ELSEVIER

Sneaking in through the back door: How category-based stereotype suppression leads to rebound in feature-based effects ☆

Sei Jin Ko [a,*], Dominique Muller [b], Charles M. Judd [c], Diederik A. Stapel [d]

[a] *Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, USA*
[b] *University of Grenoble 2, France*
[c] *University of Colorado at Boulder, USA*
[d] *Tilburg University, The Netherlands*

Abstract

Given the dramatic shifts in societal norms to curb overt stereotyping and prejudice, these biases may leak out in more subtle ways than were apparent in the past. Accordingly, we examined how the suppression of stereotypes might affect post-suppression category-based stereotyping and the more subtle feature-based stereotyping. In support of our proposition, participants in the suppression condition used more feature-based, but less category-based stereotypes in their post-suppression task than participants in the control condition. Furthermore, a relation between post-suppression category-based and feature-based stereotyping existed in the suppression condition such that decreases in category-based stereotyping were associated with increases in feature-based stereotyping. Findings as a whole suggest that norms against the expression of stereotypic biases may ironically lead people to be more vulnerable to biases as a function of within-category features.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Relationship between category- and feature-based stereotypes; Suppression; Rebound; Vocal cues; Dual effect of suppression

Over the past 50 years there has been a dramatic reduction in the overt expression of racist and sexist attitudes and stereotypes in Western societies (Campbell, 1971; Greeley & Sheatsley, 1971; Hyman & Sheatsley, 1956; Hyman & Sheatsley, 1964; Schuman, Steeh, & Bobo, 1985; Taylor, Sheatsley, & Greeley, 1978). Yet, more subtle approaches to the measurement of stereotypes and prejudice, focusing on implicit or automatic evaluations and

beliefs that are not under conscious control, suggest that racial and gender biases are still alive and well (Devine, 1989; Fazio, Jackson, Dunton, & Williams, 1995; Wittenbrink, Judd, & Park, 1997). This suggests that strong social norms have developed that encourage many people to curb the open-expression of prejudice and stereotypes based on social categories such as race and gender (Klonis, Plant, & Devine, 2005).

Certainly this effort to suppress the expression of category-based stereotypes and prejudices is not without consequences. According to the social psychological research on suppression, active efforts to suppress a thought actually increase the thought's activation and the probability that the suppressed thought may leak out on later occasions (Wegner & Erber, 1992; Wenzlaff, Wegener, & Klein, 1991). With continued practice and monitoring, of course, it may be the case that overt category-based stereotypes and prejudice never leak out, at least in their original form

* Corresponding author. Fax: +31 847 467 5700.
E-mail address: seijin-ko@northwestern.edu (S.J. Ko).

(cf. Macrae, Bodenhausen, Milne, & Jetten, 1994). Rather, the active suppression of such category-based beliefs may mean that they leak out in more subtle and seemingly harder-to-control forms.

We propose that one of these subtle ways is through within-category feature effects. Indeed, there has been recent interest in stereotyping based on physical features that vary within social categories, as well as between them (i.e., Afrocentric facial features in the case of race, Blair, Judd, Sadler, & Jenkins, 2002, and vocal features in the case of gender, Ko, Judd, & Blair, 2006). This work has demonstrated what we call feature-based stereotyping, wherein stereotypic inferences are made on the basis of such features even within social categories. So both Whites and African Americans with more Afrocentric facial features are seen as more likely to possess attributes that are stereotypic of African Americans and both males and females with more feminine voices are judged in a more stereotypically feminine manner.

As discussed below, the subtleness of feature-based stereotyping means that it may also be harder to control than category-based stereotyping. Thus, stereotype suppression may be effective at reducing category-based biases since curbing the use of category cues for judgment are relatively well practiced and therefore unlikely to show rebound effects. However, suppressing category-based stereotyping may lead to rebound in another form, with greater feature-based stereotyping following (or even during) active attempts to suppress category-based stereotyping. This is the main assertion that motivated our current work.

In line with our reasoning, recent work on racial stereotyping showed that perceivers had little trouble suppressing category-based stereotyping but were unable to suppress stereotypes arising from variations in the degree to which targets, within race, had African American (Afrocentric) features (Blair, Judd, & Fallman, 2004). Specifically, in this research, prior to a judgment task, participants were either instructed to "try avoid using the person's race as a basis for judgment" (p. 770; category suppression condition) or instructed to "try to avoid using the person's Afrocentric features as a basis for judgment" (p. 770; feature suppression condition). In a last condition, participants were given no such instructions (control condition). The judgment task results showed that in the category suppression condition participants were less likely to use targets' racial category as a basis for their judgments than in the control condition. No such effect was observed in the features suppression condition; participants' use of within-category features were unaffected by the instruction manipulation. To explain the lack of control over feature-based stereotyping, the authors argued that perceivers are largely unaware of feature-based influences. Furthermore, suppressing them is complex because feature-based judgments are literally not black or white as are category-based judgments—features are riddled with myriad shades of gray (e.g., think of

variations in nose size). This complexity is further heightened by the fact that not one but many different cues (e.g., nose size, kinkiness of hair, fullness of lips) constitute Afrocentric features, hence suppressing feature-based stereotypes would require monitoring the influence of the variation in all these cues. In other words, because within-category cues are composed of many fine-grained physical characteristics, effectively monitoring feature effects would be extremely difficult and effortful.

Recent examination of modern-day criminal sentencing also suggests a dramatic reduction in category-based outcomes such that the race of offenders seems no longer to be predictive of sentencing length, although it clearly was found to be in the past (Blair, Judd, & Chapleau, 2004). However, features impact sentencing such that offenders with more Afrocentric features received longer sentences, even when controlling for factors such as criminal record and race (Blair et al., 2004; also see Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006). Similarly, Ko, Judd, and Stapel (Submitted for Publication) found that vocal femininity features affected sex discrimination in hiring decisions such that applicants with masculine voices were rated as more competent than those with feminine voices, even while there was no discrimination based on gender itself.

In addition to its more subtle aspect, societal norms may also contribute to people's lack of control over feature effects. As discussed above, our society has exclusively focused on differences in social categories as the source of stereotyping and prejudice: societal norms that dictate against the use of within-category features as a basis for judging others are not as well defined or explicit. Consequently, people may lack the necessary practice at monitoring feature-based biases in order to control feature influences effectively.

In sum, there are clear indications that people today are quite capable of curbing blatant category-based stereotypes but not the more subtle feature-based ones. However, no research to date has examined the question of whether an actual relationship between category- and feature-based stereotyping exists. At first one could assume a positive relationship between the two, such that perceivers who do not use category-based stereotyping will not use feature-based stereotyping either. However, the literature we just reviewed leads us to doubt such a relationship. To put it another way, it does not logically follow that perceivers who find it easy to avoid using category cues will also necessarily find it easy to avoid using features as a basis for stereotyping. No relationship between these two bases of stereotyping thus seems a reasonable prediction. Based on the stereotype suppression literature, however, we propose a negative relationship. More precisely, when perceivers are made to control for the use of category-based biases then, the more they continue to suppress their category-based stereotyping, the more they will use feature-based stereotyping.

Our proposal relies on the fact that suppressing stereotypes is known to ironically increase their accessibility (Mac-

rae et al., 1994). Insofar as people continue to control their use of stereotypes, this increased accessibility remains unseen. However, once the pressure to control stereotypes is relaxed this increased accessibility manifests itself in the form of the so-called rebound effect (cf. Macrae et al., 1994). To the best of our knowledge, this decrease in the pressure to control has always been presumed to take place sometime after suppression. We believe, however, that the process of stereotype suppression and rebound does not necessarily have to follow such a sequential timeline. Indeed, if, as is the case with feature-based stereotyping, people are not aware that they are using such stereotypes, then they will not exert control over their use of these stereotypes. Hence, it follows that heightened accessibility could manifest itself (through rebound in features-based stereotyping) while people are simultaneously suppressing stereotypes (through control of category-based stereotyping). We argue that an examination of this topic is particularly timely and important, given the dual presence of strong norms in our society not to discriminate and the, nevertheless, strong and persistent experience that discrimination is alive and well. Perhaps the norms that dictate against category-based discrimination actually lead to stronger feature-based discrimination.

To test this relationship between category- and feature-based stereotyping, we adopted a particularly novel method where stereotype suppression and rebound was examined via two different modalities. Namely, we manipulated suppression of category-based stereotyping through the visual modality then measured its effects on stereotyping in the auditory modality. Traditionally, researchers have examined stereotype suppression and rebound as it relates only to visual stimuli, which perhaps represents an underlying presumption that suppression and rebound only occurs within the same modality. Furthermore, past research has solely focused on category-based suppression and rebound (e.g., Macrae et al., 1994; Wyer, Sherman, & Stroessner, 2000). As far as we are aware, we are the first to propose and show that suppression originating from one modality not only affects judgments based on another modality but also that its effect can be dual, such that suppression of category-based stereotyping in the visual modality can affect both category-based and within-category cue-based stereotyping in the auditory modality.

In sum, the present research investigated the relationship between category- and feature-based stereotyping within the auditory domain. In particular, we examined whether suppressing the use of gender-category stereotypes in one task affected stereotyping both on the basis of voice gender category and variations in within-gender vocal femininity in another task. We predicted that the suppression manipulation would have a dual effect such that it would lead participants to inhibit (as compared with a control condition) the use of category cues while increasing the use of feature cues as a basis for stereotyping. Moreover, this dual effect in the suppression condition should show a negative correlation such that the more participants control the use of category the more they rely on feature cues.

Experimental methods

Stimuli for probability

Self-descriptions

Four descriptions developed by Ko et al. (2006, Study 3) were used in this experiment. These varied on the degree to which the depicted person had gender stereotypic attributes and the degree to which the person was likeable. An example of the positively valenced female stereotypic description was "*As an elementary school teacher, I like to create an environment where students learn to cooperate and build self-confidence. An essential part of doing this is not to have favorites, but rather to give more care and attention to the children who are more shy and reticent. I make myself available even outside of the classroom if any one of them should need my help*".[1]

Recorded voices

Forty-five female and 40 male voices saying the "Rainbow Passage" (Fairbanks, 1940, p. 127) were scaled on vocal femininity.[2] Based on the results, 15 male and 15 female voices were chosen.

Participants

Thirty-eight male and 138 female University of Groningen students participated in exchange for course credit or monetary compensation.

Procedure

Participants were told that they would complete a number of short unrelated tasks. We employed a standard suppression procedure (see Macrae et al., 1994) where the first task was explained as an investigation of patterns in how perceivers write about others. Accordingly, participants were told that they would be given 10 min to write about a day in the life of a male and a female target depicted in two photographs. Furthermore, those in the suppression condition were told that they should avoid gender stereotypes and be careful that these not affect what they wrote. Those in the control condition were not given these extra instructions. Participants were then put in separate cubicles with an envelope containing the first target photograph and a sheet of writing paper. Those in the suppression condition had a written reinforcement to suppress gender stereotypes at the top of their writing paper. The gender of the first target randomly varied between participants. After 5 min, partici-

---

[1] The English descriptions were translated into Dutch and can be obtained from the first author. Further details about the descriptions can be found in Ko et al. (2006).

[2] For a detailed description of the vocal femininity scaling process, criteria used to select the final subset of voices, and procedure of the probability task see Ko et al. (2006).

pants were given a second envelope, with a photograph of the second target whose gender was always opposite to the gender of the first target, a new sheet of writing paper, and another 5 min to write. Hence, each participant saw (and wrote about) both the male and female photographs. Following this writing task, participants engaged in a probability judgment task, similar to that used by Ko et al. (2006). This task was presented as an interpersonal accuracy task, where they had to match the self-description of a target with the correct speaker. Participants were randomly assigned to read one of the four self-descriptions before listening to the 30 voices, presented in random order. For each, they estimated the probability, on a 0–100% scale, that the person speaking was the one who wrote the self-description. Since any one of the voices could be the correct speaker, participants were allowed to assign the same probability value to more than one speaker. After the probability task, participants completed a final unrelated questionnaire.

Analysis

We expected probability judgments to be related to gender category and within-gender vocal femininity to the extent that participants were making their judgments based on category membership and/or within-category vocal femininity cues. Further we were interested in whether the suppression manipulation moderated these relationships.[3] Our analysis involved a multilevel modeling procedure, estimating a separate regression model for each participant (see Blair et al., 2002; Ko et al., 2006). In these regressions, participants' probability ratings for the 30 voices were regressed onto two predictors: (a) contrast-coded gender category (+1 if female, −1 if male) and (b) degree of vocal femininity. The resulting two partial regression slopes per participant—one assessing the partial relationship between probability ratings and gender category and the other assessing the partial relationship between probability ratings and vocal femininity—became the two measures which we subsequently analyzed. These partial slopes were subjected to a 2 (self-description stereotypicality: female vs. male) × 2 (self-description valence: positive vs. negative) × 2 (condition: suppression vs. control) between-participant ANOVA.

Partial slopes for gender category in the case of female stereotypic self-descriptions should be positive to the extent that category-based stereotypes are guiding judgments because for female stereotypic inferences the probability ratings should be higher for female voices (coded +1),

regardless of vocal femininity. On the other hand, partial slopes for gender category in the case of the male stereotypic self-descriptions should be negative to the extent that category-based stereotypes are guiding judgments because for male stereotypic inferences the probability ratings should be higher for male voices (coded −1), regardless of vocal femininity. Consequently, evidence for category-based stereotyping necessitates that the average category slope for female stereotypic descriptions is positive and that the average category slope for male stereotypic descriptions is negative. Additionally, greater category stereotyping would be indicated by a greater difference between these two slopes (more positive for female stereotypic descriptions and more negative for male stereotypic ones). Accordingly, our specific prediction for category-based stereotyping would be confirmed by a significant self-description stereotypicality by condition interaction since we expected the difference between the mean category slopes for female and male descriptions to be smaller under suppression (i.e., less category-based stereotyping) than control condition.

Partial slopes for vocal femininity in the case of the female stereotypic self-descriptions should be positive to the extent that feature-based stereotypes are guiding judgments because for female stereotypic inferences the probability ratings should be higher for voices that are higher on vocal femininity, regardless of the speakers' gender category. On the other hand, partial slopes for vocal femininity in the case of the male stereotypic self-descriptions should be negative to the extent that feature-based stereotypes are guiding judgments because for male stereotypic inferences the probability ratings should be higher for voices that are lower on vocal femininity. Consequently, evidence for feature-based stereotyping necessitates that the average vocal femininity slope for female descriptions is positive and the average vocal femininity slope for male descriptions is negative. Additionally, greater feature-based stereotyping would be indicated by a greater difference between these two slopes (more positive for female stereotypic descriptions and more negative for male stereotypic ones). Accordingly, our specific prediction for feature-based stereotyping would again be confirmed by a significant self-description by condition interaction since we expected the difference between the mean femininity slopes for female and male descriptions to be larger under suppression (i.e., more feature-based stereotyping) than control conditions. Note that the direction of this interaction should be opposite to the one involving category slopes because suppression should decrease category-based stereotyping but increase feature-based stereotyping.

In sum, the two self-description by condition interactions—one with partial category slopes as the dependent variable and the other with partial vocal femininity slopes as the dependent variable—directly test our critical hypotheses concerning category- and feature-based stereotyping as a function of the suppression manipulation.

---

[3] A coder, blind to the experimental condition, rated the content of the participants' writings using a 1 (not at all gender stereotypic) to 7 (very gender stereotypic) scale. The result showed that participants' writings in the suppression condition was indeed less stereotypic ($M = 3.99$, $SD = .70$) than participants' writings in the control condition ($M = 4.31$, $SD = .63$), $t(173) = 3.07$, $p < .01$.

Results

Category use

Recall that for the measure of category use, the larger the difference between the mean category slopes for female and male descriptions, the more participants made use of category as a basis for stereotyping, controlling for vocal femininity. Presented in Fig. 1 are the mean category slopes, broken down by condition and self-description stereotypicality. As one would expect, the ANOVA revealed a description stereotypicality main effect, $F(1, 168) = 180$, $p < .001$, indicating that participants, on average, did use the speakers' gender category to make stereotypic inferences. More central to our purpose was the predicted description stereotypicality by condition interaction, $F(1, 168) = 6.97$, $p < .01$, showing that the difference between the mean positive and negative category slopes was significantly smaller in the suppression than control condition. This interaction demonstrated that participants who were told to avoid using gender stereotypes in the first writing task (suppression condition) made less use of the voices' gender category in the subsequent probability task to make stereotypic inferences than participants in the control condition.[4,5] In other words, this finding confirmed that participants did indeed curb their category-based stereotyping as a function of the suppression manipulation.

Although our critical test was the stereotypicality by condition interaction, additional simple effects tests revealed that for female stereotypic self-descriptions participants in the control condition ($M = 9.91$) used gender category marginally more than participants in the suppression condition ($M = 7.47$), $F(1, 168) = 2.99$, $p < .086$. For male stereotypic self-descriptions participants in the control condition ($M = -6.21$) used gender category significantly more than participants in the suppression condition ($M = -3.34$), $F(1, 168) = 4.04$, $p < .047$.[6]

Fig. 1. Category use as a function of condition and stereotypicality of self-descriptions (controlling for vocal femininity).

Vocal femininity use

Recall that for the measure of vocal femininity use, the larger the difference between the mean vocal femininity slopes for female and male descriptions, the more participants made use of features as a basis for stereotyping, controlling for category. Presented in Fig. 2 are the mean femininity slopes, broken down by condition and self-description stereotypicality. Again, the ANOVA revealed a stereotypicality main effect, $F(1, 168) = 166$, $p < .001$, showing that participants, on average, did make use of features to make stereotypic inferences. More central to our reasoning was the predicted description stereotypicality by condition interaction, $F(1, 168) = 5.34$, $p < .023$, showing that the difference between the mean positive and negative vocal femininity slopes was significantly larger in the suppression than control condition. This interaction demonstrates that participants who were told to suppress gender stereotypes in the first writing task made more use of vocal femininity in making their subsequent probability ratings than did participants in the control condition.[7] In other words, this finding confirmed that participants did indeed rebound in their feature-based stereotyping as a function of the suppression manipulation.

Again, in addition to the critical interaction of interest, we performed simple effects tests. These revealed that for female stereotypic self-descriptions participants in the suppression condition ($M = 3.55$) made more use of vocal fem-

---

[4] Our results also revealed effects of less theoretical interest such as a valence main effect, $F(1, 168) = 55.95$, $p < .001$. This valence effect was moderated by stereotypicality, $F(1, 168) = 19.80$, $p < .001$, revealing that the difference between female and male negative self-descriptions ($M = 14.67$ vs. $M = -3.26$) was larger than the difference between female and male positive self-descriptions ($M = 2.71$ vs. $M = -6.30$). Finally, the condition by valence interaction, $F(1, 168) = 5.43$, $p < .021$, showed that the difference between negative and positive self-descriptions was larger in the control condition ($M = 6.77$ vs. $M = -3.07$) than in the suppression condition ($M = 4.64$ vs. $M = -0.52$).

[5] Consistent with past work on vocal stereotyping (Ko et al., 2006) we found no participant gender effects here or in any of the following results.

[6] Given that one of these simple effects is significant and the other is marginally significant, one might wonder whether the suppression effect for the female descriptions (i.e., mean partial slope in control of 9.91 vs. mean partial slope in suppression of 7.47) was smaller than the suppression effect for male description (i.e., mean partial slope in control of −6.61 vs. mean partial slope in suppression of −3.34). Reversing the signs of the slopes for the male stereotypic descriptions enables us to examine this. When this analysis is conducted there is no evidence that the suppression effect is different in magnitude for the female vs. the male self-descriptions, $t(168) = .21$, $p < .84$.
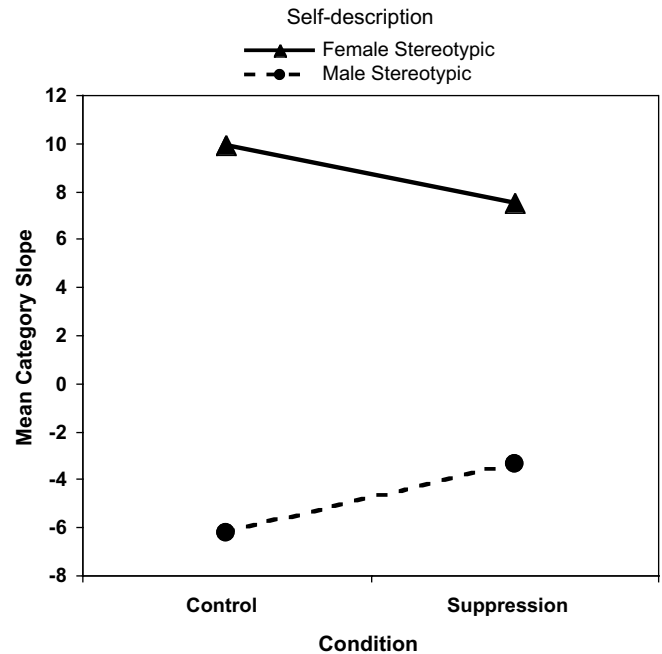
[7] Of less theoretical relevance was a valence main effect, $F(1, 168) = 4.28$, $p < .041$, indicating that more feminine voices were rated as more probable for negative self-descriptions ($M = -1.49$) than for positive ones ($M = -2.88$).
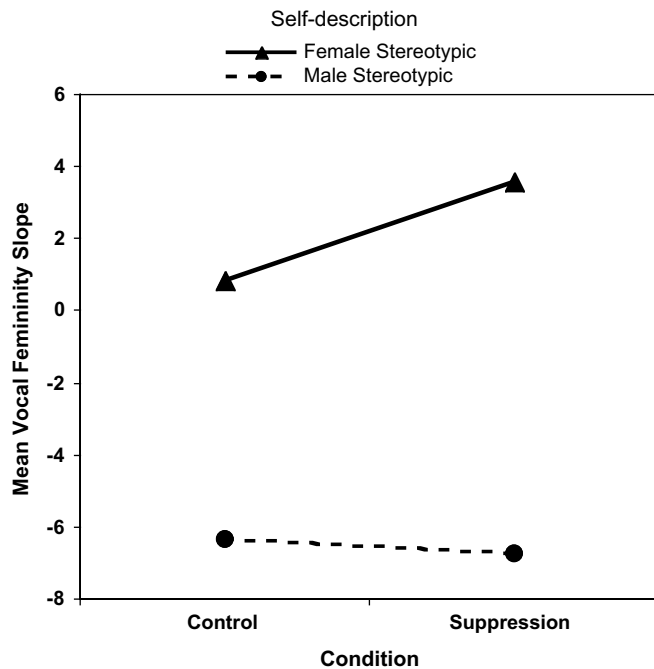
Fig. 2. Vocal femininity use as a function of condition and stereotypicality of self-descriptions (controlling for gender category).

Discussion

The goal of the current research was to investigate how suppressing category-based stereotypes might affect stereotyping due to variations in within-category cues. First, consistent with our reasoning that people may have become quite practiced at suppressing category-based stereotypes and that categorical cues are relatively simple to monitor, we found that participants who were told to suppress category-based stereotypes in the writing task continued to do so in the probability task. Second, we found strong support for our hypothesis that such suppression heightened the use of within-category cues as a basis for stereotyping.

Central to our purpose was the relation between category- and feature-based stereotyping. In support of our reasoning, the correlations revealed that when participants were told to suppress the use of stereotypes (i.e., in the suppression condition) a relation existed such that the less these participants used category cues the more they used within-category features in their judgments.

The research reported earlier by Blair and colleagues importantly outlined the ways in which stereotyping could be based not only on social categories but also on features associated with category membership that vary within those categories (Blair et al., 2002; Ko et al., 2006). They also demonstrated that these feature-based stereotyping effects were relatively subtle in the sense that participants did not seem to be able to suppress them when instructed to do so (Blair et al., 2004). The present set of results importantly extend these conclusions by examining the relationship between the two sorts of stereotyping—category-based and feature-based— and how efforts to suppress the former, while successful, may actually increase feature-based stereotyping. We would suggest that these results are particularly timely, given the widely-shared and largely effective norms that encourage people to avoid the expression of category-based stereotypes. Ironically, our work suggests that the adoption of such norm-prescribed behavior may end up increasing the extent to which one stereotypes others based on those features that are associated with category membership but that vary within categories.

A number of novel aspects of this research set it apart from past work on suppression and rebound. First, this is the first demonstration of cues involved in suppression (i.e., between-category cues) being different from the cues involved in rebound (i.e., within-category features rather than the same between-category cues involved in suppression). Second and most importantly, perceivers' use of these two cues for stereotyping allowed us to demonstrate both suppression (through the use of between-category cues) and rebound (through the use of within-category features) simultaneously. Furthermore, these two aspects provide new insight into our understanding of stereotype suppression and rebound by revealing a somewhat paradoxical effect. That is, suppressing stereotypes associated with a given category (e.g., stereotypes about females) can lead to rebound that not only affects members of that

ininity than participants in the control condition ($M = 0.83$), $F(1, 168) = 8.12$, $p < .005$. For male stereotypic self-descriptions participants in the suppression condition ($M = -6.77$) tended to make more use of vocal femininity than participants in the control condition ($M = -6.36$), although not significantly so, $F(1, 168) < 1$.[8]

Correlational analyses

As can be seen in Figs. 1 and 2, results observed on the category and femininity slopes at the mean level seem to suggest opposite effects of our suppression manipulation: suppression decreased category use but increased femininity use. As a way to investigate whether there was a relationship between the use of category and femininity, we computed partial correlations (controlling for stereotypicality and valence) between these two measures. These analyses revealed no correlation in the control condition, ($r = .10$, $p < .38$). However, in the suppression condition, the less participants used category gender, the more they used vocal femininity features ($r = -.33$, $p < .01$).

---

[8] Given that one of these simple effect is not significant, one could wonder if the rebound effect for female description (i.e., mean partial slopes of 0.83 vs. 3.55, for control and suppression, respectively) was larger than the rebound effect for male description (i.e., − 6.36 vs. − 6.77). Reversing the sign of these partial slopes for the male descriptions enables us to answer this question. Doing this shows that the suppression effect for the male descriptions is not significantly different from the suppression effect for the female descriptions, $t(168) = 1.70$, $p < .09$. In other words, the rebound effect does not significantly depend on whether the probability ratings were made for the stereotypically female descriptions or the male ones.

category (i.e., females) but, paradoxically, also affects members of the opposite category (i.e., males). Third, we showed rebound effects in a totally different physical modality from the one in which suppression was introduced. That is, we manipulated the suppression instructions through the visual modality using photographs and a writing task, whereas we presented cues and observed rebound in the auditory modality (i.e., as a function of listening to voices). Some may think that Macrae et al. (1994, Experiment 2) were the first to demonstrate suppression and rebound in two different modalities because their suppression task involved writing about a target skinhead depicted in a photograph whereas rebound was measured behaviorally. However, we would argue that this is not entirely accurate since both the stimuli used in their suppression task (i.e., photographs) and rebound task (i.e., jacket and bag supposedly belonging to the target skinhead) were visual ones. Hence, in essence, the cues used to trigger both suppression and rebound were of the same visual modality.

Affirmative action and equal opportunity laws are some examples of how greatly Western society has advanced in terms of attempting to neutralize category-based influences that may unfairly bias judgments. However, biases still exist. Our findings suggest that we can no longer rely exclusively on categorical differences to elucidate why this is the case. Instead, a more complete understanding of the biases that occur must necessarily include examinations of how the more complex yet subtle within-category cues exert their influences on judgment and behavior. Our research imparts perhaps an even more sobering message; feature-based biases do more than merely exist—they can be exacerbated by norms (and legislation) that appropriately dampen category-based stereotyping and discrimination.

References

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of afrocentric facial features in criminal sentencing. *Psychological Science, 15*, 674–679.

Blair, I. V., Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and afrocentric facial features in social judgments. *Journal of Personality and Social Psychology, 87*, 760–778.

Blair, I. V., Judd, C. M., Sadler, M. S., & Jenkins, C. (2002). The role of Afrocentric features in person perception: Judging by features and categories. *Journal of Personality and Social Psychology, 83*, 5–25.

Campbell, A. (1971). *White attitudes toward black people*. Ann Arbor, MI: Institute for Social Research.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5–18.

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.

Fairbanks, G. (1940). *Voice and articulation drillbook*. New York: Harper & Brothers, p.127.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobstrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013–1027.

Greeley, A. M., & Sheatsley, P. B. (1971). Attitudes toward racial integration. *Scientific American, 225*, 13–19.

Hyman, H. H., & Sheatsley, P. B. (1956). Attitudes toward desegregation. *Scientific American, 195*, 35–39.

Hyman, H. H., & Sheatsley, P. B. (1964). Attitudes toward desegregation. *Scientific American, 211*, 16–23.

Klonis, S. C., Plant, E. A., & Devine, P. G. (2005). Internal and external motivation to respond without sexism. *Personality and Social Psychology Bulletin, 31*, 1237–1249.

Ko, S. J., Judd, C. M., & Blair, I. V. (2006). What the voice reveals: Within- and between-category stereotyping on the basis of voice. *Personality and Social Psychology Bulletin, 32*, 806–819.

Ko, S. J., Judd, C. M., & Stapel, D. A. (Submitted for Publication). A voice in the workplace: How vocal cues impact judgments related to hiring decisions.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67*, 808–817.

Schuman, H., Steeh, C., & Bobo, L. (1985). *Racial attitudes in America: Trends and interpretations*. Cambridge, MA: Harvard University Press.

Taylor, D. G., Sheatsley, P. B., & Greeley, A. M. (1978). Attitudes toward racial integration. *Scientific American, 238*, 42–49.

Wegner, D. M., & Erber, R. (1992). The hyperaccessibility of suppressed thoughts. *Journal of Personality and Social Psychology, 63*, 903–912.

Wenzlaff, R. M., Wegener, D. M., & Klein, S. B. (1991). The role of thought suppression in the bonding of mood and thought. *Journal of Personality and Social Psychology, 60*, 500–508.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72*, 262–274.

Wyer, N. A., Sherman, J. W., & Stroessner, S. J. (2000). The roles, of motivation and ability in controlling the consequences of stereotype suppression. *Personality and Social Psychology Bulletin, 26*, 13–25.