Stereotypes as Justifications of Prejudice

Christian S. Crandall, Angela J. Bahns, Ruth Warner and Mark Schaller

Pers Soc Psychol Bull 2011 37: 1488 originally published online 9 June 2011
DOI: 10.1177/0146167211411723

Stereotypes as Justifications of Prejudice

Personality and Social Psychology Bulletin
37(11) 1488–1498

DOI: 10.1177/0146167211411723
http://pspb.sagepub.com

Christian S. Crandall[1], Angela J. Bahns[2], Ruth Warner[3], and Mark Schaller[4]

Abstract

Three experiments investigate how stereotypes form as justifications for prejudice. The authors created novel content-free prejudices toward unfamiliar social groups using either subliminal (Experiment 1, $N = 79$) or supraliminal (Experiment 2, $N = 105$; Experiment 3, $N = 130$) affective conditioning and measured the consequent endorsement of stereotypes about the groups. Following the stereotype content model, analyses focused on the extent to which stereotypes connoted warmth or competence. Results from all three experiments revealed effects on the warmth dimension but not on the competence dimension: Groups associated with negative affect were stereotyped as comparatively cold (but not comparatively incompetent). These results provide the first evidence that—in the absence of information, interaction, or history of behavioral discrimination—stereotypes develop to justify prejudice.

Keywords

stereotypes, prejudice, justification-suppression model

Received September 8, 2010; revision accepted April 29, 2011

> We therefore conclude that the rationalizing and justifying function of a stereotype exceeds its function as a reflector of group attributes.
>
> —Gordon Allport (1954, p. 196)

Stereotypes serve two primary functions: a *knowledge* function (e.g., they represent and streamline information about groups) and a *justification* function (e.g., they rationalize observed or experienced group differences). As knowledge, stereotypes represent the world; as justifications, stereotypes *explain* the world. Stereotypes can explain not only what a group is, but also *why* the group is that way, and why groups are treated the way they are (McGarty, Yzerbyt, & Spears, 2002; Yzerbyt, Rocher, & Schadron, 1997).

The knowledge function has been the primary focus of stereotype research for several decades (see Hamilton, 1981; Macrae, Stangor, & Hewstone, 1996). For instance, when stereotypes are conceptualized as a source of prejudice and/or discrimination (e.g., Dovidio, Brigham, Johnson, & Gaertner, 1996; Esses, Haddock, & Zanna, 1993; Stephan & Stephan, 2000), it is the knowledge function of stereotypes that is implicated.

Stereotypes may also be a consequence (rather than cause) of discrimination. Here the justification function of stereotypes is implicated; stereotypes may emerge (and endure) because they help to explain why different groups are treated differently. For example, following an experimental manipulation that induced Scottish students to discriminate in favor of Scotland and against England, these students (compared to those in control conditions) showed evidence of more favorable stereotypes of Scots and more negative stereotypes of the English (Rutland & Brown, 2001).

Many conceptual perspectives on stereotype formation emphasize this justification function, suggesting that stereotypes may emerge to rationalize either individual acts of discrimination or discriminatory societal norms (e.g., Eagly & Steffen, 1984; Jost & Banaji, 1994; McGarty et al., 2002; Reicher, Hopkins, & Condor, 1997; Tajfel, 1969; Yzerbyt et al., 1997). Several theories—including social dominance theory (Sidanius & Pratto, 1999) and system justification theory (Jost & Hunyady, 2003)—imply that stereotypes may develop to protect group-based inequalities. Consistent with this, high status group members who are committed to the maintenance of hierarchical social systems are particularly likely to endorse negative stereotypes about low status groups (Bigler, Brown, & Markell, 2001; Sidanius & Pratto, 1999). Similarly, social role theory (Eagly, 1987; Eagly & Steffen, 1984) states that stereotypes may develop to rationalize the

[1]University of Kansas, Lawrence, KS, USA
[2]Wellesley College, Wellesley, MA, USA
[3]St. Louis University, St. Louis, MO, USA
[4]University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Christian S. Crandall, University of Kansas, 1415 Jayhawk Blvd., Lawrence, KS 66045
Email: crandall@ku.edu

fact that human groups are differentially distributed into specific kinds of social roles. Different stereotypes emerge depending on whether a group is overrepresented among child raisers versus city workers: When overrepresented among child raisers, a group is consequently stereotyped as patient, kind, and understanding; when overrepresented among city workers, a group is stereotyped as determined and energetic (Hoffman & Hurst, 1990).

This body of research focuses almost entirely on the consequences of behavioral discrimination, not on the consequences of prejudice (which is defined not by overt behavior but by an affective experience).[1] Thus, although there is abundant evidence that stereotypes may emerge as justifications for discrimination, there is no prior evidence testing whether stereotypes emerge as justifications for mere prejudice.

Do Stereotypes Emerge to Justify Prejudice (Even in the Absence of Discrimination)?

This distinction is critical; just as behavioral discrimination may occur in the absence of prejudicial attitudes, so too prejudices may be formed in the absence of prior discriminatory behavior or knowledge about the groups (as the result, for instance, of mere association with negatively valued perceptual objects; Krosnick, Betz, Jussim, & Lynn, 1992; Livingston & Drwecki, 2007). If—even in the absence of overt discrimination—stereotypes emerge as justifications for prejudice, there are novel implications for our understanding of stereotype formation processes more generally.

The distinction also matters because overt behavior and intrapsychic affective experiences do not necessarily have identical consequences. Because behavioral discrimination is usually publicly observable, impression-management processes (designed to maintain a positive image of oneself or—as in the case of observed societal discrimination—coalitional or cultural groups with which one identifies) provide one important explanation as to why stereotypes are formed to justify discrimination (Kiesler, 1971). In contrast, prejudice (an affective evaluation of a group) is typically not publicly observable. Thus, although we can be confident that stereotypes do emerge to justify discrimination, it is entirely plausible that stereotypes do *not* emerge to justify mere prejudice. On the other hand, if stereotypes *do* emerge following the experience of prejudice, it implies the formation of stereotypes that are truly internalized and not merely constructed strategically for public image management. If such effects emerge, then one has evidence of the operation of additional psychological processes through which stereotypes emerge in response to mere affective experiences.

If So, Exactly What Stereotypes Emerge?

If indeed stereotypes are formed following the experience of mere prejudice, the phenomenon might potentially be explained by several different kinds of processes—including processes pertaining to the informational function of affective experience (Schwarz, 1990), the maintenance of cognitive consistency in general (Festinger, 1957; Heider, 1958), and the more functionally specific need to justify negative feelings about social groups (i.e., prejudices) that individuals would otherwise find personally unacceptable (Crandall & Eshleman, 2003). We will discuss these different processes—and their plausibility given the results of our experiments—in greater detail in the general discussion. In the meantime, we suggest that to help distinguish between different kinds of explanations, it is important not only to assess *whether* stereotypes emerge in response to the experience of prejudice but also to consider the *specific contents* of those stereotypes.

Much previous work on stereotypes has focused simply on the overall evaluative valence of a stereotype—whether it is positive or negative. But there are very different kinds of positive and negative traits, which have different kinds of behavioral implications. Work on the stereotype content model (Fiske, Cuddy, Glick, & Xu, 2002) reveals that the contents of group stereotypes can be usefully summarized in terms of two dimensions of social judgment: *warmth* and *competence*. Warmth stereotypes connote a group's competitive intent as friend or foe. Competence stereotypes connote a group's status or ability to enact its intent. The two dimensions are roughly equivalent in evaluative valence (i.e., both warmth and competence are positively valued); but they can be entirely independent. Group stereotypes can be highly negative along one dimension while being positive along the other (Fiske et al., 2002; Schaller & Abeysinghe, 2006).

In testing whether stereotypes emerge as justifications for mere prejudice, we assessed separately the extent to which these emerging stereotypes had specific contents relevant to the warmth dimension, the competence dimension, or both. This is important not only because connotatively distinct stereotypes have different behavioral implications, but also because the specific nature of these stereotypes can yield clues to the specific psychological processes through which they are formed. Some potential explanatory processes imply the formation of stereotypes along both dimensions, whereas other processes imply the formation of stereotypes primarily along just one dimension but not the other. We will return to this issue in the general discussion, but we first present three experiments that (a) test the hypothesis that stereotypes emerge from mere prejudice and (b) document the specific content of those stereotypes.

Overview

To test the hypothesis that stereotypes emerge from mere prejudice as rigorously as possible, we used affective conditioning methods to create novel prejudices toward novel groups. These methods drew on evidence that affectively laden attitudes can be created through associative learning processes whenever specific attitude objects are perceived

in coincidence with other positively or negatively valued stimuli (De Houwer, Thomas, & Baeyens, 2001; Krosnick et al., 1992; Livingston & Drwecki, 2007; Olson & Fazio, 2001). It has been shown, for example, that positive affective associations with a novel stimulus emerge simply from the pairing of that stimulus with a subliminally presented happy face (Edwards, 1990). In Experiment 1, we employed a subliminal conditioning paradigm. In Experiments 2 and 3, we used a supraliminal conditioning paradigm. The immediate consequence of these methods was the creation of "pure" prejudices—mere affective associations uncontaminated by prior contact, discrimination, or even prior knowledge.

Following the creation of these prejudices, we assessed stereotypic beliefs about the target groups by asking participants to indicate the extent to which different trait adjectives were descriptive of the target groups. These methods allowed us to test whether affectively conditioned prejudices consequently produced stereotypes specific to the warmth dimension, the competence dimension, or both.

Experiment 1

We employed an affective conditioning paradigm in which participants were presented with innocuous bits of geographic information about two unfamiliar countries (Moldova and Slovenia). Each bit of information was paired with the subliminal presentation of an iconic face displaying either a positive or negative emotional facial expression. Across the entire set of conditioning trials, one country (counterbalanced) was paired with subliminal presentation of the positive emotional expression, whereas the other was paired with the negative emotional expression. We expected that participants would form affective associations favoring one country over the other. More importantly, we tested whether these affective associations led participants to develop content-laden stereotypes (warmth and competence) that distinguished between Moldovans and Slovenians.

Method

Pretest to select target countries. A separate sample of nine students rated nine countries (Armenia, Belarus, Croatia, Estonia, Latvia, Lithuania, Macedonia, Moldova, and Slovenia) on 7-point scales to assess preexisting affective associations ("How do you feel toward this country?"; 1 = v*ery negative*, 7 = *very positive*) as well as preexisting familiarity ("How familiar are you with this country?"; 1 = *not at all familiar*, 7 = *very familiar*). Affect ratings for two countries, Moldova and Slovenia, were identical and very close to the midpoint of the rating scale (both *M*s = 4.22); pretest participants indicated extremely low levels of familiarity with both countries (Slovenia, $M = 1.89$, $SD = 1.36$; Moldova, $M = 1.33$, $SD = 0.71$; the mean difference was nonsignificant, $t < 1$). Moldova and Slovenia were selected as target countries for the first experiment.[2]

Participants for the main experiment. Participants were 79 undergraduates at the University of Kansas. Gender of participant was not assessed.

Affective conditioning procedure. Participants were presented with 30 consecutive screens of information on a Dell Dimension 4600 computer, with M992 CRT and a refresh rate of 85 Hz. Of these screens, 15 presented written statements containing information about Moldova and 15 presented written statements containing information about Slovenia. These statements were designed to be innocuous, focusing primarily on geography, for example, "The climate varies greatly from region to region," "The protected Dosev delta is the winter home to much of Europe's bird population," "The sun shines approximately 2000 hours per year and there is plenty of snow in the winter," and "The country has 46.6 km of seacoast or an inch per inhabitant." The pairing of information sets with country was counterbalanced. Also counterbalanced was whether the first screen presented information about Moldova or Slovenia.

As an affective prime, the computer screen presented an iconic face for 23.5 milliseconds (without masking) immediately preceding each piece of information about Moldova or Slovenia.[3] Two different faces were presented: ☺ (positive-affect face) and ☹ (negative-affect face). These iconic faces were presented in 36-point font (10 mm). Across all 30 trials, the positive-affect face was consistently paired with one country and the negative-affect face was consistently paired with the other. These pairings varied across two experimental conditions. In one condition, Moldova was paired with the positive-affect face and Slovenia was paired with the negative affect face; in the other condition, the pairings were reversed.

Pretest to ascertain success of the affective conditioning procedure. To test the hypothesis that stereotypes can emerge as justifications for prejudice, the affective conditioning paradigm must successfully create differential affective responses toward the two countries. To ascertain the success of this procedure, a separate sample of 38 participants completed the same conditioning procedure and then completed an adapted Bogardus social distance scale (Biernat & Crandall, 1999). The seven items ($\alpha = .91$) reflected a mixture of affective responses ("Moldovans are probably likable people") and affect-based action tendencies ("Moldovans are the kind of people I tend to avoid"; 0 = *no*, 1 = *yes*). Results revealed that participants responded more negatively (greater social distance) toward a country (Moldova or Slovenia, depending on condition) when it was paired with the negative-affect face ($M = 0.61$, $SD = 0.41$) rather than the positive affect face ($M = 0.30$, $SD = 0.35$), $F(1, 36) = 6.43$, $p = .016$, $\eta^2 = .15$. This indicates that the subliminal conditioning procedure was successful in creating the affective associations that define prejudice.

Measure of stereotype content. To assess stereotypes that emerged following affective conditioning, participants were presented with a list of 29 trait words (taken primarily from

Katz & Braly, 1933). Of these, 17 traits reflected the *warmth* dimension (friendly, warm, good-natured, courteous, liars, stubborn, happy, quarrelsome, threatening, quick-tempered, aggressive, generous, conceited, trustworthy, humorless, sincere, and arrogant) and 12 traits reflected the *competence* dimension (lazy, messy, capable, skillful, low in self-control, intelligent, unreliable, confident, competent, efficient, ignorant, and physically clean). Scales were verified by factor analysis, with all loadings > .50, totaling 44.1% of the variance. Participants were asked to indicate whether each trait was more descriptive of people from Moldova or Slovenia, by circling either "Moldova" or "Slovenia" on a binary response scale. After reverse scoring negatively valenced items, separate warmth ($\alpha = .93$) and competence ($\alpha = .87$) stereotype indices were computed for each target group, representing the proportion of warmth- and competence-relevant trait terms ascribed to each group.

Results

A four-factor mixed-model ANOVA was conducted on the stereotype content indices. This ANOVA included two factors that reflected conceptually uninteresting counterbalancing conditions as well as the two factors of primary conceptual interest: the face–country pairing manipulation (negative-affect face paired with Moldova or paired with Slovenia) and the within-participants assessment of stereotypes along two distinct stereotype content dimensions (warmth, competence).

Results revealed a marginally significant main effect of the face–country pairing manipulation, indicating that the country paired with the negative-affect face was generally associated with more negative traits across both stereotype content indices, $F(1, 68) = 2.84$, $p = .097$, $\eta^2 = .03$. This effect was qualified by a statistically significant interaction with the within-subjects stereotype content dimension variable, $F(1, 68) = 4.90$, $p = .03$, $\eta^2 = .06$. The nature of this interaction is illustrated in Figure 1. When a country was paired with the negative-affect face, its inhabitants were stereotyped as substantially lower in *warmth* ($M = 6.77$, $SD = 5.03$),[4] compared to the inhabitants of the country paired with the positive-affect face ($M = 10.23$, $SD = 5.05$), $F(1, 68) = 6.79$, $p = .011$, $\eta^2 = .09$. No such effect was observed on perceived *competence* ($F < 1$). Competence stereotypes were endorsed to a similar extent regardless of whether a country was paired with the negative-affect face ($M = 5.93$, $SD = 3.05$) or the positive-affect face ($M = 6.07$, $SD = 3.46$).[5]

These results are consistent with the hypothesis that stereotypes can emerge as justifications for mere prejudice or content-free affective associations with groups. They also provide clear evidence that the trait contents of these emergent stereotypes are specific to the warmth dimension, with virtually no trait contents bearing on the competence dimension.

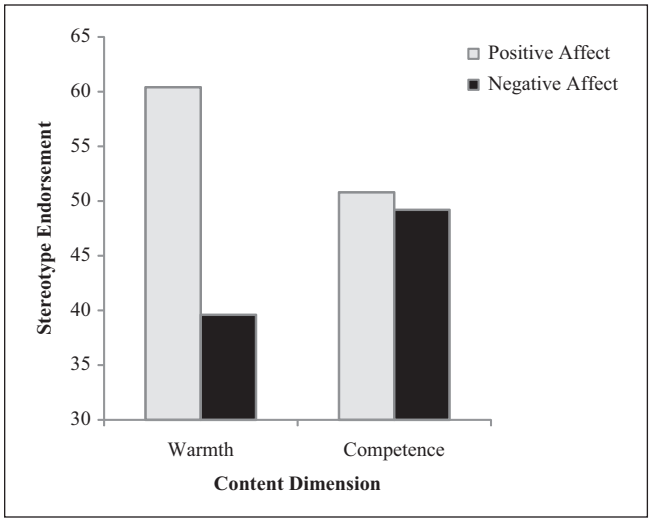

Figure 1. A subliminal conditioning procedure that creates novel prejudices (affective associations) results in the formation of new stereotypes along the warmth dimension but not the competence dimension, Experiment 1

This latter finding (about the content specificity of emerging stereotypes) is potentially of both conceptual and practical importance. However, it might be explained away as a methodological artifact if, compared to trait words connoting competence, the trait words connoting warmth are perceived to be more positive overall—and thus more readily matched to a positive affective association. To address this alternative interpretation, we conducted a follow-up study on a separate sample of 52 undergraduates. These participants rated the subjective positivity of the 29 trait words presented in Experiment 1. Ratings were made on 7-point scales (1 = *very negative*, 7 = *very positive*). After reverse scoring negatively valenced items, mean positivity ratings were computed for the 17 warmth words and for the 12 competence words. These means were nearly identical, with competence words being slightly more positive (warmth $M = 5.60$, competence $M = 5.78$), $t(51) = 0.87$, $p > .30$. These results render the alternative interpretation untenable.

Experiment 2

We do not know if the emergence of stereotypes from content-free prejudices is limited to subliminal conditioning. In addition, there is some small possibility that the smiley and frowny face icons convey some kind of content that leads to the emergence of these particular stereotypes. As a result, Experiment 2 was designed to be a conceptual replication of Experiment 1 but with different target countries and a different procedure to create affective associations with those target countries. A supraliminal affective conditioning paradigm was adapted from Olson and Fazio (2001). The target countries were Eritrea and Mauritania. The unconditioned stimuli were composed of a broad set of affectively loaded images and words.

Method

Participants. Participants were 105 undergraduates (77 women) at the University of Kansas.

Pretesting of target countries. Following a pretesting procedure (pretest $n = 32$) similar to that used in Experiment 1, Eritrea and Mauritania were chosen as the target countries. A feeling thermometer (0 = *very negative*, 100 = *very positive*) assessed general affect toward each country. Both countries received mean ratings that were very close to the midpoint of the scale (Eritrea $M = 55.19$, $SD = 21.01$; Mauritania $M = 54.07$, $SD = 20.05$; the mean difference was not significant, $p > .15$). On a rating scale (1 = *I've never heard of it*, 2 = *I've heard of it but don't know where it is*, 3 = *I've heard of it and know where it is*), both countries were rated as unfamiliar (Eritrea $M = 1.52$, $SD = 0.85$; Mauritania $M = 1.44$, $SD = 0.70$; the mean difference was not significant, $p > .15$).

Affective conditioning procedure. We told participants the study was about attention and vigilance, and we gave them a role-playing task described as testing the skills used by a security guard "to be alert and ready to respond" to intruders. Participants' task was to push a button whenever a designated country name appeared on the screen. This task was designed to distract participants from focusing on the target countries but also to ensure that participants were attending to the stimuli. The distracter country names that participants searched for were Oman, Azerbaijan, Moldova, Slovenia, and Tajikistan. Each country name appeared 10 times, and participants searched for a different country name in each block. For some trials the country name appeared on screen alone, and for other trials it was paired with either a neutral image or a neutral word.

Participants were presented with 430 screens of information, broken down into five experimental blocks. Stimulus materials were presented on the same computers as Experiment 1. Across all five blocks, there were 40 critical trials involving the target countries (20 trials involving Eritrea, and 20 involving Mauritania) in which positively valenced stimuli were consistently paired with one country and negatively valenced stimuli were consistently paired with the other country. In one condition, Eritrea was paired with positive stimuli and Mauritania was paired with negative stimuli; in the other condition, the pairings were reversed. The unconditioned stimuli paired with Eritrea and Mauritania for the critical trials were 10 positively valenced words (e.g., kittens, babies, party), 10 positively valenced images (e.g., mountains, clouds, hot fudge sundae), 10 negatively valenced words (e.g., danger, filth, pain), and 10 negatively valenced images (e.g., cigarette butts, air pollution, injured kitten). Stimulus words were selected based on pretesting at the University of Kansas ($n = 18$), and images were selected from the International Affective Picture System database (Lang, Bradley, & Cuthbert, 2005). A total of 340 trials of filler stimuli (e.g., blank screens, or affectively neutral stimuli that were not paired with either Eritrea or Mauritania) were interspersed among the critical trials and distracter country trials.

Dependent measures. After completing the affective conditioning procedure, participants were presented with a list of 25 trait words. In all, 16 traits reflected the *warmth* ($\alpha = .94$) dimension (friendly, warm, good-natured, courteous, liars, stubborn, happy, quarrelsome, threatening, quick-tempered, generous, conceited, trustworthy, humorless, sincere, and arrogant); 9 traits reflected the *competence* ($\alpha = .85$) dimension (lazy, messy, capable, skillful, intelligent, confident, competent, efficient, and physically clean). Factors were confirmed by factor analysis.[6] Participants were asked to indicate whether each trait was more descriptive of people from Eritrea or Mauritania on a binary response scale. After reverse scoring negatively valenced items, separate warmth and competence stereotype indices were computed, representing the proportion of warmth- and competence-relevant trait terms ascribed to each group.

Manipulation check. To assess the success of the affective conditioning manipulation, participants also rated their feelings toward each of the countries (including Eritrea and Mauritania) whose names appeared during the five blocks of trials (0 = *very negative*, 100 = *very positive*). Analyses on ratings of Eritrea and Mauritania revealed only the expected interaction of the affective stimuli–country pairing manipulation and country rated (i.e., a counterbalanced main effect), indicating more negative feelings toward the country that had been paired with negative stimuli, $F(1, 103) = 5.21$, $p = .025$, $\eta^2 = .05$. These results indicate that the manipulation was successful.[7]

Results

A 2 × 2 mixed-model ANOVA was conducted on the stereotype content indices. The two factors included the affective stimuli–country pairing manipulation (negative-affect stimuli paired with Eritrea or with Mauritania) and the within-participants assessment of stereotypes along two distinct stereotype content dimensions (warmth, competence).

There was a main effect of the experimental manipulation, indicating a relatively more negative stereotype of the country paired with negative stimuli, $F(1, 103) = 5.54$, $p = .020$, $\eta^2 = .05$. This main effect was qualified by an interaction between the manipulation and the stereotype content dimension variable, $F(1, 103) = 6.28$, $p = .014$, $\eta^2 = .06$.

Figure 2 reveals the nature of this interaction. When a country was paired with negative-affect stimuli, its inhabitants were stereotyped as lower in *warmth* ($M = 6.56$, $SD = 6.10$) compared to the inhabitants of the country paired with the positive-affect stimuli ($M = 9.77$, $SD = 5.23$), $F(1, 103) = 8.43$, $p = .005$, $\eta^2 = .08$. No such effect was observed on perceived *competence*, $F < 1$, $\eta^2 = .01$. Competence stereotypes were endorsed to a similar extent regardless of whether a country was paired with negative-affect stimuli ($M = 4.17$, $SD = 2.43$) or positive-affect stimuli ($M = 4.62$, $SD = 2.36$).

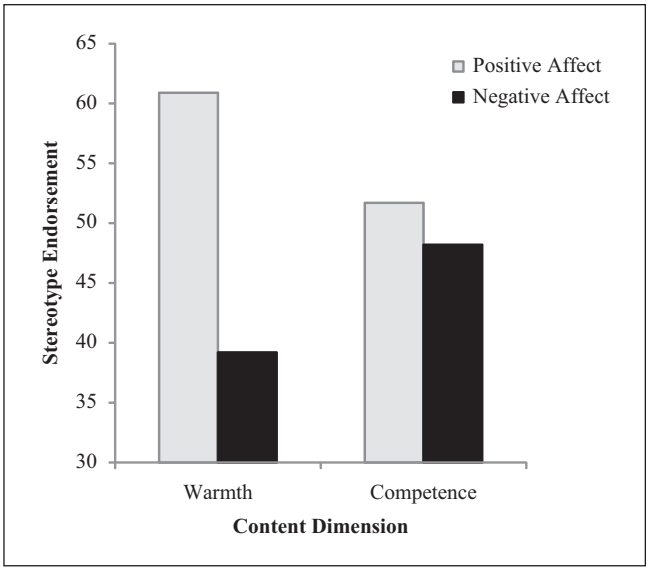

Figure 2. A supraliminal conditioning procedure that creates novel prejudices (affective associations) results in the formation of new stereotypes along the warmth dimension but not the competence dimension, Experiment 2

These results replicate those of Experiment 1. Thus, regardless of the specific target countries employed (Moldova/Slovenia or Eritrea/Mauritania), the unconditioned stimuli used (iconic faces vs. words and pictures), or whether one employs a subliminal or supraliminal affective conditioning paradigm, the creation of a mere affective association gives rise to content-based stereotypes. These emerging stereotypes arise along the warmth dimension but not the competence dimension.

Experiment 3

Experiments 1 and 2 used a forced-choice measure of stereotype endorsement, which might limit interpretability. Continuous trait scales are the most common measure of stereotypes in social psychology, and they are very effective at capturing mean differences (Biernat & Crandall, 1994, 1996). In Experiment 3 we use the same supraliminal conditioning procedure as Experiment 2, using continuous scales as dependent variables. We combined two separate samples that used exactly the same manipulation and dependent measures. The only difference was that the target groups were Eritrea and Mauritania in Sample 1 and Azerbaijan and Mauritania in Sample 2. Preliminary analyses confirmed the source of data (Sample 1 or Sample 2) had no significant effects on any of the dependent measures. Thus, in Experiment 3 we report the results for the two samples corporately.

Method

Participants. Participants were 130 (Sample 1 $n = 42$; Sample 2 $n = 88$) undergraduates (84 women) at the University of Kansas.

Target countries. Eritrea and Mauritania were again used as target countries for Sample 1. Azerbaijan and Mauritania were chosen as target countries for Sample 2 using the same pretesting sample ($n = 32$) described in Experiment 2. Both countries received mean affect ratings that were very close to the midpoint of the scale (Azerbaijan $M = 55.36$, $SD = 19.72$; Mauritania $M = 53.57$, $SD = 19.85$); the mean difference was not significant, $p > .25$. The majority of participants rated the countries as unfamiliar indicating either "I've never heard of it" or "I've heard of it but don't know where it is"; both countries were rated as unfamiliar (Azerbaijan 71.9%, Mauritania 78.2%).

Affective conditioning procedure. The conditioning procedure and cover story were the same as in Experiment 2; the set of the positive and negative words used as affective stimuli was selected from the Affective Norms for English Words database (Bradley & Lang, 1999).

Dependent measures. After the conditioning procedure, participants rated the target countries on stereotype traits using 7-point semantic differentials. Participants were asked to give their general impression of what each of the groups are like (e.g., "I would say that Mauritanians are . . ."; 1 = *unfriendly*, 7 = *friendly*). Four pairs of traits reflected the *warmth* ($\alpha = .90$) dimension (unfriendly–friendly, insincere–sincere, not warm–warm, selfish–generous); five pairs of traits reflected the *competence* ($\alpha = .82$) dimension (lazy–hardworking, messy–neat, incapable–capable, unconfident–confident, incompetent–competent). Both target countries were rated independently; separate warmth and competence stereotype indices were computed for each country.

Manipulation check. Participants completed the same affective rating scales as in Experiment 2. Analyses on affective ratings on countries revealed only the expected interaction of the affective stimuli–country pairing manipulation and country rated (i.e., a counterbalanced main effect), indicating more negative feelings toward the country that had been paired with negative stimuli, $F(1, 127) = 2.87$, $p = .09$, $\eta^2 = .02$. These results indicate that the manipulation was successful.

Results

A 2 × 2 × 2 × 2 mixed-model ANOVA was conducted on the stereotype content indices. The between-subjects factors were the affective stimuli–country pairing manipulation (negative-affect stimuli paired with Eritrea/Azerbaijan or Mauritania) and sample (Sample 1, Sample 2). The within-participants factors were assessment of stereotypes along two distinct stereotype content dimensions (warmth, competence) and country rated (Eritrea/Azerbaijan, Mauritania).

There was an main effect of stereotype content dimension, indicating relatively more endorsement of stereotypes in the competence dimension compared to the warmth dimension, $F(1, 126) = 4.23$, $p = .042$, $\eta^2 = .03$. This main effect was qualified by the predicted critical interaction between

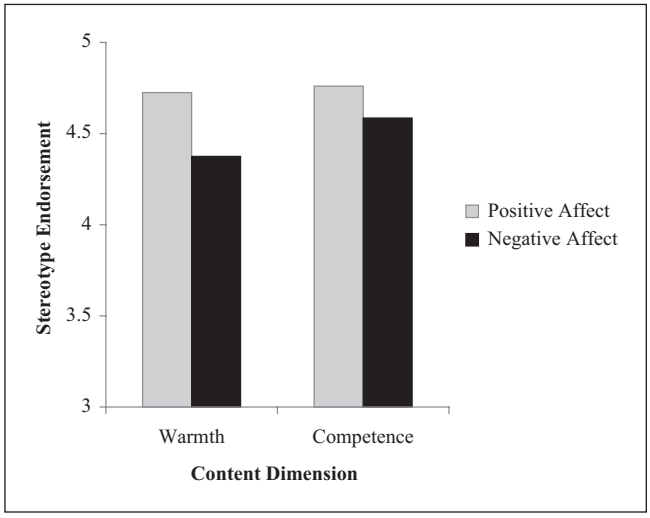

Figure 3. A supraliminal conditioning procedure that creates novel prejudices (affective associations) results in the formation of new stereotypes along the warmth dimension but not the competence dimension, Experiment 3

the manipulation, the stereotype content dimension variable, and the country rated, $F(1, 126) = 4.68$, $p = .032$, $\eta^2 = .04$.

Figure 3 reveals the nature of this interaction. When a country was paired with negative-affect stimuli, its inhabitants were stereotyped as significantly lower in *warmth* ($M = 4.38$, $SD = 1.10$), compared to the inhabitants of the country paired with the positive-affect stimuli ($M = 4.73$, $SD = 1.10$), $F(1, 126) = 12.00$, $p = .001$, $\eta^2 = .09$. In contrast, the affective conditioning manipulation resulted in a very modest (and statistically nonsignificant) effect on the *competence* dimension, $F(1, 126) = 2.91$, $p = .091$, $\eta^2 = .02$. A country was rated as about the same in competence when it was paired with negative-affect stimuli ($M = 4.59$, $SD = 0.99$) or positive-affect stimuli ($M = 4.76$, $SD = 0.96$).[8]

Thus, results from Experiments 1 and 2 are replicated using the continuous measure of stereotyping. Regardless of the specific target countries used, or whether a forced-choice or semantic differential was employed, the creation of a mere affective association gives rise to content-based stereotypes. These stereotypes emerge reliably and vigorously along the warmth dimension, but only weakly or not at all along the competence dimension.

General Discussion

Across three experiments we created "pure" prejudices (i.e., content-free affective associations) toward novel groups and tested whether stereotypes (i.e., content-laden knowledge structures) emerged as a consequence. They did. The experiments assessed the specific contents of those emergent stereotypes, thus testing the extent to which these stereotypes focused on traits connoting warmth, competence, or both. Results revealed that the stereotypes were specific to warmth, with very little evidence of stereotypes emerging on traits connoting competence. This is novel evidence that (a) stereotypes can develop from mere prejudice in the absence of any prior behavioral discrimination and (b) these stereotypes are particularly defined by the set of traits on the warm–cold dimension.

All three experiments employed affective conditioning procedures to create prejudices, but the exact procedures differed across experiments. Experiment 1 used a *subliminal* conditioning method, in which group labels were paired with iconic faces bearing either positive or negative facial expressions. Experiment 2 used a *supraliminal* conditioning method, in which group labels were paired with a wide range of nonfacial content consistent with either positive or negative affect. Experiment 3 used a supraliminal conditioning method and a continuous measure of stereotype endorsement rather than the binary response scale used in Experiments 1 and 2. Despite these procedural differences, the pattern of results replicated across studies. Although there is considerable variability in the exact means through which positive and/or negative affect comes to be associated with a group, these affective associations may consequently lead to (and be justified by) a predictably consistent set of stereotypical beliefs.

The emergent stereotypes focused on warmth, and this occurred despite the fact that warmth-connoting traits were judged slightly less positive than competence-connoting traits. Both warmth and competence are fundamental dimensions of interpersonal perception, but perceivers may place special priority on judgments of warmth (Cottrell, Neuberg, & Li, 2007; Fiske, Cuddy, & Glick, 2007; Schaller, 2008). Stereotypes of warm–cold (e.g., whether members of a group are friendly or unfriendly) are relatively unambiguous in their implications for affect and interpersonal behavior. Perceptions of warmth encourage approach-oriented behaviors; perceptions of coldness suggest avoidance. By contrast, the affective and behavioral implications of competence stereotypes depend on additional contextual information. When people are perceived to be friendly, their competence is typically valued; however, when people are perceived to have hostile intent, their competence is dismaying instead (a competent enemy is more dangerous than an incompetent one). Stereotypes that focus on warmth are likely to provide the most immediate and unambiguous justifications for affective associations.

Do Stereotypes Follow Directly From Feelings?

We have emphasized the justification function of stereotypes. Might other processes—which have little to do with justification per se—also account for the results? Discrete emotional experiences (e.g., fear, anger, disgust, etc.) have functionally specific consequences on cognition and behavior, and several perspectives on prejudice suggest that when perceivers experience specific affective states, there are predictably specific consequences on intergroup cognitions and intergroup behaviors (Cottrell & Neuberg, 2005; Mackie,

Devos, & Smith, 2000; Neuberg & Cottrell, 2002; Smith, 1999). It is not clear that these theories can be applied to our results, given that our studies were designed to create simple affective associations with novel target groups, not emotionally specific affective states. Plus, given the nature of the stimuli employed in the affective conditioning methods, the affective associations themselves are likely to have been highly diffuse (i.e., a generally negative evaluative association) rather than emotion specific. In sum, the results of our studies do not test the logical implications of these theories, nor can these theories explain our data.

There is also a body of research revealing that feelings—emotions, moods, and other affective experiences—provide useful information to the people who have those feelings (Schwarz, 1990). The feelings-as-information perspective would predict that the experience of prejudice toward a target group provides information about the likely trait characteristics of that group (e.g., negative traits). Because our methods created generally positive and generally negative affective associations and not emotionally specific affective states, a feelings-as-information process implies the emergence of stereotypes along *all kinds* of evaluative positive and negative traits. This is not what we found. The fact that stereotypes emerged only on traits connoting warmth suggests that a feelings-as-information process does not provide an adequate explanation for these results.

A third potential explanation lies in the possibility that stereotypic inferences about warmth are simply easier—that they require less informational input than stereotypic inferences about competence. If so, then stereotypes about warmth, but not competence, might be expected to emerge as the result of information-impoverished manipulations of the sort employed in our studies. How plausible is this explanation? Rothbart and Park (1986) documented the number of behavioral instances required to confirm a specific trait inference (as well as the relative number of instances required to disconfirm an existing trait inference). Using their results, we compared warmth-relevant traits and competence-relevant traits in regard to the number of instances (i.e., amount of information) required to confirm an inference and to disconfirm an inference.[9] Results revealed no meaningful differences between warmth- or competence-relevant traits (both ts < 1, ps $> .30$). These results cannot completely rule out the possibility that stereotypes about warmth are simply easier to form with less information than stereotypes about competence; but Rothbart and Park's (1986) data offer no support for the argument.

The Justification Function of Stereotypes

The most plausible alternative processes to justification do not easily explain our results. We argue that stereotypes emerge because they provide perceivers with a representation that justifies the experience of prejudice. These stereotypes emerge primarily along the warmth dimension (and weakly if at all on the competence dimension) because traits connoting warmth have logically unambiguous implications for affect and behavior and thus provide psychologically compelling justifications.

This sort of justification process fits with the broader set of psychological theories that focus on the need for cognitive consistency and its implications (e.g., Festinger, 1957; Heider, 1958). Just as specific kinds of attitudes help people rationalize and justify their past behavior, specific kinds of stereotypes might help people rationalize and justify their existing prejudices. Given their relatively unambiguous implications for affect and behavior, stereotypes that focus on warmth may be especially effective in helping individuals maintain a subjective sense of consistency between affective associations with a target group and connotative beliefs about that group.

This justification process also fits neatly within the justification-suppression model of prejudice expression (Crandall & Eshleman, 2003). According to this model, people believe that intergroup prejudices are unacceptable (because, for example, prejudice is incompatible with personal and/or societal values emphasizing egalitarianism). Consequently the experience of prejudice is unpleasant. This psychological discomfort can be reduced when perceivers have available some compelling justification for the prejudice. From this perspective, stereotypes emerge as justifications for prejudice not simply because of the general motive to maintain cognitive consistency, but because of the more specific goal of maintaining a self-image (and projecting a public image) that is consistent with value systems promoting egalitarianism. To the extent that the implications of warmth stereotypes for affect and behavior are comparatively unambiguous, stereotypes that focus on warmth are likely to serve as compelling justifications for prejudice, and thus are especially effective in maintaining (and projecting) the desired self-image.

Implications for Stereotype Change

Regardless of the processes through which they are formed, stereotypes are resistant to change (Rothbart & John, 1993; Schneider, 2004), and stereotypes that emerge as justifications for preexisting affective associations may be especially difficult to change. One reason is because the underlying affective associations are, compared to other kinds of learned associations, particularly resistant to extinction (Baeyens, Crombez, Van den Bergh, & Eelen, 1988; De Houwer, Baeyens, Vansteenwegen, & Eelen, 2000). Another reason is that when stereotypes serve a justification function rather than a purely informational function, they will be largely impervious to mere information, arguments, and counterexamples. Information-based educational approaches to stereotype erasure may therefore be little more than expensive games of Whac-A-Mole: The stereotypes and their underlying prejudice may not be eliminated at all but instead just temporarily suppressed until their pernicious heads find

another outlet. More effective interventions would focus instead on the *affective basis* of stereotypes, and on the nonrational justification processes through which they emerge.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. We adopt the definition of prejudice from Crandall, Eshleman, and O'Brien (2002), "a negative evaluation of a group or of an individual on the basis of group membership" (p. 359), with *affect as the primary component* of the negative evaluation.
2. A separate sample ($n = 102$) from an unrelated experiment from our lab also rated Moldova ($M = 54.0$, $SD = 19.75$) and Slovenia ($M = 56.73$, $SD = 18.00$) very near the midpoint on a 100-point feeling thermometer (0 = *very negative*, 100 = *very positive*); the ratings did not significantly differ from each other, $p > .15$.
3. A funneled debriefing interview probed participants for awareness of the iconic face primes, starting with vague open-ended questions and becoming increasingly more direct in regard to the contingency between target countries and affective primes. Only three participants could identify that they saw faces flashing before the country names; these participants were excluded from the analysis.
4. Descriptive statistics reported in the text indicate sums of traits ascribed to each of the groups, whereas the figures show proportions.
5. The ANOVA results revealed one additional statistically significant effect: a conceptually uninteresting interaction between stereotype content dimension and the counterbalancing factor for information set, indicating that one of the information sets led to slightly higher ratings along the warmth dimension but not the competence dimension, $F(1, 68) = 9.21$, $p < .01$. Because of counterbalancing, this effect does not change the interpretation of Figure 1.
6. Factor loadings were greater than .50 for all traits in the initial factor analysis. A reduced version of the warmth and competence scales, including only the traits with the strongest factor loadings, was also used for our primary analysis of the manipulation's effect on stereotype endorsement. The reduced warmth scale included eight traits (factor loadings all > .70): warm, courteous, generous, friendly, quick-tempered (reversed), happy, good-natured, and conceited (reversed). The reduced competence scale included seven traits (factor loadings all > .53): intelligent, confident, skillful, competent, lazy (reversed), capable, and efficient. We found the same pattern of results with the reduced scales: Stereotypes in the dimension of warmth but not competence emerged as a result of a mere (negative) affective association, interaction $F(1, 103) = 4.25$, $p = .04$, $\eta^2 = .04$.
7. A funneled debriefing interview similar to the one used in Experiment 1 assessed participants' awareness of the contingencies between target countries and affective stimuli. Ten participants expressed some awareness of the contingencies; however, awareness did not interact with the manipulation of affect or the stereotype measures, and so these cases were not excluded from the analysis.
8. The ANOVA results revealed three additional statistically significant but conceptually uninteresting effects: a main effect of country rated, indicating relatively more stereotype endorsement for Mauritania, $F(1, 126) = 6.26$, $p = .014$, $\eta^2 = .04$; an interaction of country rated and the manipulation of affect, indicating a relatively stronger effect of the manipulation on ratings of Eritrea/Azerbaijan, $F(1, 126) = 9.56$, $p = .002$, $\eta^2 = .07$; and an interaction of sample and the manipulation of affect, indicating a relatively stronger effect of the manipulation on ratings of Mauritania in Sample 1 and on ratings of Azerbaijan in Sample 2, $F(1, 126) = 4.54$, $p = .035$, $\eta^2 = .04$.
9. Rothbart and Park (1986) reported results for 19 traits that we categorized as relevant to *warmth* (arrogant, conceited, courteous, generous, good-natured, gregarious, happy-go-lucky, honest, hostile, humorless, jovial, kind, malicious, pleasant, polite, quarrelsome, quick-tempered, selfish, sincere) and 8 traits that we categorized as relevant to *competence* (brilliant, efficient, ignorant, intelligent, lazy, neat, stupid, wise).

References

Allport, G. W. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.

Baeyens, F., Crombez, G., Van den Bergh, O., & Eelen, P. (1988). Once in contact, always in contact: Evaluative conditioning is resistant to extinction. *Advances in Behaviour Research and Therapy*, *10*, 179-199.

Biernat, M., & Crandall, C. S. (1994). Stereotyping and contact with social groups: Measurement and conceptual issues. *European Journal of Social Psychology*, *24*, 1-19.

Biernat, M., & Crandall, C. S. (1996). Creating stereotypes and capturing their content. *European Journal of Social Psychology*, *26*, 867-898.

Biernat, M., & Crandall, C. S. (1999). Racial attitudes. In J. Robinson, P. Shaver, & L. Wrightsman (Eds.), *Measures of political attitudes* (2nd ed., pp. 297-411). New York, NY: Academic Press.

Bigler, R., Brown, C. S., & Markell, M. (2001). When groups are not created equal: Effects of group status on the formation of intergroup attitudes in children. *Child Development*, *72*, 1151-1162.

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings* (Technical Report C-1). Gainesville: University of Florida, Center for Research in Psychophysiology.

Cottrell, C. A., & Neuberg, S. L. (2005). Different emotional reactions to different groups: A sociofunctional threat-based approach to "prejudice." *Journal of Personality and Social Psychology*, *88*, 770-789.

Cottrell, C. A., Neuberg, S. L., & Li, N. P. (2007). What do people desire in others? A sociofunctional perspective on the importance of different valued characteristics. *Journal of Personality and Social Psychology*, *92*, 208-231.

Crandall, C. S., & Eshleman, A. (2003). The justification-suppression model of experienced and expressed prejudice. *Psychological Bulletin*, *129*, 414-446.

Crandall, C. S., Eshleman, A., & O'Brien, L. T. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology*, *82*, 359-378.

De Houwer, J., Baeyens, F., Vansteenwegen, D., & Eelen, P. (2000). Evaluative conditioning in the picture-picture paradigm with random assignment of conditioned stimuli to unconditioned stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, *26*, 237-242.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, *127*, 853-869.

Dovidio, J. F., Brigham, J. C., Johnson, B. T., & Gaertner, S. L. (1996). Stereotyping, prejudice, and discrimination: Another look. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Foundations of stereotypes and stereotyping* (pp. 276-319). New York, NY: Guilford.

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Lawrence Erlbaum.

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*, 734-754.

Edwards, K. (1990). The interplay of affect and cognition in attitude formation and change. *Journal of Personality and Social Psychology*, *59*, 212-216.

Esses, V. M., Haddock, G., & Zanna, M. P. (1993). Values, stereotypes, and emotions as determinants of intergroup attitudes. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition and stereotyping: Interactive processes in group perception* (pp. 137-166). New York, NY: Academic Press.

Festinger, L. (1957). *A theory of cognitive dissonance.* Evanston, IL: Row, Peterson.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*, 77-83.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*, 878-902.

Hamilton, D. L. (Ed.). (1981). *Cognitive processes in stereotyping and intergroup behavior*. Hillsdale, NJ: Lawrence Erlbaum.

Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: John Wiley.

Hoffman, C., & Hurst, N. (1990). Gender stereotypes: Perception or rationalization? *Journal of Personality and Social Psychology*, *58*, 197-208.

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, *33*, 1-27.

Jost, J. T., & Hunyady, O. (2003). The psychology of system justification and the palliative function of ideology. *European Review of Social Psychology*, *13*, 111-153.

Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *Journal of Abnormal and Social Psychology*, *28*, 280-290.

Kiesler, C. A. (1971). *The psychology of commitment: Experiments linking behavior to belief.* New York, NY: Academic Press.

Krosnick, J. A., Betz, A. L., Jussim, L. J., & Lynn, A. R. (1992). Subliminal conditioning of attitudes. *Personality and Social Psychology Bulletin*, *18*, 152-162.

Lang, P. J., Bradley, M., & Cuthbert, B. N. (1995). *International Affective Picture System (IAPS)*. Gainesville: NIMH Center for the Study of Emotion and Attention, University of Florida.

Livingston, R. W., & Drwecki, B. B. (2007). Why are some individuals not racially biased? Susceptibility to affective conditioning predicts nonprejudice toward Blacks. *Psychological Science*, *18*, 816-823.

Mackie, D. M., Devos, T., & Smith, E. R. (2000). Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology*, *79*, 602-616.

Macrae, C. N., Stangor, C., & Hewstone, M. (Eds.). (1996). *Stereotypes and stereotyping*. New York, NY: Guilford.

McGarty, C., Yzerbyt, V. Y., & Spears, R. (2002). *Stereotypes as explanations: The formation of meaningful beliefs about social groups*. Cambridge, UK: Cambridge University Press.

Neuberg, S. L., & Cottrell, C. A. (2002). Intergroup emotions: A biocultural approach. In D. M. Mackie & E. R. Smith (Eds.), *From prejudice to intergroup emotions: Differentiated reactions to social groups* (pp. 265-284). New York, NY: Psychology Press.

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, *12*, 413-417.

Reicher, S., Hopkins, N., & Condor, S. (1997). Stereotype construction as a strategy of influence. In P. J. Oakes & R. Spears (Eds.), *The social psychology of stereotyping and group life* (pp. 20-50). Malden, MA: Blackwell.

Rothbart, M., & John, O. P. (1993). Intergroup relations and stereotype change: A social-cognitive analysis and some longitudinal findings. In P. M. Sniderman, P. E. Tetlock, & E. G. Carmines (Eds.), *Prejudice, politics, and the American dilemma* (pp. 32-59). Stanford, CA: Stanford University Press.

Rothbart, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology*, *50*, 131-142.

Rutland, A., & Brown, R. (2001). Stereotypes as justifications for prior intergroup discrimination: Studies of Scottish national stereotyping. *European Journal of Social Psychology*, *31*, 127-141.

Schaller, M. (2008). Evolutionary bases of first impressions. In N. Ambady & J. J. Skowronski (Eds.), *First impressions* (pp. 15-34). New York, NY: Guilford.

Schaller, M., & Abeysinghe, A. M. N. D. (2006). Geographical frame of reference and dangerous intergroup attitudes: A double-minority study in Sri Lanka. *Political Psychology*, *27*, 615-631.

Schneider, D. J. (2004). *The psychology of stereotyping*. New York, NY: Guilford.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition* (Vol. 2, pp. 527-561). New York, NY: Guilford.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York, NY: Cambridge University Press.

Smith, E. R. (1999). Affective and cognitive implications of a group becoming part of the self: New models of prejudice and of the self-concept. In D. Abrams & M. A. Hogg (Eds.), *Social identity and social cognition* (pp. 183-196). Oxford, UK: Blackwell.

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 225-246). Hillsdale, NJ: Lawrence Erlbaum.

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues*, *25*, 79-97.

Yzerbyt, V. Y., Rocher, S. J., & Schadron, G. (1997). Stereotypes as explanations: A subjective essentialistic view of group perception. In R. Spears, P. Oakes, N. Ellemers, & A. Haslam (Eds.), *The psychology of stereotyping and group life* (pp. 20-50). Oxford, UK: Basil Blackwell.