Case Report

Social inferences from group size

Jack Cao *, Mahzarin R. Banaji

Department of Psychology, Harvard University, Cambridge, MA 02138, USA

ARTICLE INFO

ABSTRACT

Does the mere size of a social group influence how it is perceived? Study 1 showed that on self-report measures, smaller and larger groups are rated to be equal in *warmth/goodness*, but smaller groups are rated to be higher in *status/competence*. Self-reports indicated that the latter result stems from the small group size of the socioeconomic elite (i.e., the top 1%). Using an implicit measure (the IAT), Studies 2 and 3 demonstrated that smaller and larger groups are equally associated with warmth/goodness, showing consistency with self-reports. Surprisingly, larger groups are implicitly associated with high status/competence, not smaller groups. Two possible explanations for this explicit-implicit dissociation were ruled out. Compared to smaller groups, larger groups were not implicitly associated with greater amounts of any attribute, nor were larger groups implicitly associated with all positive attributes. Study 4 found that even when two groups are explicitly known to be the same in status/competence, the larger group is nonetheless implicitly associated with greater status/competence. To explain this result, we offer the possibility that implicit associations between larger groups and high status/competence are systematic, erroneous extensions of the association between larger groups and physical dominance. Together, these studies reveal a new dissociation between explicit and implicit cognition: although explicit ratings show that *smaller groups are elite*, implicit associations erroneously indicate that *larger groups are elite*, thereby assigning larger groups expanded power in the minds of perceivers.

© 2016 Elsevier Inc. All rights reserved.

Imagine entering a room and noticing that there are two groups of people, those dressed in red and those dressed in blue, with the blues greatly outnumbering the reds. If these two groups are equivalent on all dimensions except relative group size, what psychological properties do we, as perceivers, attribute to these groups? And on the basis of group size alone, is there consistency between the attributes that are explicitly reported and those that are implicitly revealed?

Although group size is among the most common perceptual cues that distinguish groups within societies, surprisingly little is known about how this cue shapes social perception. Contemporary research on groups has been highly instructive on wide scale intergroup conflict (Hewstone, Rubin, & Willis, 2002) and on the dynamics of small group interactions (Levine & Moreland, 1990; Worchel, Axsom, Ferris, Samaha, & Schweizer, 1978). But rarely has group size been isolated to understand whether and how it influences what perceivers infer about groups' psychological properties.

Yet early in the history of groups research, such efforts were at the forefront. Lewin (1947) theorized that a group's psychological attributes are more than simply the sum of group members' individual attributes. Allport (1954) discussed the implications of group perception for stereotyping and prejudice. Campbell (1958) imported ideas from Gestalt

psychology to posit that group members' similarity and proximity to one another influence beliefs about the group's entitativity. Later empirical work not only confirmed Campbell's claim (Dasgupta, Banaji, & Abelson, 1999; Lickel et al., 2000), but also demonstrated how entitativity influences social cognition (Hamilton, Sherman, & Castelli, 2002).

To be sure, group size has not been fully ignored as a variable of interest. Evolutionary biologists and developmental psychologists have found that across phylogeny and ontogeny, group size corresponds closely to beliefs about physical dominance (Wilson & Wrangham, 2003; Pietraszewski & Shaw, 2015; Pun, Birch, & Baron, 2016). The larger a group is, the greater the belief in the group's ability to impose costs on competing groups and for the group to prevail during conflict (Maynard Smith & Parker, 1976; McComb, Packer, & Pusey, 1994). In the domain of group performance, majority groups have been shown to sway individuals to conform to an obviously wrong position (Asch, 1951). And members of larger groups have been found to cooperate less, exhibit greater amounts of misbehavior, and be less satisfied (Latane, 1981; Kerr, 1989; Markham, Dansereau, & Alutto, 1982; Pinto & Crow, 1982).

But still the question remains of how group size affects which psychological properties perceivers attribute to different groups. With the exception of gender, all other major social groups – different races, ethnicities, religions, castes, ages, and sexualities – vary greatly in size. Given how much groups within a society vary in size – especially as group size varies markedly over time (Craig & Richeson, 2014) – it is

important to understand this cue's influence on social inferences concerning fundamental qualities of warmth, goodness, status, and competence. The present research aims to fill this gap and is rooted theoretically in the Stereotype Content Model, which has identified two primary dimensions of group perception: warmth and competence (Fiske, Cuddy, Glick, & Xu, 2002; Fiske, Cuddy, & Glick, 2006). However, the current work expands the warmth dimension to include attributes such as *nice*, *pleasant*, and *honest* – leading to the label *warmth/goodness*. The competence dimension is expanded even more to include attributes related to socioeconomic status – such as *rich*, *classy*, and *business class* – leading to the label *status/competence*.

In the first study, participants were shown two groups that differ solely in size and were asked to rate each group on attributes related to warmth/goodness and status/competence. Thus, we collected ratings of two minimal groups, with relative size rather than group membership manipulated, as Tajfel and Turner (1986) did before. In later studies, the Implicit Association Test (IAT; Greenwald, McGhee, & Swartz, 1998) was used to assess whether implicit associations are consistent with explicit ratings, something early researchers were unable to pursue.

In conducting this research, it was not immediately clear how group size would, if at all, influence perceptions of warmth/goodness, which are typically based on a group's motives and its past and present behaviors (Reeder, Kumar, Hesson-McInnis, & Trafimow, 2002). On the one hand, a larger group's greater numbers may signal less familiarity and closeness and therefore less warmth/goodness, leading the smaller group to be viewed as warmer/better. But on the other hand, a smaller group may be viewed as more isolated while a larger group may be seen as offering greater opportunities for social interaction, leading the larger group to be viewed as warmer/better. Finally both groups may be viewed as equally warm/good. At present, how group size affects perception along this dimension is unknown, and the current studies are the first to suggest an answer.

Status/competence is the second dimension of interest. Measures utilized here primarily focus on status-related attributes (e.g., *rich*, *classy*, *business class*) because of the importance of status to social hierarchies and to perceptions of competence (Magee & Galinsky, 2008; Fiske et al., 2002). In fact, Fiske et al. (2002) found a near perfect correlation between ratings of status and ratings of competence, which might reflect a belief that high status cannot emerge without high competence. There is also a strongly held consensus that socioeconomic elites are few in number, even though the degree of economic inequality is underestimated (Norton & Ariely, 2011). Furthermore, media coverage of events like the Occupy movement and the top 1% of earners (e.g., Moynihan, 2015) may have increased the salience of how small the socioeconomic elite are compared to the larger group of people of more modest means. Together, these reasons lead to the hypothesis that high status/competence attributes will be associated with smaller rather than larger groups.

The primary focus here is on whether smaller groups will be linked with high status/competence on both explicit and implicit measures. On explicit ratings, smaller groups may be rated as high in status/competence. Whether implicit associations are consistent with explicit ratings remains to be seen. In past social cognition work, explicit and implicit attitudes and beliefs towards the same target have been shown to dissociate (Rydell, McConnell, Strain, Claypool, & Hugenberg, 2007; Nosek et al., 2007) and influence group-based thinking (Macrae & Bodenhausen, 2000). Frequently, these dissociations appear in domains where the demand to appear egalitarian is strong (Nosek, 2005). But the targets in the forthcoming studies are unnamed, abstract groups that differ solely in size. As such, any pressure to appear egalitarian need not apply.

Therefore, the extent to which we observe an explicit-implicit dissociation on the dimension of status/competence will be instructive for theories of group perception. For if implicit associations indicate that high status/competence accrues to larger groups while explicit ratings indicate otherwise, then we must attempt to understand this dissociation in ways other than the demand to respond in a socially acceptable manner. New

theorizing would be necessary to explain why conscious inferences from group size align with the social reality that smaller groups tend to hold high socioeconomic status, but why less conscious inferences reflect the view that such high status is held by larger groups.

1. Study 1a

A preliminary study was conducted to test explicit, self-reported ratings of warmth/goodness and status/competence.

1.1. Method

1.1.1. Participants

Based on sample sizes from past research on group perception, 250 participants were recruited from Amazon Mechanical Turk (M_{age} = 36.30 years, SD = 12.80; 130 males, 120 females). For all studies in this paper, sample sizes were determined before the start of data collection, and analyses began after all data were collected.

1.1.2. Procedure

Participants were shown an image of a smaller group of 6 stick figures and a larger group of 12 stick figures. Below the image were two opposing attributes (e.g. light/heavy). The task was to match each attribute with one of the groups in a forced-choice task. All participants completed a total of twenty trials. In each trial, the attributes differed but the groups remained the same.

The first three trials consisted of practice attributes (e.g., light/heavy) to acclimate participants to the task. After completing the practice trials, participants completed two trials whose attributes pertained to physical dominance (i.e., strong/shaky; powerful/helpless), five trials whose attributes pertained to warmth/goodness (e.g., warm/cold; pleasant/unpleasant), and six trials whose attributes pertained to status/competence (e.g., smart/stupid; rich/poor).[1] Thus, all participants matched the smaller and larger groups with all attributes in a within-subjects design. Trial order and the location of the groups in the image were randomized.

If the larger group is seen as physically dominant – consistent with intuition and past work demonstrating the physical power of larger groups over smaller groups – then we are provided with a reassuring baseline verification that allows us to interpret more confidently the results from the dimensions of warmth/goodness and status/competence. All stimuli – for this study and subsequent studies – are in the supplemental materials. For all studies in this paper, we report all measures, manipulations, and exclusions.

1.2. Results and discussion

Data were analyzed with a logistic mixed effects model, implemented using R statistical computing's lme4 package (Bates, Machler, Bolker, & Walker, 2015). Participants' responses were regressed on the fixed effect of dimension (physical dominance vs. warmth/goodness vs. status/competence) and two random effects: one for attribute pair since multiple attribute pairs were tested, and another for participant given the within-subjects design. The intercept was removed from the model so that the fitted probability for each dimension could be compared to chance-level responding of 50%. Fig. 1 displays the fitted probabilities of the larger group being matched with each dimension (see supplemental materials for results of each attribute pair; Fig. S2). Three results emerged.

---

[1] Four other pairs were also tested (envied/disgusting; admired/pitied; envied/admired; pitied/disgusting). These four pairs consist of emotions that arise from specific high-low combinations of warmth and competence. The focus of this paper is not on these second-order emotions. Thus, the results of these four pairs, along with the three practice trials, are in the supplemental materials (Fig. S1). It is unlikely that these four pairs biased the results because attribute pairs were tested in a random order.
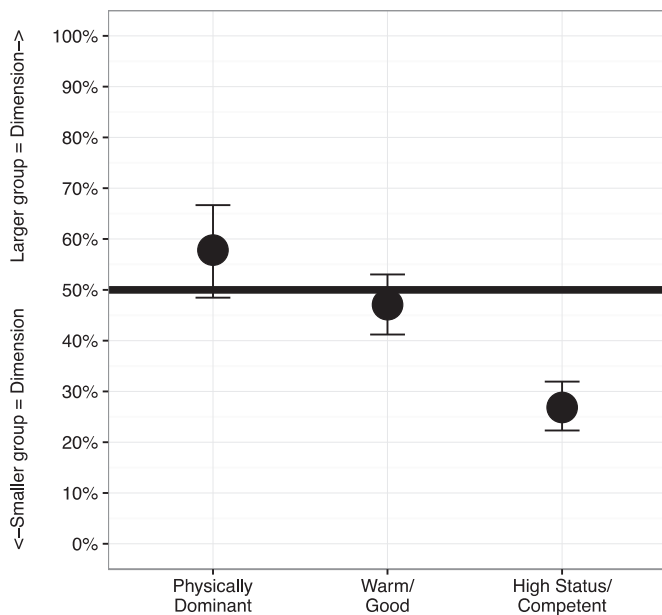
Fig. 1. Study 1a ($N = 250$): Fitted probabilities of the larger group being matched with attributes from the dimension of physical dominance, warmth/goodness, and status/competence. The horizontal line at 50% indicates chance level responding. Values above 50% indicate that the larger group was perceived as physically dominant, warm/good, or high status/competent. Values below 50% indicate that the smaller group was perceived as physically dominant, warm/good, or high status/competent. Error bars are 95% *CIs*.

First, the larger group was rated as physically dominant over the smaller group (fitted probability = 57.82%, 95% *CI* [48.45%, 66.66%]), although not significantly so (log odds ratio $b = 0.32$, $SE = 0.19$, $z = 1.64$, $p = 0.10$). This result was not as strong as expected based on the generally accepted belief that larger groups will trounce smaller groups in physical conflict. Therefore, this result may reflect attempts to imagine when smaller groups, through their high status/competence, achieve greater strength and power (e.g., minority rule over majorities), leading to less extreme ratings of larger groups as physically dominant. We return to this point in Study 2 where an implicit measure is used.

Second, neither the smaller nor larger group was rated as warmer/better than the other (fitted probability = 47.08%, 95% *CI* [41.20%, 53.04%], log odds ratio $b = -0.12$, $SE = 0.12$, $z = -0.96$, $p = 0.34$). That warmth/goodness did not appear to explicitly differentiate between smaller and larger groups is striking in comparison to the third result, which is that, as predicted, smaller groups were rated as higher in status/competence than larger groups (fitted probability = 26.85%, 95% *CI* [22.31%, 31.93%], log odds ratio $b = -1.00$, $SE = 0.13$, $z = -8.01$, $p < 0.0001$). Just as socioeconomic elites are indeed few in number, the smaller group in this study was perceived to be the more elite group.

These initial results show that on the dimension of warmth/goodness, smaller and larger groups are rated equally. This is an interesting result because a fundamental feature of groups, their size, seems not to have any influence on perceived warmth/goodness. However, smaller groups were rated as higher in status/competence. We assume this latter result stems from common knowledge that wealth and status tend to be concentrated in smaller groups of people. In later studies that use the IAT, we test if this common knowledge also influences implicit associations between different sized groups and status/competence, or if implicit associations deviate substantially from explicit ratings.

2. Study 1b

Study 1b was conducted to replicate the finding that smaller groups are explicitly rated to be higher in status/competence and to solicit explanations for why such a rating would be given.

2.1. Method

2.1.1. Participants

Three hundred participants were recruited from Amazon Mechanical Turk ($M_{age} = 31.74$ years, $SD = 9.64$; 179 males, 119 females, 2 unspecified).

2.1.2. Procedure

Participants were randomly assigned to one of three conditions: 1) rich/poor, 2) smart/stupid, or 3) professional class/working class. In each condition, participants answered ten questions of the form, "Which group is *X*?" Five questions tested the positively-valenced attribute (e.g., Which group is rich?) and its synonyms (e.g., Which group is wealthy?) and five questions tested the negatively-valence attribute (e.g., Which group is poor?) and its synonyms (e.g., Which group is penniless?). The two answer choices, which were counterbalanced, were a smaller group of 6 stick figures or a larger group of 12 stick figures. The first question in each condition always queried the positive attribute (e.g., Which group is rich?). The order of the remaining nine questions was then randomized.

After answering these ten questions, participants were asked to select a reason for why they answered in the way they did. Participants were shown two reasons and were able to select one reason, both reasons, or neither reason. One reason was consistent with our assumption that common knowledge was brought to bear (e.g., In the real world, wealth is concentrated in smaller groups of people). A second reason was consistent with the opposite notion that the larger a group is, the greater its resources (e.g., More people in a group means the group has more collective wealth).

2.2. Results and discussion

Data were again analyzed with a logistic mixed effects model. Participants' responses were regressed on the fixed effect of attribute (rich vs. poor vs. smart vs. stupid vs. professional class vs. working class) and two random effects: one for synonym (e.g., wealthy) since multiple synonyms were tested, and another for participant given that each participant provided multiple responses. As before, the intercept was removed from the model so that the fitted probability for each attribute could be compared to chance-level responding of 50%. Fig. 2 displays the fitted probabilities of the larger group being matched with each attribute (see supplemental materials for the results of each synonym; Fig. S3).[2]

Replicating the previous result, smaller groups were higher in status/competence while larger groups were lower in status/competence. Smaller groups were rated as rich, smart, and professional class (fitted probabilities < 27%, log odds ratios $bs < -1.01$, $SEs < 0.17$, $zs < -6.89$, $ps < 0.0001$), while larger groups were rated as poor, stupid, and working class (fitted probabilities > 72%, log odds ratios $bs > 0.98$, $SEs < 0.17$, $zs > 6.70$, $ps < 0.0001$).

Participants' explanations for their responses are consistent with the assumption that the small size of the socioeconomic elite was brought to bear: 87% of participants in the rich/poor condition selected the reason that wealth is concentrated in smaller groups; 59% of participants in the smart/stupid condition selected the reason that education is concentrated in smaller groups; and 83% of participants in the professional class/working class condition selected the reason that socioeconomic status is concentrated in smaller groups. As one participant commented, "It's the whole 1% thing – the higher you get in terms of pay and education, and the smaller the group is."

---

[2] For each participant, a Cohen's kappa was calculated comparing his or her responses to a hypothetical set of responses in which the smaller group was always rated as high in status/competence and the larger group was always rated as low in status/competence. The average Cohen's kappa in the rich, smart, and professional class conditions was 0.65, 0.46, and 0.65, respectively, all of which indicate moderate to substantial agreement.
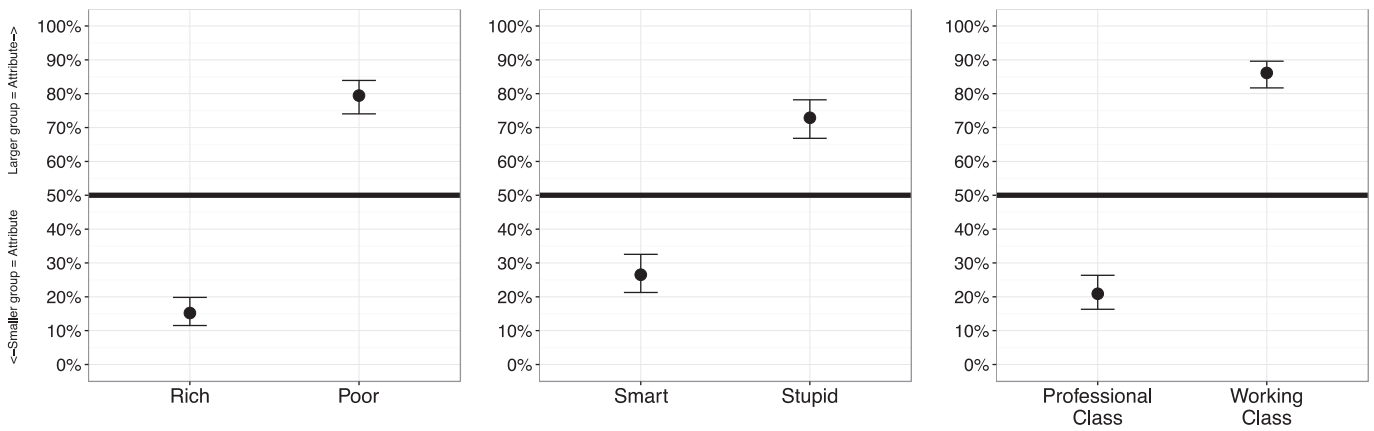
Fig. 2. Study 1b (*N* = 300): Fitted probabilities of the larger group being matched with attributes from the dimension of status/competence. The horizontal line at 50% indicates chance level responding. Values above 50% indicate that the larger group was perceived as rich, poor, smart, stupid, professional class, or working class. Values below 50% indicate that the smaller group was perceived in that way. Error bars are 95% *CI*s.

Studies 1a and 1b demonstrate that smaller and larger groups are rated as equally warm/good. But smaller groups are rated as higher in status/competence than larger groups. This latter finding stems from the fact that socioeconomic status tends to be concentrated in smaller, not larger, groups of elites.

3. Study 2

Study 2 was conducted to test the main question of interest: are implicit associations between groups of different sizes and warmth/goodness and status/competence consistent with explicit ratings? Both association and dissociation are possible; we sought to establish which is the case and explore reasons for why this may be.

3.1. Method

3.1.1. Participants

Participants were volunteer visitors to Project Implicit (implicit.harvard.edu). Of the 2778 participants who completed the procedure, 45 were excluded for going faster than 300 ms on 10% or more of IAT trials, in accordance with the scoring algorithm established by Greenwald, Nosek, and Banaji (2003) and well within expected rates. The distribution of excluded participants across conditions for this study and all others can be found in the supplemental materials. The final sample consisted of 2733 participants (M_{age} = 32.24 years, *SD* = 13.62; 1720 females, 984 males, 29 unspecified).

3.1.2. Procedure

In five independent rounds of data collection, participants took one IAT in which the concepts were 6 and 12 in number form to represent the smaller and larger groups. The attributes consisted of all the attribute pairs shown in Fig. 1 (e.g., White Collar/Blue Collar), with each participant randomly receiving only one attribute pair in a between-subjects design.[3] Stimuli for the concepts were images of stick figure groups with either 6 or 12 individuals. Stimuli for the attributes were synonyms or related words (e.g., CEO for White Collar; mailman for Blue Collar).

The IAT consisted of five blocks of trials. In blocks 1, 2, and 4, only group size or only attribute stimuli were sorted, but not both. In blocks 3 and 5 – the critical blocks – stimuli were sorted according to both group size and attribute. In one critical block, stimuli were sorted in a manner that was congruent with explicit ratings of smaller groups as higher in status/competence (e.g., 6 + White Collar/12 + Blue Collar). In the next critical block, stimuli were sorted in a manner that was incongruent with explicit ratings (e.g., 12 + White Collar/6 + Blue Collar). The order in which participants completed the critical blocks was randomly assigned.
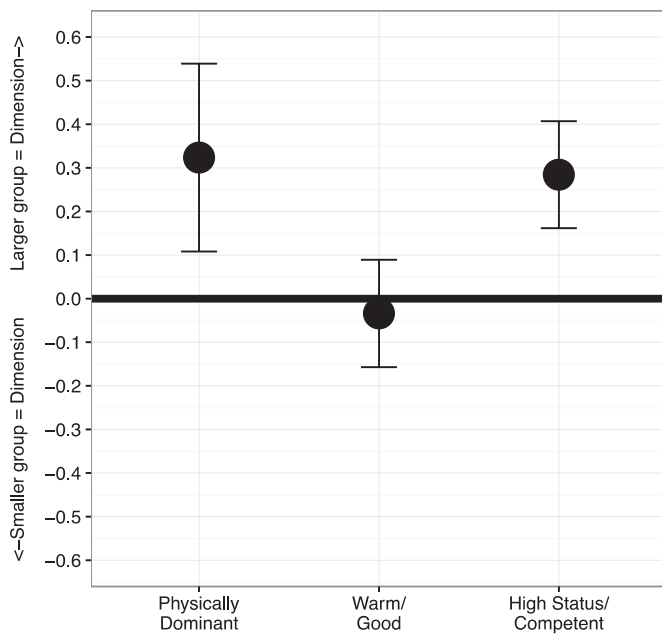
3.2. Results and discussion

Data were analyzed with a mixed effects model, implemented using R statistical computing's nlme package (Pinheiro et al., 2016). IAT *D* scores were regressed on the fixed effect of dimension (physical dominance vs. warmth/goodness vs. status/competence) and a random effect for attribute pair given that multiple attribute pairs were tested. The intercept was removed from the model so that the *D* score for each dimension could be compared to a neutral score of zero. Fig. 3 displays *D* scores for each dimension (see supplemental materials for results of each attribute pair; Fig. S4). Three results emerged.

First, larger relative to smaller groups are implicitly associated with attributes related to physical dominance (*M* = 0.32, 95% *CI* [0.11, 0.54], $t(11)$ = 2.94, *p* = 0.01, *r* = 0.66). This finding, which is consistent with past work demonstrating the physical dominance of larger over smaller groups, provides face validity for using the IAT to test attributes related to warmth/goodness and status/competence. Furthermore, it is notable that in Study 1a, only a marginally significant effect emerged for physical dominance. But here in Study 2, a significant and medium-sized effect emerged. Whereas the high status/competence of smaller groups may have influenced explicit ratings of physical dominance, leading to reduced effects in Study 1a, this was not the case on the implicit measure, where larger groups are clearly physically dominant over smaller groups.

Previously, smaller groups were rated as high in status/competence. But the second result shows that implicitly, larger groups are associated with high status/competence (*M* = 0.28, 95% *CI* [0.16, 0.41], $t(11)$ = 4.54, *p* = 0.0008, *r* = 0.81). This implicit association between larger group size and high status/competence presents a stark contrast to the explicit ratings observed before. The same unnamed, abstract groups were represented differently depending on whether the measure was explicit or implicit. Explicitly, smaller groups were elite. But implicitly, larger group were elite.

One explanation for this dissociation may be that larger groups are implicitly associated with more of any dimension. If this were the

---

[3] For the attribute pairs of Healthy/Sick, Generous/Stingy, Powerful/Helpless, and Truth Tellers/Lie Tellers, 9 and 21 stick figures were used instead of 6 and 12. In Study 1a, the attribute pair of Honest/Dishonest was used. In Study 2, we used Honest/Liar and Truth Tellers/Lie Tellers instead. It is extremely unlikely that these slight differences affected the results. The stimuli used to represent Liar and Lie Tellers were synonyms for Dishonest (e.g., Deceitful; see supplemental materials), and a larger group of 21 was associated with Powerful, just as a larger group of 12 was associated with Strong. We also tested the attribute pair of Admire/Pity in Study 2. We present the results of this attribute pair in the appendix for the same reason discussed in footnote 1.

Fig. 3. Study 2 ($N = 2733$): IAT D scores collapsed by dimension. The horizontal line at zero indicates no implicit association. Values above zero indicate that the larger group was implicitly associated with physical dominance, warmth/goodness, or status/competence. Values below zero indicate that the smaller group was implicitly associated with physical dominance, warmth/goodness, or status/competence. Error bars are 95% *CIs*.

case, then larger groups would also be associated with more warmth/goodness. However, the third result rules out this possibility, as neither the smaller nor larger group was implicitly associated with warmth/goodness ($M = -0.03$, 95% CI [-0.16, 0.09], $t(11) = -0.54$, $p = 0.60$, $r = 0.16$), a finding that is consistent with explicit self-reports.

Another possible explanation is that on the IAT, larger groups are simply better than smaller groups. However, not only was the larger group *not* associated with more warmth/goodness, but neither sized group was associated with the specific attribute of pleasant ($M = -0.04$, $SD = 0.50$, one-sample $t(212) = -1.31$, $p = 0.19$, Cohen's $d = 0.09$). This null finding may indicate that the implicit association between larger groups and high status/competence reflects meaningful semantic content rather than general positivity.

In sum, these results show that larger groups are not perceived as having "more" of all attributes nor are they perceived to be merely better than smaller groups. Larger groups are robustly associated with high status/competence, contrary to what is expressed on self-report measures.

4. Study 3

The dissociation between explicit ratings and implicit associations on the status/competence dimension was found from two different participant pools: Amazon Mechanical Turk (Studies 1a and 1b) and Project Implicit (Study 2). These different samples leave open the possibility that the result does not represent a dissociation within one and the same mind. But if participants from the same sample report that smaller groups are high in status/competence but implicitly associate larger groups with high status/competence, then such a result would be far more striking.

Study 3 is both a replication and a test of the explicit-implicit dissociation in a far more robust within-subjects design. Here, we focus on the attribute pairs of rich/poor, professional class/working class, and business class/economy class. Study 1b showed that participants are well aware that the rich and the professional class are groups of smaller size. And as anyone who has traveled on a commercial airline can report, business class is obviously smaller than economy class. Thus, we should be least likely to observe an explicit-implicit dissociation with these three attribute pairs.

4.1. Method

4.1.1. Participants

Participants were volunteer visitors to Project Implicit. Of the 1127 participants who completed the IAT, 20 were excluded for going faster than 300 ms on 10% or more of trials, in accordance with the Greenwald et al. (2003) and well within expected rates. An additional 90 participants were excluded for not completing all explicit measures. The final sample consisted of 1017 participants ($M_{age} = 33.19$ years, $SD = 14.12$; 590 females, 415 males, 12 unspecified).

4.1.2. Procedure

First, participants completed and IAT in which the concepts were Group of 6 People and Group of 12 People and the attributes were Rich/Poor, Professional Class/Working Class, or Business Class/Economy Class. One of these three attribute pairs was randomly assigned to each participant in a between-subjects design. As before, stimuli for the concepts were images of a group of 6 or 12 stick figures and stimuli for the attributes were synonyms or related words. The order in which participants completed the critical blocks of the IAT was randomly assigned. D scores were calculated such that negative values indicate an implicit association between the smaller group and the positive attribute (i.e., rich, professional class, or business class) and positive values indicate an implicit association between larger group and the positive attribute.

After completing the IAT, participants answered ten explicit questions. Each question displayed an image of a group of 6 stick figures and a group of 12 stick figures and asked which group is best characterized by a description (e.g., Which group is impoverished?). Participants could select either the group of 6 people or the group of 12 people, but not both. Each queried description was also a stimulus on the IAT. The first question always queried the positive attribute (i.e., Which group is rich? vs. Which group is professional class? vs. On an airplane, which group sits in business class?). The order of the remaining nine questions was randomly determined.[4]

Participants' answers to the ten questions were transformed to a single variable to allow for direct comparison with IAT D scores. For each question, participants were given a score of zero if they selected the group of 6 and a score of one if they selected the group of 12. Scores for the positive attribute (e.g., *rich*) and its synonyms (e.g., *wealthy*) were summed, as were the scores for the negative attribute (e.g., *poor*) and its synonyms (e.g., *impoverished*). The sum for the negative attribute and its synonyms was then subtracted from the sum for the positive attribute and its synonyms. This process yielded a value from -5 to 5. Negative values indicate an explicit rating that the smaller group is the positive attribute, whereas positive values indicate an explicit rating that the larger group is the positive attribute.

4.2. Results and discussion

In a within-subjects design, an explicit-implicit dissociation emerged. The smaller group was explicitly rated as rich ($M = -1.17$, $SD = 3.73$; Wilcoxon signed rank test $V = 18934$, $p < 0.0001$, $r = -0.29$), professional class ($M = -2.74$, $SD = 3.37$; Wilcoxon signed rank test $V = 9384$, $p < 0.0001$, $r = -0.58$), and business class ($M = -3.77$, $SD = 2.69$; Wilcoxon signed rank test $V = 3468$, $p < 0.0001$, $r = -0.79$).[5] But on the IAT, the opposite effect emerged:

---

[4] In a previous version of this study, question order was not randomized and two additional attribute pairs were included (Smart/Stupid and Educated/Ignorant). We present and discuss these data in the supplemental materials.

[5] Explicit ratings were not normally distributed, so this nonparametric version of a one-sample *t*-test was appropriate. A Cohen's kappa was also calculated for each participant comparing his or her responses to a hypothetical set of responses in which the larger group was always rated as low in status/competence and the smaller group was always rated as high in status/competence. The average Cohen's kappa in the rich, professional class, and business class conditions was 0.23, 0.55, and 0.75 respectively, all of which indicate fair to substantial agreement.

the larger group was associated with rich ($M = 0.37$, $SD = 0.48$; one-sample $t(360) = 14.76$, $p < 0.0001$, $r = 0.61$), professional class ($M = 0.37$, $SD = 0.45$; one-sample $t(336) = 15.09$, $p < 0.0001$, $r = 0.64$), and business class ($M = 0.29$, $SD = 0.48$; one-sample $t(318) = 10.95$, $p < 0.0001$, $r = 0.52$). In Fig. 4, bootstrapped effect sizes for explicit ratings and implicit associations are plotted.

Additional evidence for the explicit-implicit dissociation comes from correlations between explicit ratings and implicit associations within each condition. In all three conditions, these correlations range from extremely small to modest ($r_{Rich} = 0.08$, $p = 0.14$; $r_{Professional\ Class} = 0.10$, $p = 0.07$; $r_{Business\ Class} = 0.16$, $p = 0.004$), which suggests that these different measures tap into relatively separate constructs. There is little hesitancy in reporting that smaller groups are rich, professional class, and business class. And yet implicitly, larger groups instead of smaller groups are robustly associated with these attributes. Possible reasons for this bias are presented in the General discussion.

5. Study 4

Thus far, smaller groups have been explicitly rated as higher in status/competence than larger groups. But this is not always the case. Consider Australia and New Zealand, two countries that are widely known to differ in size but not necessarily in status/competence. Australia and New Zealand are extremely similar in culture, health outcomes, and educational attainment. Both countries were British colonies. Life expectancies are virtually identical: 81.4 years in New Zealand, 82.3 in Australia (World Bank, 2016). And literacy rates in both countries exceed 96% (UNESCO, 2016).

Given these similarities, we expect that on explicit ratings, participants will indicate that Australia and New Zealand are equally classy, healthy, and smart – three attributes from the status/competence dimension. However, if it is indeed the case that larger groups accrue greater status/competence on the IAT, then Australia should be implicitly associated with classy, healthy, and smart. If this result is observed, it would demonstrate how pervasively dissociated group perception can be.

5.1. Method

5.1.1. Participants

Participants were volunteer visitors to Project Implicit. Only U.S. citizens were recruited for this study so that direct experience with Australia or New Zealand would be roughly similar and minimized. Of the 179 participants who completed the IAT, 1 was excluded for going faster than 300 ms on 10% or more of trials, in accordance with Greenwald et al. (2003) and well within expected rates. An additional 7 participants were excluded for not completing all explicit measures. The final sample consisted for 171 participants ($M_{age} = 40.35$ years, $SD = 15.65$; 106 females, 65 males).

5.1.2. Procedure

First, participants were shown a map of Australia and New Zealand as a visual reminder of the size difference between the two countries. Next, participants completed an IAT in which the concepts were Australia and New Zealand and the attributes were Classy/Crude, Healthy/Sick, or Smart/Stupid. One of these three attribute pairs was randomly assigned to each participant in a between-subjects design. As before, the order in which participants completed the critical blocks was randomly assigned. D scores were calculated such that negative values indicate an implicit association between New Zealand and the positive attribute (i.e., classy, healthy, or smart) and positive values indicate an implicit association between Australia and the positive attribute.

After completing the IAT, participants answered one explicit question about how classy, healthy, or smart the people of Australia vs. New Zealand are, depending on what condition they had been assigned to. Explicit ratings were provided on a Likert-type scale ranging from
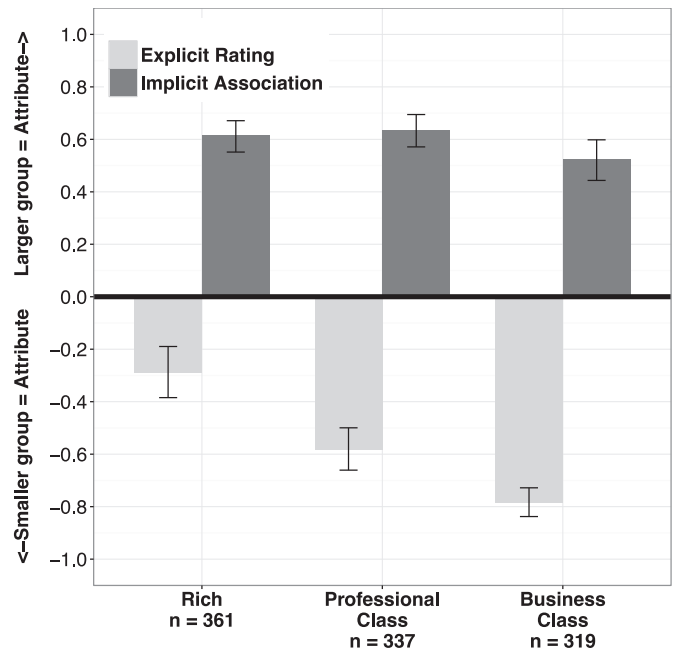

Fig. 4. Study 3 ($N = 1017$): Average effect sizes (Pearson's rs) for explicit ratings and implicit associations. The horizontal line at zero indicates no effect. Values greater than zero indicate that the larger group was perceived as rich, professional class, or business class. Values less than zero indicate that the smaller group was perceived as rich, professional class, or business class. Average effect sizes and error bars (95% CIs) were calculated from 10,000 bootstrapped samples.

−3 to 3, where negative values indicate that New Zealand is classier, healthier, or smarter and where positive values indicate that Australia is classier, healthier, or smarter.

5.2. Results and discussion

As expected, neither group was explicitly rated as classier, healthier, or smarter than the other. Both Australia and New Zealand were rated as equally classy ($M = -0.12$, $SD = 0.59$; $t(57) = -1.55$, $p = 0.13$, $r = -0.20$), equally healthy ($M = -0.13$, $SD = 0.71$; $t(61) = -1.43$, $p = 0.16$, $r = -0.18$), and equally smart ($M = -0.04$, $SD = 0.69$; $t(50) = -0.40$, $p = 0.69$, $r = -0.06$). By contrast, on the IAT, participants differentiated between these two groups in a manner consistent with previous findings. The larger group of Australia was implicitly associated with classy ($M = 0.14$, $SD = 0.38$; $t(57) = 2.76$, $p = 0.008$, $r = 0.34$), healthy ($M = 0.21$, $SD = 0.33$; $t(61) = 5.09$, $p < 0.0001$, $r = 0.55$), and smart ($M = 0.13$, $SD = 0.38$; $t(50) = 2.51$, $p = 0.02$, $r = 0.33$). In Fig. 5, bootstrapped effect sizes for explicit ratings and implicit associations are plotted.

Furthermore, within each condition, correlations between explicit ratings and implicit associations range from extremely small to modest ($r_{Classy} = 0.27$, $p = 0.04$; $r_{Healthy} = 0.06$, $p = 0.62$; $r_{Smart} = -0.06$, $p = 0.66$), which again supports an explicit-implicit dissociation in group perception. These findings show that this explicit-implicit dissociation in status/competence can arise even when two groups are explicitly rated to be the same, as it is not always the case that smaller groups are more elite. Explicit cognition recognizes that Australia and New Zealand are extremely similar in terms of culture, health outcomes, and educational attainment. But implicitly, the larger group of Australia is associated with classy, healthy, and smart.

6. General discussion

Group size plays a complex role in eliciting social inferences. While smaller and larger groups were rated as equally warm/good, smaller groups were rated as higher in status/competence than larger groups
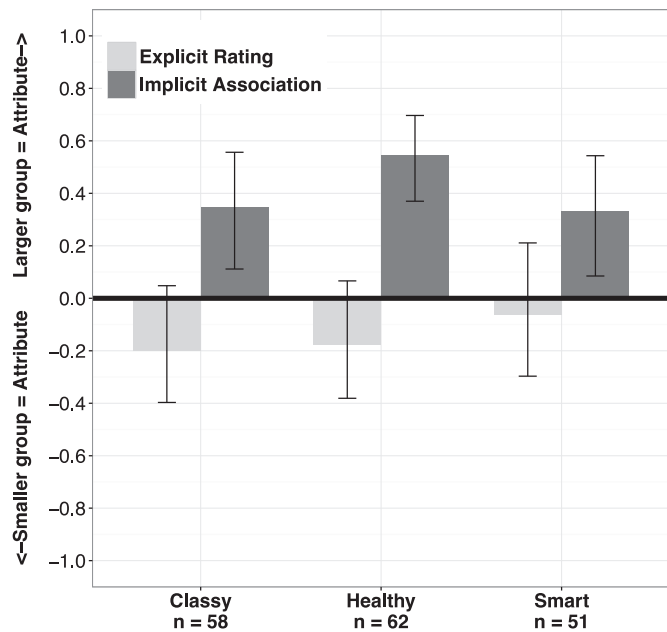
Fig. 5. Study 4 (*N* = 171): Average effect sizes (Pearson's *r*s) for explicit ratings and implicit associations. The horizontal line at zero indicates no effect. Values greater than zero indicate that the larger group was perceived as classy, healthy, or smart. Values less than zero indicate that the smaller group was perceived as classy, healthy, or smart. Average effect sizes and error bars (95% CIs) were calculated from 10,000 bootstrapped samples.

because socioeconomic elites are few in number. A quite different picture emerged when the measure was implicit. Smaller and larger groups were equally associated with warmth/goodness, mirroring explicit ratings. However, in direct contrast with explicit ratings, larger instead of smaller groups were implicitly associated with high status/competence. Even when two groups are known to be equal in status/competence, the larger group was still implicitly associated with greater status/ competence.

What might account for this particular explicit-implicit dissociation? First, we ruled out the possibility that the dissociation reflects an artifact. On the IAT, larger groups may be associated to a greater degree with any attribute. But larger groups were not associated with greater warmth/goodness compared to smaller groups. Alternatively, larger groups might be associated with more positive attributes than smaller groups. But this cannot be the case since both sized groups were equally associated with *pleasant*. Finally, it is also unlikely that social desirability led participants to conceal any explicit beliefs that larger groups are actually high in status/competence. In Study 1b, participants offered explanations for their ratings that were aligned with the reality that smaller groups tend to hold elite status.

With these possibilities ruled out, we are left with the clear result that automatic social cognition demonstrates a robust implicit association between larger groups and high status/competence. In the context of these studies, this implicit association contradicts both participants' self-reports and the widely known fact that socioeconomic elites are fewer in number, making the "large = elite" association an error. Of course, in certain situations, larger groups can be higher in status/competence than smaller groups. But when groups are unnamed and abstract, as was the case in Studies 1–3, an implicit association between larger groups and high status/competence is incorrect. To understand this result and guide future research, we put forward three possibilities.

First, implicit associations between larger groups and high status/ competence may represent an unwarranted extension of the association between larger groups and physical dominance. Since we used abstract stimuli to represent the smaller and larger groups, there was no clear basis to infer status/competence. It is for this very reason that we

did not initially use real social groups, which can signal status/competence. Although notably, a dissociation still emerged in Study 4 when real groups were used. However, when abstract stimuli were used, there was a clear basis to infer that the larger group is physically stronger than the smaller group (Maynard Smith & Parker, 1976; McComb et al., 1994), leading to the possibility that implicit associations between larger groups and high status/competence may emerge from an over-generalization of the association between larger groups and physical dominance.

On the IAT, participants may have fallen prey to a type of halo effect. Perhaps similar to how perceptions of physical beauty can lead to biased perceptions of dependability, loyalty, and other personality characteristics (Thorndike, 1920; Dion, Berscheid, & Walster, 1972), perceptions of physical strength might lead to biased perceptions of status/competence. It is not necessarily the case that groups with greater physical strength have the most money and status, and participants indicated as much in their explicit ratings.

A second interpretation is that the extension from physical dominance to high status/competence may be sensible instead of unwarranted. To the extent that larger groups were able to use their physical dominance to accrue resources and gain status/competence, these larger groups may have forged a privileged association with high status/ competence. Furthermore, the notion of a small group (e.g., the top 1%) holding a disproportionate amount of resources is a recent phenomenon that would have been highly unlikely in our evolutionary past, as smaller groups would have been unable to consistently overpower larger groups or have a way to safely store resources for long periods of time. So while such an extension from physical dominance to high status/competence may have been valid in ancestral environments, it is not in today's environment where people are well of aware of the fact that the socioeconomic elite are generally few in number and high in status/competence compared to larger groups of people of more modest means.

A third possibility to consider is that implicit associations between group size and status/competence may be a function of one's sociocultural experience. Indeed, IAT results have been interpreted as reflecting traces of such experience (Banaji, Nosek, & Greenwald, 2004). If one's sociocultural experience includes larger groups of rich, smart, and professional class individuals who fly business class, then the implicit associations between larger groups and these attributes may be facilitated, even though explicit ratings by the same person will accord with the reality that smaller groups typically embody these attributes. Conversely, if one's cultural experience includes larger groups of low socioeconomic status individuals, then both explicit ratings and implicit associations may converge. Replicating these studies in special populations where one's sociocultural experience includes smaller vs. larger groups of high status/competence individuals would provide a clearer test of the generality of the results reported here.

These results have implications for consequential behaviors of helping and harming. Given two groups that vary in size and the decision to help one group and harm the other, perceptions of status/competence may come into play. First, consider the identifiable victim effect: single victims (e.g., baby Jessica; Variety, 1989) are more likely to be helped than larger groups of victims. Explanations include single victims evoking 1) more vivid emotion, 2) greater certainty of death in the absence of help, and 3) more positive impact if help is offered (Jenni & Loewenstein, 1997). Underlying these explanations may be the perception that a single victim is low in status/competence while a larger group is high in status/competence. The implicit association between larger groups and high status/competence may lead larger groups to be viewed as either needing less help or less likely to benefit substantially from help. Thus, identifiable victims, who are smaller in number, may receive a disproportionate share of help.

Second, consider the trolley problem. Will one person be killed to save five? Or will nothing be done, resulting in the deaths of five people? People who choose the latter generally decide faster than those

who choose the former (Baron, Gurcay, Moore, & Starcke, 2012), a result typically interpreted in a dual systems framework: people quickly heed an affective aversion to intentionally inflicting harm (Greene, 2013). But this quick decision may also stem, in part, from an implicit association between group size and status/competence. If one group must be harmed, then harming a group perceived to be low in status/competence (i.e., the lone individual) might be seen as worse than harming a group perceived to be in high in status/competence (i.e., the larger group of five).

Implicit associations have predicted various behaviors above and beyond explicit ratings (Dovidio, Gaertner, & Kawakami, 2002; Penner et al., 2009; Spence & Townsend, 2007). These implicit associations might also predict helping and harming behaviors towards groups that vary in size. While this possibility awaits future research, the replicated results reported here based on large samples of participants suggest that within the same perceiver, two conflicting mental representations co-exist. Explicit representations of the social hierarchy coincide with reality, for smaller groups that tend to be high in status/competence occupy the apex and larger groups that tend to be low in status/competence are below. Implicit mental representations appear to invert this. When more automatic social cognition is engaged, larger groups are viewed as holding more status/competence, contradicting both social reality as well as consciously reported explicit beliefs.

Acknowledgements

Appendix A. Supplementary materials

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jesp.2016.11.005. All data are avaiable on the Open Sciene Framework (osf.io/tgn44).

References

Allport, G. W. (1954). *The nature of prejudice.* Cambridge, MA: Perseus Books.
Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men* (pp. 177–190). Pittsburg, P. A: Carnegie Press.
Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2004). No place for nostalgia in science: A response to Arkes and Tetlock. *Psychological Inquiry, 15*(4), 279–310.
Baron, J., Gurcay, B., Moore, A. B., & Starcke, K. (2012). Use of a Rasch model to predict response times to utilitarian moral dilemmas. *Synthese, 189*(1), 107–117.
Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects model using lme4. *Journal of Statistical Software, 67*(1), 1–48.
Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science, 3*(1), 14–25.
Craig, M. A., & Richeson, J. A. (2014). On the precipice of a "majority-minority" America: Perceived status threat from the racial demographic shift affects White Americans' political ideology. *Psychological Science, 25*(6), 1189–1197.
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), 991–1033.
Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*(3), 285–290.
Dovidio, J. F., Gaertner, S. L., & Kawakami, K. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology, 82*(1), 62–68.
Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2006). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77–83.
Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A mode of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878–902.
Greene, J. D. (2013). *Moral tribes: Emotion, reason, and the gap between us and them.* New York: Penguin Books.
Greenwald, A. G., McGhee, D. E., & Swartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464–1480.
Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197–216.
Hamilton, D. L., Sherman, S. J., & Castelli, L. (2002). A group by any other name: The role of entitativity in group perception. *European Review of Social Psychology, 12*(1), 139–166.
Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53*, 575–604.
Jenni, K. E., & Loewenstein, G. (1997). Explaining the "identifiable victim effect.". *Journal of Risk and Uncertainty, 14*, 235–257.
Kerr, N. L. (1989). Illusions of efficacy: The effects of group size on perceived efficacy in social dilemmas. *Journal of Experimental Social Psychology, 25*(4), 287–313.
Latane, B. (1981). The psychology of social impact. *American Psychologist, 36*(4), 343–356.
Levine, J. M., & Moreland, R. L. (1990). Progress in small group research. *Annual Review of Psychology, 41*, 585–634.
Lewin, K. (1947). Concept, method and reality in social science: Social equilibria and social change. *Human Relations; Studies Towards the Integration of the Social Sciences, 1*(1), 5–41.
Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology, 78*, 223–246.
Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51*(1), 93–120.
Magee, J. C., & Galinsky, A. D. (2008). Social hierarchy: The self-reinforcing nature of power and status. *The Academy of Management Annals, 2*(1), 351–398.
Markham, S. E., Dansereau, F., & Alutto, J. A. (1982). Group size and absenteeism rates: A longitudinal analysis. *The Academy of Management Journal, 25*(4), 921–927.
Maynard Smith, J., & Parker, G. A. (1976). The logic of asymmetric contests. *Animal Behaviour, 24*(1), 159–175.
McComb, K., Packer, C., & Pusey, A. (1994). Roaring and numerical assessment in contests between groups of female lions, Panthera leo. *Animal Behaviour, 47*(2), 379–387.
Moynihan, C. (2015, April 2). Occupy wall street, the tour. *The New York times* Retrieved from http://www.nytimes.com
Norton, M. I., & Ariely, D. (2011). Building a better America – One wealth quintile at a time. *Perspectives on Psychological Science, 6*(1), 9–12.
Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134*(4), 565–584.
Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., ... Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology, 18*(1), 1–53.
Penner, L. A., Dovidio, J. F., West, T. V., Gaertner, S. L., Albrecht, T. L., Dailey, R. K., & Markova, T. (2009). Aversive racism and medical interactions with black patients: A field study. *Journal of Experimental Social Psychology, 46*(2), 436–440.
Pietraszewski, D., & Shaw, A. (2015). Not by strength alone: Children's conflict expectations follow the logic of the asymmetric war of attrition. *Human Nature, 26*, 44–72.
Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Development Core Team (2016). Linear and Nonlinear Mixed Effects Models. Available at http://CRAN.R-project.org/package=nlme
Pinto, L. J., & Crow, K. E. (1982). The effect of size on other structural attributes of congregations within the same denomination. *Journal for the Scientific Study of Religion, 21*, 304–316.
Pun, A., Birch, S., & Baron, A. S. (2016). Infants use relative numerical group size to infer social dominance. *Proceedings of the National Academy of Sciences of the United States of America, 113*(9), 2376–2381.
Reeder, G. D., Kumar, S., Hesson-McInnis, M. S., & Trafimow, D. (2002). Inferences about the morality of an aggressor: The role of perceived motive. *Journal of Personality and Social Psychology, 83*(4), 789–803.
Rydell, R. J., McConnell, A. R., Strain, L. M., Claypool, H. M., & Hugenberg, K. (2007). Implicit and explicit attitudes respond differently to increasing amounts of counterattitudinal information. *European Journal of Social Psychology, 37*(5), 867–878.
Spence, A., & Townsend, E. (2007). Predicting behavior towards genetically modified food using implicit and explicit attitudes. *British Journal of Social Psychology, 46*, 437–457.
Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worschel, & W. G. Austin (Eds.), *The social psychology of intergroup relations* (pp. 33–37). Brooks/Cole: Monterey, CA.
Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4*(1), 25–29.
UNESCO (2016). Effective literacy practice, Asia and the pacific. Available at http://www.unesco.org/uil/litbase/?menu=14 (Accessed August 12, 2016)
Variety (1989). *TV reviews—Network. Everybody's baby. Vol 335:7.* (May 31).
Wilson, M. L., & Wrangham, R. W. (2003). Intergroup relations in chimpanzees. *Annual Review of Anthropology, 32*, 363–392.
Worchel, S., Axsom, D., Ferris, F., Samaha, G., & Schweizer, S. (1978). Determinants of the effect of intergroup cooperation on intergroup attraction. *Journal of Conflict Resolution, 22*(3), 429–439.
World Bank (2016). Life expectancy at birth, total (years). Available at http://data.worldbank.org/indicator/SP.DYN.LE00.IN (Accessed August 12, 2016)