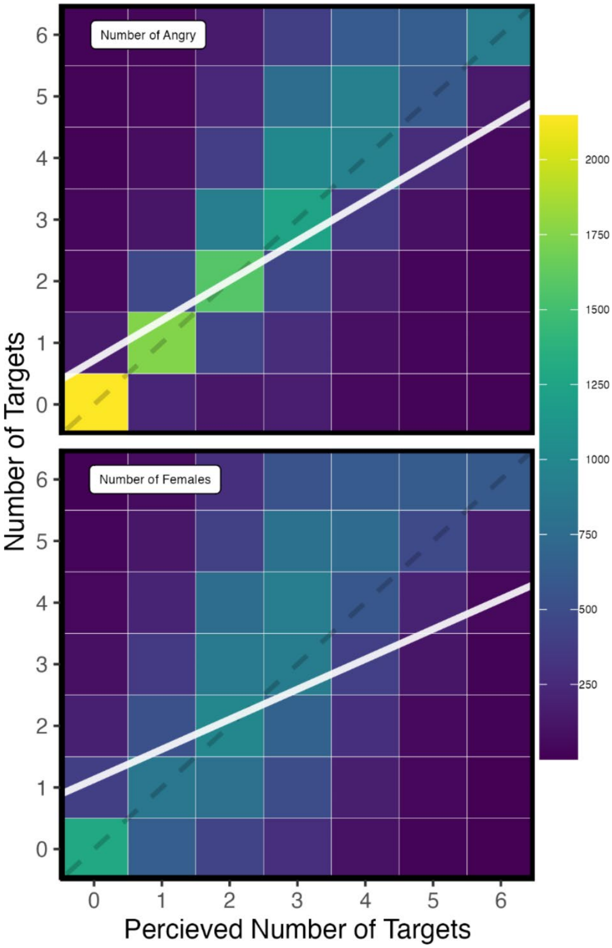

Fig. 4 Heatmap distribution plots of concatenated individual trials for each participant show participants' perceived number of targets (Top panel: Number of angry bodies; Bottom panel: Number of females) as a function of the actual number of targets. Warmer colors reflect higher density of responses. Dashed diagonal is reference line reflecting perfect responding; Solid white line is predicted regression line based on participant data

How Accurately can Abstract Representations from Human Body Ensembles be Extracted?

We found that participants were relatively accurate at identifying the exact emotional expression (M=47.68%, SD=14.68%) and gender composition (M=29.59%, SD=8.90%) of human body ensembles. Importantly, these accuracy scores were significantly above chance level (14.28%) responding for both emotion (p(correct)=0.476, 95% CI [0.47, 0.48], $p<.001$) and gender (p(correct)=0.295, 95% CI [0.29, 0.30], $p<.001$) compositions[2]. Moreover, participants' accuracy in identifying the exact number of reference targets (i.e., angry bodies) was negatively associated with the number of targets for emotional (b =

[2] We also evaluated the impact of learning across time in ensemble coding accuracy. We found no effect of time for gender ($b=-0.0001$, $se=0.0001$, 95%CI [-0.0003, 0.0001], $p=.35$), but there was a significant linear effect of time for emotion ($b=0.0009$, $se=0.0001$, 95%CI [0.0006, 0.001], $p<.001$). However, a logarithmic trend fit the data better than a linear trend (AIC=27.29, $p=.03$). Specifically, there was rapid learning particularly early (i.e., within the first 10 trials) with diminishing returns across later trials.

-0.41, $se=0.009$, 95% CI [-0.43, -0.40], $p<.001$) and gender (b = -0.21, $se=0.009$, 95% CI [-0.023, -0.20], $p<.001$) body arrays. These results suggest that participants are particularly good at extracting the emotional and gender composition of groups of bodies when the number of targets was low (e.g., 0 angry people; see Fig. 4). These results did not significant differ for all-male or all-female arrays ($b=-0.02$, $se=0.02$, 95% CI [-0.05, 0.01] $p=.23$).

We did find a significant effect of participant gender on gender ensemble perception accuracy ($b=0.17$, $se=0.02$, 95% CI [0.14, 0.20], $p<.001$), but not on emotion ensemble perception accuracy (b = -0.02, $se=0.02$, 95% CI [-0.05, 0.02], $p=.37$). Relative to women, men showed better accuracy when the number of male bodies was high, and worse when the number of female bodies was higher.

We also evaluated the correspondence of ensemble perception across conditions. Specifically, we found a significant positive correlation between ensemble perception accuracy for emotion and gender compositions ($r=.51$, $p<.001$). Participants who demonstrated higher accuracy for ensemble coding of emotion also demonstrated higher accuracy for ensemble coding of gender. Finally, we evaluated the degree of error in extracting emotional and gender composition from human body ensembles. For emotion expression, 47.68%, 78.74%, and 92.37% of participant responses were ±1, ±2, and ±3 bodies from the actual number of reference targets (i.e., angry bodies). For gender, 29.59%, 65.53%, and 86.84% of participant responses were ±1, ±2, and ±3 bodies from the actual number of reference targets (i.e., female bodies; see Fig. 5). This discrepancy in accuracy between gender and emotion ensembles could be due to features of the BEAST; specifically, the BEAST depicts both males and females in various outfits, most of which are not form-fitting. Heterogenous clothing, in addition to variability in other typically gendered features (e.g., hairstyles), may obscure gender perception.

General Discussion

In two studies, we tested the ability of human observers to ensemble code human bodies. Specifically, we examined whether observers could extract summary statistics of both concrete (BMI, WHR, WCR) and abstract (emotion, gender) characteristics from groups of bodies. Though existing evidence indicates clearly how the perception of individual bodies matters – for example, by influencing attractiveness judgements (Fan et al., 2005; Singh, 1994), prejudice (Rubino et al., 2020), and even discriminatory police behavior (Hester & Gray, 2018) – our findings suggest that perceptual sensitivity to bodies is not limited to individuals.

Across our two studies, participants accurately, distinctively, and rapidly perceived average WHR, WCR, BMI, emotion, and gender from groups of bodies. Accuracy in ensemble judgements was remarkably consistent among WHR, WCR, and BMI (see Supplemental Table 1 for standardized coefficients), yet participants did not appear to use a common size heuristic to estimate body size/shape. That is, perceivers achieved consistently strong accuracy via *distinctive* ensemble codes for WHR, WCR, and BMI, supporting the view that ensemble representations of objectively-defined features reflect lower-level, concrete construals. Additionally, the functional advantages conveyed by these lower-level ensemble codes may have supported ensemble coding of more abstract and dynamic body character-

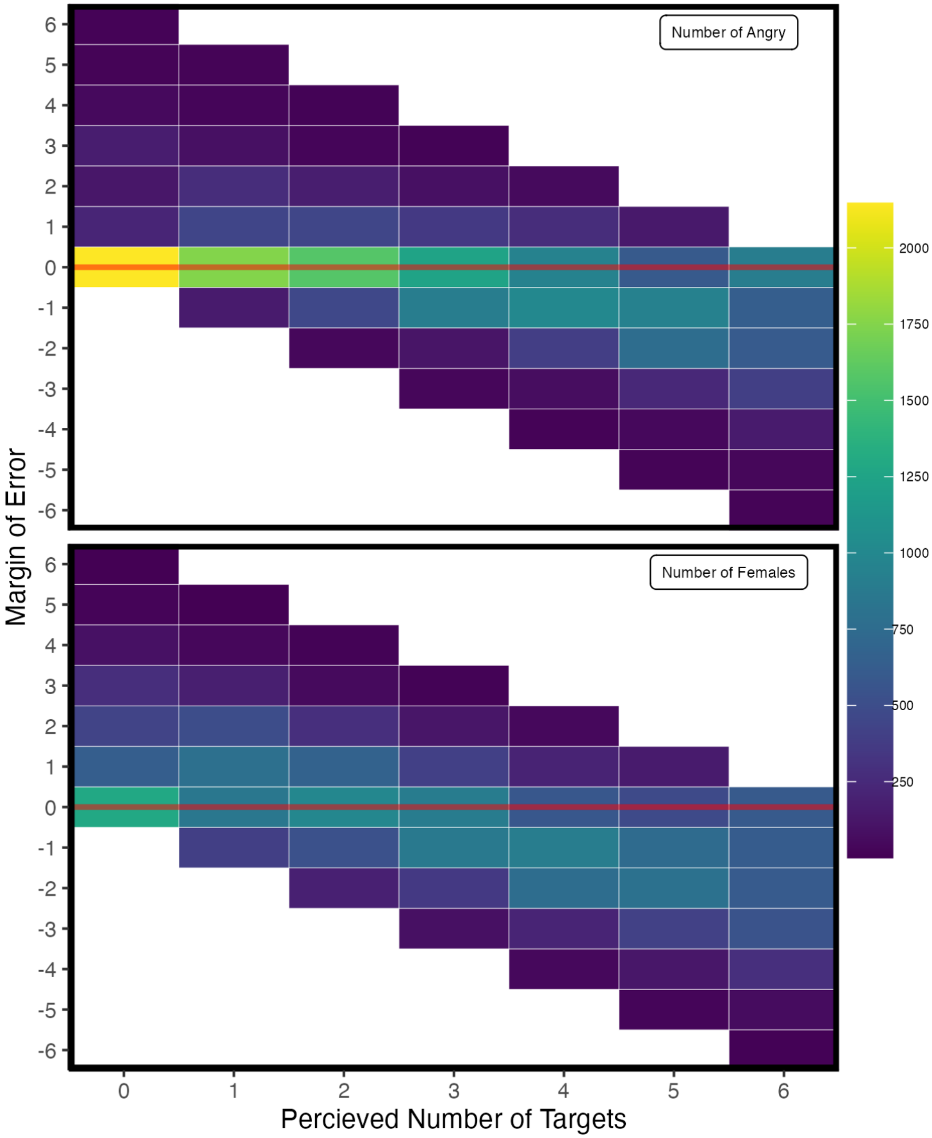

Fig. 5 Heatmap distribution plots of concatenated individual trials for each participant show participants' perceived number of targets (Top panel: Number of angry bodies; Bottom panel: Number of Females) as a function of margin of error. Warmer colors reflect higher density of responses

istics in Study 2, which demonstrated heightened accuracy relative to judgments in Study 1 (see Supplemental Table 1 for standardized coefficients). Despite some evidence that bodies provide lower-fidelity information about *primary* emotion than do faces (e.g., App et al., 2011), we found that participants were particularly accurate at ensemble coding primary emotion from bodily expressions. These findings lend additional support to the domain-generality of ensemble coding but as reviewed below, the findings also have unique implications for our understanding of body perception, particularly in group contexts.

Where perceptions of individual bodies may shape judgments of individuals' attributes, perceptions of groups of bodies shape judgments of group norms. Rapid ensemble coding of bodies enables perceivers to identify normative (average) body characteristics in a group, and our findings suggest that people are capable of accurately discerning several group norms (e.g., WHR, emotion) after only ½ second of exposure to an entire group of human bodies. These results point to a robust role for ensemble coding mechanisms in how people navigate group life. For example, rapid perceptions of group BMI could inform perceivers about whether a group favors one's body identity (e.g., as a fat person; see Oswald et al., 2022); rapid perceptions of group gender-ratio could inform perceivers about their likely belonging in a group and desire to enter the group (see Goodale et al., 2018); rapid perceptions of the average WCR or bodily anger in a group could inform perceivers about whether a group presents a threat to their welfare. These downstream consequences have yet to be explored, but the results of the current studies suggest that ensemble coding may play a (heretofore) underappreciated role in group and intergroup behavior.

We additionally examined potential effects of learning across trials and of gender on body ensemble perception accuracy. Learning across trials did not appear to play a significant role in accuracy of ensemble judgements in either study, possibly with the exception of the emotion task[3]. However, that learning overall did not appear to play a significant role suggests that observers are innately prepared to make ensemble judgements about a number of characteristics relevant to groups of bodies. Additionally, we found inconsistent effects of participant gender on ensemble judgement accuracy; overall, men and women exhibited similar effects.

The current studies also address an issue of ecological validity (Brunswik, 1955, 1956) in social psychology studies of ensemble coding. When encountering a crowd, body information is likely to be more visually available than facial information and yet most ensemble coding research in social psychology has prioritized face crowds (e.g., Alt et al., 2019; Haberman & Whitney, 2007; Neumann et al., 2013). In the "real world", the ability to see 6–8 faces simultaneously is limited by one's visual field and the fact that distances between faces must account for should width and personal space concerns—to see all those faces at once, the perceiver must be relatively distant from most of the faces in their visual field. Under these circumstances (at a distance), body structures and movements are likely to be more salient than facial structures and movements. Although this may not be a problem for those scientists interested in the basic visual mechanisms involved in ensemble coding, it is a problem for those scientists interested in the role of ensemble coding in social life. A focus on disembodied-faces in ensemble coding research may strongly underestimate the role of ensemble coding in social phenomena. Given the processing conditions required for seeing an entire group of people (a larger visual field than for perceiving an individual), and given the current evidence that participants formed precise and distinctive ensemble codes from human bodies alone, it seems likely that effects of ensemble coding on social cognition and behavior would be most pronounced when both face *and* body information is available dur-

[3] We suspect that learning played a heightened role in initial trials for the emotion task given the abstract nature of the response variable ("angry" and "neutral" posed bodies) relative to other tasks where the response variable was more concrete or stereotypically paired with the judgement (e.g., pink and blue male and female bodies are more commonly observed).

ing group perception. Future studies should clarify how summary perception of faces and bodies co-occur as we naturalistically perceive the two in tandem. It is possible that faces attract our visual attention more so than do bodies, but that bodies are also attended to in summary perception processes (like in person identification, e.g., O'Toole et al., 2011); it is also possible that bodies may contribute more to perceptions of certain characteristics (e.g., emotion) than faces (see Aviezer et al., 2012; but see also Albohn et al., 2022).

Future work on summary perception should also draw connections between the perceptual process of ensemble coding of bodies and downstream implications for social judgements, particularly those pertinent to relevant marginalized groups. Ensemble coding of faces is recognized as a potential mechanism for understanding social disparities (e.g., Kardosh et al., 2022); Ensemble coding of bodies is likely to be relevant for understanding how people with marginalized identities – especially those marginalized on the basis of their bodies, such as fat people – perceive their social worlds (see Oswald & Adams, 2022; Oswald et al., 2022).

Limitations

A primary limitation of the current work is the lack of external validity; when we encounter natural groups, body information is available alongside facial and contextual information. While much existing work has focused on face ensemble coding in the absence of body information, in these studies we do the opposite by focusing on body ensemble coding in the absence of facial information. We do so in order to understand body perception processes as distinct from confounding facial perception processes, as is often done in work examining individual body perception (see Griffin & Oswald, 2022); Future research should however examine ensemble coding in naturalistic contexts with full information availability to better understand how results from these paradigms might generalize to real-world contexts.

Additionally, future studies should examine whether these effects replicate with different stimuli. We utilized two different sets of controlled body stimuli – a self-developed set of 3D modelled bodies and the BEAST (de Gelder & Van den Stock, 2011). There are not many standardized sets of realistic body stimuli available, however, it would be advantageous to examine whether the effects described here extend to more diverse stimuli and to different body characteristics (e.g., muscularity, race/ethnicity), particularly given what we hypothesize to be stimulus-specific effects for ensemble coding of gender using the BEAST.

Conclusion

The current findings are consistent with the notion that visual processes rapidly summarize group characteristics from both concrete and abstract bodily information. Like ensemble coding of faces, it is likely that summary precepts of groups of bodies has important social meaning for perceivers; for example, these precepts might cue belonging in a given group or activate processes of social identity threat. Embracing bodies in ensemble coding research could enhance our understanding of both the visual processes and downstream social inferences that color our experiences in social groups.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10919-023-00443-8.

Author Contributions F.O conceptualized the project and wrote the initial draft of the manuscript. J.W.G. programmed the experiment and conducted data analysis and data visualization. M.W. and R.B.A. contributed to conceptual development and writing. All authors reviewed and edited the final manuscript.

Declarations

Competing Interests The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. Data and materials can be accessed at: https://osf.io/whqgk/?view_only=65ea0a2edb654727bbb290e6965ed064.

References

Albohn, D. N., Brandenburg, J. C., Kveraga, K., & Adams, R. B. Jr. (2022). The shared signal hypothesis: Facial and bodily expressions of emotion mutually inform one another. *Attention Perception & Psychophysics*, *84*, 2271–2280. https://doi.org/10.3758/s13414-022-02548-6.

Alt, N. P., & Phillips, L. T. (2021). Person perception, meet people perception: Exploring the social vision of groups. *Perspectives on Psychological Science*. https://doi.org/10.1177/17456916211017858. (Advance online publication).

Alt, N. P., Goodale, B., Lick, D. J., & Johnson, K. L. (2019). Threat in the company of men: Ensemble perception and threat evaluations of groups varying in sex ratio. *Social Psychological and Personality Science*, *10*(2), 152–159. https://doi.org/10.1177/1948550617731498.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131. https://doi.org/10.1016/j.tics.2011.01.003.

App, B., McIntosh, D. N., Reed, C. L., & Hertenstein, M. J. (2011). Nonverbal channel use in communication of emotion: How may depend on why. *Emotion*, *11*(3), 603–617. https://doi.org/10.1037/a0023164.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157–162. https://doi.org/10.1111/1467-9280.00327.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, *338*(6111), 1225–1229. https://doi.org/10.1126/science.1224313.

Bannerman, R. L., Milders, M., De Gelder, B., & Sahraie, A. (2009). Orienting to threat: Faster localization of fearful facial expressions and body postures revealed by saccadic eye movements. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1662), 1635–1641. https://doi.org/10.1098/rspb.2008.1744.

Bates, D., Maechler, 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

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497.

Bratovcic, V., Mikic, B., Kostovski, Z., Teskeredzic, A., & Tanovic, I. (2015). Relations between different dimensions of self-perception, self-esteem and body mass index of female students. *International Journal of Morphology*, *33*(4), 1338–1342.

Brewer, M. B. (2007). The social psychology of intergroup relations: Social categorization, ingroup bias, and outgroup prejudice. In A. W. Kruglanski, & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (pp. 695–715). The Guilford Press.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, *62*(3), 193–217. https://doi.org/10.1037/h0047470.

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

Centers for Disease Control and Prevention (2021). *Body measurements*https://www.cdc.gov/nchs/fastats/body-measurements.htm.

Chang, T. Y., & Gauthier, I. (2022). Domain-general ability underlies complex object ensemble processing. *Journal of Experimental Psychology: General*, *151*(4), 966–972. https://doi.org/10.1037/xge0001110.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*(4), 393–404. https://doi.org/10.1016/S0042-6989(02)00596-5.

Corbett, J. E., Utochkin, I., & Hochstein, S. (2023). *The pervasiveness of ensemble perception: Not just your average review*. Cambridge University Press.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Coy, A. E., Green, J. D., & Price, M. E. (2014). Why is low waist-to-chest ratio attractive in males? The mediating roles of perceived dominance, fitness, and protection ability. *Body Image*, *11*(3), 282–289. https://doi.org/10.1016/j.bodyim.2014.04.003.

Dasgupta, N., Banaji, M. R., & Abelson, R. P. (1999). Group entitativity and group perception: Associations between physical features and psychological judgment. *Journal of Personality and Social Psychology*, *77*(5), 991–1003. https://doi.org/10.1037/0022-3514.77.5.991.

de Fockert, J., & Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces. *Quarterly Journal of Experimental Psychology*, *62*(9), 1716–1722. https://doi.org/10.1080/17470210902811249.

de Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3475–3484. https://doi.org/10.1098/rstb.2009.0190.

de Gelder, B., & Van den Stock, J. (2011). The bodily expressive action stimulus test (BEAST). Construction and validation of a stimulus basis for measuring perception of whole body expression of emotions. *Frontiers in Psychology*, *2*, 181. https://doi.org/10.3389/fpsyg.2011.00181.

Fan, J., Dai, W., Liu, F., & Wu, J. (2005). Visual perception of male body attractiveness. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1560), 219–226. https://doi.org/10.1098/rspb.2004.2922.

Fink, B., Neave, N., & Manning, J. T. (2003). Second to fourth digit ratio, body mass index, waist-to-hip ratio, and waist-to-chest ratio: Their relationships in heterosexual men and women. *Annals of Human Biology*, *30*(6), 728–738. https://doi.org/10.1080/03014460310001620153.

Fox, E., Russo, R., & Georgiou, G. A. (2005). Anxiety modulates the degree of attentive resources required to process emotional faces. *Cognitive Affective & Behavioral Neuroscience*, *5*(4), 396–404. https://doi.org/10.3758/CABN.5.4.396.

Freeman, J. B., Ma, Y., Han, S., & Ambady, N. (2013). Influences of culture and visual context on real-time social categorization. *Journal of Experimental Social Psychology*, *49*(2), 206–210. https://doi.org/10.1016/j.jesp.2012.10.015.

Goldenberg, A., Sweeny, T. D., Shpigel, E., & Gross, J. J. (2020). Is this my group or not? The role of ensemble coding of emotional expressions in group categorization. *Journal of Experimental Psychology: General*, *149*(3), 445–460. https://doi.org/10.1037/xge0000651.

Goodale, B. M., Alt, N. P., Lick, D. J., & Johnson, K. L. (2018). Groups at a glance: Perceivers infer social belonging in a group based on perceptual summaries of sex ratio. *Journal of Experimental Psychology: General*, *147*(11), 1660–1676. https://doi.org/10.1037/xge0000450.

Griffin, J. W., & Oswald, F. (2022). A multilevel bayesian meta-analysis of the body inversion effect: Evaluating controversies over headless and sexualized bodies. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-022-02067-3. Advance online publication.

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*(17), R751–R753. https://doi.org/10.1016/j.cub.2007.06.039.

Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 718–734. https://doi.org/10.1037/a0013899.

Haberman, J., Brady, T. F., & Alvarez, G. A. (2015). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General*, *144*(2), 432–446. https://doi.org/10.1037/xge0000053.

Hartmann, A., Borgers, T., Thomas, J. J., Giabbiconi, C. M., & Vocks, S. (2020). Faced with one's fear: Attentional bias in anorexia nervosa and healthy individuals upon confrontation with an obese body stimulus in an eye-tracking paradigm. *Brain and Behavior*, *10*(11), e01834. https://doi.org/10.1002/brb3.1834.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0.

Hester, N., & Gray, K. (2018). For Black men, being tall increases threat stereotyping and police stops. *Proceedings of the National Academy of Sciences*, *115*(11), 2711–2715. https://doi.org/10.1073/pnas.1714454115.

Hugenberg, K., & Sacco, D. F. (2008). Social categorization and stereotyping: How social categorization biases person perception and face memory. *Social and Personality Psychology Compass*, *2*(2), 1052–1072. https://doi.org/10.1111/j.1751-9004.2008.00090.x.

Im, H. Y., Albohn, D. N., Steiner, T. G., Cushing, C., Adams, R. B. Jr., & Kveraga, K. (2017). Differential hemispheric and visual stream contributions to ensemble coding of crowd emotion. *Nature Human Behaviour*, *1*, 828–842. https://doi.org/10.1038/s41562-017-0225-z.

Ip, G., Chiu, C. Y., & Wan, C. (2006). Birds of a feather and birds flocking together: Physical versus behavioral cues may lead to trait- versus goal-based group perception. *Journal of Personality and Social Psychology*, *90*(3), 368–381. https://doi.org/10.1037/0022-3514.90.3.368.

Johnson, K. L., Lick, D. J., & Carpinella, C. M. (2015). Emergent research in social vision: An integrated approach to the determinants and consequences of social categorization. *Social and Personality Psychology Compass*, *9*(1), 15–30. https://doi.org/10.1111/spc3.12147.

Johnstone, L. T., & Downing, P. E. (2017). Dissecting the visual perception of body shape with the Garner selective attention paradigm. *Visual Cognition*, *25*(4–6), 507–523. https://doi.org/10.1080/13506285.2017.1334733.

Joiner, T. E. Jr., Schmidt, N. B., & Singh, D. (1994). Waist-to-hip ratio and body dissatisfaction among college women and men: Moderating role of depressed symptoms and gender. *International Journal of Eating Disorders*, *16*(2), 199–203. https://doi.org/10.1002/1098-108X(199409)16:2%3C199::AID-EAT2260160211%3E3.0.CO;2-M.

Kardosh, R., Sklar, A. Y., Goldstein, A., Pertzov, Y., & Hassin, R. R. (2022). Minority salience and the overestimation of individuals from minority groups in perception and memory. *Proceedings of the National Academy of Sciences*, *119*(12), e2116884119. https://doi.org/10.1073/pnas.2116884119.

Kramer, R. S. S., Young, A. W., Day, M. G., & Burton, A. M. (2017). Robust social categorization emerges from learning the identities of very few faces. *Psychological Review*, *124*(2), 115–129. https://doi.org/10.1037/rev0000048.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13.

Lamer, S. A., Sweeny, T. D., Dyer, M. L., & Weisbuch, M. (2018). Rapid visual perception of interracial crowds: Racial category learning from emotional segregation. *Journal of Experimental Psychology: General*, *147*(5), 683–701. https://doi.org/10.1037/xge0000443.

Max Planck Institute (2011). Perceiving Systems Body Visualizer. https://bodyvisualizer.com.

McElvaney, T. J., Osman, M., & Mareschal, I. (2021). Perceiving threat in others: The role of body morphology. *PloS One*, *16*(4), e0249782. https://doi.org/10.1371/journal.pone.0249782.

Mihalache, D., Lamer, S. A., Allen, J., Maher, M., & Sweeny, T. D. (2021). Anger bias in the evaluation of crowds. *Journal of Experimental Psychology: General*, *150*(9), 1870–1889. https://doi.org/10.1037/xge0001025.

Miller, A. L., & Sheldon, R. (1969). Magnitude estimation of average length and average inclination. *Journal of Experimental Psychology*, *81*(1), 16–21. https://doi.org/10.1037/h0027430.

Molarius, A., Seidell, J. C., Sans, S., Tuomilehto, J., & Kuulasmaa, K. (1999). Waist and hip circumferences, and waist-hip ratio in 19 populations of the WHO MONICA Project. *International Journal of Obesity*, *23*(2), 116–125. https://doi.org/10.1038/sj.ijo.0800772.

Neumann, M. F., Schweinberger, S. R., & Burton, A. M. (2013). Viewers extract mean and individual identity from sets of famous faces. *Cognition*, *128*(1), 56–63. https://doi.org/10.1016/j.cognition.2013.03.006.

O'Toole, A. J., Phillips, P. J., Weimer, S., Roark, D. A., Ayyad, J., Barwick, R., & Dunlop, J. (2011). Recognizing people from dynamic and static faces and bodies: Dissecting identity with a fusion approach. *Vision Research*, *51*(1), 74–83. https://doi.org/10.1016/j.visres.2010.09.035.

Oswald, F., & Adams, R. B. Jr. (2022). Feminist social vision: Seeing through the lens of marginalized perceivers. *Personality and Social Psychology Review*. https://doi.org/10.1177/10888683221126582. Advance online publication.

Oswald, F., Stevens, S., Kruk, M., Murphy, C. I., & Matsick, J. L. (2022). Signaling sizeism: An assessment of body size-based threat and safety cues. *Analyses of Social Issues and Public Policy*. Advance online publication. https://doi.org/10.1111/asap.12301.

Oyserman, D., Brickman, D., Bybee, D., & Celious, A. (2006). Fitting in matters: Markers of in-group belonging and academic outcomes. *Psychological Science*, *17*(10), 854–861. https://doi.org/10.1111/j.1467-9280.2006.01794.x.

Paganini, G. A., McConnell, A. A., Deska, J. C., Almaraz, S. M., Hugenberg, K., & Lloyd, E. P. (2023). Waist-to-hip ratio predicts sexual perception and responses to sexual assault disclosures. *Personality and Social Psychology Bulletin*. Advanced online publication. https://doi.org/10.1177/01461672221148008.

Pazhoohi, F., Hosseinchari, M., & Doyle, J. F. (2012). Iranian men's waist-to-hip ratios, shoulder-to-hip ratios, body esteem and self-efficacy. *Journal of Evolutionary Psychology*, *10*(2), 61–67. https://doi.org/10.1556/jep.10.2012.2.2.

Phillips, L. T., Weisbuch, M., & Ambady, N. (2014). People perception: Social vision of groups and consequences for organizing and interacting. *Research in Organizational Behavior*, *34*, 101–127. https://doi.org/10.1016/j.riob.2014.10.001.

Phillips, L. T., Slepian, M. L., & Hughes, B. L. (2018). Perceiving groups: The people perception of diversity and hierarchy. *Journal of Personality and Social Psychology*, *114*(5), 766–785. https://doi.org/10.1037/pspi0000120.

Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological Science*, *14*(4), 302–308. https://doi.org/10.1111/1467-9280.14431.

Rubino, F., Puhl, R. M., Cummings, D. E., Eckel, R. H., Ryan, D. H., Mechanick, J. I., Nadglowski, J., Salas, X. R., Schauer, P. R., Twenefour, D., Apovian, C. M., Arrone, L. J., Batterham, R. L., Berthoud, H., Boza, C., Busetto, L., Dicker, D., De Groot, M., Eisenberg, D., & Dixon, J. B. (2020). Joint international consensus statement for ending stigma of obesity. *Nature Medicine*, 1–13. https://doi.org/10.1038/s41591-020-0803-x.

Schmader, T., & Sedikides, C. (2018). State authenticity as fit to environment: The implications of social identity for fit, authenticity, and self-segregation. *Personality and Social Psychology Review*, *22*(3), 228–259. https://doi.org/10.1177/1088868317734080.

Singh, D. (1994). Ideal female body shape: Role of body weight and waist-to-hip ratio. *International Journal of Eating Disorders*, *16*(3), 283–288. https://doi.org/10.1002/1098-108X(199411)16:3%3C283::AID-EAT2260160309%3E3.0.CO;2-Q.

Staiano, A. E., Reeder, B. A., Elliott, S., Joffres, M. R., Pahwa, P., Kirkland, S. A., & Katzmarzyk, P. T. (2012). Body mass index versus waist circumference as predictors of mortality in canadian adults. *International Journal of Obesity*, *36*(11), 1450–1454. https://doi.org/10.1038/ijo.2011.268.

Sweeny, T. D., Haroz, S., & Whitney, D. (2013). Perceiving group behavior: Sensitive ensemble coding mechanisms for biological motion of human crowds. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(2), 329–337. https://doi.org/10.1037/a0028712.

Thierry, G., Pegna, A. J., Dodds, C., Roberts, M., Basan, S., & Downing, P. (2006). An event-related potential component sensitive to images of the human body. *Neuroimage*, *32*(2), 871–879. https://doi.org/10.1016/j.neuroimage.2006.03.060.

Thompson, J. K., & Thompson, C. M. (1986). Body size distortion and self-esteem in asymptomatic, normal weight males and females. *International Journal of Eating Disorders*, *5*(6), 1061–1068. https://doi.org/10.1002/1098-108X(198609)5:6%3C1061::AID-EAT2260050609%3E3.0.CO;2-C.

Trope, Y., Liberman, N., & Wakslak, C. (2007). Construal levels and psychological distance: Effects on representation, prediction, evaluation, and behavior. *Journal of Consumer Psychology*, *17*(2), 83–95. https://doi.org/10.1016/S1057-7408(07)70013-X.

Volpert-Esmond, H. I., & Bartholow, B. D. (2021). A functional coupling of brain and behavior during social categorization of faces. *Personality and Social Psychology Bulletin*, *47*(11), 1580–1595. https://doi.org/10.1177/0146167220976688.

Webster, J., Kay, P., & Webster, M. A. (2014). Perceiving the average hue of color arrays. *Journal of the Optical Society of America A*, *31*(4), A283–A292. https://doi.org/10.1364/JOSAA.31.00A283.

Weisbuch, M., & Ambady, N. (2011). Thin-slice vision. In Adams Jr., R. B., Ambady, N., Nakayma, K., & Shimojo, S. (Eds.), *The science of social vision* (pp. 228–247). Oxford.

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology*, *69*(1), 105–129. https://doi.org/10.1146/annurev-psych-010416-044232.

Wilson, J. P., Hugenberg, K., & Rule, N. O. (2017). Racial bias in judgments of physical size and formidability: From size to threat. *Journal of Personality and Social Psychology*, *113*(1), 59–80. https://doi.org/10.1037/pspi0000092.

Woolhouse, M. H., & Lai, R. (2014). Traces across the body: Influence of music-dance synchrony on the observation of dance. *Frontiers in Human Neuroscience*, *8*, 965. https://doi.org/10.3389/fnhum.2014.00965.