HANDBOOK OF PREJUDICE, STEREOTYPING, AND DISCRIMINATION

EDITED BY TODD D. NELSON
When Mary Shelley published her novel, *Frankenstein*, in 1818, she did so anonymously, perhaps to disguise her gender. When it was later discovered that the author was, indeed, a young woman, one reviewer wrote about the novel, “For a man, it was excellent, but for a woman, it was wonderful” (*Blackwood’s*, 1823, as cited in Hindle, 1985). This comment reflects the explicit acknowledgment that gender can affect the standards against which a work product is evaluated. In this case, the standard for women is likely lower than the standard for men (work product is expected to be less good in women than in men), but it is also qualitatively different. Although both “excellent” and “wonderful” signify greatness, “wonderful” suggests something astonishing—perhaps especially so given the philosophical and violent nature of the novel.

Subjective descriptors such as “wonderful” or “excellent” are always used with reference to some standard (Kraut & Higgins, 1984). When we describe a package as “huge” or a day as “hot,” there is an assumption that the audience of our communication will understand roughly what those terms mean, objectively. The huge package is presumably not so huge that it does not fit in a car’s trunk, and the hot day is hot relative to expectations; hot in Tucson is very different than hot in Anchorage.1 This slipperiness of subjective language is particularly interesting as it applies to descriptions of people, as one likely referent for these descriptions is the social category membership of the individual being described. As in Mary Shelley’s case, gendered expectations can provide the framework against which descriptions can be interpreted, as can expectations based on race, age, or any number of other features of the person. Although we typically do not include the tag line, “ . . . for a woman” or “ . . . for an African American,” our impressions and descriptions of others are likely to be based, in part, on reference to the group stereotype as a judgment standard.

This is the basic premise behind the shifting standards model of stereotype-based judgment (Biernat, 2003; Biernat & Manis, 1994; Biernat, Manis, & Nelson, 1991). According to this model, stereotypes—generally defined as expectations about the attributes of a group—provide judgment standards for evaluating individual group members. Because different groups have different stereotypes associated with them, standards shift depending on the social category membership of the individual being judged. Thus, a woman and a man who engage in identical behavior might be described differently because different standards have been invoked, or the same subjective description might mean something substantively different because it was made with reference to shifting standards. For example, because of shifting standards, a woman who steals a parking space might be rated a 6 on a 7-point scale of aggressiveness whereas a man who does the same might be described as a 5. Additionally, a man rated a 5 in aggressiveness would likely be judged to be more objectively aggressive than a woman rated identically.

In this chapter, I first review evidence documenting the tendency to shift standards when judging individual group members on stereotyped dimensions. I then consider how this tendency plays out in communication and the translation of others’ judgments. Next, I examine the effects of stereotypes on the setting of standards, and how those standards affect behavior toward individual members of stereotyped groups. The chapter concludes with a consideration of the link between shifting

---

1 As I write this during a Kansas summer, a day with an expected high of only 87 degrees is being described as “cool.”
standards and outgroup prejudice, and highlights the complexity inherent in charting the impact of stereotypes on social life.

**SHIFTING STANDARDS IN SOCIAL JUDGMENT**

A schematic depiction of the basic idea behind the shifting standards model is presented in Figure 7.1. The model is particularly concerned with understanding how stereotypes affect judgments of individual members of stereotyped groups on stereotyped dimensions. In this example, we assume that a perceiver holds the stereotype that men are better leaders than women. The perceiver also knows some information about an individual male or female target, in this case, that the target is a manager (i.e., in a leadership role). Now the perceiver is asked, “How skillful a leader is Katherine/Kenneth?”

One key assumption of the model is that the stereotype provides a standard against which the perceiver can make this judgment. Stereotypes include representations of the mean level of an attribute that members of a given group possess, as well as a likely range that members of the group will exhibit (Judd & Park, 1993). In this way, stereotypes serve an “endpoint setting” function: They allow perceivers to fix the endpoints of a subjective rating scale to reflect the expected distribution of a class of targets on the dimension of interest (see this theme in the classic judgment models of Parducci, 1963; Upshaw, 1962; and Volkmann, 1951). Thus, in Figure 7.1, the standard for women’s leadership is lower than that for men, some range of leadership skill is expected within each group (the degree of variability within men and women is equated in this example), and there is some overlap in the expected distributions.

The result of this differential standard setting, along with a stereotyped expectation about Katherine’s and Kenneth’s leadership skill, is that Katherine is judged a subjectively better leader than Kenneth (she receives a mean rating above 5; he receives a 4.9). This is a contrast effect, as it reflects an apparently counterstereotypical pattern of judgment. However, the shifting standards argument suggests this contrast is more apparent than real—it is based on the male and female target

![Schematic depiction of stereotypes influencing judgment standards and evaluations of male and female targets.](image_url)

**FIGURE 7.1** Schematic depiction of stereotypes influencing judgment standards and evaluations of male and female targets.
being judged in units with very different, gender-specific, meaning. One way we can determine that this is the case is by comparing judgments made on these “slippery” subjective rating scales to those made on a rating scales that are anchored in some external reality. We have typically referred to such scales as objective or common rule in nature, in that the units of judgment mean the same thing, regardless of attributes of the target being judged. In the example of heat discussed earlier, “hot” and “cool” judgments are subjective, but estimated temperature in degrees Fahrenheit is an objective or common rule index. In the example presented in Figure 7.1, a common rule judgment of leadership skill might include assigning a “letter grade” to indicate the target’s performance. By these metrics, Tucson would pretty solidly be judged “hotter” than Anchorage, and Kenneth would be judged a better leader than Katherine. These are classic examples of assimilation to stereotypes, and such effects tend to emerge most strongly on objective or common rule response scales that do not allow for shifts in meaning from one target to another (Tucson to Anchorage; Katherine to Kenneth).

Based on these examples, the “signature” shifting standards effect is evidence of assimilation to stereotypes on common rule response scales, but reductions or reversals of this pattern when judgments are made in subjective units. Statistically, this is a target category X response scale interaction, and research from my laboratory has documented such a pattern for a number of social groups on a number of stereotyped dimensions. For example, in one study, participants viewed 40 photographs of men and women and were asked to judge their “financial success.” Half of the participants made these estimates in objective units (dollars earned per year), whereas the other half made subjective judgments of financial success on a scale ranging from 1 (financially unsuccessful) to 7 (financially successful). Objective judgments clearly revealed assimilation to the (accurate) stereotype that men earn more money than women, but subjective judgments revealed a reliable contrast effect: Women were judged more financially successful than men (Biernat et al., 1991, Study 2).

The scatter plot depicted in the top panel of Figure 7.2 presents the mean objective and subjective financial success ratings attributed to each of the 40 photographs. It is clear that separate regression lines can be fit to the sets of male and female targets, such that a woman could earn objectively less money than a man to achieve any given subjective rating. For example, for a woman to be rated a 4 on financial success, she could earn about $9,000 less per year than a man with the same rating. For comparison purposes, the bottom panel of Figure 7.2 presents a scatter plot based on the same 40 targets, in this case judged on age in either objective (years) or subjective (young–old) units. There is no stereotype that “men are older than women,” so standards need not shift when one is judging men versus women on this dimension. Indeed, as the scatter plot indicates, a single regression line captures the relationship between subjective and objective age judgments, across female and male targets. This finding supports the idea that differential group stereotypes are necessary to trigger the use of shifting standards to judge individual group members (Biernat et al., 1991).

Beliefs about men’s better financial standing relative to women’s may strike some as only tangential to understanding stereotyping. After all, this belief is based in reality, and it does not have the pernicious quality that many social stereotypes have. Additional research from our laboratory, however, has demonstrated that a comparable pattern of shifting standards emerges when one considers the kinds of stereotypes that are of greater social concern. For example, evidence of shifting standards emerges in judgments of men’s and women’s job-related competence, verbal ability, writing quality, athleticism, and leadership competence (as well as height, weight, and income; Biernat, Crandall, Young, Kobrynowicz, & Halpin, 1998; Biernat & Manis, 1994; Biernat et al., 1991; Biernat & Vescio, 2002). Similarly, evidence of assimilation to racial stereotypes is stronger when judgments are made in common rule rather than subjective units in the domains of verbal and math ability, athleticism, and job-related competence (Biernat & Kobrynowicz, 1997; Biernat & Manis, 1994; Kobrynowicz & Biernat, 1997). For example, when judging the athleticism of Black and White targets, rankings (a common rule indicator that invites cross-category judgment) produced stronger evidence than subjective ratings that Blacks were perceived as more athletic than Whites. Furthermore, when Black and White targets were “tied” in terms of the subjective athleticism ratings they received, rankings clearly indicated that the Black targets were seen as objectively more athletic than the similarly rated
White targets (Biernat & Manis, 1994, Study 3). In sum, across a variety of domains, gender and racial stereotypes prompt the use of shifting standards to judge individual targets. Recent research has also documented that the tendency to shift standards remains stable across the life span (at least in the domain of height judgments; Hoessler & Chasteen, 2008), but is enhanced under conditions of cognitive busyness (Biernat, Kobrynowicz, & Weber, 2003).

Evidence consistent with the shifting standards perspective has also emerged in research on judgments of “real” people, as in the literature on workplace performance appraisal. In a recent meta-analysis examining the effects of employee race on the evaluations they receive from supervisors, Roth, Huffcutt, and Bobko (2003) found that race bias is greater on objective than subjective indicators. For example, for measures of job quality, the effect size (Whites evaluated more favorably than Blacks) was $d = .24$ for objective measures (objective indicators of work product, errors, complaints) and $d = .20$ for subjective measures (subjective ratings of quality). For measures of job quantity, the objective $d = .32$ (e.g., number of units produced, sales volume), and the subjective $d = .09$. On measures of job knowledge (which included ratings [subjective] or tests [objective] of mastery of training material) the objective $d = .55$; subjective $d = .15$. On measures of absenteeism, the objective $d = .23$; subjective $d = .13$. Several studies also indicated comparable patterns with regard to White–Hispanic

![Figure 7.2](image-url)
differences (for job knowledge, objective $d = .67$, subjective $d = .04$). Furthermore, in the two studies reviewed in which supervisors made both objective and subjective judgments of the same employees, the same pattern emerged (Campbell, Crooks, Mahoney, & Rock, 1973).

In short, judgments of real employees also indicate that evaluation standards may shift based on the target’s race, and that the observed pattern is consistent with the shifting standards model predictions about the differences between objective (common rule, cross-category) and subjective (within-category) appraisals. In research on gender-based stereotypes of leadership competence, U.S. Army captains’ judgments of each other also demonstrated a pattern consistent with shifting standards—evidence that men were perceived as better leaders than women was stronger in rankings (a cross-category judgment) than subjective ratings (Biernat, Crandall, et al., 1998). Both race and gender stereotypes seem to invoke the use of different, shifting standards to make subjective judgments.

**TRANSLATION AND THE COMMUNICATION OF SUBJECTIVE LANGUAGE**

When we talk to other people about other people, our conversation is peppered with subjective language. We might discuss the “really tall” woman we saw, or the “obnoxious” sales clerk, or the “aggressive” driver we encountered. Given our facility with making within-category adjustments of meaning for object descriptions (e.g., large cats vs. large elephants; fast cars vs. fast bicycles), one question is whether we do the same for evaluations of different groups of people. Do listeners “decode” subjective language in a manner that takes into account stereotype-based shifting standards?

We know from studies on height estimation that when asked “How tall is tall?” respondents indicate a height of about 6’3” for men, but only 5’9” for women (Roberts & Herman, 1986). Does the same within-category translation occur with respect to nonphysical attributes of the sort that characterize most social stereotypes? In one relevant study, participants listened to an audiotape of a man or woman describing himself or herself as a “very good” or “all right” parent (Kobrynowicz & Biernat, 1997, Study 2). Participants were asked to “decode” those descriptions by estimating the frequency with which the parent engaged in a wide variety of parenting behaviors (including physical care, emotional care, engagement in play, etc.). At both levels of quality (“very good” and “all right”), mothers were perceived to have objectively more involvement than fathers. That is, “very good” for a mom translated into more than twice as much physical care (diaper changes, baths, meal preparations) as “very good” for a dad; in fact, “all right” mothers were estimated to engage in slightly more physical care of children than “very good” fathers. In another study, “good at math” translated into a higher objective grade point average (GPA) performance for Asian than for White or Black students (Kobrynowicz & Biernat, 1997, Study 3). Identical subjective language in these cases was interpreted, or decoded, to mean objectively more evidence of the attribute (involved parenting, math skill) among individuals stereotyped as having the attribute (women, Asian students).

Subjective language is also prevalent in an important means of written communication—the letter of recommendation. Such letters are often key to admission and hiring decisions (e.g., Brelend, 1983; Lopez, Oehler, & Moberly, 1997), and a number of studies have examined whether their content may or may not be biased against women (e.g., Bronstein, Black, Pfennig, & White, 1986; Colarelli, Hechanova-Alampay, & Canali, 2002; Lunneborg & Lillic, 1973). Yet little is known about how these letters are interpreted by others, and whether equivalent content used to describe women and men is decoded differentially. In a series of studies, Biernat and Eidelman (2007) exposed participants to a favorable letter of recommendation supposedly written for a man or a woman applying to a graduate program in physics (a male domain in which women would likely be stereotyped as less competent than men). Participants were asked to translate what the professor writing the letter meant about the student’s qualifications, and what they themselves thought about the student’s qualifications (estimates were made in objective units such as estimated GPA and Graduate Record Exam (GRE) scores, as well as subjective ratings). Consistent with the shifting standards model, the woman about whom favorable things were written was assumed to be less
academically accomplished than the comparable man (both in translating the professor’s thoughts and estimating one’s own).

However, in two other conditions of the study, we provided participants with additional information about the male professor—his gender attitudes. Specifically, in one condition, participants were led to believe the professor was sexist and that he did not think women were as competent in science as men; in another condition, the professor was portrayed as antisexist, and he viewed women and men as equally competent in physics. We expected these manipulations would be suggestive of lower or higher standards for women, respectively. That is, a sexist professor would be assumed to have low expectations for women; a nonsexist professor would be expected to have higher expectations. Thus, according to a standards-based prediction, positive comments about a woman would be translated to mean lower academic standing when provided by a sexist, and equal or higher academic standing when provided by a nonsexist writer (relative to the male student). These predictions are at odds with what might be predicted according to attributional rules of augmenting and discounting (Kelley, 1971). According to this perspective, if a sexist says something nice about a woman, this must mean she is really good (having overcome the writer’s negative inclinations), and if an antisexist says something nice, she might not be all that good (the positive views of the writer provide a discounting cue).

Results were consistent with the standards-based interpretation: Relative to the male student, the female student was assumed to be significantly less objectively competent in the sexist writer condition, and nonsignificantly more competent in the antisexist writer condition (Biernat & Eidelman, 2007). That is, knowledge of the writer’s sexism enhanced the pattern of translation that emerged in the control condition of the study. At the same time, however, respondents seemed to distance their own views from those of the sexist professor: Although they “translated” the positive letter to mean less objective quality for the female than male student, they themselves thought more highly of the female student. This suggests that individuals may understand others’ communications by referencing the likely standards the communicator employed to make judgments; however, they may reject those judgments when the perspective of the communicator is viewed as inappropriate, or as a “mental contaminant” (Wilson & Brekke, 1994).

Recent work in my lab has also explored dyadic communication chains, in which one partner receives “objective” information about a target individual, then communicates a subjective impression to another individual, who then “back-translates” the subjective impression into (perceived) objective standing of the target. In a doctoral dissertation, Collins (2006) reported a series of studies in which “communicators” viewed academic performance information (e.g., a college transcript) that was attributed to a Black or White male student. Their task was to communicate their impression of the student, in writing, to another study participant. Consistent with the shifting standards model, subjective impressions were more favorable in the case of the Black than the White student. Yoked participants were then asked to read the communications and to estimate the objective academic attributes of the student (e.g., GPA, American College Testing [ACT] scores, etc.); that is, to reproduce key parts of the college transcript. Despite the greater positivity of the impressions formed of the Black student, back-translations either revealed no difference in perceived Black–White standing, or in two studies, judgments that the Black student was academically worse than the White student.2

That is, equal objective standing was communicated in more favorable terms when the student was Black than White. This could be based on shifting standards (“for a Black student, this record is pretty good”), but admittedly could also reflect the operation of “political correctness” norms or the desire to appear unbiased on the part of communicators. However, the fact that interpreters

2 Consistent with the general finding that stereotyping effects are stronger when the target to be judged is ambiguous as opposed to clear-cut or extreme in his or her attributes (e.g., see Kunda & Thagard, 1996, for a review), Black targets were back-translated to be less objectively good than White targets when academic credentials were mediocre in quality (as opposed to very good).
understood more favorable subjective language to mean less strong academic credentials for the Black student argues against the “norms” interpretation. Instead, it seems more consistent with the central argument of the shifting standards model—that stereotypes affect the standards we use to judge individual members of stereotyped groups, and that these judgments are interpreted by others (who presumably share the same cultural knowledge of stereotypes) with reference to those differential standards.

**SETTING STANDARDS**

In the research discussed thus far, the role of standards in judgment has been assumed rather than directly assessed. Gender stereotypes about aggression, financial standing, or competence in masculine work domains are assumed to mean that standards are lower for women than men on these dimensions; racial stereotypes about academic ability similarly mean lower standards for Blacks than Whites. As Figure 7.1 indicates, lower standards for a group indicate lower expectations, which lead to anchoring of within-group subjective rating scales at lower levels of a stereotyped dimension. In turn, this pattern of standard setting can give rise to contrast effects in social judgment. This idea is stated eloquently in a book by law professor Stephen Carter (1993), writing about Black achievement in academics and the workplace: “Like a flower blooming in winter, intellect is more readily noticed where it is not expected to be found” (p. 54).

At the same time, Carter (1993) recognized an opposite pattern of standard setting, in which members of devalued or negatively stereotyped groups are held to higher standards than members of valued or positively stereotyped groups: “Our parents’ advice was true: We really do have to work twice as hard to be considered half as good [as Whites]” (p. 58). This idea is well-articulated in a theoretical perspective on double standards for evaluating competence (Foddy & Smithson, 1989; Foschi, 1992, 2000; Foschi & Foddy, 1988). According to this sociological perspective, gender and race (as well as other category memberships) are status characteristics that implicate different standards for evaluating competence:

Those who are considered to be of lower status will have their performances scrutinized and then assessed by a stricter standard than those who are of higher status; the latter, on the other hand, will be given the benefit of the doubt and will be treated with a more lenient standard than the former. (Foschi, 1998, p. 63)

Furthermore, “the application of a more lenient standard to the higher status person ensures that more ability is assigned to him or her than to the lower status person with the same record” (Foschi, 2000, p. 26). In other words, holding lower status groups to higher standards of competence than high-status groups is conducive to assimilation effects in judgment.

My guess is that most readers of this chapter would agree, based on experience or intuition (or the research literature), that both patterns of standard setting are likely to occur. Stereotypes do mean that standards are lower for groups devalued on a given attribute, but stereotypes also lead us to require more evidence of an attribute if a target is stereotyped as not possessing the attribute. Indeed, the latter pattern is consistent with related findings that individuals are ready to perceive in others the attributes they expect (Brewer, 1988; Bruner, 1957; von Hippel, Sekaquaptewa, & Vargas, 1995), that information is sought in a way that may confirm stereotyped expectations (Trope & Thompson, 1997), and that expected behaviors are believed to have higher diagnostic value for inferring dispositions than unexpected behaviors (Trope & Liberman, 1996).

To recognize and reconcile the possibility that stereotypes can implicate both lower and higher standards for members of negatively stereotyped groups, Biernat and Kobrynowicz (1997) made a distinction between two types of standards: minimum standards—those indicating the low-end expectations about a group’s standing on an attribute—and confirmatory (or ability) standards—those relevant for conclusively demonstrating that a group member has the attribute in question. When judges are asked to define minimum standards, these should be lower for groups stereotyped as deficient on the attribute in question (a prediction consistent with the shifting standards model).
When asked to define confirmatory standards, however, these should be higher for those same groups (as predicted by the double standards/status characteristics perspective). That is, a member of a negatively stereotyped group may be held to low minimum standards (consistent with low expectations) but high confirmatory standards (consistent with the idea that a judge will require stronger evidence that an unexpected outcome is due to an underlying disposition).

To test this prediction, Biernat and Kobrynowicz (1997) oriented judges toward either minimum or confirmatory standards and asked them to consider the résumé of a male or female (Study 1) or Black or White male (Study 2) job applicant. To measure standards, judges were asked to indicate the level of performance they would require of the applicant before considering him or her for a job. For example, in the minimum standards condition, judges were asked “how many examples [of job-relevant skills] would you require” for the applicant to “meet the minimum standard to perform the skill?” In the confirmatory standards condition, the question about examples concluded, “before feeling confident that [the applicant] has the ability to perform the skill?” Consistent with predictions, when minimum standards were assessed, women were held to lower standards than men, and Black applicants to lower standards than White applicants. Among those who indicated their confirmatory standards, however, the opposite pattern emerged—White women and Black men were held to higher standards to confirm their ability to perform job-related skills than White men.

A similar pattern emerged in research examining the standards for female and male targets to “qualify” as possessing a variety of personality attributes (Biernat, Ma, & Nario-Redmond, 2008). For example in one study, participants were provided with a list of 20 behaviors relevant to the trait “emotional.” They were asked to consider a male or female target person, and to check off either (a) “the minimum number of behaviors necessary to detect whether or not the target is emotional; to give you some inkling” that the target is emotional, or (b) “the total number of behaviors that are necessary to confirm whether or not the target is emotional; to demonstrate to you” that the target is emotional. Because men are stereotyped as “deficient” in emotionality, the prediction was that they should be held to lower minimum emotionality standards, but higher confirmatory emotionality standards than women. The number of required behavioral examples checked was the index of the standard (minimum or confirmatory) the participant had in mind. Results confirmed that fewer behaviors were required to suspect that a man might be emotional (to meet minimum standards), but more behaviors were required to confirm that he was emotional relative to women (see also Biernat & Ma, 2005; Maass, Montalcini, & Biciotti, 1998).

In short, the minimum–confirmatory standards distinction seems useful for understanding how perceivers use and interpret trait terms. Perceivers apply different evidentiary rules depending on trait stereotypicality, target group membership, and the judgment standard (minimum or confirmatory) at hand. These different types of standards may be particularly relevant in decision-making contexts where, for example, one sets an initial screening standard (e.g., establishment of a “short list” of candidates for a job), followed by a final choice (e.g., a hiring decision). The short list may be akin to a minimum standard, and should potentially favor those group members who are stereotyped as deficient in job-related competence, whereas the hiring decision is akin to a confirmatory standard, and should favor those stereotyped as competent. In two studies involving simulated hiring decisions, female judges were indeed more likely to place female than male job applicants on short lists, but were also less likely to hire them (Biernat & Fuegen, 2001). These findings, like those based on distinctions between subjective and objective (or common rule) response scales, highlight the subtlety and complexity of the stereotyping process: Stereotypes can lead to either leniency or stringency in the evaluation of individual group members.

**BEHAVING TOWARD MEMBERS OF STEROTYPED GROUPS**

The shifting standards approach has highlighted the complex ways in which stereotypes guide our perceptions and judgments of individual members of stereotyped groups. Judgment is crucially
important, but understanding whether and how that judgment translates into behavior matters as well. When and how do people act on their judgments of individual targets?

From a shifting standards perspective, this question becomes more complex because patterns of judgment vary depending on whether subjective or objective rating scales are used. For example, a woman might be subjectively judged as a “very good” candidate for an executive chief of staff position—better than a comparable man—but objectively judged as the less strong candidate (Biernat & Kobrynowicz, 1997). But whom would the judge hire in this situation? Would hiring behavior follow from the subjective sense of the female as strong, or the common rule judgment of the male as stronger than the female?

The shifting standards model suggests that the latter will be the case—that the hiring decision would favor the male candidate. One of the assumptions of the model is that judgments on objective scales, because they invoke a cross-category framework, better reflect perceivers’ mental representations of target (see Figure 7.1). In this case, the male target is represented as the better of the two candidates, and a decision to hire calls for a cross-category choice; it will therefore reflect the representation. However, this is not to suggest that subjective judgments of individuals will never predict behavior. Indeed, consider the following scenario: You are the manager of a coed softball team, and one of your players steps up to bat and hits a single. How will you respond to this event? Will you cheer? Pat the player on the back? Do nothing in particular? Will your response depend on the gender of the player?

In research establishing precisely such a role-playing scenario, Biernat and Vescio (2002) found that female players were more likely than male players to be the recipients of effusive cheers and praise following a hit. This kind of behavioral response—praise, cheers, and so on—seemed not to follow from the stereotyped perception of men as better athletes than women, but rather from the subjective perception of a female hitting a single as “pretty good . . . for a woman.” At the same time, role-playing managers favored male over female players on every other behavioral indicator: Male players were more likely than female players to be chosen for the team, placed in the top of the batting lineup, and assigned to valuable infield positions.

These two very different forms of behavior—team assignment and selections versus cheering and praise—capture an important distinction that may account for the different pattern of gender bias on each. Specifically, behavioral choices such as hiring decisions or position assignments can be characterized as having a zero-sum quality, in that behavior toward one individual constrains the behavioral options available toward another; scarce resources are involved. On the other hand, behaviors such as cheering, delivery of praise, and a variety of nonverbal acts have a nonzero-sum quality, in that they are in (relatively) endless supply and can be bestowed on any number of targets. Cheering Player 1 does not prevent me from cheering Player 2, but assigning Player 1 to shortstop means I cannot place Player 2 in that position (Biernat & Vescio, 2002; Biernat, Vescio, & Manis, 1998).

We have argued that zero-sum behaviors will tend to reveal assimilation to group stereotypes, whereas nonzero-sum behaviors will tend toward contrast (Biernat et al., 1998), and that objective judgments will better predict zero-sum behaviors, whereas subjective judgments will better predict nonzero-sum behaviors. The basis for these predictions lies in the fact that zero-sum behaviors require a cross-category frame of reference, comparable to the frame invoked when judgments are made on a common rule response scale (see Figure 7.1). To select the shortstop or hire the employee, one looks at the whole field of possible candidates. On the other hand, nonzero-sum behaviors such as provision of feedback or praise are likely made with reference to within-category expectations. Effusive praise, for example, may result when a target person exceeds expectations. If expectations are lower for one group than another (as is likely the case when considering the hitting power of female compared to male athletes), praise may be more likely in the former case. A similar finding can be seen in research on White judges’ responses to Black versus White authors of poorly written essays (Harber, 1998). Feedback (marginal comments on these essays) was more favorable when White “graders” thought the author was Black than White. This positivity toward Blacks may have stemmed from low standards, compared to which the essays seemed “better” or more deserving of
positive commentary. Not surprisingly, this pattern of praise (based on comparison to low expectations), is likely to be perceived as patronizing in nature (Foschi, 1992; Jackman, 1994).

The distinction between zero-sum and nonzero-sum behaviors can also be seen in a recent study by Vescio, Gervais, Snyder, and Hoover (2005, Study 1). Participants were placed in leadership roles and had authority over female and male subordinates in a work group described in masculine terms (“requiring strong strategic planning and competitive skills”). Leaders were asked to assign their subordinates to a number of valued tasks or roles (e.g., being the team captain or on the first string of a team competing in an “academic challenge” contest), a form of zero-sum behavior, and to offer comments to their “workers” (praise or criticism), a form of nonzero-sum behavior. Consistent with predictions derived from the shifting standards model, zero-sum behavioral choices favored male over female subordinates (assimilation to gender stereotypes), whereas praise (e.g., “Your answers during the first phase of the experiment were excellent!”) was more likely to be offered to female than male subordinates. This pattern of behavior was not universal, however. It only occurred among male leaders, and only when these men had been oriented (through task instructions) to “avoid weaknesses” in their team. Among female leaders, and among men oriented to focus on strengths, gender bias on both types of behavior was reduced. Conditions invoking high-status group members’ attention to stereotypical weaknesses of others may be particularly likely to instantiate assimilation to stereotypes in zero-sum behavioral choices, but contrast in nonzero-sum behaviors (e.g., a patronizing pattern of high praise).

COMPLEXITY AND MORE COMPLEXITY

If there is one theme apparent in the research generated by the shifting standards perspective, it is that stereotyping effects are complex and varied. Judgments of and behavior toward members of stereotyped groups may show evidence of assimilation or contrast—indeed the standards evoked by stereotypes may be conducive to assimilation or contrast—depending on the nature of the judgment or behavior at hand. The shifting standards research reviewed here has distinguished between judgments made in common rule and subjective units, standards based on confirmatory and minimum evidentiary criteria, and behaviors that have a zero-sum or nonzero-sum quality. In general, assimilation to stereotypes is more likely when judgments are made on common rule scales, with reference to confirmatory standards, and when the behaviors at hand are zero-sum. Contrast effects (or null effects) are more likely when judgments are made in subjective units, with reference to minimum standards, and when a judgment or a behavior is nonzero-sum in nature. This complexity means that we may sometimes underestimate the extent to which stereotypes guide judgment—we may be too quick to take an apparently null effect of a stereotype as evidence that the stereotype is no longer operative (e.g., Locksley, Borgida, Brekke, & Hepburn, 1980). When women and African Americans are placed on short lists in greater numbers, or when they are praised for their work, we may remark on the positively changing climate for these groups. However, if we recognize that at the same time, hiring or pay raises do not follow, we can continue to see evidence of stereotype operation in everyday life.

Certainly a large number of questions remain about when and why stereotypes will guide judgment and behavior in an assimilative versus contrastive direction, and the shifting standards model can provide only a partial perspective on these issues. For example, the model has little to say about the role of motivation in patterns of stereotyping effects, although it is clear that such factors matter (Fein & Spencer, 1997; Govorun, Fuegen, & Payne, 2006; Kunda & Sinclair, 1999). The model also says little about the normative context in which judgment occurs, although stereotyping and discrimination can be enhanced or reduced depending on situational factors and salient norms (Blanchard, Crandall, Brigham, & Vaughn, 1994; Gaertner & Dovidio, 1986; Pettigrew, 1959; Stangor, Sechrist, & Jost, 2001).

Some issues well within the realm of the shifting standards model seem important to address, however. One concerns the fact that people simultaneously belong to multiple categories, and thus
Stereotypes and Shifting Standards

which category standard will be activated in any given setting is an open question. In research in my lab, we have typically focused on one target group at a time (e.g., studying gender stereotyping of White targets, or racial stereotyping of male targets), although occasionally target race and gender have been crossed (e.g., Biernat & Manis, 1994). For example, in one study, judges rated the verbal ability of Black and White, male and female targets in either subjective or objective units (Biernat & Manis, 1994, Study 2). Two independent “shifting standards effects” emerged, such that White targets were judged more verbally able than Black targets in objective ratings, more so than subjective ratings, and women were judged more verbally able than men in objective but not subjective ratings. There was no evidence of a race and gender interaction, but perhaps this was the case because stereotypes about verbal skill exist with regard to both gender and race. One prediction might be that when a stereotype is relevant to one social category but not another, standards based on only the relevant category will be activated. Category expectations may also conflict, as they do for Black women on traits such as “aggressive” or “athletic” (Blacks are stereotyped as more aggressive and athletic than Whites, but women as less aggressive and athletic than men). It is unclear what standards might be invoked in these settings, although one category may dominate depending on context (e.g., Macrae, Bodenhausen, & Milne, 1995), or unique subtype standards may apply.

Additionally, standards may be determined by other features besides a target’s category membership. Standards may be established by the context—either explicitly (as when criteria for hiring are specified), or implicitly (e.g., via priming effects; Mussweiler & Englich, 2005). Another commonly used standard for judging others is the self (Dunning & Cohen, 1992; Lambert & Wedell, 1991; Mussweiler, Epstude, & Rüter, 2005; Ross, Greene, & House, 1977). Additional work is clearly needed to identify which standards matter, and how different sources of standards are integrated to influence social judgment.

A final issue that research on the shifting standards model raises is whether the tendency to shift standards is a good or bad thing—more specifically, is the shifting standards effect a marker of prejudice? There is certainly reason to suspect that it may be. At the crux of the model is the assumption that perceivers hold stereotypes of social groups. Indeed, Biernat et al. (1991) found that standards did not shift on judgment dimensions (e.g., movie-going frequency, age) for which no gender stereotypes existed, and Biernat and Manis (1994) found that only individuals who explicitly endorsed the stereotype that women have greater verbal ability than men showed evidence of shifting standards in their judgments of individual men’s and women’s verbal ability. Endorsement of the stereotype that Blacks are more athletic than Whites was also associated with the tendency to shift athleticism standards for judging individual Blacks and Whites (Biernat & Manis, 1994, Study 3).

To the extent stereotypes are related to prejudice (albeit imperfectly), shifting standards may also be linked to prejudice. In several studies, Biernat and Manis (1994) examined whether racial and gender-role attitudes moderated the tendency to shift standards. Results were inconsistent. In one study only those high in racism showed evidence of shifting standards; in another, scores on the Modern Racism Scale (McConahay, Hardee, & Batts, 1981) had no effect on this tendency. Similarly, traditional attitudes toward women (Spence & Helmreich, 1972) were associated with shifting standards in judgments of male and female competence in one study, but not with judgments of verbal ability in another (Biernat & Manis, 1994).

This inconsistency may reflect, in part, the fact that sometimes using shifting standards can be nice. When I judge my 8-year-old’s math skills relative to a lower standard than I use to judge the math skills of the college students in my classes, I am probably being appropriate and kind. When judging our assistant professor colleagues on their “service to the department” contributions, it may again be reasonable to use a lower standard than that used to judge our senior colleagues. When we consider the Verbal GRE scores of nonnative English speakers, we may be well advised to use a lower standard for evaluation than we do for native English speakers. In each of these examples, the use of shifting standards seems to suggest kindness, or even fairness, rather than prejudice. Of course, the use of lower standards for some groups may be patronizing in nature; it has the “benevolent” quality that may typify at least certain kinds of prejudice (Glick & Fiske, 2001; Jackman,
1994). Nonetheless, the question of whether shifting standards is linked to indicators of prejudice is an open one.

In one recent line of research, we focused on racial attitudes and the tendency for individuals to shift standards when judging Black versus White targets on academic ability (Biernat, Collins, Katzarska-Miller, & Thompson, in press). A measure of individual differences in the tendency to shift standards was created by subtracting the race difference (White–Black) in subjective ratings of academic ability (very poor to very good) from the race difference in objective ratings of academic ability (estimated ACT scores). By this metric, individuals would score high in the tendency to shift standards if they judged Whites more academically competent than Blacks on the objective index and Blacks more competent than Whites on the subjective index. We then correlated this measure with an explicit measure of racial attitudes (the Pro-Black/Anti-Black Attitudes Questionnaire; Katz & Hass, 1988), and with a race implicit associations test (IAT) effect (Greenwald, McGhee, & Schwartz, 1998). We found little relationship between the tendency to shift standards and anti-Black attitudes \(r = .09\), pro-Black attitudes \(r = -.13\), or the IAT \(r = -.14\).

Similarly, in two other studies, the tendency to shift standards was uncorrelated with an evaluative priming measure of prejudice and stereotyping (Wittenbrink, Judd, & Park, 1997; \(rs = .02-.20, ns\)). Thus, there was no indication that race-based shifting standards overlapped with commonly used indicators of explicit or implicit race prejudice.

However, there was some evidence that the indicator of shifting standards predicted an important behavioral outcome. Participants were asked to allocate funds to various student organizations, one of which was the “Black Student Union.” This funding task was based on a procedure used by Haddock, Zanna, and Esses (1993), and required participants to cut the budget to all listed student organizations by 20% (from $10,000 to $8,000). In three studies, the tendency to shift race standards when judging academic ability predicted lower funding allocation to the Black Student Union. That is, shifting standards was associated with a more negative behavioral response to a Black organization, and this behavior was not consistently predicted by any other prejudice indicator.  

Thus, although it may be unclear whether shifting standards can be conceptualized as a marker of prejudice, it seems to matter for at least one form of behavioral response. Interestingly, this behavior (funding allocation) can be conceptualized as zero-sum in nature—it involved allocation of a scarce resource in which behavior toward one group restricted the options available toward another. Further research is needed to examine the ability of the tendency to shift standards (for a variety of groups across a variety of domains) to predict a variety of behavioral responses. I suspect that the pattern of association will be complex and varied.

**CONCLUSION**

This chapter has provided an overview of questions and research generated from the perspective of the shifting standards model (Biernat et al., 1991). The basic idea behind the model is a simple one—that judgments of others are often based on a frame of reference provided by social category membership(s). It is because of group stereotypes that category memberships provide a frame of reference; stereotypes create a context of group expectations against which an individual group member is evaluated. The shifting standards model therefore suggests that instead of stereotypes solely guiding judgments in an assimilative fashion (“I expect that men are more aggressive than women and therefore I judge individual men as more aggressive than individual women”), they may lead to contrast effects, particularly on subjective rating scales or in the production of subjective language (Biernat, 2003).

---

3 It was the shifting standards index and not its subcomponents (e.g., objective and subjective judgments of Blacks and Whites) that predicted funding decisions in all three studies. Additionally, the negative relationship between shifting standards and funding was strongest among those high in implicit prejudice.
One message that emerges from the work reviewed here is that whether stereotypes result in assimilative or contrastive effects depends in large part on the nature of the judgment or behavior at hand. As noted throughout this chapter, assimilation is the more likely outcome when judgments are rendered in common rule units, when confirmatory standards are invoked, and when behaviors are zero-sum in nature, whereas contrast effects or null effects are more likely in subjective judgments, when minimum standards are invoked, and when behaviors have a nonzero-sum quality. Thus, a female applicant for a masculine job may find herself judged favorably subjectively (largely because she is held to a lower minimum standard), but not objectively, compared to a comparable male applicant. More evidence will be required of her to meet confirmatory standards, and perhaps because of this, she is may be less likely to be hired. Nonetheless, she may find herself praised during her interview. In the event she is hired, she may find herself lauded for her work, but nonetheless assigned the less valuable tasks, positions, or resources. Stereotyping effects are complex, and perhaps downright confusing from the perspective of the recipient of these effects.

Of course, other factors may moderate the patterns predicted. For example, we assume equivalence in the qualifications of the female and male applicant in the preceding scenario. The shifting standards model has also typically assumed an “average” target—in this example, not extremely bad nor extremely good—as stereotyping effects are generally more evident when targets are “neutral,” “ambiguous,” or “average” in quality (e.g., Biernat & Vescio, 2002; Hodson, Dovidio, & Gaertner, 2002; Kunda & Thagard, 1996; Stapel & Winkielman, 1998). Strong motivations (e.g., to appear unprejudiced) may also overcome some of these tendencies. Nonetheless, ceteris paribus, stereotyping effects are likely to be revealed in the complex pattern of effects described in this chapter.

Research from the shifting standards model suggests a view of stereotyping beyond simple assimilation. A broader view of social stereotyping suggests a subtlety and complexity of effects that might be missed by studying judgment and behavior through a single lens or method.

REFERENCES


Stereotypes and Shifting Standards


