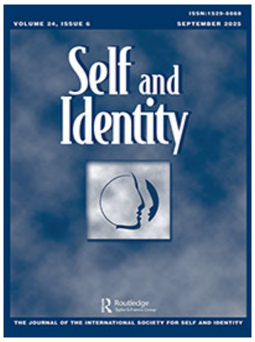

Self and Identity

Routledge
Taylor & Francis Group

ISSN: 1529-8868 (Print) 1529-8876 (Online) Journal homepage: www.tandfonline.com/journals/psai20

The self as a representational base: my sense of 'me' shapes what I think 'us' looks like

Youngki Hong & Kyle G. Ratner

To cite this article: Youngki Hong & Kyle G. Ratner (31 Aug 2025): The self as a representational base: my sense of 'me' shapes what I think 'us' looks like, Self and Identity, DOI: 10.1080/15298868.2025.2555463

To link to this article: https://doi.org/10.1080/15298868.2025.2555463

Published online: 31 Aug 2025.

SELF AND IDENTITY
https://doi.org/10.1080/15298868.2025.2555463

Routledge
Taylor & Francis Group

The self as a representational base: my sense of 'me' shapes what I think 'us' looks like

Youngki Hong[a] and Kyle G. Ratner[b]

[a]Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, USA; [b]Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA, USA

ABSTRACT

Previous research shows that people mentally represent novel ingroup and outgroup faces differently in favor of their ingroup. How are people able to mentally represent faces of others whom they have never met before? The current research tests the self-as-a-representational-base hypothesis: people use their self-evaluation and self-image as a basis for visualizing novel ingroup faces. Study 1 showed that individuals with high self-esteem generated more favorable ingroup images, indicating that their self-evaluations relate to visual representations of ingroup members. Study 2 found that an independent sample of participants could match photographs of individuals who generated ingroup images significantly better than chance, suggesting that self-image directly relates to visual representations of ingroup members. Study 3 used OpenFace to show that individuals' photographs objectively resemble their own visual representations of ingroup more than outgroup members. Together, we show that the sense of self can explain how we visually represent novel ingroup members.

ARTICLE HISTORY
Received 12 December 2024
Revised 26 August 2025
Accepted 27 August 2025

KEYWORDS
Faces; minimal group paradigm; reverse correlation; self-image; self-evaluation

How do people imagine what others look like? Mental representations of others shape how we evaluate, trust, and interact with others (Kunda & Thagard, 1996; Lloyd et al., 2020; Ratner et al., 2014). Previous research has shown that people form mental images of others based on both personal and social information, a process Warr and Knapper (1968) described as "person conception." These images may draw on exemplar memories or prototypical traits stored in memory (Nosofsky & Zaki, 2002). One powerful way to study these mental images is through reverse correlation image classification methods (e.g., Dotsch & Todorov, 2012). This technique presents participants with many pairs of noisy, face-like images and asks them to repeatedly choose which one looks more like a member of a target group (e.g., "Which face looks more American?"; Hong et al., 2024). By aggregating these choices, researchers can generate composite images that reflect the participant's mental representation of the group in question (Dotsch & Todorov, 2012). Reverse correlation thus provides a way to make the "invisible" contents of mental representations visible.

CONTACT Youngki Hong youngki.hong@colorado.edu Department of Psychology and Neuroscience, University of Colorado Boulder, Muenzinger D244, 345 UCB, Boulder, CO 80309-0345, USA

This method has been especially useful in studying minimal group contexts, where individuals are tasked with visualizing faces of others with whom they have never met or interacted before. The minimal group paradigm (Tajfel et al., 1971) is an experimental framework used to study intergroup behavior by assigning participants to groups based on arbitrary or trivial criteria (e.g., preference for a painting). Despite the lack of meaningful differences or prior interaction, participants reliably show ingroup favoritism, such as allocating more resources to ingroup members than to outgroup members. Pertinent to the current research, previous studies have also used the minimal group paradigm to separate participants into novel groups and then used the reverse correlation image classification technique to visualize what they thought faces of novel ingroup and outgroup members look like (Hong & Ratner, 2021; Ratner et al., 2014). They find that renderings of novel ingroup face representations are associated with more favorable attributes than those of outgroup, most notably trustworthiness and competence (Hong et al., 2023). However, it is unclear whether such "ingroup favoritism" is the only source of visual representation or if other factors may also contribute to the construction of mental imagery of novel group members. For instance, in cases of individuals showing little to no preference for their own group over the outgroup and thereby not relying on facial cues that signal desirable traits as much, how can we explain the distinct visual representations of ingroup and outgroup members? Indeed, there is considerable variation in the extent to which individuals express ingroup favoritism (e.g., Brewer, 1999; Hewstone et al., 2002).

One possible source of visual representation in such cases is the sense of self. The self as intrinsically linked to one's social identity, such that the self is always part of the ingroup (the group to which one belongs) rather than the outgroup. Humans have a fundamental need to belong (Baumeister & Leary, 1995), which in turn influences how they perceive others and themselves (Brewer, 1991; Gramzow & Gaertner, 2005). According to Social Identity Theory and Self-Categorization Theory (Hornsey, 2008; Tajfel & Turner, 1979), individuals derive their sense of self from their group memberships and desire to see themselves and their groups as distinct from others. This may lead people to assimilate their self-concept to the characteristics of their ingroup (i.e., self-stereotyping; Hogg & Turner, 1987), even at the cost to oneself (Elder et al., 2022). Additionally, the sense of self can influence intergroup behavior and cognition. For example, individuals may believe that ingroup members share their views (Holtz & Miller, 1985; Krueger & Zeiger, 1993). Similarly, self-evaluation can affect evaluations of ingroup members (Crocker & Luhtanen, 1990; Gramzow & Gaertner, 2005). These studies indicate a strong link between the self and the ingroup, opening up the possibility that the self could influence visual representations of novel group members. However, these studies typically focus on preexisting groups making it difficult to isolate the effects of the sense of self from the influences of knowledge about the groups (e.g., stereotypes) and prior interactions with group members. Thus, the question remains: do people's sense of self influence their visual representations of novel group members when group distinctions are impoverished, such as in minimal group contexts?

Several hypotheses can be drawn regarding how the self may influence visual representations of novel ingroup and outgroup members. First, how people evaluate themselves may guide what they think about what their ingroup members should look like. Prominent theories suggest that people tend to favor ingroups to protect and maintain their self-esteem (Hogg & Abrams, 1990; Tajfel & Turner, 1979). Indeed, previous research shows that self-esteem is related to intergroup discrimination even when group distinctions are novel (Gramzow & Gaertner, 2005). Relatedly, physical appearance is a source of self-esteem because people who are trustworthy looking and attractive are valued by society more than those who are not (Little et al., 2011). Thus, it is possible that those who evaluate themselves favorably (i.e., those with high self-esteem) may rely more heavily on facial cues that convey desirable attributes when visualizing members of their ingroups, compared to those who may have lower self-esteem. Alternatively, people may directly rely on their own self-image when visualizing their ingroup members. Prominent social groupings, such as family units, friend networks, and racial categories, are often bonded through a shared cultural and genetic heritage that results in members of a group sharing appearance cues ranging from how they dress and cut their hair to the contours and features of their faces (DeBruine, 2004; Hehman et al., 2018). Thus, the notion that ingroup members physically resemble the self could be so ingrained in social category representations that it persists even when groups have no shared social history and are not defined by physical attributes.

In this research, we examine how self-evaluation and self-image may contribute to visual representations of novel ingroup and outgroup members. To this end, we use the reverse correlation image classification technique and the minimal group paradigm. Study 1 uses an economic trust game to investigate whether self-evaluation, as measured by self-esteem, relates to how favorable people's visual representations of novel ingroup and outgroup members appear. Study 2 uses a photo-matching game to examine how similar the ingroup and outgroup visual representations are to the photographs of the individuals who generated those images (self-image). Study 3 uses a convolutional neural network approach to examine the objective similarity between individuals' self-images and their visual representations of ingroup versus outgroup members, and to parse the relative contributions of self-evaluation and self-image in visual representations of novel ingroup and outgroup members. See Figure 1 for the overview of the studies. Together, this research demonstrates that at the earliest stage of social category formation (i.e., when only the "us" versus "them" distinction is definitively knowable), ingroup representations are influenced by a person's sense of self. All data, study materials, and analysis scripts are publicly available at https://osf.io/tma5x/?view_only=0a6cc474962f4812b72fd4023dc4de6e.

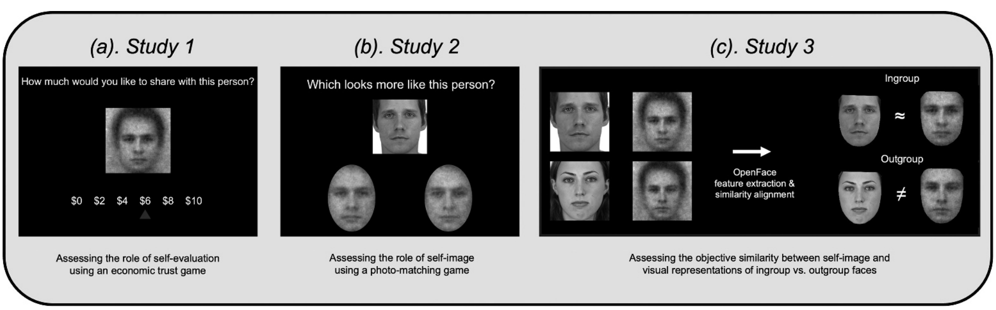

Figure 1. Schematic overview of the current research. To keep the identities of the actual participants anonymous, we used face images from the chicago face database (Ma et al., 2015) for this illustration only.

Study 1

Study 1 examined whether people's evaluations of themselves influence how they visualize novel ingroup and outgroup members. Specifically, we tested the hypothesis that individuals with higher self-esteem would generate ingroup face representations that appear more trustworthy to others.

Method

Stimuli

The present and subsequent studies used novel ingroup and outgroup face images from Hong and Ratner (2021) as stimuli. This set of stimuli included two samples of reverse correlation participant-level classification images (CIs) of novel ingroup and outgroup faces. These two samples were identical except for their sample sizes ($N = 362$ vs. $N = 200$) and the versions of the minimal group paradigm used to create the ingroup/outgroup distinctions (dot-estimation style vs. artistic preference from Tajfel et al., 1971). In the studies, participants completed a 450-trial face categorization task after getting assigned to a novel group. On each trial, they selected either an ingroup or an outgroup face from two adjacent grayscale, noisy face images. Half of the participants were instructed to choose the ingroup face on every trial, while the other half were instructed to choose the outgroup face. Finally, classification images for ingroup and outgroup faces were generated for each participant by averaging each participant's responses from the face categorization task. For more details on how these CIs were produced, please refer to Studies 1 and 2 in Hong and Ratner (2021).

After generating novel ingroup and outgroup face images, self-esteem of each participant who generated those images was assessed using the Rosenberg self-esteem scale (RSE; Rosenberg, 1979). Lastly, a subset of participants who generated the ingroup and outgroup images gave their permission to have their photographs taken. They were asked to maintain a neutral facial expression. Participants who declined to consent to have their photograph taken were not penalized. The photographs and their associated data are analyzed in Studies 2 and 3.

Participants

We recruited two convenient (Sample 1: $N = 108$, $M_{age} = 18.77$, $SD = 1.37$; 67 female, 41 male; Sample 2: $N = 148$, $M_{age} = 19.16$, $SD = 1.44$; 94 female, 54 male) to participate in an economic trust game study in exchange for course credit. Participants were undergraduate students a large public research university on the California coast. The student body is predominantly traditional-aged (18–22), with diverse racial and ethnic representation (approximately 29% Hispanic/Latinx, 19% Asian, and 35% White). The racial and ethnic breakdown of our samples was Sample 1: 40 White, 32 Asian, 20 Latinx, 1 Black, 10 multiracial, and 5 other; Sample 2: 50 Asian, 38 Latinx, 38 White, 16 multiracial, 3 Black, 1 Pacific Islander/Hawaiian, 2 other, and 1 unidentified. Up to four participants were run simultaneously.

Procedure

Participants played an economic trust game with different interaction partners. The interaction partners were the ingroup and outgroup face CIs. Sample 1 played the game with 362 CIs from Study 1 and Sample 2 with 200 CIs from Study 2 of Hong and Ratner (2021). Participants in the current studies were naïve to the image generation phase and therefore were not aware whether face images appeared in the economic trust game depicted ingroup members or outgroup members. Participants were instructed to imagine they had $10 on each trial and could choose to either keep this money or share a certain amount with their interaction partners, represented by the ingroup and outgroup CIs. During each trial, participants made a decision to share a portion of the $10 (i.e., $0, $2, $4, $6, $8, or $10). They were informed that any money they shared would be quadrupled and given to the interaction partner, who would then have the option to return half of the total amount back to the participant who shared. Thus, participants could earn more money than if they had kept the money to themselves. However, participants indicated how much money they wanted to share with each partner without receiving actual feedback. The amount of money shared was used as an indicator of how much the participants trusted their interaction partners.[1]

Results

First, we averaged the amount of money each participant-level CI received. We then used a linear mixed-effects model to predict the average amount of money received in the trust game, using the self-esteem scores of those who generated participant-level CIs (centered), group membership of the CI (outgroup = −0.5, ingroup = 0.5), and their interaction. We included sample (Sample 1 vs. Sample 2) as a random intercept in the linear mixed-effects model to account for potential variability across samples. This random effect accounted for minimal variance and did not meaningfully influence the results, indicating that findings were consistent across samples.

The main effect of group membership was significant, b = .43, β = .29, SE = .06, z = 7.28, $p < .001$, 95% CI = [.31, .54], Cohen's d = .61, indicating that the ingroup faces received more money (*M* = 3.02, SD = .71) than the outgroup faces (*M* = 2.59, SD = .71). The main effect of self-esteem was not significant, b = .00, β = .04, SE = .00, z = .95, *p* = .34, 95% CI = [−.00, .01]. Critically, there was a significant interaction, b = .02, β = .11, SE = .01, z = 2.76, *p* = .006, 95% CI = [.00, .03]. That is, the effect of self-esteem was significant for the ingroup faces, simple b = .11, β = .17, SE = .00, z = 2.72, *p* = .007, 95% CI = [.00, .02], indicating that individuals with higher self-esteem generated ingroup face images that received more money (Figure 2). However, self-esteem was not predictive of the amount of money the outgroup faces received, simple b = −.01, β = −.10, SE = .00, z = 1.24, *p* = .22, 95% CI = [−.02, .00].

A sensitivity power analysis was conducted to determine the minimum detectable effect size with a sample size of 562 (the total number of CIs), with a power of .80 and an alpha level of .05. The analysis indicated that the smallest detectable effect size was Cohen's d = .24. The observed effect size in the study was Cohen's d = .61, which is above the threshold of detectability established by the power analysis.

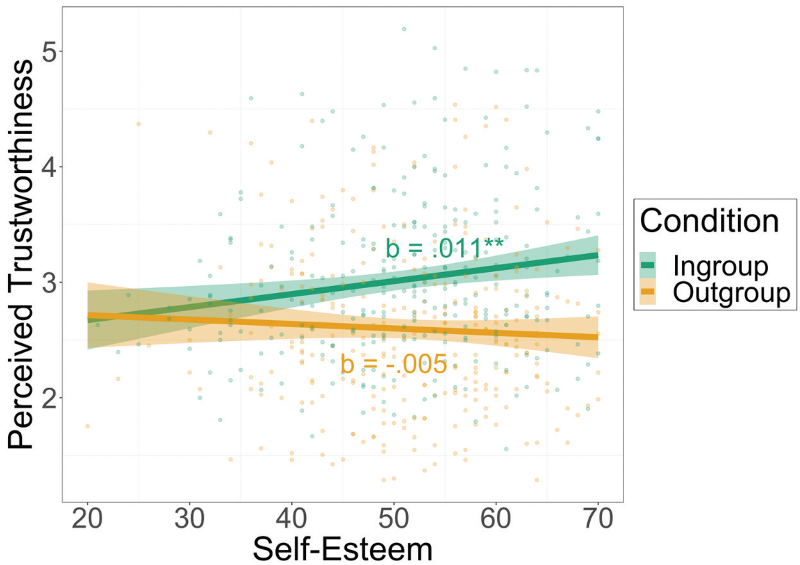

Figure 2. Perceived trustworthiness (average amount of money received in the trust game) of novel ingroup and outgroup face representations as a function of self-esteem. Significance codes:** < .01

Discussion

In Study 1, we found that group membership affected the visual representations of novel group members, with ingroup faces appearing more trustworthy than their outgroup counterparts (i.e., received more money in the trust game). Further, we found that such differences in perceived trustworthiness were moderated by the self-evaluations of those who generated the images. Those with higher self-esteem generated ingroup faces that appeared more trustworthy compared to those with lower self-esteem. Self-esteem did not moderate the perceived trustworthiness of outgroup faces.

Study 2

Study 2 examined whether people's own self-image directly shapes how they visualize novel group members. While Study 1 focused on self-evaluation as a potential influence on ingroup representations, Study 2 tested whether the physical appearance of the self might serve as a cue in constructing these mental images. We predicted that ingroup visual representations would resemble the individuals who generated them more than outgroup representations, suggesting a direct role of self-image in shaping early social category representations.

Method

Participants

We recruited two convenient samples of American college students[2] (Sample 1: $N = 171$, $M_{age} = 18.89$, $SD = 1.44$; 118 female, 53 male) and (Sample 2: $N = 148$, $M_{age} = 19.16$, $SD = 1.44$; 94 female, 54 male) to participate in research about matching faces in exchange for course credit. We did not predetermine our sample size, but instead, we ran as many participants as possible in a single 10-week quarter. All the analyses were conducted after data collection concluded. The racial and ethnic breakdown of our Sample 1 was 64 White, 32 Latinx, 50 Asian, 1 Black, 13 multiracial, 10 other, and 1 unidentified. For Sample 2, the racial and ethnic breakdown was 50 Asian, 38 Latinx, 38 White, 16 multiracial, 3 Black, 1 pacific islander/Hawaiian, 2 other, and 1 unidentified. Up to four participants were run simultaneously in each session.

Stimuli

We had a subset of participants who generated classification images of ingroup and outgroup give permission to have their photographs taken: 418 photographs, 289 from Study 1 and 129 from Study 2 of Hong and Ratner (2021). Out of 418 whose photographs were taken, 210 had generated ingroup faces and 208 had generated outgroup faces.

Procedure

Participants completed a task where they matched faces that looked most like each other. On each trial, participants saw three images: the top image was a photograph of a person who generated a CI of either ingroup or outgroup, and two bottom participant-level classification images, masked to only show face regions (i.e., background and hair removed), one of which was generated by the person in the photograph and the other was randomly chosen from the pool of participant-level classification images generated by the remaining participants whose pictures were used in other trials (Figure 1(b)). Participants were then asked to indicate which of the bottom two images looked most like the photograph on top regardless of the gender of the photograph because all participant-level classification images were male faces (generated from a male base face image); they pressed the "d" key if they thought the image on the left looked most like the photograph and the "k" key if they thought the image on the right looked most like the photograph. Sample 1 played the matching game with 289 CIs from Study 1 and Sample 2 with 129 CIs from Study 2 of Hong and Ratner (2021).

Results

First, we coded correctly matched trials as 1 and incorrect trials as 0. We then averaged these numbers for each CI, generating an average matching accuracy score. We used a linear mixed-effects model to predict the accuracy score of each CI, using the group membership of the CI (ingroup or outgroup). We included sample (Sample 1 vs. Sample 2) as a random intercept in the linear mixed-effects model to account for potential variability

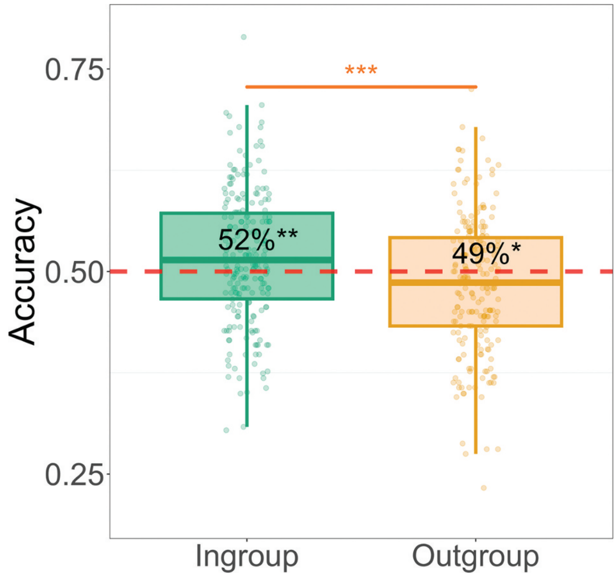

Figure 3. Matching accuracy of ingroup and outgroup faces with photographs of their generators. Significance codes: *** < .001 ** < .01 * < .05

across samples. This random effect accounted for minimal variance and did not meaningfully influence the results, indicating that findings were consistent across samples. The main effect of group membership was significant, b = .03, β = .18, SE = .01, z = 3.83, *p* < .001, 95% CI = [.01, .05], Cohen's d = .37, indicating that the ingroup faces were more accurately matched with the photographs of their generators (*M* = 51.42%, SD = 8.20) than the outgroup faces (*M* = 48.58%, SD = 8.37) (Figure 3). Additionally, we conducted one-sample t-tests to determine if our participants performed significantly better or worse than chance (50%) for ingroup and outgroup images. Ingroup faces were correctly matched with the photographs of people who generated them significantly better than chance, t(209) = 2.93, *p* = .004, 95% CI [50.54, 52.77], Cohen's d = .20. Conversely, outgroup faces were matched significantly worse than chance, t(207) = 2.45, *p* = .02, 95% CI [47.43, 49.72], Cohen's d = .17.

A sensitivity power analysis was conducted to determine the minimum detectable effect size with a sample size of 418 (the total number of CI-photograph pairs), with a power of .80 and an alpha level of .05. The analysis indicated that the smallest detectable effect size was Cohen's d = .28. The observed effect size in the study was Cohen's d = .37, which is above the threshold of detectability established by the power analysis.

Discussion

Study 2 provided evidence that self-image directly relates how people visualize novel group members. Specifically, we found that ingroup images were more accurately matched with the photographs of their generators than outgroup images, suggesting that ingroup images more closely resembled their generators.

Study 3

One possible alternative explanation for the finding that ingroup images were matched with photographs of their generators more accurately is that ingroup images appeared more trustworthy than outgroup images, and since trustworthiness is considered a uniquely human trait (Wilson et al., 2018), participants in our studies might have simply matched photographs to more human-looking (i.e., more trustworthy-looking) classification images, leading to higher accuracy scores for ingroup faces. Therefore, in Study 3, we sought to rule out such an alternative explanation. Specifically, we used OpenFace (Amos et al., 2016) to assess more objective similarities between CIs and photographs of their generators. We then examined the relationships among the objective similarity index, perceived trustworthiness (Study 1), and photo-matching accuracy scores (Study 2). This allowed us to explore how each CI's perceived trustworthiness and its objective similarity to the generator's photograph contributed to photo-matching accuracy.

Method

Procedure

To assess objective similarities between classification images (CIs) and photographs of their generators, we utilized the OpenFace 2.0 toolkit. This open-source tool uses a Convolutional Neural Network (CNN) trained on 500,000 face images and is designed for facial analysis, including facial landmark detection, head pose estimation, facial action unit recognition, and eye-gaze estimation (Amos et al., 2016). OpenFace first estimates the face's shape using a standard face model and refines this by examining local image gradients near each landmark for precise adjustments. It also masks irrelevant areas and aligns facial features across different images by normalizing head orientation and scale. Our focus was specifically on parameters related to 3D eye landmarking and facial features, detected using an ensemble of regression trees with gradient boosting, which predicts landmark positions directly from image pixel intensities. The detected landmarks, which encompass key facial areas such as the jawline, eyebrows, eyes, nose, and mouth contours, enabled us to create an index for objective similarities. This index was generated by computing the Euclidean distance between each CI – photograph pair. We then used it to examine the relative contributions of CIs' perceived trustworthiness and objective similarity to photo-matching accuracy. Eight CI – photograph pairs were discarded because OpenFace was unable to detect faces in the photographs.

Results

First, we used a linear mixed-effects model to predict the objective similarity between CIs and the photographs of their generators, using the group membership of the CI (ingroup or outgroup). We included sample (Sample 1 vs. Sample 2) as a random intercept in the linear mixed-effects model to account for potential variability across samples. This random effect accounted for minimal variance and did not meaningfully influence the results, indicating that findings were consistent across samples. The main effect of group membership was significant, $b = 33.80$, $\beta = .11$, $SE = 13.86$, $z = 2.29$, $p = .02$, 95%

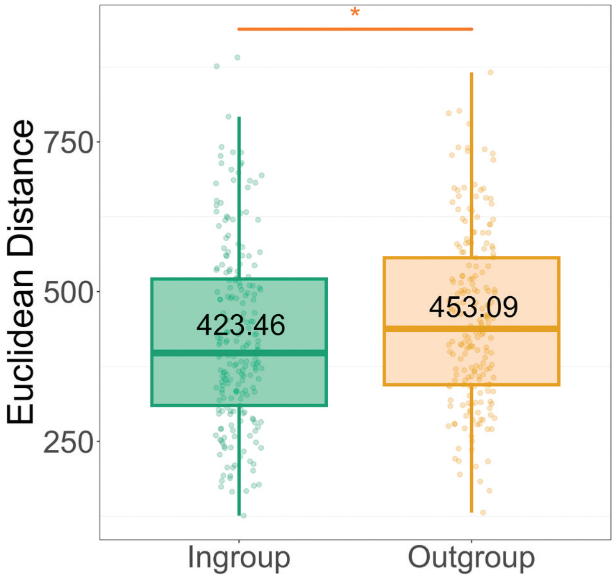

Figure 4. Euclidean distance (objective similarity) between ingroup and outgroup CIs and photographs of their generators based on OpenFace parameters. Significance codes: * < .05

CI = [4.55, 58.95], Cohen's d = .20, indicating that the ingroup faces were more objectively similar to the photographs of their generators (*M* = 423.46, SD = 152.30) than the outgroup faces were to their generators (*M* = 453.09, SD = 142.28) (Figure 4).

A sensitivity power analysis was conducted to determine the minimum detectable effect size with a sample size of 410 (the total number of CI-photograph pairs after OpenFace processing), with a power of .80 and an alpha level of .05. The analysis indicated that the smallest detectable effect size was Cohen's d = .28. The observed effect size in the study was Cohen's d = .20, which is under the threshold of detectability established by the power analysis. Thus, our study was underpowered to reliably detect such an effect, and these results should be interpreted with caution.

Next, we computed correlations among the photo-matching accuracy scores from Study 2, the average amount of money CIs received in the trust game (perceived trustworthiness) from Study 1, and the Euclidean distance between CIs and the photographs of their generators (objective similarity). The correlation between matching accuracy and perceived trustworthiness was significant, r(416) = .34, $p < .001$, 95% CI = [.25, .42], indicating that faces perceived as more trustworthy were more likely to be accurately matched with the photograph of their generators. The correlation between matching accuracy and objective similarity was also significant, r(408) = –.11, $p = .03$, 95% CI = [–.20, –.01], indicating that more objectively similar CI-photograph pairs were also more accurately matched (Figure 5). However, the correlation between perceived trustworthiness and objective similarity was not significant, r(408) = –.02, $p = .74$, 95% CI = [–.11, .08].

Lastly, we used a linear mixed-effects model to predict the photo-matching accuracy scores using perceived trustworthiness and objective similarity. We included sample (Sample 1 vs. Sample 2) as a random intercept in the linear mixed-effects model to account for potential variability across samples. This random effect accounted for minimal

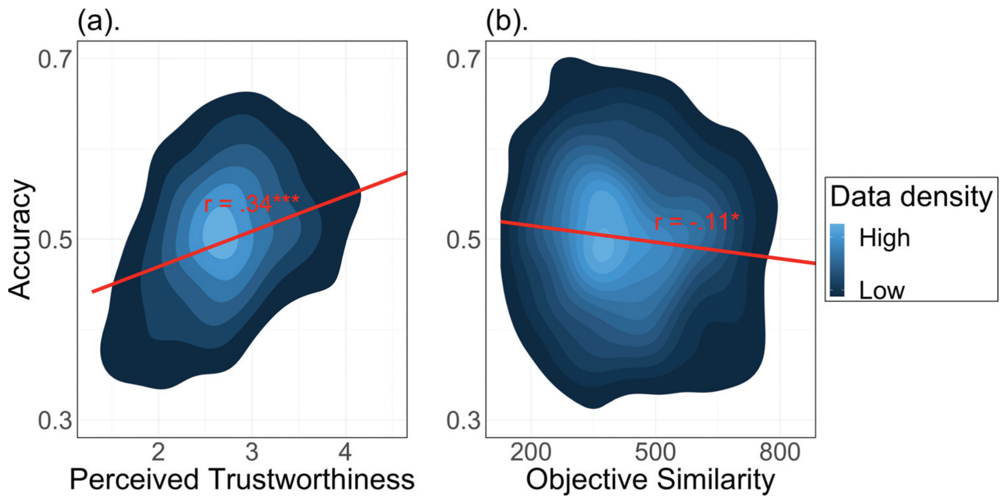

Figure 5. The relationships between photo matching accuracy and (a). perceived trustworthiness and (b). objective similarity. Although images perceived as more trustworthy were more likely to be accurately matched to photographs of their generators, images with greater objective similarity to photographs of their generators were also more likely to be accurately matched. Significance codes: * < .05 *** < .001

variance and did not meaningfully influence the results, indicating that findings were consistent across samples. Both predictors remained significant after controlling for each other: the main effect of perceived trustworthiness, b = .04, β = .37, SE = .01, z = 8.05, $p < .001$, 95% CI = [.03, .05], and the main effect of objective similarity, b = −.00, β = −.10, SE = .00, z = 2.22, $p = .03$, 95% CI = [−.00, −.00].

Discussion

Study 3 did not definitively rule out the alternative explanation that faces perceived as more trustworthy were more accurately matched with the photographs of their generators. However, we demonstrated that the objective similarities between classification images and the photographs of their generators could explain the photo-matching accuracy above and beyond the perceived trustworthiness of the CIs. This finding suggests that self-image does contribute, albeit to a smaller extent than self-evaluation, to people's visual representations of novel group members.

General discussion

Across three studies, we tested whether visual representations of novel ingroup members are influenced by their self-evaluation and self-image. In Study 1, we demonstrated that participants' self-evaluations were reflected in how others perceived the ingroup faces they had generated: those with higher self-esteem generated ingroup classification images that were perceived more trustworthy (i.e., received more money in the trust game) compared to those associated with lower self-esteem. In Study 2, we showed that participants perceived classification images as more similar to the photographs of individuals who generated them if the images were that of an ingroup member rather than an outgroup member, suggesting that self-image directly relates to visual representations of novel ingroup members. In Study 3, we found that ingroup images being more trustworthy-appearing contributed to their perceived similarity to the photographs of the individuals who generated them. However, we also demonstrated that self-image still contributes to how people visually represent faces of novel ingroup members. Overall, these findings suggest that the self may serve as a representational base on which people visualize what a novel ingroup member should look like, relying on different aspects of the self, either self-evaluation or self-image, to do so.

Our research demonstrates that the self plays an important role in shaping social category representations in contexts where people lack exemplar or prototype information retrieved from memory about the groups, such as in minimal group settings. These findings are consistent with, and further extend, the theoretical frameworks of Social Identity Theory and Self-Categorization Theory (Tajfel & Turner, 1979), by demonstrating a strong visual link between the self and social identities. For instance, individuals may be motivated to perceive ingroups positively to maintain a favorable self-esteem (Hogg & Abrams, 1990; Tajfel & Turner, 1979), or they may view themselves favorably and, by extension, assume that their ingroup members also possess similar desirable qualities. Alternatively, individuals may be motivated to perceive outgroups more negatively to enhance their ingroups and, by extension, themselves (Tajfel & Turner, 1979), or they may strive to maximize differentiation from outgroups (Leonardelli et al., 2010). Although our results are more consistent with the interpretation that the ingroup is assimilated to the self (rather than the self is contrasted with the outgroup), given the null relationship between self-esteem and perceived trustworthiness of outgroup images in Study 1, our findings cannot fully disentangle these possibilities. Regardless, our research shows that rather than ingroup and outgroup representations simply drawing on existing social categories, people actively use their self-concept to shape how they visualize and interpret novel ingroup members.

Although our work provides a first demonstration that the self provides a representational base on which people visually represent their group members, there are several limitations of this work. First, the reliance on minimal group paradigms, while useful for isolating specific variables, might not capture the full complexity of real-world social identities and interactions. Future research should explore these dynamics in natural settings where group memberships are preexisting and carry real-world significance. Furthermore, it is possible that we underestimated the true effect size for our findings because the base image used to create classification images was an average White male face when many of our participants who generated the images were not. This aspect of the design could have reduced the sensitivity of our paradigm to detect the use of self-image for non-White male participants, although the findings did not change even after controlling for race and gender of those who generated the images.

Beyond our research focus, the methods we used can benefit other reverse correlation researchers. One ongoing critique of the technique is that results can be difficult to interpret and may vary depending on analytic decisions (Kevane & Koopmann-Holm, 2021). To address some of these concerns, we analyzed participant-level classification images rather than relying on group-level composites, which have been argued to inflate Type I errors (Cone et al., 2020). We also incorporated OpenFace to assess objective similarities between reverse correlation images and photographs of our participants, which can be easily applied to other face processing research (Amos et al., 2016; Hong et al., 2023).

In conclusion, the present work provides evidence that the self plays a formative role in how people represent novel group members. By showing that both self-evaluation and self-image contribute to the construction of social category representations, our findings suggest that the self is not only reflective of group membership but also constructive in shaping it. This perspective advances theories of self and identity by positioning the self as a representational base for social categorization – a process through which people make sense of "us" versus "them" even in the absence of shared history or stereotypes. By examining visual processes and the role of the self in them, this work contributes to our understanding of how interactions with physical and social worlds both shape and are shaped by the self.

Notes

1. While we used the economic trust game to assess the perceived trustworthiness of the face images in a more engaging context than simple face ratings, we do not believe the results would have differed had we instead used simple trustworthiness ratings.
2. They are from the same university population described in Study 1.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Amos, B., Ludwiczuk, B., & Satyanarayanan, M. (2016). *OpenFace: A general-purpose face recognition library with mobile applications*. Technical report. CMU-CS-16-118. Carnegie Mellon University School of Computer Science.

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

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality & Social Psychology Bulletin*, *17*(5), 475–482. https://doi.org/10.1177/0146167291175001

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *The Journal of Social Issues*, *55*(3), 429–444. https://doi.org/10.1111/0022-4537.00126

Cone, J., Brown-Iannuzzi, J. L., Lei, R., & Dotsch, R. (2020). Type I error is inflated in the two-phase reverse correlation procedure. *Social Psychological & Personality Science*, *12*(5), 760–768. https://doi.org/10.1177/1948550620938616

Crocker, J., & Luhtanen, R. (1990). Collective self-esteem and ingroup bias. *Journal of Personality & Social Psychology*, *58*(1), 60–67. https://doi.org/10.1037/0022-3514.58.1.60

DeBruine, L. M. (2004). Resemblance to self increases the appeal of child faces to both men and women. *Evolution and Human Behavior*, *25*(3), 142–154. https://doi.org/10.1016/j.evolhumbehav.2004.03.003

Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological & Personality Science*, *3*(5), 562–571.

Elder, J., Davis, T., & Hughes, B. (2022). Self-derogating to align with group norms: Similarity-based learning drives assimilation of the group to the self. *PsyArxiv*. https://doi.org/10.31234/osf.io/38bf7

Gramzow, R. H., & Gaertner, L. (2005). Self-esteem and favoritism toward novel in-groups: The self as an evaluative base. *Journal of Personality & Social Psychology*, *88*(5), 801–815. https://doi.org/10.1037/0022-3514.88.5.801

Hehman, E., Flake, J. K., & Freeman, J. B. (2018). The faces of group members share physical resemblance. *Personality & Social Psychology Bulletin*, *44*(1), 3–15. https://doi.org/10.1177/0146167217722556

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *The Annual Review of Psychology*, *53*(1), 575–604. https://doi.org/10.1146/annurev.psych.53.100901.135109

Hogg, M. A., & Abrams, D. (1990). Social motivation, self-esteem and social identity. In D. Abrams & M. A. Hogg (Eds.), *Social identity theory: Constructive and critical advances* (pp. 28–47). Harvester Wheatsheaf.

Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self-stereotyping and the salience of social categories. *British Journal of Social Psychology*, *26*(4), 325–340. https://doi.org/10.1111/j.2044-8309.1987.tb00795.x

Holtz, R., & Miller, N. (1985). Assumed similarity and opinion certainty. *Journal of Personality & Social Psychology*, *48*(4), 890–898. https://doi.org/10.1037/0022-3514.48.4.890

Hong, Y., Maitner, A. T., Ratner, K. G. (2024). Psychological effects of anti-Arab politics on American and Arab peoples' views of each other. *PLOS ONE*, *19*(5), e0301282. https://doi.org/10.1371/journal.pone.0301282

Hong, Y., & Ratner, K. G. (2021). Minimal but not meaningless: Seemingly arbitrary category labels can imply more than group membership. *Journal of Personality & Social Psychology*, *120*(3), 576–600. https://doi.org/10.1037/pspa0000255

Hong, Y., Reed, M., & Ratner, K. G. (2023). Facial stereotypes of competence (not trustworthiness and dominance) most resemble facial stereotypes of group membership. *Social Cognition*, *41*(6), 562–578. https://doi.org/10.1521/soco.2023.41.6.562

Hornsey, M. J. (2008). Social identity theory and self-categorization theory: A historical review. *Social and Personality Psychology Compass*, *2*(1), 204–222. https://doi.org/10.1111/j.1751-9004.2007.00066.x

Kevane, M., & Koopmann-Holm, B. (2021). Improving reverse correlation analysis of faces: Diagnostics of order effects, runs, rater agreement, and image pairs. *Behavior Research Methods*, *53*(4), 1609–1647. https://doi.org/10.3758/s13428-020-01499-w

Krueger, J., & Zeiger, J. S. (1993). Social categorization and the truly false consensus effect. *Journal of Personality & Social Psychology*, *65*(4), 670–680. https://doi.org/10.1037/0022-3514.65.4.670

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, *103*(2), 284–308. https://doi.org/10.1037/0033-295X.103.2.284

Leonardelli, G. J., Pickett, C. L., & Brewer, M. B. (2010). Optimal distinctiveness theory: A framework for social identity, social cognition, and intergroup relations. In M. P. Zanna & J. M. Olson (Eds.), *Advances in experimental social psychology*, (Vol. 43, pp. 63–113). Academic Press. https://doi.org/10.1016/S0065-2601(10)43002-6

Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: Evolutionary based research. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1638–1659. https://doi.org/10.1098/rstb.2010.0404

Lloyd, E. P., Sim, M., Smalley, E., Bernstein, M. J., & Hugenberg, K. (2020). Good cop, bad cop: Race-based differences in mental representations of police. *Personality & Social Psychology Bulletin*, *46*(8), 1205–1218. https://doi.org/10.1177/0146167219898562

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavioral Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Nosofsky, R. M., & Zaki, S. R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *28*(5), 924–940. https://doi.org/10.1037//0278-7393.28.5.924

Ratner, K. G., Dotsch, R., Wigboldus, D. H. J., Van Knippenberg, A., & Amodio, D. M. (2014). Visualizing minimal ingroup and outgroup faces: Implications for impressions, attitudes, and behavior. *Journal of Personality & Social Psychology*, *106*(6), 897–911. https://doi.org/10.1037/a0036498

Rosenberg, M. (1979). *Conceiving the self*. Basic Books.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149–178. https://doi.org/10.1002/ejsp.2420010202

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–37). Brooks/Cole.

Warr, P. B., & Knapper, C. (1968). *The perception of people and events*. John Wiley & Sons.

Wilson, J. P., Young, S. G., Rule, N. O., & Hugenberg, K. (2018). Configural processing and social judgments: Face inversion particularly disrupts inferences of human-relevant traits. *Journal of Experimental Social Psychology*, *74*, 1–7. https://doi.org/10.1016/j.jesp.2017.07.007