

When the Association Between Appearance and Outcome Contaminates Social Judgment: A Bidirectional Model Linking Group Homogeneity and Collective Treatment

Adam L. Alter
New York University

John M. Darley
Princeton University

Group formation is an inevitable consequence of social life, and the tendency to perceive people as a collective unit persists once they have been categorized as a group. Drawing on the concept of homogeneity, the authors propose a model suggesting that groups may endure in part because people who are perceived as homogeneous attract collective treatment (e.g., monetary rewards and punishment), and such treatment further reinforces the perception that the group's members are homogeneous. In support of this model, more homogeneous groups attracted collective treatment and collectively treated groups seemed to be more homogeneous thereafter. The authors suggest that these effects arise in part because people intuitively believe that group homogeneity is associated with collective treatment, and they present evidence suggesting that this applies to at least one policy-relevant real-world setting.

Keywords: homogeneity, collective punishment, group perception, attribution, social judgment

Almost a half century has passed since Martin Luther King, Jr., lamented that his children were treated according to the color of their skin, rather than the content of their character. During the same 50-year period, social psychologists have shown that skin color is only one of many superficial dimensions along which people categorize others. People who share salient features are treated as interchangeable members of a group defined by those features, and this effect persists across domains including, but not limited to, race, ethnicity, socioeconomic status, religion, nationality, physical appearance, and disability (e.g., Brewer & Harasty, 1996; Campbell, 1958; Dasgupta, Banaji, & Abelson, 1999; Denson, Lickel, Curtis, Stenstrom, & Ames, 2006; Hamilton, 2007; Hamilton & Sherman, 1996; Ip, Chiu, & Wan, 2006; Lambert, Barton, Lickel, & Wells, 1998; Levy, Freitas, & Salovey, 2002; Lickel, Hamilton, & Sherman, 2001; Luttmner, 2001; McConnell, Sherman, & Hamilton, 1997; Rothbart & Taylor, 1992; Yzerbyt, Corneille, & Estrada, 2001; Yzerbyt, Rocher, & Schadron, 1997). Research similarly suggests that people make inferences in the opposite direction, as perceivers seem to assume that people who experience similar outcomes in life are similar on other

dimensions (e.g., Corrigan, 2007; Hamilton, 2007; Ip et al., 2006; Lambert et al., 1998). For example, people who are arbitrarily labeled as *mentally ill* tend to seem similar on personality dimensions that are independent of mental illness (e.g., Corrigan, 2007).

A Bidirectional Model of Collective Treatment

Integrating evidence for this bidirectional relationship between homogeneity and collective treatment, we propose a bidirectional model of collective treatment. According to this model, homogeneous groups attract more collective treatment (Link 1) and collectively treated groups simultaneously seem more homogeneous on dimensions unrelated to that collective treatment (Link 2; see Figure 1).

In this paper, we define collective treatment as the act of behaving toward more than one individual uniformly, whether that behavior is positive (e.g., a monetary reward) or negative (e.g., criminal punishment). Collective treatment is distinguished from individualized treatment, in which individuals are treated differently from one another according to one or more relevant criteria (e.g., rewarding each member of a group for his or her relative contribution; punishing a gang of offenders individually according to their relative contributions to the crime). Collective treatment can occur at the hands of individuals (e.g., a government official who rewards a group within the population; a judge who sentences a gang of criminals) or by luck or natural processes (e.g., a natural disaster that affects more than one person; a lottery that rewards a group of people). In this paper, we attempt to provide empirical evidence for the existence of both links in the model (Studies 1–4), identify one mechanism for the effect (Studies 4 and 5), and, finally, document the model's impact in a real-world policy-making context (Studies 6a and 6b).

Locating Homogeneity Among Related Concepts

Our model focuses on the bidirectional relationship between homogeneity and collective treatment. However, perceived *homo-*

Adam L. Alter, Stern School of Business, New York University; John M. Darley, Department of Psychology and Woodrow Wilson School of Public and International Affairs, Princeton University.

We benefited from funding from the Woodrow Wilson Society of Scholars at Princeton University. Adam L. Alter was additionally funded by a Charlotte Elizabeth Procter Fellowship at Princeton University, and John M. Darley received additional funding from the Woodrow Wilson School of Public and International Affairs at Princeton University. We thank Elana Broch, Joel Cooper, Susan Fiske, Margaret Gerbasi, Geoff Goodwin, Dena Gromet, Lily Jampol, Valerie Loehr, Alison Newman, Danny Oppenheimer, Debbie Prentice, and Hulda Thorisdottir for their invaluable assistance.

Correspondence concerning this article should be addressed to Adam L. Alter, Stern School of Business, 40 West 4th Street, Room 818, New York, NY 10012. E-mail: aalter@stern.nyu.edu

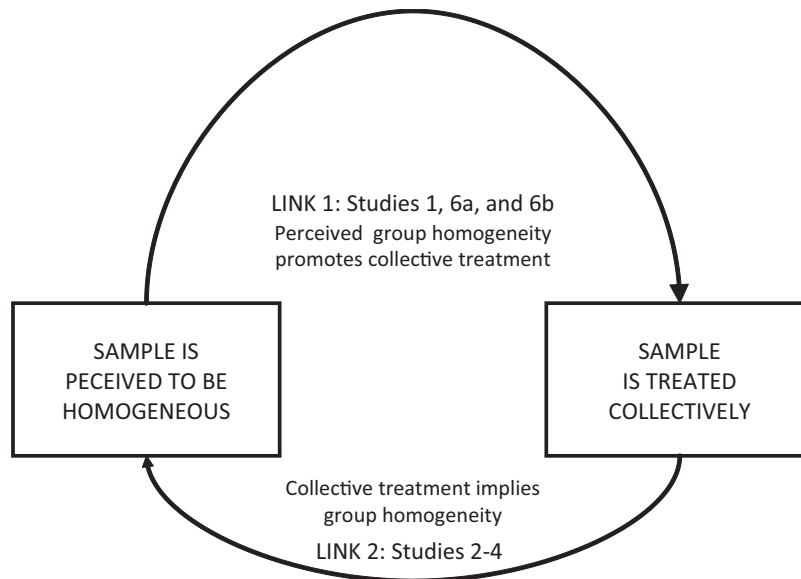


Figure 1. A model of the bidirectional relationship between group homogeneity and collective treatment. The model suggests that groups that appear more homogeneous are more likely to be treated collectively (Link 1: Studies 1, 6a, and 6b) and groups that are treated collectively appear more homogeneous (Link 2: Studies 2–4).

geneity—the extent to which a group’s members appear uniform on one or more salient dimensions—is closely related to two important constructs, *entitativity* and *essentialism* (for a comprehensive treatment of this issue, see Yzerbyt, Judd, & Corneille, 2004). Of the three concepts, homogeneity is the most basic and pervasive characteristic of group membership (Rothbart & Park, 2004). Homogeneity refers to lower order similarities between the members of a group, such as shared skin color, religious affiliation, nationality, or language (e.g., Hamilton, 2007). Some groups that are homogeneous also acquire the property of entitativity. A group is entitative to the extent that its members are perceived “as having the nature of an entity, of having real existence” (Campbell, 1958, p. 17). However, not every homogeneous sample of people forms an entity, and several relevant variables aside from homogeneity influence the perception of entitativity. For example, a social aggregate is more likely to appear entitative if its members are homogeneous and interact with one another and if the group’s members share common goals and common outcomes (Lickel et al., 2000). Moreover, some but not all entitative groups acquire an essence—the sense that the group is so fundamental that its members share an “underlying reality or true nature . . . that one cannot observe directly, but that gives [the group] its identity and is responsible for other similarities that the category members share” (Gelman, 2003, p. 8; see also Haslam, 1998). Essentialized groups can be distinguished from artificial groups, which exist only because a society imbues them with that label.

We focus on homogeneity, rather than entitativity or essentialism, for several reasons. First, homogeneity is the broadest criterion of group perception (e.g., Hamilton, 2007; Rothbart & Park, 2004). Whereas individuals who form a group tend to appear more homogeneous than individuals who do not form a group, not all groups are entitative or essentialized. Accordingly, a model that emphasizes homogeneity characterizes a larger set of groups than

would a model that focuses on either entity or essence. For example, every group’s members are alike on some salient dimension (e.g., all members of a knitting circle like to knit), but this does not mean that knitting lovers represent an entity or possess a common essence. Second, homogeneity is the weakest marker of group identity (see, e.g., Brewer, Hong, & Li, 2004), so it affords a more conservative test of our bidirectional model than would entity or essence. Third, homogeneity is objective and therefore experimentally tractable, whereas reliably manipulating the relatively subjective constructs of entitativity and essence is more difficult. To test the two links in our model empirically, we therefore chose to focus on homogeneity rather than entitativity and essence. Fourth, and perhaps most important, the results would be somewhat trivial were we to focus on entitativity and essence; groups that take on the subjective property of groupness will inevitably attract more collective treatment, because they are indeed collective units. Thus, although it might make sense for perceivers to label three people who are treated collectively as an entity, but if they also believe those people thereafter seem more homogeneous on unrelated dimensions, such as race, ethnicity, or nationality, these effects move beyond the realm of triviality and demonstrate an important bias in social perception.

The First Link: Homogeneous Groups Attract Collective Treatment

The first link of the model proposes that a homogeneous sample of individuals should attract more collective treatment than should a relatively heterogeneous sample. Because the model relies on subjective perceptions of the targeted individuals, homogeneity can take the form of any dimensions that are salient to an interested perceiver. For example, researchers have most commonly shown that others are perceived as forming groups when they share salient

physical characteristics (Campbell, 1958; Dasgupta et al., 1999; Kashima, 2004; Hamilton, 2007; Park & Hastie, 1987; Rothbart & Park, 2004). Physical similarity is a logical basis on which to perceive groups, as physical attributes are often the first pieces of available information about others. Not surprisingly, then, sports team members wear uniforms and sports fans wear clothes that identify them as supporters of a particular team.

There are good historical reasons to believe that people prefer to treat uniform groups collectively. During wars, each side construes the other as a single organism, often compared to lower order animals (e.g., Glover, 2001). In one World War II account, a Nazi officer recalled that he could only exterminate children, the elderly, and the sick by recasting them as a set of identical deindividuated vermin (Browning, 1992). Similarly, people might prefer to reward uniform and deserving groups, and some researchers have argued that citizens of collectivistic cultures treat their fellow citizens collectively because they perceive the country's citizens as physically and culturally homogeneous (Hui, Triandis, & Yee, 1991; Leung & Bond, 1982).

Consistent with these anecdotes, decades of social psychological research has shown that people assume that physically homogeneous groups are also homogeneous on other dimensions. Once physically similar individuals seem to share more meaningful character traits, the group's identity is more likely to become a salient category (Gelman, 2003; Prentice & Miller, 2006; Rothbart & Taylor, 1992). Groups in turn come to have the property of agency, and the actions of one group member are seen to represent the will of the group at large (e.g., Abelson, Dasgupta, Park, & Banaji, 1998; Dasgupta et al., 1999; Morris, Menon, & Ames, 2001). In one study, Dasgupta et al. (1999, Experiment 2) found that people perceived computer-generated creatures called Greebles as more psychologically similar when they looked homogeneous, either because they stood in a collective formation or because they were similarly colored. When people conflate superficial physical similarity with character-based psychological similarity, they might more readily justify treating a collection of similar-looking individuals as a collective unit. Indeed, people perceive greater fairness in punishing homogeneous groups for transgressions that were perpetrated by the group's members, including the Holocaust and suicide bombings (e.g., Denson et al., 2006).

Accordingly, we expected participants to perceive a sample of superficially homogeneous people as a group and to treat this sample more collectively than they did a sample of superficially heterogeneous people. We examined this effect in a lab study (Study 1) and in two naturalistic policy-relevant studies that capitalized on archival data (Studies 6a and 6b).

The Second Link: Collectively Treated People Thereafter Appear More Homogeneous

The second link in our model proposes the converse relationship between group-member homogeneity and collective treatment: Instead of expecting homogeneous people to be treated more collectively, the second link suggests that perceivers will interpret collectively treated people as being more homogeneous than people who are not treated collectively.

As with the first link, there are many reasons to expect that collectively treated people will appear more homogeneous. In his

proposed early theoretical account of group perception, Campbell (1958) proposed the term *entitativity* to describe social aggregates that have "the nature of entity, having real existence" (p. 17). This enduring theoretical construct draws on the Gestalt principles of perceptual organization (for reviews, see, e.g., Hamilton, 2007; Hamilton & Sherman, 1996), notably for our purposes including the principle of common fate (Wertheimer, 1938). The law of common fate, applied to object perception, suggests that objects moving in a uniform direction will be perceived as a collective unit. However, Campbell (1958) broadened the principle and noted that people who experience the same fate might also seem to form a collective unit. For example, the survivors of a disaster sometimes acquire group status (e.g., the Oceanic Six in television's *Lost*; the Jena Six, six black teenagers who were controversially charged with the beating of a white teenager in Jena, Louisiana, in 2006), and people who live in a particular region acquire regional nicknames based on the conditions they experience (e.g., Carter, 1944).

Common fate has been used as a group marker across many domains and extending through time (see, e.g., Brewer, Weber, & Carini, 1995; Insko et al., 1988; Gaertner & Schopler, 1998; Wilder & Simon, 1998). Holocaust survivors, lottery winners, and medical patients coalesce around their experiences; universities and local communities similarly sponsor support groups for sexual assault victims, grief sufferers, and even people who have come into newfound wealth and are not quite sure how to manage their windfall. The people who occupy these groups appear more homogeneous following their common experiences, in part because the experiences that bind them together are relatively more salient than the differences that distinguish them from each other. Similarly, Fiske and Neuberg's (1990) continuum model of impression formation suggests that people are cognitive misers who prefer to categorize others according to salient group membership markers, such as a common fate. The more salient and consistent the group marker, the less likely the perceiver is to engage individuating processes for distinguishing the individual from his or her fellow group members. Brewer's (1988) dual-process model of impression formation makes a similar observation, suggesting that people intuitively identify and "type" others in a holistic and superficial manner and engage controlled individuation processes only when this typing process inadequately describes the perceived individual. Given research suggesting that people rely on their intuitions unless those intuitions are somehow challenged (e.g., Alter, Oppenheimer, Epley, & Eyre, 2007; Forgas, 1992, 1993), we expected participants to more readily perceive commonly treated individuals than their individually treated counterparts as a group.

Seeking a Mechanism for the Bidirectional Relationship

One further possibility that we sought to test empirically is that people form an association between collective treatment and homogeneity across time and that this subsequently leads them to associate the two constructs in novel settings. So, all else being equal, people who look similarly attractive or share an ethnic background are probably more likely to attract similar treatment than are people who look different from each other. Decades of research in social psychology supports this contention, because group markers, such as race, ethnicity, and religion, and individual differences, such as attractiveness, have profound effects on how

people are treated and the outcomes they tend to experience (e.g., Dion, 1972; Dion, Berscheid, & Walster, 1972; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006; Fiske, 1998; Frieze, Olson, & Russell, 1991).

Considerable evidence in the developmental and social-cognitive literatures suggests that people are liable to form associations between homogeneity and collective treatment across time. As children develop, they adopt attitudes toward people who experience lucky and unlucky outcomes (Olson, Banaji, Dweck, & Spelke, 2006) and increasingly favor their in-group over out-groups (e.g., Zosuls et al., 2009). Adults also form associations between neutral concepts in the environment and subsequently incorporate those associations into their behaviors and judgments (e.g., Briñol, Petty, & Tormala, 2006; Linville, Fischer, & Yoon, 1996; Spencer-Rodgers, Williams, Hamilton, Peng, & Wang, 2007; Unkelbach, 2006). Accordingly, in two studies, we examined whether participants perceived a stronger link between homogeneity and collective treatment when they generally endorsed the principle that homogeneous people tend to be treated more similarly (Study 4) and showed that people generally endorse this naive theory (Study 5).

The Present Studies

We conducted seven studies to examine the two links in our bidirectional model. The first five studies were conducted in the lab, and the remaining two studies (Studies 6a and 6b) comprised analyses of two data sets examining the implications of the model for real-world policy making. Studies 1, 6a, and 6b examined Link 1: whether homogeneous samples of people would seem more likely to attract collective treatment. Studies 2–4 considered the second link in the model: the tendency for collectively treated individuals to seem more homogeneous than individually treated group members. These studies examined the effects of collective treatment on memories of group homogeneity (Study 2), immediate perceptions of the homogeneity of the group's members (Study 3), and whether people imagined collectively treated group members to be more homogeneous than individually treated but otherwise identical group members (Study 4). We sought to demonstrate the effect using both positive (Studies 2–4) and negative (Study 3) forms of collective treatment that were carried out by human agents (Study 2), natural agents (Study 3), and processes governed by chance (Study 4). Seeking a mechanism for the effect, we conducted Study 5 to show that people expect homogeneous groups to attract collective treatment relative to heterogeneous groups, whose members they expect will be treated as individuals.

Study 1: Sentencing Criminal Gang Members on the Basis of Visual Information

Participants in Study 1 suggested sentences for the members of a five-member criminal gang. The gang's members, who had allegedly robbed a jewelry store, were depicted as silhouettes against an arrest-chart background. Participants responded to one of two versions of the questionnaire: a homogeneous version, in which the members appeared to be of similar height and weight, and a heterogeneous version, in which the members differed markedly in their height and weight. Participants assigned a sentence to each member of the gang and afterward estimated the defendant's

height and weight based on the image. We expected participants to recommend a more uniform array of sentences for the homogeneous gang members than for the heterogeneous gang members.

Method

Participants

Thirty-five Princeton University undergraduates (22 women; $M_{\text{age}} = 19.62$ years, $SD = 1.22$) participated in this study in exchange for partial course credit. Of participants, 70% identified themselves as White, 4% as Black, 18% as Asian, 4% as Hispanic, and 4% as belonging to one of several other ethnic groups.

Materials, Design, and Procedure

Participants completed a questionnaire titled "Punishing Criminals." The questionnaire began by explaining that five men had recently been convicted of robbing a jewelry store in Trenton, New Jersey. Each man played a part in the robbery, and the men were later recognized by a bystander who identified them in a lineup. Following this introduction, participants were shown a stylized version of the image that allowed the bystander to identify the five defendants by their silhouettes.

In fact, there were two versions of the questionnaire that differed according to how similar the five gang members appeared. The homogeneous gang members had markedly more uniform silhouettes than did the heterogeneous gang members (see Figure 2). Participants were asked to assign sentences to each of the five gang members based on the part that each member played in the crime. Defendant A purchased the equipment for the robbery, Defendant B drove the getaway car, Defendant C masterminded the plan, Defendant D looked out for the police, and Defendant E broke into the jewelry store and stole the jewelry. The 7-point sentencing scale was labeled with the options of "no punishment," "\$500 fine," "\$2,000 fine," "\$5,000 fine," "1-month jail term," "1-year jail term," and "3-year jail term" (adapted from, e.g., Alter, Kernochan, & Darley, 2007; Erikson & Gibbs, 1979; Gescheider, Catlin, & Fontana, 1982). The standard deviation and range of these sentences served as the primary dependent measures, whereby a greater standard deviation and wider range of sentences indicated greater variation in sentencing. Finally, participants reported their age, ethnicity, and gender, and the experimenter explained the purposes of the experiment.

Results and Discussion

Primary Analyses

To examine how uniformly participants sentenced the defendants in each condition, we conducted independent samples *t* tests comparing the standard deviation and range of the sentences imposed on the five defendants in each condition. As expected, the sentences imposed on the homogeneous defendants ($M_{\text{sentenceSD}} = 0.34$, $SD = 0.41$) were less varied than the sentences imposed on the heterogeneous defendants ($M_{\text{sentenceSD}} = 0.83$, $SD = 0.79$), $t(33) = 2.31$, $p < .03$, $\eta_p^2 = .14$. Moreover, participants assigned a narrower range of sentences to the homogeneous defendants ($M = 0.76$, $SD = 0.97$) than to the heterogeneous defendants ($M = 1.88$, $SD = 1.89$), $t(33) = 2.21$, $p < .04$, $\eta_p^2 = .13$.

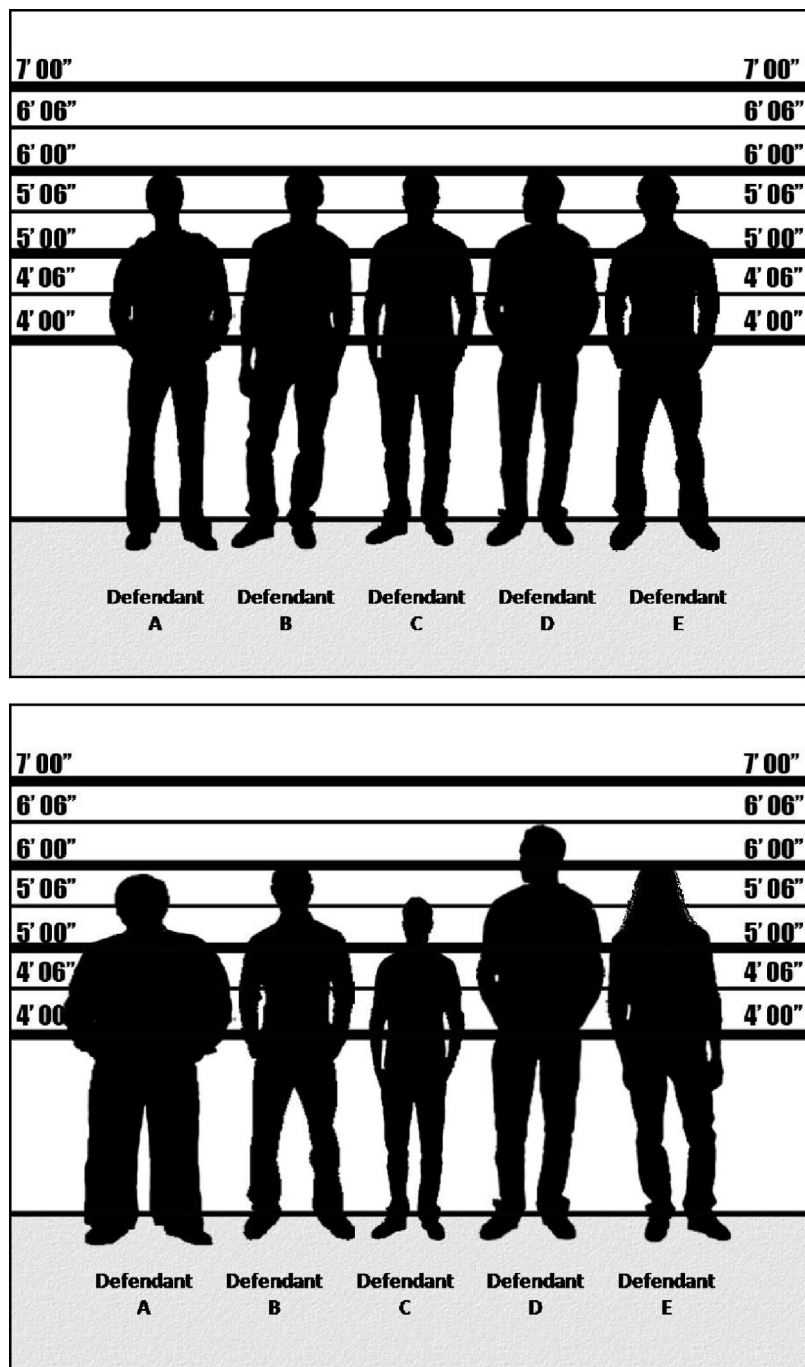


Figure 2. Homogeneous criminal gang (top panel) and heterogeneous criminal gang (bottom panel) from Study 1.

Secondary Analyses

These results are consistent with our expectations, but simple comparisons leave open the possibility that participants merely assigned more erratic sentences when confronted with a gang of motley criminals and ignored the relative severity of the gang members' roles in the robbery altogether. To rule out this possibility, we examined the mean sentences assigned to each gang

member in the two conditions. As Figure 3 shows, participants ranked the gang members' crimes identically in both conditions: The equipment purchaser received the lightest sentence; the police lookout the second lightest; the getaway driver the median sentence; the mastermind the second heaviest; and the gang member who broke into the jewelry shop the heaviest sentence. Notably, and consistent with our hypothesis, these sentences were spread more widely across the 7-point scale in the heterogeneous condi-

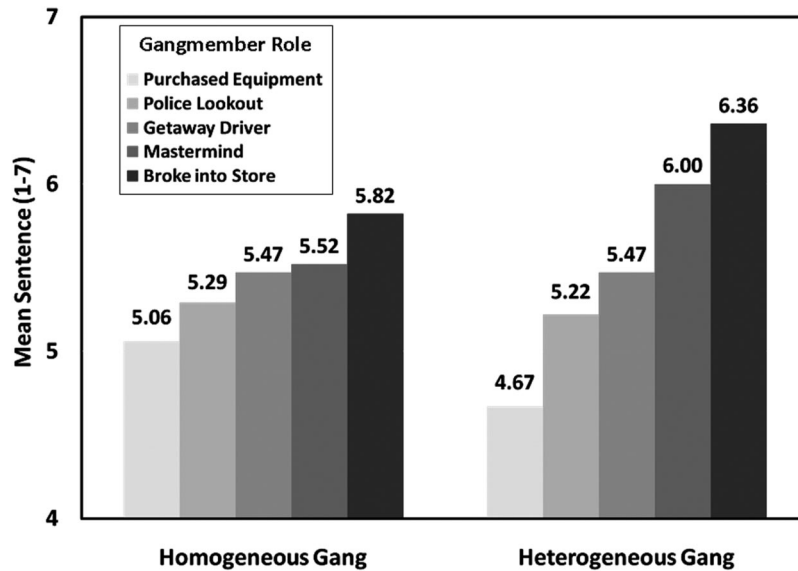


Figure 3. Mean sentences assigned to each of the five gang members in Study 1.

tion than in the homogeneous condition. Thus, participants were not merely more erratic when assigning sentences to heterogeneous gang members; rather, their more varied sentences remained tethered to each gang member's role in the robbery.

We were also concerned that participants may have assigned higher sentences, on average, to defendants in one condition relative to the other, thereby spuriously elevating the variance in those sentences. However, as Figure 3 suggests, the mean sentence imposed on the five defendants was very similar in the two conditions ($M_{\text{homogeneous}} = 5.44$, $SD = 1.32$ vs. $M_{\text{heterogeneous}} = 5.52$, $SD = .90$), $t(33) < 1$, $p > .82$, $\eta_p^2 < .01$. This suggests that the variance of the sentences imposed on the defendants in each condition was not an artifact of the absolute sentences imposed on the five defendants in each condition.

A third concern with Study 1 was that the defendants' height and weight were salient visual cues that may have influenced sentencing for reasons unrelated to our hypothesis. For example, participants may have believed that larger defendants were more capable of sustaining longer sentences and therefore imposed harsher sentences on those defendants relative to the more diminutive defendants. Consistent with this position, Kolber (2009) has recently suggested that judicial decision makers attempt to assign punishments across an array of defendants, so that they experience a punishment of roughly equivalent magnitude. Thus, a larger defendant may be punished more harshly than a smaller defendant who might suffer more greatly in prison. In fact, the sentencing data suggest that participants did not use this alternative approach to sentencing. The two most strikingly different defendants were C and D: In the homogeneous condition, participants believed those defendants were on average 180 cm tall and weighed approximately 80 kg; in contrast, their counterparts in the heterogeneous condition were judged to be 168 cm tall and to weigh 66 kg and to be 193 cm tall and to weigh 96 kg, respectively. Nonetheless, these notably larger and smaller defendants in the heterogeneous condition received sentences very similar to those of their moderately

sized counterparts in the homogeneous condition ($t < 1$, $p > .88$ and $t(33) = 1.29$, $p > .20$, respectively). Indeed, the larger defendant was sentenced slightly less heavily than his moderately sized counterpart ($M = 5.22$ vs. $M = 5.29$), and the smaller defendant was sentenced slightly more heavily than his moderately sized counterpart ($M = 6.00$ vs. $M = 5.53$). Thus, participants did not appear to use each gang member's size as a cue when assigning sentences for the robbery.

Study 1 therefore suggests that a homogeneous collection of individuals attracts a more consistent array of treatments than does a heterogeneous collection of individuals who have otherwise behaved similarly. Having shown evidence for the first limb in the bidirectional model (see Figure 1), we sought evidence for the converse effect in Studies 2–4: that collectively treated targets thereafter seem more homogeneous on unrelated dimensions than do individually treated targets.

Study 2: Remembering Collectively Treated Tribes as More Homogeneous

Participants in Study 2 read a fabricated story describing the origin of affirmative action interventions. According to the story, two downtrodden tribes in the Comoros Islands approached the national government, seeking special leniency in gaining employment. The government handed down a ruling in which one tribe was granted affirmative action, but the other tribe was told that its members would have to fend for themselves as individuals and that they could not benefit from their membership in the tribe. After reading about the demographic composition of the two tribes, participants were given an unexpected memory test, in which they were asked to recall the number of different races, religions, languages, and dialects represented in each tribe. We chose to use a memory test because misremembering a group as more homogeneous than it actually is might affect how the group is perceived and treated in the long term. This method has also been used successfully in similar studies, where misremembering a group's

members might have long-term consequences for the group's well-being (e.g., Norton, Vandello, & Darley, 2004). Consistent with the second link in the model, we expected participants to misremember the tribe that received affirmative action as more homogeneous than the tribe whose members continued to compete for employment as individuals.

Method

Participants

Forty Princeton University undergraduates participated in this study in exchange for partial course credit.

Materials, Design, and Procedure

Participants were asked to complete a questionnaire ostensibly seeking their views on affirmative action policies. The questionnaire began with a bogus explanation of the origin of affirmative action policies. The story recounted how two tribal groups in the Comoros Islands, the Mwali and Quola tribes, had been downtrodden historically and their members had struggled to find employment on the islands. The story was written to depict the two tribes as quite similar to one another in all material respects, and we refrained from explicitly distinguishing them on any material dimensions like income, health, and poverty. Both tribes had approached the island's government seeking what amounted to affirmative action—a government-sanctioned collective boost in their prospects of gaining employment. According to the story, the government ultimately decided to award this version of affirmative action to one tribe, but not the other, stating that the successful tribe's members would be treated as a group when they applied for employment, whereas the unsuccessful tribe's members would continue to be treated as individuals, independently of their tribal membership. To eliminate tribe-name effects, half the questionnaires stated that the Mwali tribe was successful and the remaining questionnaires stated that the Quola tribe was successful.

With the ostensible aim of familiarizing participants with each tribe, we showed participants a table depicting six demographic characteristics of each tribe, four of which were related to homogeneity and two of which were not related to homogeneity. The four homogeneity-related demographics were the number of racial groups and religious groups represented among the tribes' members and the number of languages and dialects spoken in each tribe. Like the tribe names, these numbers were counterbalanced so the successful tribe comprised either 18 racial groups, 15 religious groups, 19 languages, and 41 dialects or 19 racial groups, 14 religious groups, 17 languages, and 43 dialects (in both cases, these numbers summed to 93). The number of variations within each dimension were chosen to be quite similar between the two tribes, and each tribe was more diverse on two of the four dimensions and more uniform on the remaining two dimensions. For half the participants, the successful tribe's demographic information was presented in the left-hand column of the table, whereas for the remaining participants the successful tribe's information was presented in the right-hand column.

Participants were also shown two demographic variables that were not related to the tribes' relative homogeneity and therefore should not have been influenced by the collective treatment ma-

nipulation: the population of each tribe (12,000 and 12,500, counterbalanced) and each tribe's average income (4,000 and 4,200 Comorian francs, also counterbalanced). We did not expect participants to remember the groups differently on these dimensions, in contrast to the four diversity-related demographic variables.

The experimenter collected the first page of the questionnaire when participants were ready to continue and handed them an empty-celled replica of the demographic table from the first page. The experimenter asked participants to replicate the table they had seen on the previous page as accurately as they could. This memory test was designed to determine whether participants would erroneously remember the collectively treated tribe members as more demographically homogeneous than the individually treated tribe members.

Finally, to probe participants' memories explicitly, we concluded by asking participants to recall whether one of the tribes was more diverse than the other on a 7-point scale (anchored at 1 = *Tribe X was more diverse*; 4 = *Both tribes were equally diverse*; and 7 = *Tribe Y was more diverse*). We counterbalanced which tribe appeared at each pole and which side of the scale was labeled with the collectively and individually treated tribes. We recoded participants' responses so that lower scores indicated a belief that the collectively treated tribe was more homogeneous.

None of the participants appeared suspicious of the purpose of the experiment, and all reported being unaware of the hypotheses when asked at the conclusion of the debriefing process.

Results and Discussion

Preliminary Analyses

None of the counterbalancing variables (tribe name, which tribe was successful, which demographic characteristics were associated with the successful tribe, which tribe's demographic characteristics were depicted in the left-hand side of the table) interacted with participants' responses. These variables are therefore not discussed in the remaining analyses ($F_s < 1$).

Because the absolute numbers associated with the demographic characteristics varied widely, we began by standardizing participants' estimates across the six demographics so they were represented on a uniform scale. We combined the four homogeneity-related demographics and, separately, the two homogeneity-unrelated demographics to form two indices, the first of which should have been influenced by the tribes' treatment and the second of which should have been insensitive to the manipulation. Lower scores on the indices indicated lower estimates.

Across the sample of 40 participants, eight of the 480 estimates were left blank. These empty cells were not disproportionately associated with either the individually or the collectively treated tribes, so they were treated as missing data in all further analyses.

Primary Analyses

We began by analyzing participants' memories of the tribes on the homogeneity-related demographics index (which comprised religions, races, languages, and dialects). As expected, participants perceived fewer demographic variations within the collectively treated tribe ($M = -0.05$, $SD = 0.63$) than within the individually treated tribe ($M = 0.04$, $SD = 0.62$), $t(39) = -2.67$, $SEM = .03$,

$p = .01$, $\eta_p^2 = .16$. This result suggests that participants remembered the collectively treated tribe as more homogeneous than the individually treated tribe, despite the fact that the four demographic characteristics were counterbalanced between the two tribes. Simply put, this study suggests that when a group of individuals experiences a uniform outcome, people misremember the group as more homogeneous than it actually is.

Participants should have remembered one tribe as more homogeneous on two of the four demographic characteristics and the other tribe as more homogeneous on the other two dimensions. Instead, a one-sample t test suggested that, on average, participants remembered the collectively treated group as relatively more homogeneous than the individually treated group on statistically more than two of the four dimensions ($M = 2.34$, $SD = 0.89$), $t(39) = 2.41$, $p = .02$, $\eta_p^2 = .13$.

Finally, we examined whether participants misremembered the collectively treated tribe as more homogeneous than the individually treated tribe on the explicit 7-point scale. Because we coded participants' responses so lower scores indicated a belief that the collectively treated tribe was more homogeneous, we expected participants to generally respond with scores below the midpoint of 4. Indeed, on average, participants' estimates were significantly lower than the midpoint of 4 ($M = 3.38$, $SD = 1.00$), $t(39) = 3.93$, $p < .001$, $\eta_p^2 = .28$. This result suggested that participants perceived the collectively treated tribe to be more homogeneous.

Secondary Analyses

Although these results are consistent with our prediction that collectively treated tribes should be remembered as more homogeneous, these differences might also reflect greater accuracy in recalling the demographics of one tribe rather than the other. For example, having learned of a tribe's outcome, participants may have paid greater attention to that tribe than the other. However, this alternative mechanism for the effect seems unlikely because participants did not remember the tribes differently on the collapsed population-income index, which was not related to the tribes' diversity ($M_{\text{homogeneous}} = 0.04$, $SD = 0.64$ vs. $M_{\text{heterogeneous}} = -0.04$, $SD = 0.55$), $t = 0.76$, $SEM = 0.10$, $p = .45$, $\eta_p^2 = .02$. Thus, participants were not merely more accurate in recalling the demographic characteristics of one of the two tribes; rather, they perceived differences only when recalling the demographic characteristics associated with homogeneity.

These analyses suggest that participants misremembered the collectively treated tribe as more homogeneous than its individually treated counterpart. In Study 3, we sought to replicate this effect, while addressing a number of concerns with the design of Study 2.

First, although we labeled the tribes "collectively treated" and "individually treated," an alternative interpretation might be that one tribe received collective treatment whereas the other did not receive any form of treatment at all. Although we attempted to emphasize the collective and individual nature of the treatment tribe members would expect depending on the outcome of the decision, participants may not have adopted that interpretation and may have seen the individual treatment condition as an absence of treatment altogether. Accordingly, in Study 3, we ensured that the sample of targets unambiguously received treatment in both conditions that differed only in its collectivity.

Second, the design of Study 2 left open the possibility that participants misremembered the collectively treated group as more homogeneous because they assumed that more homogeneous groups are naturally more deserving of collective treatment in the first place. Although this interpretation is psychologically interesting, it obfuscates the causal direction of the effect. Whereas we argued that collectively treated groups subsequently seemed more homogeneous, this interpretation leaves open the possibility that participants believed the successful group was granted affirmative action because its members were originally more homogeneous.

We adopted two methodological approaches in Study 3 to eliminate this concern. First, the collectively treated and individually treated targets in Study 3 were presented between subjects, so participants could not compare them to one another and devise naive theories about why one group was treated collectively whereas the other was treated individually. Second, and more important, the targets in Study 3 were "treated" collectively or individually by natural processes: a tornado or the discovery of oil. Because neither tornadoes nor oil deposits selectively affect homogeneous groups of people, participants could not logically infer that the collective treatment followed from some preexisting difference in homogeneity that distinguished the two groups of people.

A further distinction between the collectively and individually treated tribes in Study 2 was the relative advantage enjoyed by the collectively treated tribe. We eliminated the unlikely possibility that lucky groups seem more homogeneous by having the target individuals in Study 3 experience either a positive event (a shared oil discovery) or a negative event (destruction arising from a tornado), so we could examine whether the valence of the event interacted with participants' responses.

Finally, Study 2 documented a memory effect, whereby participants selectively overestimated the homogeneity of the collectively treated group relative to the individually treated group. In Study 3, we sought to extend the effect to the domain of real-time person perception and showed that people perceive photographs of a group of collectively treated individuals as more homogeneous than the same photographs attributed to individually treated individuals.

Study 3: Perceiving Collectively Treated Apartment Owners as More Homogeneous

Participants in Study 3 read a fabricated newspaper article describing an event that affected the owners of 40 apartments in an apartment block. The event—either the discovery of oil on the apartment block's land parcel or the destruction of one or all the apartments by a tornado—affected either one of the apartment owners (non-collective treatment) or all 40 owners equally (collective treatment). Participants saw a picture of 40 people who were supposedly the apartment owners and estimated the number of different nationalities, languages, and ethnicities represented among the 40 apartment owners. We expected participants to perceive greater homogeneity when the apartment owners were treated collectively than when they were not treated collectively.

Method

Participants

Thirty-eight adults (20 women; $M_{\text{age}} = 22.05$ years, $SD = 3.52$) volunteered to complete this study while eating at the Princeton University student campus center.

Materials, Design, and Procedure

Participants completed one of four versions of a one-page questionnaire in which they were asked to respond to a fabricated newspaper article. There were four versions of the article that differed according to whether the apartment owners did or did not experience collective treatment (the event affected all the owners vs. just one owner) and whether that treatment was positive or negative (a financial gain stemming from the accidental discovery of oil on the property vs. the destruction of all or part of the complex by a tornado). Thus, the four versions were collective positive (each person benefited from an accidental discovery of oil on the apartment block's land); noncollective positive (oil was discovered but only the apartment owner who discovered the oil benefited); collective negative (a tornado destroyed all 40 apartments); and noncollective negative (the tornado destroyed just one apartment and spared the remaining 39). Thus, collective and noncollective treatment differed according to whether all the apartment owners experienced the same positive or negative fate or, rather, whether one owner was singled out relative to the others.

Forty small headshot photographs, described as a headshot of each of the 40 apartment owners, followed the paragraph describing the event. The same 40 photographs were used in all four conditions. The people depicted in the photographs were actually amateur dancers who were taking a class that was publicized on the Internet. They appeared to be quite ethnically diverse and ranged in age from approximately 20 to 40 years.

After participants had read the story and while they looked at the photographs, we asked them to make various ratings of the apartment owners, most of which were fillers designed to obfuscate the purpose of the study (e.g., on average, how long do you think the apartment owners have lived in the building?). We included the three dependent measures among the fillers, in which participants were asked to estimate how many different nationalities, ethnic groups, and languages were represented among the 40 apartment owners. To reiterate, we expected participants to perceive greater homogeneity among the apartment owners when they were all affected equally by the oil discovery or the tornado.¹ Moreover, we expected the effect to emerge regardless of whether the event was beneficial (oil discovery) or detrimental (tornado destruction). Finally, participants reported their age, gender, and ethnicity and were debriefed about the purpose of the experiment. None of the participants claimed to be suspicious about the purpose of the experiment during the debriefing process.

Results and Discussion

As in Study 2, we began by collapsing participants' estimates on the nationality, ethnic group, and languages demographic variables to form a single homogeneity index. Using this index as a dependent measure of perceived homogeneity, we conducted a 2 (event type: oil discover, tornado) \times 2 (collectivity: all affected, one

affected) between-subjects analysis of variance (ANOVA) to determine whether participants perceived the 40 apartment owners differently across the four conditions. As expected, participants perceived the apartment owners as more homogeneous when they all experienced the same outcome ($M = 5.13$, $SD = 1.43$) than when one apartment owner experienced a different outcome ($M = 6.72$, $SD = 2.53$), $F(1, 34) = 5.31$, $p < .03$, $\eta_p^2 = .14$ (see Figure 4). Neither the main effect of event type nor the interaction between event type and collectivity was significant ($F_s < 1$, $p_s > .5$), and this suggested that the effect was not qualified by the positivity or negativity of the event that affected one or all of the apartment owners.

Studies 2 and 3 demonstrated support for the second link of the model (see Figure 1), such that people perceived a group of individuals as more homogeneous after they uniformly experienced the same outcome. The effect persisted regardless of whether that outcome was positive or negative and whether it had occurred at the hands of nature, chance, or a human decision maker. Collective treatment appeared to influence how homogeneous the group seemed both in retrospective memory and in immediate perception. We sought to extend and replicate these results in Study 4 by investigating whether participants would imagine the members of collectively treated groups as more homogeneous than their counterparts from individually treated groups. To test the possibility that these effects are driven in part by a naive association between homogeneity and collective treatment, we also examined whether the effect was strongest among participants who most strongly endorsed the naive theory that homogeneous groups generally attract more uniform treatment.

Study 4: Imagining People Who Travel to the Same Town as More Homogeneous

In Study 4, participants were told that the U.S. government had sponsored a program in which six American students were given the opportunity to travel abroad, all expenses paid. Each week for 6 weeks, one student was randomly selected in a lottery broadcast on local television stations from a larger pool of applicants. In one condition, all six winning students visited the same small town in a European country, whereas in the other condition, each student winner went to a different country. Participants were shown a photograph of the first winner and were asked to guess which five of 10 students depicted in a set of photographs were randomly selected to participate in the program over the remaining 5 weeks. Half the photographs were morphed with the photograph of the first winner, so the students in those photographs looked somewhat like the first winner, whereas the remaining students' photos were not morphed and were therefore independent from the first winner's photograph. Because we expected participants to imagine that the winners would be more homogeneous when they were

¹ To minimize the possibility that participants merely paid more attention to each of the 40 faces in the noncollective treatment conditions, thereby elevating how diverse the faces appeared to be, we did not indicate which of the 40 individuals whose faces were shown experienced the target outcome (striking oil or losing his or her apartment in a tornado). We also refrained from suggesting that each of the 40 owners experienced a different outcome, as this might have led participants to devote more attention to individual faces.

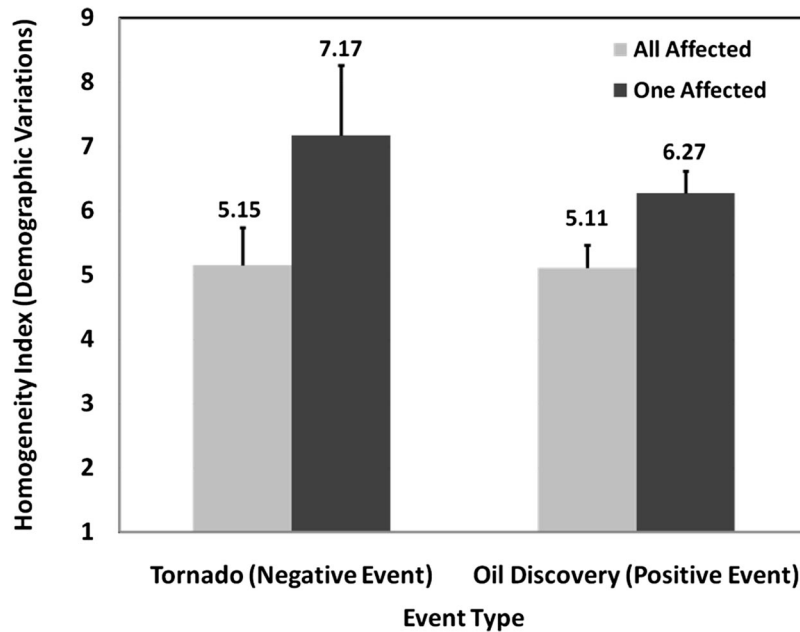


Figure 4. Estimated variance in demographic characteristics (ethnicity, language, and nationality) among 40 apartment owners who collectively or individually experienced the benefits of an oil discovery or the destruction of a tornado in Study 3. Bars represent standard error.

treated collectively (i.e., traveled to the same destination), we expected the morphed faces to be selected as winners more often when the six students traveled to the same town. We also measured how strongly participants generally believed that homogeneous groups attract similar treatment to determine whether the effect was driven in part by how strongly participants endorsed the naive theory that collectively treated groups tend to be more homogeneous.

Method

Participants

Ninety Princeton University undergraduates and staff (46 women; $M_{\text{age}} = 22.21$ years, $SD = 6.87$) volunteered to complete this study while eating lunch at the university campus center.

Materials, Design, and Procedure

Participants completed a one-page questionnaire, titled "Travel for Students: A Government Program." The questionnaire described a recent government initiative, in which six lucky high school students were randomly selected from a larger pool of applicants to travel overseas on a cultural exchange program. Under the program's rules, for each of 6 weeks one student from the pool was randomly selected to take part in the program. Participants were shown a photograph of the first winner, and their job was to guess which five of 10 other students (half of each gender) depicted in separate photographs won the remaining five places in the program.

There were four versions of the questionnaire, conforming to a 2×2 design: The first winner was either a male student or a female student, and the six winners embarked on trips either to six

small towns in six different countries around the world (individual treatment) or to the same small town in France (collective treatment). We varied the gender of the first winner to eliminate the possibility that participants merely assumed that the subsequent winners shared the first winner's gender and also to ensure that the results generalized beyond one set of photographs and one winner.

Five of the 10 target photographs were created by morphing the first winner's face with one of five other faces, so they all shared similar features with each other and with the first winner's face. The remaining five faces were not morphed and were therefore completely distinct from each other and from the five morphed faces. We took the faces from a University of St. Andrews *Face of the Future* computer science webpage (see <http://morph.cs.st-andrews.ac.uk/averager/index.html>) and used a morphing program embedded in the same webpage.

We expected participants to imagine that the winning students were relatively more homogeneous (leading them to select the morphed faces) when the students were collectively offered a trip to the same small town in France than when they visited six different small towns. As a test of this hypothesis, participants completed two dependent measures: First, they were asked to guess which five of the 10 target photos won a trip. From these estimates, we calculated the percentage of winning students that participants selected from the set of five morphed faces. Second, participants rated the likelihood that each of the 10 faces would win a place on the trip (1 = *not at all likely* to 10 = *very likely*). We averaged their responses to the five morphed faces and the five nonmorphed faces to create indices that represented how likely participants believed each of the 10 faces were to win a place on the trip.

Finally, we examined whether participants' responses reflected naive theories about the relationship between homogeneity and

collective treatment. To measure their endorsement of this naive theory, we asked participants how strongly they agreed with the statement that “people who experience similar outcomes in life tend to look similar, whereas people who experience different events in life tend to look different from each other” (1 = *strongly disagree* to 7 = *strongly agree*).

After they had completed the questionnaire, participants were thanked for their participation, debriefed about the nature of the study, and told that the student travel program was actually fabricated for the purposes of the experiment.

Results and Discussion

Primary Analyses

Participants’ responses did not differ depending on whether they were exposed to the male-target or the female-target version of the questionnaire, so we collapsed their responses on this dimension. Accordingly, we were left with two independent variables of interest: whether the students won vacations to the same destination (collective treatment) or to six different destinations (individual treatment; a between-participants variable) and whether the students’ faces were morphed with the target face depicting the winner of the first of the six weekly prizes (morphed vs. nonmorphed; a within-participant variable).

Binary win variable. As expected, participants believed that a greater proportion of the students with morphed faces won the trip in the collective treatment condition ($M = 55.48\%$, $SD = 22.70\%$) than in the individual treatment condition ($M = 42.50\%$, $SD = 19.41\%$), $t(88) = 2.92$, $p < .01$, $\eta_p^2 = .09$.

Continuous win likelihood variable. A 2×2 mixed-design ANOVA revealed the expected interaction between treatment collectivity and whether the faces were morphed or nonmorphed on

participants’ win-likelihood judgments, $F(1, 88) = 11.44$, $p = .001$, $\eta_p^2 = .12$ (see Figure 5). Follow-up simple-effects analyses showed that when all six students went to same destination, participants rated the morphed (and therefore more homogeneous) students as more likely to share the prize with the target than were the nonmorphed (and therefore more heterogeneous) students, $F(1, 41) = 4.86$, $p < .04$, $\eta_p^2 = .11$. In contrast, when the students went to six different destinations, participants imagined that the non-morphed students were more likely to share the prize with the target than were the morphed students, $F(1, 47) = 6.86$, $p < .02$, $\eta_p^2 = .12$.

These results confirmed our expectations that participants would generally expect homogeneous students to have won trips to the same country and heterogeneous students to have won trips to six different countries.

Naive Theory Mechanism Analyses

One possible explanation for these results is that participants generally endorse the naive theory that similar-looking people experience similar outcomes in life, whereas different-looking people experience more dissimilar outcomes. As a test of this explanation, at the end of the questionnaire all participants indicated how strongly they agreed (1 = *not at all* to 7 = *very strongly*) with the statement “People who experience similar events (e.g., tragedies, triumphs, daily events) during their lives tend to look similar to each other.” We expected participants who endorsed this statement more strongly to believe that the collectively treated students were relatively more homogeneous in appearance and the individually treated students were relatively more heterogeneous in appearance.

First, we calculated how much more strongly participants expected the students with morphed faces (homogeneous targets)

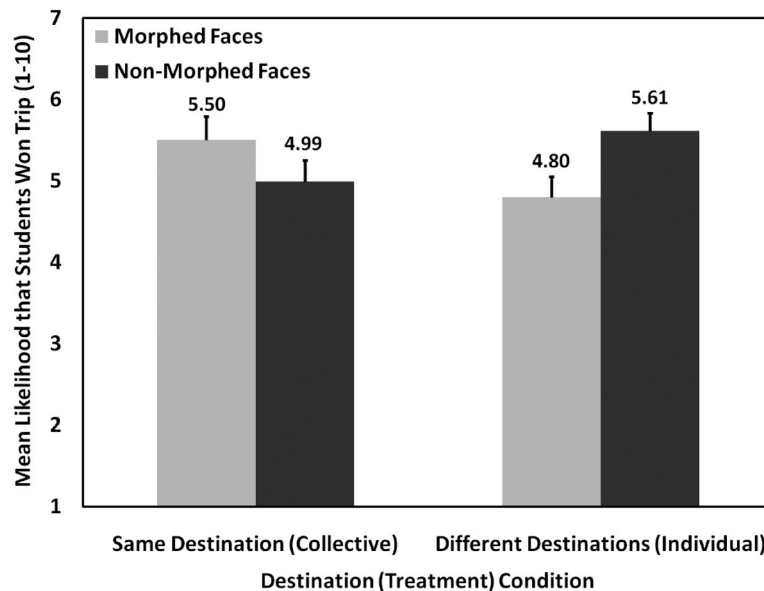


Figure 5. Likelihood that morphed and nonmorphed students won a trip as a function of whether the six students won a trip to the same destination or to six different destinations in Study 4. Bars represent standard error.

rather than the students without the morphed faces (heterogeneous targets) to have won the vacations. Second, we examined the correlation between this difference score and participants' endorsement of the earlier statement in each condition of the experiment.

We found the expected interaction between participants' endorsement of the statement, whether the students went to the same or different destinations, and whether participants expected a greater proportion of morphed or nonmorphed student to win the trip, $\beta = .26$, $t(88) = 2.54$, $p < .02$. As expected, participants who were told that the students vacationed in the same town believed that a higher proportion of the winning students were similar to the target (i.e., were morphed with the target) when they more strongly endorsed the statement that similar-looking people tend to experience similar outcomes in life, $\beta = .33$, $t(40) = 2.22$, $p < .04$. In contrast, participants who were told that the students vacationed in six different towns believed that a higher proportion of the winning students were similar to the target when they less strongly endorsed the statement that similar-looking people tend to experience similar outcomes in life, $\beta = -.37$, $t(46) = -2.66$, $p < .02$. These results suggest that participants in both conditions responded according to the naive theories they held about the relationship between physical appearance and life outcomes. Participants who more strongly perceived a general relationship between homogeneity and collective treatment believed that the collectively treated students were homogeneous and the individually treated students were heterogeneous.

Summary of Studies 1–4

The first four studies in this paper suggest that more homogeneous groups attract collective treatment (Study 1) and that collectively treated groups are in turn perceived as more homogeneous (Studies 2–4). Study 4 provided preliminary evidence that these effects are more pronounced among people who generally believe that more homogeneous groups will attract more collective treatment, and we examined further this naive theory-based mechanism for the effects in Study 5. The effects in Studies 1–4 may have arisen because people have come to form an ecologically driven association between homogeneity and collective treatment. One pertinent example is race: Across time, people from similar racial backgrounds have received similar treatment, whereas people on different sides of arbitrary racial boundaries have tended to experience quite different forms of treatment (e.g., Eberhardt et al., 2006; Fiske, 1998). Accordingly, it is possible that people have come to form an association between homogeneity and collective treatment across time. This might explain why they assume that collectively treated groups are more homogeneous, even in contexts where that assumption is not strictly logical (e.g., in Study 3, depending on whether all the members of a group have benefited from the discovery of oil on their land or suffered the loss of their apartments in a tornado). Accordingly, in Study 5 we examined participants' lay assumptions to show that people generally expect homogeneous groups to attract collective treatment from a government agency, whereas they expect the government agency to treat members of more heterogeneous groups more individually.

Study 5: Naive Associations Between Research Laboratory Homogeneity and Funding

In Study 5, we sought to show that people intuitively expect more homogeneous lab groups to receive relatively more collective forms of government funding and more heterogeneous groups to receive relatively more individualized forms of funding. Participants viewed photographs of science lab groups and rated the likelihood that each group would attract group funding (funding for the lab at large) or individualized funding (funding according to the performance of each individual in the lab). They also rated the uniformity of each group's members. Because group funding is a form of collective treatment and individual funding is a form of individual treatment, we expected participants to believe that more homogeneous labs would attract relatively more group funding than would heterogeneous labs (and vice versa when the funding was individualized).

Method

Photograph Selection

From a Google Images search for "lab group," we selected the first 16 photos of science labs that contained 10–14 clearly discernible members.

Photograph Ratings

Participants. One hundred Princeton University undergraduates (57 women; $M_{\text{age}} = 19.25$ years, $SD = 1.14$) participated in this study in partial fulfillment of a course requirement.

Group homogeneity ratings. Twenty participants rated the 16 lab groups according to how homogeneous their members appeared to be to each other based on the accompanying picture (1 = *members are not at all similar to each other* to 7 = *members are very similar to each other*). These ratings functioned as a proxy for group homogeneity, the independent measure in this study.

Collective funding ratings. A second sample of 20 participants rated the 16 lab groups according to how likely they were to receive group funding from the government (1 = *not at all likely* to 7 = *very likely*). The questionnaire explained that this form of funding was allocated to the lab as a whole rather than calculated according to the contributions of each lab member separately. These ratings formed the first dependent measure, which functioned as a proxy for the extent to which each group attracted a collective treatment.

Individual funding ratings. A third sample of 20 participants rated the 16 lab groups according to the likelihood that they would attract individualized funding from the government (1 = *not at all likely* to 7 = *very likely*). The questionnaire explained that individualized funding was calculated according to the contributions of each lab member separately. These ratings formed the second dependent measure, a proxy for the extent to which participants believed each group would attract individualized treatment.

Competence and photograph clarity covariate ratings. Two other samples of 20 participants each rated the 16 lab groups according to how competent they seemed (1 = *not at all competent* to 7 = *very competent*) and how clear the photos were (1 = *not at all clear* to 7 = *very clear*). These ratings functioned as covariates, as it was plausible that more competent lab groups or those depicted in clearer

photos might have attracted different forms of funding independently of their homogeneity. For example, lab groups depicted in less clear photographs might have seemed more homogeneous and also less worthy of funding in general, and more competent lab groups might have generally attracted more funding of one or both types. Regardless, to rule out these effects, we included these ratings as covariates in the analyses examining the effects of group homogeneity on collective and individualized funding.

Results and Discussion

Preliminary Analyses

We began by calculating a mean rating for each of the 16 lab groups on each of the five measures by averaging the ratings from the 20 participants. Thus, each lab group was assigned a homogeneity rating, a collective funding rating, an individual funding rating, a competence rating, and a photo clarity rating.

We also ruled out several potentially confounding factors before commencing the primary analyses. First, we found no significant association between the number of racial or ethnic groups represented in the photographs and their homogeneity ratings and, similarly, no significant relationship between the number of lab members in each photograph and their collective funding, individual funding, or homogeneity ratings ($ps > .15$).

Primary Analyses

We conducted all analyses at the item (or lab group) level, averaging participants' ratings of each of the 16 labs and examining the relationships between those averaged ratings in our analyses. Table 1 contains the zero-order correlations between the five variables of interest.

To demonstrate the divergent effect of lab group homogeneity on collective and individual funding predictions, we conducted the within-subjects moderation analysis suggested by Judd, Kenny, and McClelland (2001). The conditions for moderation are satisfied when the independent measure significantly predicts scores that represent the difference between the two dependent measures (collective funding and individual funding). Accordingly, we calculated a funding difference index score for each lab group by subtracting the group's mean individual funding rating from its mean collective funding rating. A hierarchical regression analysis, controlling for the effect of photo clarity and lab-group competence ratings, showed that lab-group homogeneity significantly moderated the effect of funding type, $\beta = .59$, $\Delta R^2 = .32$, $\Delta F(1, 12) = 6.29$, $p < .03$.²

To examine the simple effects of group homogeneity on collective funding and individual funding, we examined the relationship between lab-group homogeneity and each funding type separately. Participants anticipated significantly greater collective funding for more homogeneous groups, $\beta = .65$, $t(12) = 3.00$, $p = .01$, and marginally (but not significantly) greater individual funding for more heterogeneous groups, $\beta = -.41$, $t(12) = -1.59$, $p < .14$.³

In sum, despite the relatively small sample size of 16 lab groups in this study, participants appeared to anticipate greater collective funding for the more homogeneous groups and greater individual funding for the more heterogeneous groups. Moreover, this effect was independent of the apparent competence of the lab groups and the relative

clarity of the photograph depicting those groups. This study therefore suggests that people naively assume that homogeneous groups will receive relatively collective treatment and that heterogeneous groups will receive relatively individualized treatment.

Studies 1–5 establish a bidirectional link between homogeneity and collective treatment and suggest one potential mechanism for that link. We chose to conduct controlled laboratory studies in the first five experiments for several reasons. First, we were able to isolate the mechanisms responsible for the observed effects, while controlling extraneous factors. Second, by manipulating one variable (perceived group homogeneity in Study 1 and treatment collectivity in Studies 2–4) and measuring the other (treatment collectivity in Study 1 and perceived group homogeneity in Studies 2–4), we were able to show that the relationship between perceived group homogeneity and treatment collectivity is bidirectional. Having shown these effects in lab studies, we sought to identify similar effects in the real world. In Studies 6a and 6b, we sought to show that consequential policy decisions appear to reflect the same relationship between group homogeneity and treatment collectivity. It is important to note that these field studies use real data sets, and although we attempted to control for as many confounds as possible, it is impossible to exert the same degree of experimental control as we attempted in the lab studies. We believe this loss of experimental control is more than justified by the benefits of being able to show that the effect persists in the real world, beyond the confines of the lab. Thus, in Studies 6a and 6b, we sought evidence that the relationship between homogeneity and collective treatment in our lab studies also influenced the government policy decisions recorded in two archival data sets.

Study 6a: Homogeneous Nations Attract More U.S. Disaster Aid

When U.S. government policymakers send aid to a struggling nation in the wake of a natural disaster, they face an imposing decision. They might assign aid based on whether the country is poverty stricken, how many people are affected by the disaster, and the likelihood that the aid will benefit the nation's citizens, among other issues. However, decisions such as these might be influenced by other, unexpected criteria. For example, the first five studies in this paper suggest that a policymaker who knows how homogeneous a nation's population is relative to the population of other nations might

² The same interaction between lab-group homogeneity and funding type held when we did not control for competence and photo clarity ratings, $\beta = .51$, $t(14) = 2.19$, $p < .05$.

³ The extent to which lab members vary in age is one form of homogeneity that may be confounded with experience and, therefore, the extent to which some members deserve more or less funding than others. Accordingly, we asked three blind raters to estimate the age of each lab member in each of the 16 labs and calculated the standard deviation of the ages within each lab as a measure of age-related homogeneity. If anything, partialing out the effects of age-related homogeneity strengthened our effect, and this fact suggests that it was not solely driven by this potentially confounding cue. In particular, lab-group homogeneity significantly moderated the effect of funding type, $\beta = .82$, $\Delta R^2 = .51$, $\Delta F(1, 11) = 14.13$, $p < .01$. Follow-up simple effects analyses showed that participants anticipated significantly greater collective funding for more homogeneous groups, $\beta = .80$, $t(11) = 3.43$, $p < .01$, and significantly greater individual funding for more heterogeneous groups, $\beta = -.50$, $t(11) = -2.45$, $p < .04$.

Table 1
Zero-Order Correlations Between Independent Variables, Dependent Variables, and Control Variables Included in Study 5

Variable	Collective funding (DV1)	Individual funding (DV2)	Photo clarity (control)	Group competence (control)
Group homogeneity (IV)	.50*	-.35**	.29	.05
Collective funding (DV1)	—	-.53*	-.02	-.09
Individual funding (DV2)		—	.33	.42**
Photo clarity (control)			—	.23

Note. $n = 16$. DV = dependent variable; IV = independent variable.

* $p < .05$. ** $p < .10$, one-tailed.

unintentionally allow that information to contaminate the assessment of how much aid that nation deserves. Of course this is not the only factor that influences aid decisions, but it may emerge as a significant criterion when aid decisions and demographic homogeneity are compared. Accordingly, we examined the relationship between the homogeneity of numerous countries and the amount of aid they attracted from the U.S. government.⁴

Method

Disaster Aid Data Collection

We collected U.S. disaster aid figures from U.S. Agency for International Development (USAID) annual reports from 2005 to 2007. The USAID disaster aid reports list the donations by U.S. government agencies to foreign nations that have experienced natural or anthropogenic disasters. Although hundreds of such disasters occur each year, we focused on the costliest disasters because they attracted over 80% of the total aid and tended to affect similarly less-developed countries. The remaining aid was widely spread among highly developed countries that differed from the less-developed countries on many dimensions that may have confounded the relationship between population homogeneity and aid. This approach yielded 18 disasters. The reports listed the number of individuals affected, each country's Human Development Index (HDI) score (a composite of educational, economic, and health indicators), and the U.S. government aid contribution for each disaster.

Homogeneity Data Collection

We retrieved various demographic characteristics for the country affected by each disaster from the online CIA World Factbook (www.cia.gov/cia/publications/factbook/). In particular, we recorded the number of religions, ethnicities, and languages represented and the proportion of the population falling within each category.

To quantify each country's homogeneity, we adopted Blau's (1977) index of heterogeneity. A popular measure among sociologists, Blau's index is designed to quantify demographic homogeneity according to the following formula: heterogeneity index = $1 - (\sum P_i^2)$, where P is the proportion of the population represented by the i th category within that demographic. Higher index scores represent greater heterogeneity within the population for the demographic of interest. We averaged the language, ethnicity, and religion heterogeneity indices to form a single heterogeneity score for each country.

Results and Discussion

As expected, heterogeneous countries attracted less aid than did more homogeneous countries, $r(16) = -.68$, $p = .002$. This result also held when we partialled out the effect of the number of affected individuals per country and the country's HDI score, $r(14) = -.52$, $p = .04$ (see scatter plot in Figure 6).⁵

These results provide preliminary evidence that homogeneous countries attract greater government aid. However, several features of Study 6a prompted us to replicate the analysis in Study 6b, this time examining the relationship between the homogeneity of a nation and its propensity to attract general foreign aid (vs. disaster aid in Study 6a). In particular, Study 6b addressed several limitations in Study 6a. First, to show that this effect was not limited to the small sample of nations in Study 6a, we investigated a larger sample of 50 recipient nations in Study 6b. It contained 43 countries that were not included in Study 6a and thus was quite different. Second, with the larger sample we were able to eliminate a range of potential confounds that we discuss below.

Study 6b: Homogeneous Nations Attract U.S. Financial Aid

Method

Foreign Aid and Homogeneity Data Collection

We collected U.S. foreign aid figures from USAID reports (www.usaid.gov/). In our analysis we extracted aid obligations for the

⁴ It is important to note that we operationalized collective treatment as funding sent to an entire country. Rather than individual treatment, the alternative to this form of collective treatment is no treatment at all. This definition distinguished this study from the others in which we contrasted collective and individual treatment.

⁵ We conducted a multivariate outlier analysis (Mahalanobis distance analysis) to determine whether there were any outlying data points that may have skewed our results. In neither this nor the following study was there a single outlier (Mahalanobis scores < 3). However, as the data were positively-skewed, we conducted all analyses using untransformed and log-transformed data. The results held for both the log-transformed and the untransformed data, but the results we present here use the log-transformed data. Moreover, all results held both for aid per capita and for absolute aid per country and also when we controlled for the size of each country. This rules out the possibility that these effects were driven by population or country size.

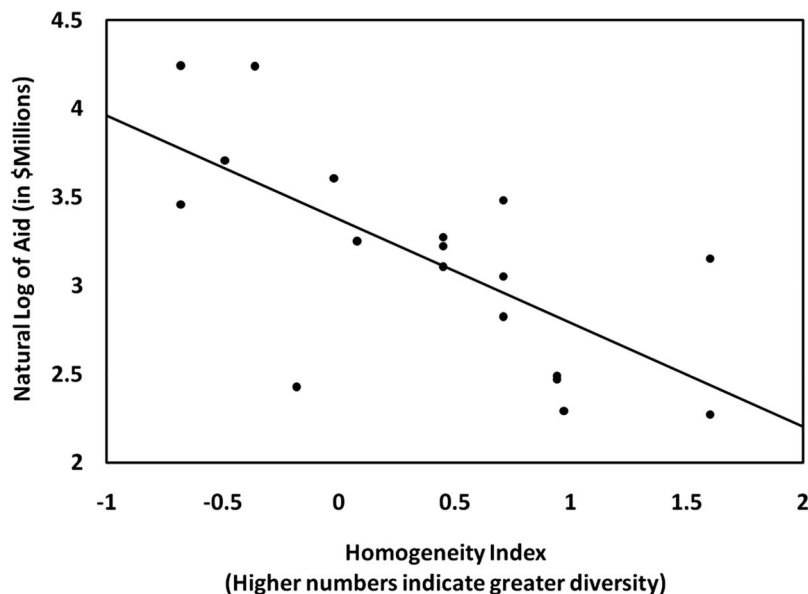


Figure 6. A scatter plot of the relationship between each nation's homogeneity and the natural log of the charity aid that nation received (\$ millions) in Study 6a.

top 50 recipient nations for the years 2002 and 2003.⁶ As in Study 6a, we used CIA World Factbook data to compile the same index of heterogeneity (a composite of linguistic, ethnic, and religious diversity).

Eliminating Alternative Explanations

Many factors potentially influence both the demographic homogeneity of a country and its propensity to attract foreign aid. Accordingly, we controlled several factors in measuring the correlation between demographic heterogeneity and aid. Many of these factors were suggested in a similar analysis by Apodaca and Stohl (1999), who analyzed the relationship between a country's human rights record and that country's ability to attract U.S. aid, but we also included factors that were not included in Apodaca and Stohl's analyses. From the Penn World Tables (<http://pwt.econ.upenn.edu>), we compiled each country's level of democracy, corruption, and openness to trade with the U.S.A. (a proxy for the quality of its relationship with the United States), and from the CIA World Factbook we recorded each country's HDI. We controlled for the effects of all these variables in our model to eliminate the possibility that they alone explained the relationship between demographic homogeneity and tendency to attract U.S. foreign aid.

Results and Discussion

The various data sets failed to include demographic data for four countries, so these four countries were not included in the remaining analyses.

We began by calculating the simple correlation between heterogeneity scores and the amount of U.S. foreign aid donated to each country.⁷ As we expected, the less diverse the country, the more aid it attracted, $r(44) = -.33, p = .025$. The same result held when we controlled for the effect of the numerous covariates, $r(38) = -.36, p = .024$ (see Table 2 for zero-order correlations; see scatter plot in Figure 7).⁸

These results replicate those in Study 6a, showing that the U.S. government donates greater foreign aid to more homogeneous nations. In addition, we controlled for a number of alternative explanations that were difficult to eliminate, given the small sample of recipient countries in Study 6a. Studies 6a and 6b therefore replicated the controlled lab results in Study 1, and this suggested that the link between group homogeneity and collective treatment may have important and unexpected consequences for real-world decision making.

General Discussion

In seven studies, we found that people perceived a relationship between a group's homogeneity and that group's tendency to

⁶ The report contained data for 185 nations, but many of those nations received no aid and the top 50 received more than 90% of all aid. We also limited our analysis to the top 50 nations, as many other factors might differ between countries that receive substantial aid and those that receive relatively little aid. Nonetheless, analyses that included all 185 nations produced similar results.

⁷ As in Study 6a, the distribution of aid was positively skewed, so we also calculated all analyses using log-transformed aid figures. Again, the results were identical, so we present the log-transformed data to maintain consistency with the approach we adopted in Study 6a.

⁸ Because the effect should be driven by the perception of homogeneity, rather than each country's true homogeneity, we sought to confirm that a person with expertise akin to a government policymaker was aware of these homogeneity statistics. We asked a "population and demography" sociology graduate student at Princeton University to estimate the homogeneity of various characteristics of 20 countries. Her ratings correlated highly with the actual homogeneity statistics, $r(18) = .74, p = .0001$. As was consistent with the objective homogeneity indices, the more diverse the student believed the country to be, the less aid it received, $r(18) = -.60, p = .005$.

Table 2
Zero-Order Correlations Between All Variables Included in Study 6b

Variable	Total foreign aid	Democracy score	Openness to trade	Corruption score	HDI	Foreign debt	Population of nation
Homogeneity index	−.33*	.02	.11	−.02	−.39**	−.09	.04
Total foreign aid	—	.22	.21	−.22	.21	.45**	.15
Democracy score		—	.25	−.31*	.39**	.21	.21
Openness to trade			—	.02	−.14	.19	.29
Corruption score				—	−.28	−.08	.12
HDI					—	.40**	.01
Foreign debt						—	.40**

Note. $n = 46$. Bold font denotes critical correlation between homogeneity and total foreign aid. All other variables were covaried out of the model to eliminate the possibility that they, alone, contributed to the relationship between homogeneity and aid. HDI = Human Development Index.

* $p < .05$. ** $p < .01$.

experience collective treatment. This relationship was bidirectional, whereby homogeneous groups attracted collective treatment (Studies 1, 6a, and 6b) and collectively treated groups thereafter appeared more homogeneous (Studies 2–4). Moreover, these effects arose in part because people naively associate homogeneity with collective treatment (Study 5), and the relationship between homogeneity and collective treatment is more robust the more strongly people endorse this naive theory (Study 4).

Implications of the Bidirectional Collective Treatment Model

Theoretical Implications

This model illuminates a new psychological route by which groups are created and their identities are strengthened and perpetuated. Treating a sample of people uniformly leads them to be perceived as more homogeneous, and this strengthens and perpetuates their identity as a group worthy of attracting collective

treatment. These results suggest that an entity forms not just when the entity's members move together spatially but also when they share a common fate (cf. Ip et al., 2006).

Among other theoretical implications, these findings suggest one reason why group stereotypes are remarkably stubborn and often persist in the face of disconfirming evidence (e.g., Crawford, Sherman, & Hamilton, 2002; Eiser, Fazio, Stafford, & Prescott, 2003; Fazio & Olson, 2003; Fiske, 1998; Garcia-Marques & Mackie, 1999; Hewstone, Hopkins, & Routh, 1992; Johnston & Coolen, 1995; Moreno & Bodenhausen, 1999; Kunda, 1990; Yzerbyt et al., 2004). Recognizing the insidiousness of stereotypes and the racial divisions they sometimes foment, government and judicial institutions have come to ascribe many societal problems to group-based stereotyping. Their attempts to eliminate stereotypes may reinforce those stereotypes in the process of attempting to eradicate them. For example, the U.S. Supreme Court continues to support the application of affirmative action enrollment policies. Delivering the majority judgment in *Grutter v. Bollinger* (2003),

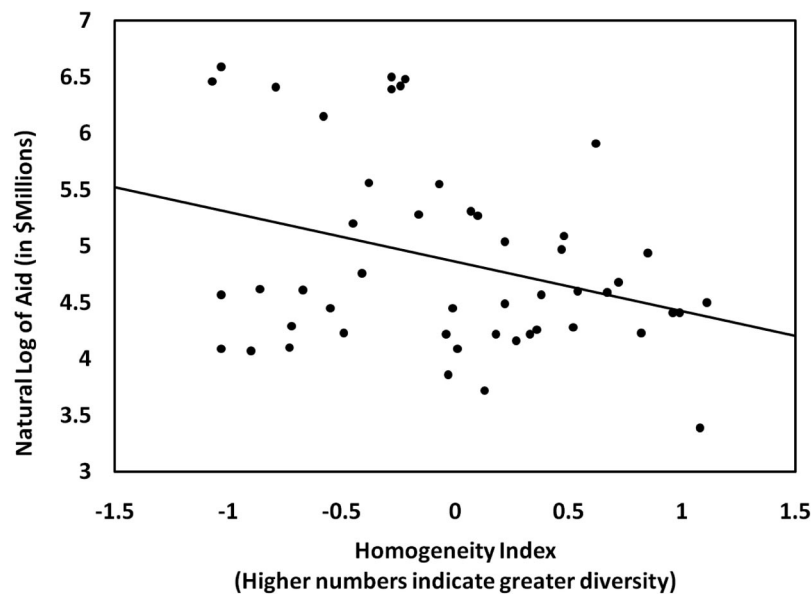


Figure 7. A scatter plot of the relationship between each nation's homogeneity and the natural log of the financial aid that nation received (\$ millions) in Study 6b.

Justice Sandra Day O'Connor noted, "[Although we support affirmative action policies] the Court expects that 25 years from now, the use of racial preferences will no longer be necessary" (p. 343). This example illustrates an institutional tendency to correct imbalances by using collective treatment to reward a historically subjugated group. Our findings suggest that, despite the benefits of affirmative action policies, one unintended effect of such forms of collective treatment is to accentuate the division between the minority and majority groups.

One related and nontrivial question is whether people who behave according to the model are behaving rationally or normatively. Sometimes, perhaps, homogeneous people should be treated collectively (e.g., testing people from one particular ethnic group for a disease that selectively affects that group), but racial profiling and prejudice represent the destructive flip side of the same coin. The same outcome probably applies to the second link in the model. Collectively treated people may be homogeneous on certain dimensions (e.g., people who are selectively tested for a disease that affects one ethnic group might be similar in appearance), but there is no reason to believe that people who are affected by chance events (e.g., winning the lottery or losing their houses in a tornado) are similar in other respects. Sometimes, therefore, people behave rationally when they associate homogeneity and collective treatment, though, as several studies in this paper show, they tend to inappropriately assume the relationship applies in situations where there is no logical link between homogeneity and collective treatment.

Practical Implications

The practical implications of this model are best illustrated in Studies 6a and 6b, which suggest that homogeneity might affect collective economic decisions on a national scale. If we focus on positive treatment, nations that seem homogeneous appear to attract greater foreign and disaster aid, which might in turn reinforce perceptions that they are demographically homogeneous. However, negative collective treatment might have similar consequences. Weiner (1997) has argued that Japanese citizens are considerably more diverse than many laypeople believe, and our model suggests that inflated perceptions of Japan's homogeneity may have facilitated postwar sanctions. In turn, these sanctions may have perpetuated the notion that Japanese society is homogeneous.

However, these macro-level effects should not overshadow the application of our model to smaller scale intergroup interactions. People belong to myriad groups that might attract levels of collective treatment that accord with their perceived degree of homogeneity. Minority groups are often perceived as more homogeneous than their apparently diverse majority counterparts (Yzerbyt et al., 2004). For example, women in academia are perceived as considerably more homogeneous than their male counterparts, which might in part explain why female academics generally experience greater group-based treatment (Brown & Smith, 1989). Indeed, Brown and Smith found that even female academics believe they are more homogeneous than are male academics.

In sum, our model proposes a novel explanation for the perpetuation of collective treatment and group distinctiveness. We suggest that groups are distinguishable in part by the different forms of treatment their members attract, so even well-intentioned re-

wards have the ironic consequence of reaffirming group divisions. These findings suggest several avenues of future research that we discuss in turn below.

Limitations and Future Directions

In this paper we construe homogeneity quite narrowly, as physical similarity. We specifically focus on physical homogeneity, because physical features are often the most salient forms of information available to perceivers, yet in the wake of a decades-long backlash against prejudice, people prefer to believe that their decisions are not based on superficial physical characteristics (cf. Devine, 1989; Fiske, 1998). Accordingly, manipulations and dependent measures that tapped into physical homogeneity have provided a conservative test of our hypotheses. Nonetheless, because recent research has paid great attention to disentangling the components of homogeneity, entitativity, and essence (e.g., Hamilton, 2007; Hamilton & Sherman, 1996), future research might examine whether these effects generalize beyond physical forms of similarity and whether physically homogeneous groups come to take on the properties of entities over time. For example, groups whose members share a common goal might similarly appear more homogeneous and attract more collective treatment than groups whose members appear to hold disparate goals (Ip et al., 2006).

Our proposed model is not the first to suggest that social psychological processes are bidirectional. Similar bidirectional models have been used to explain expectancy confirmation (Darley & Gross, 1983), the relationship between familiarity and liking (Monin, 2003; Zajonc, 1980), the relationship between scarcity and valuation (Lynn, 1992; Dai, Wertenbroch, & Brendl, 2008), and self-fulfilling prophecies in the classroom (e.g., Rosenthal & Jacobson, 1968). Such models are useful because they propose mechanisms that explain why certain processes continue doggedly once they begin. The present model suggests a possible mechanism for stereotype perpetuation and argues that some attempts to eradicate stereotype-based group differences ironically reinforce those differences. The lab studies in this paper establish a mechanism by which stereotype perpetuation might occur, but they do not directly examine the process in a stereotype-maintenance context. If the model were applied to stereotyping more directly, future research might examine how policies could be designed to avoid reinforcing group divisions while attempting to attenuate prejudice.

Although this paper examined the two links in our model separately, there is good reason to believe that they feed into one another. Because Studies 1, 6a, and 6b suggested that homogeneous groups attract more collective treatment and Studies 2–4 suggested that collectively treated individuals thereafter seem more homogeneous, it seems plausible that the two links in the model are iterative. These studies in concert suggest that as a group attracts collective treatment across time, it is likely to appear increasingly homogeneous, which should in turn increase its tendency to attract collective treatment. Future research might, however, directly test the possibility that the model is indeed iterative.

Our model also suggests ironic effects that might extend to antiprejudice and antistereotyping education. Considerable evidence over the past 20 years has shown that children learn to stereotype as soon as they recognize basic perceptual categories (Gelman, 2003; Mervis & Rosch, 1981). They subsequently adopt a host of category-based stereotypes, including gender (Cowan &

Hoffman, 1986), national (Rutland, 1999), and racial stereotypes (Aboud, 2005; Bigler & Liben, 2007). The iterative nature of group treatment and group essentialism might explain, in part, why children adopt category boundaries that cannot be explained by basic perceptual processes alone. For example, children might perceive that girls are treated differently from boys, poor children are treated differently from rich children, and black people are treated differently from white people. These processes—many of which occur in the service of educating children on the principles of equality—are ironically also responsible for promoting stereotypes. Future research might examine the effectiveness of educational alternatives that preserve the integrity of distinct groups (e.g., ethnic groups) while emphasizing the primacy of superordinate groups (e.g., humanity at large). Indeed, the model suggests that treating all humans collectively may be the best approach to emphasizing their homogeneity across racial, ethnic, linguistic, and national boundaries.

Although emphasizing the superordinate group might generally diminish the model's insidious effects, a related question is whether certain individual differences might exacerbate these effects. For example, people with low need for cognition scores (Cacioppo & Petty, 1982) may be less likely to individuate the members of an apparent group and may thus more readily perceive those individuals as a homogeneous unit. Similarly, people with a high score on the belief in a just world construct (e.g., Lerner, 1980; Lerner & Simmons, 1966) might more willingly believe that a collectively treated group of individuals deserve to be treated as a single entity. These individuals might therefore show a particularly strong tendency to conflate an instance of collective treatment with the homogeneity of the group's members on other dimensions.

Conclusion

Group formation is an inevitable consequence of social life, and the model supported by our research explains in part why group identities are so resistant to change. Once a sample of individuals is perceived as a group, these individuals attract collective treatment, which strengthens the group's identity. Returning to the Supreme Court's continued support of affirmative action, there is good reason to believe that affirmative action policies do not entirely benefit minority students in the long term. Students who accept a form of treatment reserved for those experiencing long-term disadvantages may be consigned to a group both distinct from and perceived as inferior to the majority.

References

- Abelson, R. P., Dasgupta, N., Park, J., & Banaji, M. R. (1998). Perceptions of the collective other. *Personality and Social Psychology Review*, 2, 243–250.
- Aboud, F. E. (2005). The development of prejudice in childhood and adolescence. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 310–326). New York, NY: Blackwell.
- Alter, A. L., Kernochan, J., & Darley, J. M. (2007). Transgression wrongfulness outweighs its harmfulness as a determinant of sentencing severity. *Law and Human Behavior*, 31, 319–335.
- Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, 136, 569–576.
- Apodaca, C., & Stohl, M. (1999). United States human rights policy and foreign assistance. *International Studies Quarterly*, 43, 185–198.
- Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science*, 16, 162–166.
- Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure*. New York, NY: Free Press.
- Brewer, M. B. (1988). A dual-process model of impression formation. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.
- Brewer, M. B., & Harasty, A. S. (1996). Seeing groups as entities: The role of perceiver motivation. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition* (Vol. 3, pp. 347–370). New York, NY: Guilford Press.
- Brewer, M. B., Hong, Y.-Y., & Li, Q. (2004). Dynamic entitativity: Perceiving groups as actors. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity, and essentialism* (pp. 25–38). New York, NY: Psychology Press.
- Brewer, M. B., Weber, J. G., & Carini, B. (1995). Person memory in intergroup contexts: Categorization versus individuation. *Journal of Personality and Social Psychology*, 69, 29–40.
- Brin  ol, P., Petty, R. E., & Tormala, Z. L. (2006). The malleable meaning of subjective ease. *Psychological Science*, 17, 200–206.
- Brown, R., & Smith, A. (1989). Perceptions of and by minority groups: The case of women in academic. *European Journal of Social Psychology*, 19, 61–75.
- Browning, C. R. (1992). *Ordinary men: Reserve Police Battalion 101 and the Final Solution in Poland*. New York, NY: Harper Collins.
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116–131.
- Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, 3, 14–25.
- Carter, W. A. (1944). Nicknames and minority groups. *Phylon*, 5, 241–245.
- Corrigan, P. W. (2007). How clinical diagnosis might exacerbate the stigma of mental illness. *Social Work*, 52, 31–39.
- Cowan, G., & Hoffman, C. D. (1986). Gender stereotyping in young children: Evidence to support a concept-learning approach. *Sex Roles*, 14, 211–224.
- Crawford, M. T., Sherman, S. J., & Hamilton, D. L. (2002). Perceived entitativity, stereotype formation, and the interchangeability of group members. *Journal of Personality and Social Psychology*, 83, 1076–1094.
- Dai, X., Wertenbroch, K., & Brendl, M. C. (2008). The value heuristic in judgments of relative frequency. *Psychological Science*, 19, 18–19.
- Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, 44, 20–33.
- 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, 991–1003.
- Denson, T. F., Lickel, B., Curtis, M., Stenstrom, D. M., & Ames, D. R. (2006). The roles of entitativity and essentiality in judgments of collective responsibility. *Group Processes and Intergroup Relations*, 9, 43–61.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5–18.
- Dion, K. (1972). Physical attractiveness and evaluations of children's

- transgressions. *Journal of Personality and Social Psychology*, 24, 207–213.
- Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, 24, 285–290.
- Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science*, 17, 383–386.
- Eiser, J. R., Fazio, R. H., Stafford, T., & Prescott, T. J. (2003). Connectionist simulation of attitude learning: Asymmetries in the acquisition of positive and negative evaluations. *Personality and Social Psychology Bulletin*, 29, 1221–1235.
- Erikson, M., & Gibbs, J. (1979). On the perceived severity of legal penalties. *Journal of Criminal Law and Criminology*, 70, 102–116.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology*, 54, 297–327.
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, 4th ed., pp. 357–411). Boston, MA: McGraw-Hill.
- Fiske, S. T., & Neuberg, S. L. (1990). The continuum of impression formation from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). New York, NY: Academic Press.
- Forgas, J. P. (1992). On bad mood and peculiar people: Affect and person typicality in impression formation. *Journal of Personality and Social Psychology*, 62, 863–875.
- Forgas, J. P. (1993). On making sense of odd couples: Mood effects on the perception of mismatched relationships. *Personality and Social Psychology Bulletin*, 19, 59–71.
- Frieze, I. H., Olson, J. E., & Russell, J. (1991). Attractiveness and income for men and women in management. *Journal of Applied Social Psychology*, 21, 1037–1039.
- Gaertner, L., & Schopler, J. (1998). Perceived ingroup entitativity and intergroup bias: An interconnection of self and others. *European Journal of Social Psychology*, 28, 963–980.
- Garcia-Marques, L., & Mackie, D. M. (1999). The impact of stereotype-incongruent information on perceived group variability and stereotype change. *Journal of Personality and Social Psychology*, 77, 979–990.
- Gelman, S. A. (2003). *The essential child*. New York, NY: Oxford University Press.
- Gescheider, G., Catlin, E., & Fontana, A. (1982). Psychophysical measurement of the judged seriousness of crimes and severity of punishments. *Bulletin of the Psychonomic Society*, 19, 275–278.
- Glover, J. (2001). *Humanity: A moral history of the 20th century*. New Haven, CT: Yale University Press.
- Grutter v. Bollinger, 539 U.S. 306 (2003).
- Hamilton, D. L. (2007). Understanding the complexities of group perception: Broadening the domain. *European Journal of Social Psychology*, 37, 1077–1101.
- Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological Review*, 103, 336–355.
- Haslam, N. (1998). Natural kinds, human kinds, and essentialism. *Social Research*, 65, 291–314.
- Hewstone, M., Hopkins, N., & Routh, D. A. (1992). Cognitive models of stereotype change: (1) Generalization and subtyping in young people's view of the police. *European Journal of Social Psychology*, 22, 219–234.
- Hui, C. H., Triandis, H. C., & Yee, C. (1991). Cultural differences in reward allocation: Is collectivism the explanation? *British Journal of Social Psychology*, 30, 145–157.
- Insko, C. A., Hoyle, R. H., Pinkley, R. L., Hong, G., Slim, R. M., Dalton, B., . . . & Schopler, J. (1988). Individual-group discontinuity: The role of consensus rule. *Journal of Experimental Social Psychology*, 24, 505–519.
- Ip, G. W.-M., Chiu, C.-Y., & Wan, C. (2006). Birds of a feather and birds flocking together: Physical versus behavioral cues may lead to trait-versus goal-based group perception. *Journal of Personality and Social Psychology*, 90, 368–381.
- Johnston, L., & Coolen, P. (1995). A dual-process approach to stereotype change. *Personality and Social Psychology Bulletin*, 21, 660–673.
- Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. *Psychological Methods*, 6, 115–134.
- Kashima, Y. (2004). Culture, communication, and entitativity: A social psychological investigation of social reality. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity and essentialism* (pp. 257–273). Philadelphia, PA: Psychology Press.
- Kolber, A. J. (2009). The subjective experience of punishment. *Columbia Law Review*, 109, 182–236.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480–498.
- Lambert, A. J., Barton, L. M., Lickel, B., & Wells, J. (1998). The influence of group variability and processing goals on the ease of making judgments about social categories. *Personality and Social Psychology Bulletin*, 24, 807–820.
- Lerner, M. (1980). *The belief in a just world*. New York, NY: Plenum Press.
- Lerner, M., & Simmons, C. H. (1966). Observers' reactions to the "innocent victim": Compassion or rejection? *Journal of Personality and Social Psychology*, 4, 203–210.
- Leung, K., & Bond, M. H. (1982). How Chinese and Americans reward task-related contributions: A preliminary study. *Psychologia*, 25, 32–39.
- Levy, S. R., Freitas, A. L., & Salovey, P. (2002). Construing action abstractly and blurring social distinctions: Implications for perceiving homogeneity among, but also empathizing with and helping, others. *Journal of Personality and Social Psychology*, 5, 1224–1238.
- Lickel, B., Hamilton, D. L., & Sherman, S. J. (2001). Elements of a lay theory of groups: Types of groups, relational styles, and the perception of group entitativity. *Personality and Social Psychology Review*, 5, 129–140.
- Lickel, B., Hamilton, D. L., Wierzchowska, 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.
- Linville, P. W., Fischer, G. W., & Yoon, C. (1996). Perceived covariation among the features of ingroup and outgroup members: The outgroup covariation effect. *Journal of Personality and Social Psychology*, 70, 421–436.
- Luttmer, E. F. P. (2001). Group loyalty and the taste for redistribution. *Journal of Political Economy*, 109, 500–528.
- Lynn, M. (1992). Scarcity's enhancement of desirability: The role of naive economic theories. *Basic and Applied Social Psychology*, 13, 67–78.
- McConnell, A. R., Sherman, S. J., & Hamilton, D. L. (1997). Target entitativity: Implications for information processing about individual and group targets. *Journal of Personality and Social Psychology*, 72, 750–762.
- Mervis, C., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, 32, 89–113.
- Monin, B. (2003). The warm glow heuristic: When liking leads to familiarity. *Journal of Personality and Social Psychology*, 85, 1035–1048.
- Moreno, K. N., & Bodenhausen, G. V. (1999). Resisting stereotype change: The role of motivational capacity in defending social beliefs. *Group Processes and Intergroup Relations*, 2, 5–16.
- Morris, M. W., Menon, T., & Ames, D. R. (2001). Culturally conferred conceptions of agency: A key to social perception of persons, groups,

- and other actors. *Personality and Social Psychology Review*, 5, 169–182.
- Norton, M. I., Vandello, J. A., & Darley, J. M. (2004). Casuistry and social category bias. *Journal of Personality and Social Psychology*, 87, 817–831.
- Olson, K. R., Banaji, M. R., Dweck, C. S., & Spelke, E. S. (2006). Children's biased evaluations of lucky versus unlucky people and their social groups. *Psychological Science*, 17, 845–846.
- Park, B., & Hastie, R. (1987). Perception of variability in category development: Instance- versus abstraction-based stereotypes. *Journal of Personality and Social Psychology*, 53, 621–635.
- Prentice, D. A., & Miller, D. T. (2006). Essentializing differences between women and men. *Psychological Science*, 17, 129–135.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York, NY: Rinehart & Winston.
- Rothbart, M., & Park, B. (2004). The mental representation of social categories: Category boundaries, entitativity, and stereotype change. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity and essentialism* (pp. 79–100). Philadelphia, PA: Psychology Press.
- Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In G. R. Semin & K. Fiedler (Eds.), *Language, interaction, and social cognition* (pp. 11–36). London, England: Sage.
- Rutland, A. (1999). The development of national prejudice, in-group favoritism and self-stereotypes in British children. *British Journal of Social Psychology*, 38, 57–70.
- Spencer-Rodgers, J., Hamilton, D. L., & Sherman, S. J. (2007). The central role of entitativity in stereotypes of social categories and target groups. *Journal of Personality and Social Psychology*, 92, 369–388.
- Unkelbach, C. (2006). The learned interpretation of cognitive fluency. *Psychological Science*, 17, 339–345.
- Weiner, M. (Ed.). (1997). *Japanese minorities: The illusion of homogeneity*. London, England: Routledge.
- Wertheimer, M. (1938). Laws of organization in perceptual forms. (Trans.). In W. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 71–88). London, England: Routledge & Kegan Paul. (Reprinted from *Psychologische Forschung*, 4, 301–350)
- Wilder, D., & Simon, A. F. (1998). Categorical and dynamic groups: Implications for social perception and intergroup behavior. In C. Sedikides, J. Schopler, & C. A. Insko (Eds.), *Intergroup cognition and intergroup behavior* (pp. 27–44). London, England: Erlbaum.
- Yzerbyt, V. Y., Corneille, O., & Estrada, C. (2001). The interplay of subjective essentialism and entitativity in the formation of stereotypes. *Personality and Social Psychology Review*, 5, 141–155.
- Yzerbyt, V. Y., Judd, C. M., & Corneille, O. (2004). *The psychology of group perception: Perceived variability, entitativity, and essentialism*. New York, NY: Psychology Press.
- Yzerbyt, V. Y., Rocher, S., & Schadron, G. (1997). Stereotypes as explanations: A subjective essentialism view of group perceptions. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 20–50). Cambridge, MA: Blackwell.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151–175.
- Zosuls, K. M., Ruble, D. N., Tamis-LeMonda, C. S., Shrout, P. E., Bornstein, M. H., & Greulich, F. K. (2009). The acquisition of gender labels in infancy: Implications for gender-typed play. *Developmental Psychology*, 45, 688–701.

Received January 27, 2009

Revision received June 17, 2009

Accepted June 18, 2009 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://notify.apa.org/> and you will be notified by e-mail when issues of interest to you become available!