Journal of Personality and Social Psychology
1990, Vol. 59, No. 4, 664-675

Everyday Base Rates (Sex Stereotypes): Potent and Resilient

Thomas E. Nelson, Monica R. Biernat, and Melvin Manis
University of Michigan and Ann Arbor Veterans Administration Medical Center

Undergraduate Ss studied photographs of students and estimated the heights of the pictured models. Contrary to reports of base-rate neglect, sex stereotypes regarding height (the implicit recognition that men are normally taller than women) significantly affected these estimates, even when the targets' actual height was statistically controlled. Base rates were especially influential when information about targets was ambiguous, that is when targets were pictured seated. These base-rate effects were *robust*, remaining significant and substantial despite efforts to lessen their magnitude. Attempts to reduce base-rate effects by encouraging Ss to strive for accuracy, discouraging their reliance on the target's sex (as a cue), or offering cash rewards for accuracy did not succeed. Informing Ss that for the sample to be judged, sex would not predict targets' heights attenuated the base-rate effect, although it remained highly significant.

Broadly speaking, research on stereotypes and stereotyping tends to follow two tracks: (a) cataloging the attributes or mental images that (rightly or wrongly) are associated with different social groups and (b) examining how group stereotypes affect people's assessments of individual group members (Ashmore & Del Boca, 1981). The present research focused on the second of these concerns, showing that stereotype (or base-rate) effects could be very stable and difficult to eradicate, especially everyday or widely known stereotypes. In support of this claim, we present two experiments in which efforts were made to lessen or eliminate subjects' reliance on stereotypes when evaluating individual group members. The results show substantial evidence of continuing base-rate influence, despite these debiasing efforts.

Classic discussions of social stereotypes (Allport, 1954; Katz & Braly, 1933) have assumed a normative model that deems judging individuals solely on the basis of their membership in one or another social group as unfair and unwise. Research has shown that stereotypes can indeed affect the assessments of individual group members, often to their detriment. For example, Sagar and Schofield (1980) reported that ambiguous behaviors were seen as more aggressive when enacted by a Black person than by a White person. Similarly, Darley and Gross (1983) showed that ambiguous performance on a school test was thought to reflect weaker academic ability if it was observed in a lower class child, rather than in a child from a middle-class, suburban background.

Judgment patterns of this sort clearly betray the influence of the judge's prior beliefs (stereotypes) about the various groups. Such beliefs, when quantified, are often referred to as *base rates* by Bayesian theorists. Respondents in the Darley and Gross (1983) and Sagar and Schofield (1980) studies apparently held stereotypes in which a given trait (e.g., aggressiveness) was considered more prevalent in the stereotyped group (Blacks) than in the comparison group (Whites; Sagar & Schofield, 1980). These assumed differences between groups could explain why subjects judged members of the different groups in a biased manner after examining identical individuating information.

The discovery that stereotypes can significantly affect judgments of individuals seems inconsistent with the broad implications of Kahneman and Tversky's (1973) early work, which showed that people largely neglected base rates when presented with individuating information. More recent work has shown, however, that although base rates are not always influential, they can exert a substantial effect on the judgment process under many circumstances. Some investigators have shown, for example, that base-rate effects are enhanced when there is an implied causal relationship between the category in question and the judgment that is involved (Ajzen, 1977; Tversky & Kahneman, 1980). Ginossar and Trope (1980, 1987) demonstrated that base rates are effective when individuating information is inconsistent, irrelevant, or otherwise inappropriate for the judgment task or when experimenters discourage respondents from "psychologizing," by presenting the judgment task as one not involving the identification of personality types (see also Krueger & Rothbart, 1988). Hilton and Fein (1989) confirmed that people generally ignore clearly irrelevant individuating information, although they sometimes neglect categorical information in the presence of *pseudorelevant* individuating information, or information that is frequently, but not always, appropriate for judgments of individuals. More generally, Nisbett, Krantz, Jepson, and Kunda (1983) showed that the common social perceiver displays a surprising amount of sophistication when he or she applies statistical principles to some everyday issues, such as sports.

In the present studies, we sought to deactivate a base-rate effect we had previously observed in a simple height estimation task. In this earlier work, subjects were presented with full-length photographs of different target individuals, some of whom were shown in a sitting pose and others of whom were shown standing beside a familiar object, such as a door or a table. Subjects were asked to estimate the height of each target in feet and inches, including the shoes or boots worn by the

Monica Biernat is now at the University of Florida, Gainesville.

Correspondence concerning this article should be addressed to Melvin Manis, Research Center for Group Dynamics, Institute for Social Research, University of Michigan, Ann Arbor, Michigan 48106-1248.

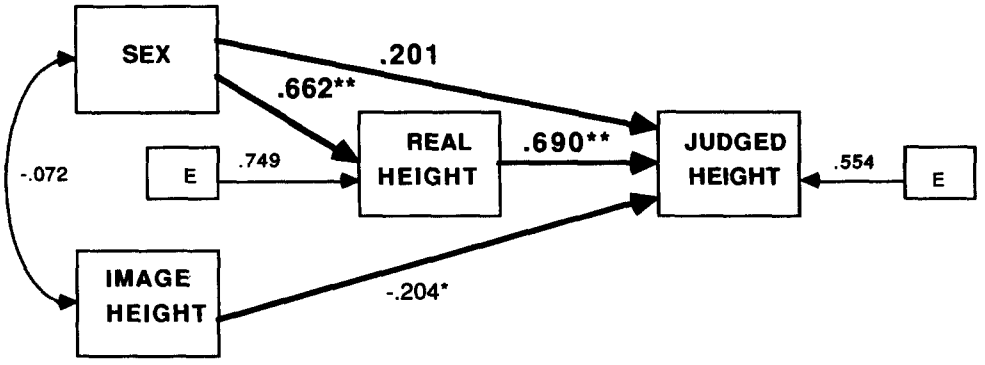

Figure 1. Path model for judgments of standing targets. (*$p < .05$. **$p < .01$.)

target. Because the camera angle and distance to target varied randomly from one photograph to the next, respondents could not rely on image size as a cue to the targets' height.

The results showed clear evidence of a base-rate effect, in that the male targets were judged to be taller, on average, than the female targets, even after the actual difference in height between the male and female targets had been statistically controlled. This stereotype effect was substantially enhanced when the targets were shown in a sitting pose, in which presumably the stimuli were more ambiguous with respect to target height, thus forcing subjects to rely more heavily on base rates, as prescribed by Bayes's theorem.

Figures 1 and 2 present path analytic models of these data, using the average judged height of each target as our focal endogenous variable. The gender of the individual targets, the target's true height, and the size of the different photographic images serve as predictor variables. In Figures 1 and 2, note the substantial coefficient linking sex to real height, or the actual height of the target models ($\beta = .662$, $p < .01$ for standing targets; $\beta = .519$, $p < .01$ for sitting targets[1]). This path reflects the real-life kernel of truth that on average, men are indeed taller than women. Note also that subjects were properly attentive to the real height of the standing targets (Fig. 1), as reflected by the significant coefficient linking real height to judged height. Nevertheless, subjects also relied on the target's sex to guide their judgments ($\beta = .201$, $p < .10$), although this effect was relatively weak, as compared with the impact of the target's actual height. For the sitting targets (Fig. 2), the linkage between target sex and judged height was substantially stronger, with the path coefficient ($\beta = .624$) nearly twice as strong as the coefficient that reflected the impact of targets' real height ($\beta = .317$).

These results suggest that our respondents were relatively accurate in assessing the height of standing targets, for they were strongly affected by the actual heights of the models. Even here, however, we found evidence of stereotyping, because with other factors held constant statistically, our subjects' assessments were influenced to some extent by the sex of the target. For the sitting targets (for which the assessment task was more difficult), the sex of the target became an even more important cue in guiding the respondents' height estimates.

The finding that sex stereotypes regarding height substantially influenced subjects' judgments of individual men's and women's height is consistent with a body of work demonstrating the impact of categorical beliefs (stereotypes) on judgments of particular category members (Darley & Gross, 1983; Sagar & Schofield, 1980). However, a growing battery of studies outlines the conditions under which normally powerful stereotype effects may be circumvented. In a series of provocative studies, Fiske, Neuberg, and their colleagues (Fiske, Neuberg, Beattie, & Milberg, 1987; Neuberg & Fiske, 1987) showed that both informational and motivational factors can lead subjects to pay closer attention to specific qualities of a target, overriding their exclusive concern with the target's social category. In brief, subjects paid closer attention to individuating information (a) when categorical information was uselessly broad (e.g., "a person"), (b) when individuating information was inconsistent with categorical information (e.g., "an uneducated doctor"), (c) when the subjects' outcome depended on the performance of the target person—namely, when they had an opportunity to win a prize for a joint project, and (d) when subjects were implored to strive for accuracy. Neuberg and Fiske (1987) presented reaction time data consistent with their theory that "attribute-based" (as opposed to categorical) responses are associated with greater attention allocated to the specific characteristics of the individual, rather than to their social group alone. It is important to note that, in contrast with other research within this tradition, Neuberg and Fiske's subjects did not judge any particular trait of the target persons but rather estimated how much they thought they would like the target persons. That is, rather than report on any of the cognitive attributes that composed their stereotypes, subjects in these studies described instead the affect attached to these stereotype and trait collections. Though Fiske and her colleagues reported success with these debiasing strategies in studies of the stereotyping process, Fischhoff (1982) concluded that such efforts were largely ineffectual in eliminating other judgmental errors, in particular the hindsight bias and the ubiquitous overconfidence effect.

In the present work, we too sought to deactivate the strong

[1] The modest differences across the two path models in this particular coefficient are probably due to sampling error.

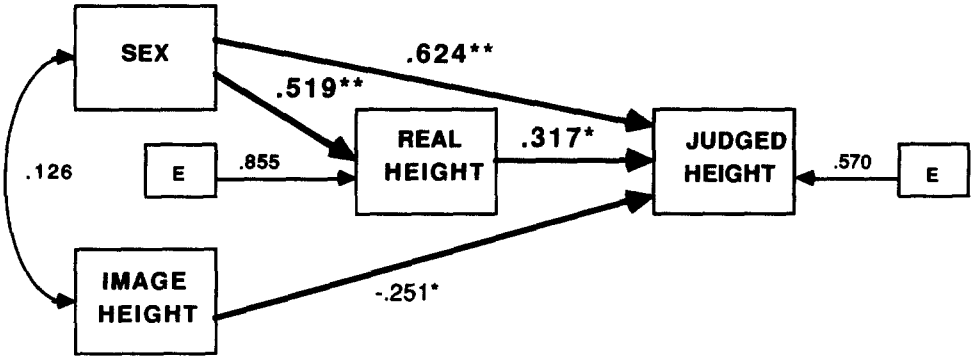

Figure 2. Path model for judgments of sitting targets. (*$p < .05$. **$p < .01$.)

stereotype effects we had discovered in the domain of height judgments. We devised three separate lines of approach, borrowing in part from tactics that had succeeded in prior research.

1. *Typicality.* Several lines of research suggest that atypical group members might be ascribed different traits and qualities from those associated with the more typical members. Fiske et al. (1987) showed that subjects were less inclined to evaluate a target solely on the basis of his or her membership in a stereotyped group when overt qualities of the person were inconsistent with the modal group member. Similarly, Rothbart and Lewis (1988) reported a series of experiments indicating that atypical exemplars were not adequately represented in memory as category members and thus were not effectively taken into account in judgments and inferences about the category as a whole.

Starting with this background, we hypothesized that the dominant stereotype of a given social group might seem substantially less relevant when assessing an atypical member; the atypical member might, for example, be seen as part of an unusual "subtype" (Weber & Crocker, 1983), in which the more general stereotype is inapplicable. Deaux and Lewis (1984) and Shapiro (1986) independently proposed related models that suggest that people who are different from the prototypic group member in one domain are generally expected to have other atypical traits as well.

In Experiment 1, respondents attempted to estimate the heights of different target persons, whose purported interests were presented along with their photographs. We hypothesized that when judging targets who were atypical in their interests (e.g., a woman who wanted to become a physicist or a man who wanted to be a nurse), our respondents might be relatively unaffected by the prevailing base rates, namely, by the fact that in general, a male target is likely to be taller than a female target.

2. *Motivation.* Stereotypes enable an economy of judgment. Their power derives, in part, from an inattentiveness to the unique characteristics of a given target because of a less effortful reliance on general beliefs. Starting from this premise, we reasoned that by providing a substantial financial incentive—$50—for accurate performance, we might reduce the importance of the target's gender as a determinant of the height estimates and encourage subjects to expend more effort to discern the actual height of each unique target. The underlying reasoning here is analogous to the idea that the prejudiced manager of a baseball team may be relatively unaffected by the ethnicity of a potential team member (and more sensitive to the player's individual athletic ability) if a wise addition to his roster will provide the team with a good chance to win the championship. When the manager's choice seems unlikely to affect the team's chances, however, his or her decision may be less critically determined by the player's ability and may instead be more substantially affected by ethnic considerations. This hypothesis is consistent with the findings of Neuberg and Fiske (1987), who showed that subjects were less attentive to a target's social category when they needed to work collaboratively toward winning a shared prize.

3. *Forewarnings.* Social stereotypes play a significant role in a variety of appraisal settings (e.g., in the screening of job applicants). To reduce the impact of these stereotypes, personnel managers and other judges are sometimes alerted to the unfairness and inefficacy of appraisals that rely on group-level information when evaluating individual applicants and are therefore admonished to ignore group stereotypes. To evaluate the effectiveness of such admonitions, some of our respondents were explicitly discouraged from relying on the target's sex as a cue to his or her height. These people were told that past research had shown that height judgments were often inaccurate because the respondents were unduly reliant on the target's sex (as a cue). These respondents were forewarned to avoid this stereotype error when making their own height estimates. Neuberg and Fiske (1987) reported great success with simple verbal instructions of this kind. Subjects who were merely told "it's extremely important that you make every effort to form as accurate an impression as possible" (Neuberg & Fiske, 1987, pg. 441) virtually ignored the label *schizophrenic* when evaluating target individuals.

Experiment 1

Method

Seventy-five students at the University of Michigan participated in Experiment 1. They were recruited to take part in a study on social judgment and were paid $5 for their efforts. Subjects were randomly

assigned to one of three groups: a control group, a financial incentives group (in which a $50 reward was offered to the best height judge in the study), and a verbal forewarning (don't stereotype) group that was cautioned against an undue reliance on the sex of the target when making their height judgments. In addition to these between-subjects conditions, the individual targets in each test booklet varied with respect to sex (male versus female), posture (sitting versus standing), and typicality (described later).

Subjects were told that the experiment was concerned with "how accurately people like yourself can judge the physical and personal characteristics of individuals based on a small amount of information." Subjects were presented with photographs of University of Michigan students, along with a small amount of background information, including the student's ostensible hometown, age, and a favorite summer job, hobby, or career goal. We had conducted pretests in which a separate group of subjects were queried as to what kinds of summer jobs, hobbies, and career goals were either typical or atypical of the average college man and woman. Those characteristics mentioned most frequently were used to vary the putative typicality of the targets.

Materials. Stimulus photos were photocopied reproductions of 3.5 × 5 inch black-and-white prints. Our models were selected in a more-or-less haphazard fashion from libraries and cafés around campus. A model's height was measured on the spot, including their footwear. Male heights ranged from 65 inches to 81 inches, with an average height of 70.3 inches. Female heights ranged from 59 inches to 74.5 inches, with an average of 65.8 inches. At least one standing and one sitting pose was photographed for each model, who was always posed on or near a familiar reference object, such as a chair, door, or car. Models were photographed from different distances and angles, to ensure that the size of the photographic image would not be a reliable indicator of the target's true height.

Instructions. All respondents were instructed to estimate the height of each target in feet and inches (to the nearest half inch). In making these estimates, they were told to "be sure that you are focusing on each person's real height, wearing the shoes or boots in which he or she is shown." Instructional conditions were varied through introductory paragraphs that appeared on the first page of each test booklet.

The money incentive was introduced with these instructions:

> Take special care with your height judgments. In previous studies we have found that people are often poor judges of height. We would like you to try your best to be as accurate as possible when making your judgments. To motivate you further, we will pay a $50 cash reward to the person who makes the most accurate judgments.

Subjects in the control and don't stereotype conditions did not receive these instructions. The don't stereotype condition was introduced with this paragraph:

> Take special care with your height judgments. We all know that men are, in general, taller than women, and people often use this fact when judging the heights of strangers. However, we also know that some women are taller than many men, and that some men are shorter than many women. Therefore, in order to make as accurate a judgment as possible, try to judge each case as an individual; do not rely on the person's sex (female or male).

Subjects in the control and cash conditions did not receive these instructions.

Subjects judged 100 photographs, each on a separate page, working at their own pace. For each photo, subjects estimated the model's height (to the nearest half inch) and, as a manipulation check, evaluated each target's typicality on a 1-to-7 scale.

The three target factors (sex, typicality, and posture) combined to form an eight-celled classification of photographs (e.g., standing atypical men, sitting typical women, and so on). There were 10 photos within each of these eight categories—along with 20 filler photographs that contained stereotypically neutral information about the target (e.g., she or he aspired to be a psychologist), which were included to conceal the typicality manipulation—for a total of 100 photos. Within each of the instructional conditions of this experiment, two counterbalanced subgroups were developed, so that the targets who were labeled as typical for half the subjects were labeled atypical for the other half.

Results

Path analyses. Our initial approach to the data consisted of a series of path analyses similar to those summarized in Figures 1 and 2. Three variables (sex of target, real height of the target, and size of photographic image) were entered as predictors of the endogenous variable, the mean height judgment associated with each target over all subjects. Altogether, 12 path models were estimated, corresponding to the four types of targets (standing typical, standing atypical, sitting typical, and sitting atypical) within each of the three motivational conditions (control, don't stereotype, and cash).

Although the resulting path coefficients differed modestly from model to model, we found in general a clear replication of the results shown in Figures 1 and 2. Despite our typicality manipulation and despite our attempts to motivate some subjects to strive for greater accuracy, we again found that estimates of height were strikingly affected by sex stereotypes when the target was sitting (please see Table 1 for a summary of the path coefficients). When the target was standing, base rates were less prominent in relation to the target's true height, as was previously discovered. Note, however, that in four out of the six path models for standing targets, the direct path linking target sex to height judgment was highly significant ($p < .01$) and that in the other two models, it was of borderline significance ($p < .10$). In fact, base rates apparently played a more prominent role in this set of analyses than in the data set summarized in Figures 1 and 2. In our earlier work (see Figure 1), the direct effect of the target's sex was only marginally important when our subjects evaluated standing targets; moreover, this effect was grossly overshadowed by the impact of the target's actual height ($\beta = .201$ vs. $\beta = .662$). In the present experiment, by contrast, target sex had a clearer impact on the judgments evoked by standing targets; indeed for three out of the six path models, the impact of the category (sex) variable rivaled the impact of the target's actual height. When the targets were seated, the stereotype clearly assumed the dominant role, and the effect of the target's true height fell to insignificance in half of the path models. Overall then, although we found modest variations in the path coefficients across the different models (probably because of the vagaries of sampling), the pattern of results consistently replicated the results of our earlier work, showing if anything an enhanced role for stereotypes, despite our attempts to weaken their impact.[2]

[2] Although we believe that the widespread association between gender and height was responsible for the results that are summarized in Table 1, we were concerned about the possibility that some aspect of

Table 1
Summary of Path Coefficients for Experiment 1

Condition	Standing targets: Sex to est. ht.	Standing targets: Real ht. to est. ht.	Sitting targets: Sex to est. ht.	Sitting targets: Real ht. to est. ht.
Control				
Typical	.44***	.45***	.72***	.23**
Atypical	.33***	.51***	.75***	.16
Don't stereotype				
Typical	.25*	.47***	.51***	.37***
Atypical	.26*	.56***	.69***	.28**
Cash				
Typical	.46***	.45***	.79***	.13
Atypical	.45***	.44***	.74***	.16
Mean	.36	.48	.68	.22

Note. Est. = estimated; Ht. = height.
* $p < .10$. ** $p < .05$. *** $p < .01$.

Group comparisons. Because the use of standardized path coefficients makes comparisons across different path models somewhat problematic (Asher, 1983), we also examined the effect of our contextual manipulations in an analysis of variance (ANOVA) format. We first included all experimental manipulations in a 3 (instructional set) × 2 (target sex) × 2 (target posture) × 2 (target typicality) mixed design ANOVA; the subject's assigned instruction condition was the between-subjects manipulation, and the three aspects of the photos (sex, posture, and typicality) were the within-subject manipulations. For each subject, we calculated eight mean judgments corresponding to all possible combinations of the three manipulated photo characteristics (e.g., sitting typical men, standing atypical women, and so on). Ten different height estimates (10 photos) contributed to each of these eight judgment categories.

The results of this analysis showed that the manipulation of judgment instructions (through monetary incentives or verbal admonitions) did not make a reliable difference in our subjects' height judgments. The interaction between condition and target sex was far from significant, $F(2, 72) = .23$, $p > .79$, nor was the three-way interaction between condition, sex, and posture significant, $F(2, 72) = 1.54$, $p > .20$. In other words, the difference between judgments of male and female targets was neither attenuated nor exaggerated by our debiasing instructions. As a more conservative test, we performed a second ANOVA on just the control and don't stereotype conditions, as these seemed to be the most different on the basis of the results of the path analysis (see Table 1). Again the two critical interactions, condition by sex, $F(1, 44) = .25$, $p > .61$, and condition by sex by posture, $F(1, 44) = 2.26$, $p < .14$, did not reach conventional levels of statistical significance.

Although our instructional manipulations apparently had little effect on height judgments, the group-level analysis supplied more evidence for the importance of target posture. There was a significant interaction between sex and posture, $F(1, 72) = 132.55$, $p < .001$, that confirmed the results obtained from the path analyses. That is, in all three instructional conditions, the height judgments were more markedly influenced by sex stereotypes when the targets were seated than when they were standing. Last, our manipulation of target typicality yielded a significant Sex × Typicality interaction, $F(1, 72) = 8.04$, $p < .01$. Table 2 shows that our subjects thought atypical men were slightly shorter and atypical women were slightly taller than the typical members of their respective sex category. Although this effect is highly significant, the means in Table 2 indicate that the absolute difference between the judged heights of typical and atypical targets is actually rather small (less than one quarter of an inch).

Discussion

Experiment 1 showed not only that base rates exerted a significant influence on height judgments but also that these effects were quite robust. The impact of base rates remained strong despite our efforts to diminish their power by altering the experimental instructions and the social characteristics of the targets. Two between-subjects manipulations were included in hopes of heightening the subjects' desire to provide accurate judgments, thus presumably enhancing their attention to each target. In one condition, we explicitly warned subjects that undue reliance on the target's sex as a clue to height would result in less accurate judgments. The second manipulation made no

Table 2
Effects of Sex and Typicality on Height Judgments (Experiment 1)

Sex of target	Ascribed typicality: Typical	Ascribed typicality: Atypical
Male	69.99	69.91
Female	66.17	66.36

Note. Entries are mean estimated heights, in inches.

personal appearance that was correlated with gender might be responsible for the apparent stereotype effect. In particular, it seemed possible that our female targets were more likely than the male targets to wear shoes with high heels or hairstyles that rose substantially above the model's scalp, thereby adding significantly to their perceived height. If this were the case, subjects might have corrected for these gender-related height enhancers, thereby giving height estimates for the female targets that were significantly lower than they should have been. Though we asked subjects to estimate the heights of the models including their footwear, we were cognizant of the possibility that subjects might have ignored or forgotten these instructions; furthermore, we failed to include instructions concerning hairstyle. To better address these problems, we asked two female judges who were naive about the hypotheses of this research to examine the photographs and identify the targets whose hairstyles or footwear might have made a half-inch or more difference in someone's estimate of their height. As a conservative test, we eliminated any target chosen by either or both of these judges (13 female and 8 male targets) and repeated the analyses. The results of this purified sample of photographs clearly replicated the data patterns that are summarized in Table 1 and in the preceding paragraphs.

mention of the association between sex and height but instead offered a valuable inducement (a significant cash reward) for accuracy. Our results indicated that these manipulations did not reduce the stereotype effect we had previously observed, for similar judgment patterns were observed in our control, don't stereotype, and cash conditions. We tentatively conclude from these results that base-rate effects in this domain are quite strong and resistant to instructional and motivational manipulation.

In contrast, our alterations of target characteristics (i.e., their career goals, hobbies, and favorite summer jobs) did reliably affect height judgments, albeit to a small degree. These changes were in the expected direction: Atypical men were seen as shorter than typical men; atypical women were seen as taller than typical women. With the wisdom of hindsight, however, we recognize that our manipulation of typicality was imprecise. That is, we cannot be certain if it is unusualness per se that reduced the impact of the sex stereotype or if our atypical targets were judged differently because they were seen as more typical of the other gender category. Our results do not allow us to determine if atypical men (in general) or feminine men (in particular) are seen as shorter than typical men and if atypical women or masculine women are seen as taller than typical women. Extrapolating from Rothbart and Lewis's (1988) work, we anticipated that a target who was atypical of its category (for whatever reason) might not be as strongly affected by the category stereotype as a typical member. Deaux and Lewis (1984) and Shapiro (1986) suggest further that when subjects assess a target who seems unusual in some regard, they feel free to infer that other aspects of the target (e.g., physical dimensions) may take on unusual values as well. Our data do not allow us to choose between these models, or between a generalized unusualness model and a more restricted model, on the basis of the possibility that a man with feminine interests might have a correspondingly feminine physique (i.e., be shorter), whereas women with masculine interests might also possess relatively masculine bodies and be judged as taller as a consequence.

Experiment 2

Experiment 1 testified to the robustness of our respondents' sex stereotypes. Stereotypes about men and women continued to exert a highly significant effect on the respondents' height judgments despite our attempts to eliminate this effect by changing the instructions and incentives associated with the judgment task and by varying the social characteristics of the targets. Reflecting on our instructions, however, we recognized that we might have sent mixed signals to our subjects. On the one hand, we admonished or enticed them to ignore their stereotypes when judging the heights of men and women, but on the other hand, we presented stimuli that continued to show a clear association between height and gender (i.e., on average, our male targets were taller than the females, as in the everyday world). To eliminate this problem, in Study 2, we changed the stimuli that were presented for judgment and reexamined the effects produced by monetary incentives and by don't stereotype forewarnings. In brief, we presented our subjects with a collection of stimuli in which sex was no longer a cue of diagnostic value. This was accomplished by selecting a series of target photos in which the distribution of height was identical for the male and female targets.

In Study 2, the male and female targets were matched for height, so that for every woman of a certain height, there was a corresponding man of the same height. Some subjects were explicitly informed on this manipulation, so they knew that the sex of the various targets would not constitute a valid cue for height. Half of the informed subjects were told, in addition, that a $50 prize would be awarded to the person whose height judgments proved to be the most accurate. A third (control) group was not informed of the matched targets, nor were they told of the reward for superior performance.

Method

Seventy-two University of Michigan students participated in Study 2 for a $5 reimbursement. Subjects were recruited in class and by means of posters in popular student haunts.

The photos in Experiment 2 were selected from the same pool we had used in our earlier studies. Test booklets included 22 male and 22 female single photos (11 standing and 11 sitting for each sex), the first 4 of which were considered practice trials. For single photographs (which appeared at the beginning of the booklets), respondents estimated the height of each target to the nearest half inch. As in Experiment 1, respondents were to estimate the height of each target, "wearing the shoes or boots in which he or she is shown." The singles series was followed by 16 pages of male–female pairs, consisting of side-by-side photographs of models who were equal in height, but whose pictures had been taken in different settings and at different distances. There were eight standing pairs and eight sitting pairs; subjects had to choose which of the models was taller and indicate their confidence in that choice, using a 7-point scale. Our models ranged in height from 64.5 inches to 74.5 inches. Subjects were told simply that we were interested in their ability to make physical judgments that were based on photographs.

We included two instructional manipulations in an effort to reduce the effect of the sex stereotype on the height judgments. In the informed condition, subjects were told the following:

> In this booklet, the men and women are actually of equal height. We have taken care to match the heights of the men and women pictured. That is, for every woman of a particular height, somewhere in the booklet there is also a man of that same height. Therefore, in order to make as accurate a height judgment as possible, try to judge each photograph as an individual case; do not rely on the person's sex.

The informed plus cash condition included this warning along with the promise of a $50 prize to the best judge of height. Finally, control subjects were told nothing about the distribution of height among our models, nor were they aware of the cash prize for the best judge.[3]

Results

Single photographs. The focal analysis of the study involved a 3 × 2 × 2 × 2 mixed design ANOVA; instructional condition and respondent's sex were between-subjects variables, and tar-

[3] In this experiment, as in the previous study, all subjects were eligible for the cash prize, though only some were informed of the prize prior to making their judgments.

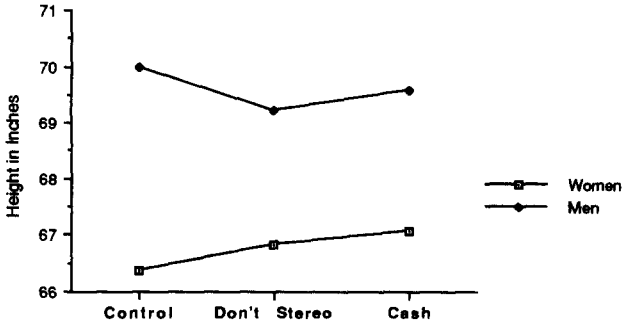

Figure 3. Average height judgments of male and female targets in three motivational conditions—control, don't stereotype, and informed plus cash: Experiment 2.

get sex and posture were within-subject variables.[4] For each subject, mean height judgments were calculated for the four different classifications of targets: standing men, sitting men, standing women, and sitting women. These data were entered as repeated measures in our analysis.

The results indicated a significant main effect that was due to the sex of the different targets, $F(1, 66) = 277.91$, $p < .001$; as in our earlier work, men were judged taller than the women, even though in this case there was no actual difference in the mean height of these targets. Study 2 also replicated the Target × Posture interaction that was observed in our earlier work, $F(1, 66) = 102.01$, $p < .001$. As before, respondents were particularly reliant on the sex stereotype when assessing seated (versus standing) targets.

Our ANOVA also showed a significant interaction between the sex of the target and the contrasting instructional conditions, $F(2, 66) = 5.32$, $p < .01$. Inspection of Figure 3 reveals that the difference between the height judgments assigned to male and female targets was attenuated in the two conditions in which the subjects had been alerted about our matching procedure for selecting targets (i.e., subjects in the informed and the informed plus cash conditions).

Although Figure 3 shows a reduced sex-of-target effect among subjects assigned to the two informed conditions, the sex-stereotype effect was clearly far from eradicated in these groups. In fact, for the two informed groups, the difference between the height estimates normally associated with the male and female targets was reduced by less than 50%. In other words, most of the stereotype effect remained.

Subsequent analyses that were based on data from just the two informed groups continued to yield highly significant effects of target sex on height judgments, $F(1, 44) = 118.77$, $p < .001$, despite our detailed comments regarding the invalidity of this cue. The Sex × Posture interaction was also significant, $F(1, 44) = 78.39$, $p < .001$, replicating our previous results and showing that even among informed subjects, height judgments tended to be more stereotypic when the targets were sitting rather than standing (i.e., the perceived height difference between male and female targets was amplified when the models were presented in a sitting posture).[5]

Reflecting on these results, we considered the possibility that our initial instructions regarding the equal heights of the male and female targets might have been forgotten over the course of the judgment series. That is, although we found a clear difference between the judgments prompted by male and female targets in the two informed conditions, these subjects may have initially responded similarly to the two target groups (as called for in the instructions) but then showed the familiar base-rate effect in later trials as the experimenter's introductory remarks faded from memory. To address this possibility, a final ANOVA was performed, this time entering the judgments for male and female targets, standing and sitting, for each quarter of the test series, namely, the first 10 photos, the second 10, and so on. Thus, a total of 16 means were entered for each subject, using trial block (1–4) as an additional within-subjects factor.

If the effect of our instructions had diminished over time for the informed respondents, because of their forgetting, the assessments of male and female targets should have become increasingly disparate during the course of the experiment for these informed subjects but not for control subjects. This process would be reflected in a triple interaction: Sex of Target × Trial Block × Experimental Condition. This interaction was far from significant, however, suggesting that over time, the pattern of differences between the judgments associated with male and female targets did not differ across instructional conditions, $F(6, 198) = .45$, $p > .84$. Although the more general two-way interaction between sex of target and trial block was significant, $F(3, 198) = 40.47$, $p < .001$, this simply means that the difference in judged height between the male and female targets did not remain constant over the four trial blocks. Figure 4 shows these results by quarter for subjects in the informed and informed plus cash conditions. Clearly the subjects in these two conditions did not come to see a greater divergence in height between male and female targets as the experiment progressed. Instead there appeared to be a gradual convergence in height ratings, perhaps because of our respondents' increased exposure to the matched targets. In any event, Figure 4 suggests that the informed subjects did not simply forget that the male and female targets had been matched with respect to height.

Pair judgments. In the last part of the test booklet, we sought to determine if implicit base rates regarding the relative height of men and women would affect our respondents' choices in a series of paired-comparison trials involving male and female targets who were matched with respect to height. In previous work with this paradigm, we found a significant tendency to overchoose the male models. We wanted to replicate

[4] Subject sex was found to play no role as a main effect or in interaction with any other variable. Thus, it was not considered further.

[5] As in Study 1, we were concerned about the potential role that high heels or unusually voluminous hairstyles may have played in creating a spurious relationship between the gender of a target and his or her judged height (see Footnote 2). Using the procedure we had developed in Study 1, we therefore eliminated targets whose hair or shoes seemed problematic. For Study 2, we eliminated not only the troublesome photos (three men and six women) but also their matches in height, so that we ended up with a reduced set of targets that nevertheless retained equal height distributions for the male and female targets. We repeated the analyses and successfully replicated our initial patterns of results.

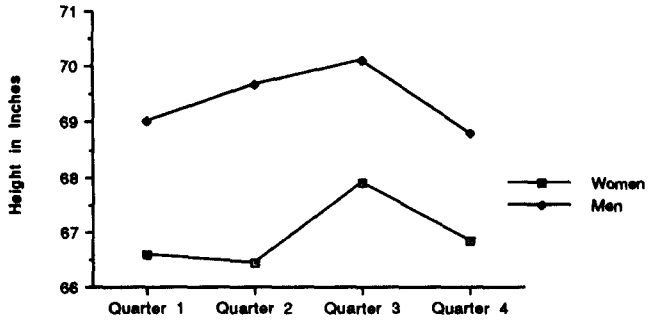

Figure 4. Average height judgments of female and male targets, over time, for subjects in informed and informed plus cash conditions: Experiment 2.

this bias and, more important, to see if it could be moderated by the instructional and incentive manipulations that we had introduced. To answer these questions, we conducted a 3 × 2 × 2 ANOVA; condition (control, informed, informed plus cash) and sex of subject were between-subjects variables, and posture of the target pair was a within-subject variable. The dependent variable was the proportion of the target pairs (standing and sitting pairs were scored separately) in which each subject chose the man as the taller of the two pictured individuals. The range for this variable ran from 0 (subject never chose the man as taller) to 1.0 (subject always chose the man as taller than the woman with whom he was paired).

The three instructional conditions yielded significantly different results, $F(2, 66) = 5.98$, $p < .01$. The mean choice proportions for the two experimental conditions (informed and informed plus cash) hovered around the chance level of .50 (M = .492 for the informed condition; M = .492 for the informed plus cash condition), whereas the mean proportion in the control condition was substantially higher (M = .592).[6]

The difference between the choice results observed in the control condition and the average of the two experimental conditions was then subjected to a t test. The results were highly significant, $t(70) = 3.20$, $p < .01$. Finally, the observed choice results for each group were compared with the null expectation of .50. For the control group, there was a significant bias in favor of the male targets, $t(70) = 3.66$, $p < .01$; by contrast, neither of the experimental groups differed from chance expectations.[7]

Discussion

Study 2 provides additional evidence regarding our subjects' robust reliance on gender as a cue to the height of different targets. We did, however, succeed in reducing the magnitude of this effect. By telling our subjects that sex would be of no value in diagnosing target height, we managed to reduce the effect of the sex stereotype on our respondents' height estimates. Nonetheless, there was a significant residual difference between the height estimates associated with the male and female targets in our two informed groups, a lingering stereotype effect that was not reduced further by the introduction of a substantial monetary incentive for superior performance. Indeed, the overall reduction of the sex stereotype effect was relatively modest, amounting to less than half of the effect size that was observed in the control condition.

We did finally succeed in eliminating the stereotype effect in our subjects' judgments of the target pairs. Control subjects replicated an effect we had previously obtained by overselecting the men as the taller targets in a series of paired-comparison judgments. This bias disappeared in our two experimental conditions, leaving no residual stereotype effects. However, we cannot be sure that our instructions altered subjects' perceptions of relative target height for mixed-sex pairs. Instead, our informed subjects may have adopted a strategy of choosing the man as taller for roughly half of the test pairs, regardless of his apparent height in relation to the woman, because subjects had been forewarned that the average height for the male and female targets was the same. The smaller number of test pairs (16) in relation to the number of singles (44) combined with the simple dichotomy of judgment options (man taller vs. woman taller) would make such an accounting strategy plausible for test-pair photos, but not for single photos.

Conclusion

The main conclusion we derived from these experiments was that group stereotypes have a continuing, robust effect on the evaluations of individual group members. In spite of earnest efforts to convince our subjects that the sex of the different targets should not influence their height estimates and regardless of the substantial monetary reward that could have been won through superior sensitivity to the height of the individual targets, the subjects in these experiments continued to show a significant reliance on group stereotypes when evaluating individual targets. The stereotype effect was particularly marked when subjects evaluated sitting targets, presumably because of the ambiguous height cues that these pictures provided. These results are provocative with respect to several theoretical issues.

Automaticity

In part, the present results parallel Devine's (1989) findings that some stereotype effects are involuntary and difficult to overcome, despite the good intentions of the evaluator. Devine found that by presenting subliminal cues that are widely associated with American Blacks, she was able to exert a significantly negative impact on her respondents' assessments of neutral target behaviors. This effect was interpreted as a type of priming phenomenon, in which the presentation of a subliminal Black prime increased the availability of associated (largely negative) stereotypic attributes. What was most disturbing was that positive attitudes toward Blacks did not moderate the ste-

[6] Contrary to the data from the single photo test trials, for reasons that are presently unclear, this tendency to choose the men as taller was not qualified by an interaction with posture, $F(2, 66) = .26$, *ns*.

[7] We conducted a similar set of analyses using a dependent variable that included both the respondents' observed pair choices and their confidence in these choices. The results were essentially unchanged.

reotype effect, presumably because all members of American society are familiar with the cultural stereotype of Blacks, whether or not they consciously accept these images as valid.

In the present work, in contrast to Devine's (1989) experiments, we used supraliminal stereotype cues and attempted to lessen their impact by warning the subjects that their everyday assumptions (stereotypes) would interfere with their ability to judge the targets, because the normal association between sex and height no longer applied. Although these instructions significantly reduced the magnitude of the effects that we observed, there was substantial evidence of residual stereotyping, in that the perceived difference between the height of male and female targets remained highly significant ($p < .001$). We concluded, therefore, that people may be largely unable to control the influence of real-life base rates (e.g., the stimulus–response association between sex and height) that have been built up over a lifetime of experience, despite their best attempts to do so. In a similar vein, Manis and Ruppe (1969) demonstrated the persistence of learned behavior patterns, regardless of people's conscious efforts to suppress them. In the Manis and Ruppe experiment, respondents first learned to use plural (or singular) nouns as subject words in a sentence construction task. Subjects then completed a series of test trials in which they were instructed to discover an unspecified new rule for constructing sentences and were furthermore explicitly told that the previous rule would be irrelevant from that point on. All subjects were conscious of the old rule and its inappropriateness for the new task; nonetheless the old rule concerning plural versus singular subject words continued to affect their behavior, just as the informed subjects in Experiment 2 continued to rely on the sex of the various targets to guide their judgments even though they had been explicitly told that sex was an invalid cue.

Motivation

Our attempts to reduce the respondents' reliance on sex-stereotypes by providing extra monetary incentives or by imploring subjects to strive for accuracy were completely ineffective. We had anticipated that by providing a special incentive for superior performance we might induce subjects to overcome the routinized, stereotype-driven judgments that we had observed in our earlier work (see Figures 1 and 2) and to become more responsive to the height cues contained in the individual photographs. These motivational inducements were conceptual replications of procedures that had previously led Neuberg and Fiske's (1987) subjects to override category-based judgment strategies. The role of motivation in the stereotype process is thus unclear at the present time and may depend on complex interactions between the type of motive, the content of the prevailing stereotypes, and the domain of judgment.

To account for their results, Fiske and her colleagues (Fiske et al., 1987; Fiske & Pavelchak, 1986) have proposed a continuum model in which the perceiver's judgments are said to reflect some combination of category-based and individual- (or "piecemeal")-based strategies. The relative strength of categorical and individual variables depends in part on the demands placed on the perceiver by the judgment context. Presumably, the greater the competing task demands on the perceiver, or the lower his or her motivation to produce precise judgments, the more likely it becomes that judgments will reflect a largely category-based strategy. This model would lead us to anticipate a rather different pattern of results from the robust stereotype effect that we observed here. A skeptical reader might, as a consequence, attribute our results to mundane features of this particular experimental paradigm. For example, perhaps the subjects felt rushed and thus paid little attention to the specific features of the individual targets. Or perhaps our subjects were simply bored with the task, or fatigued because of the considerable number of judgments required of them.

Although such interpretations are conceivable, they strike us as doubtful for the following reasons: (a) Subjects were encouraged to work at their own pace, (b) the available data suggest that the magnitude of the stereotype effect was reduced rather than enhanced as subjects proceeded through the test booklet (see Figure 4 and associated text), despite the fact that careless responding was probably increasingly prevalent as the experiment progressed, and (c) the introduction of a substantial cash incentive presumably inspired greater care and thoughtful responding. Nonetheless, as noted earlier, the monetary incentive had no effect on our respondents' judgments. For these reasons, we believe that the gender effect observed here reflects the power of the stereotype and cannot simply be attributed to the demands of this particular experimental paradigm.

Base Rates and Judgment

In contrast to the bulk of the base-rate literature, which indicates that base rates may (or may not) influence the judgment process, depending on a variety of situational factors, the present results show clear and robust evidence of stereotype attentiveness across a variety of informational and motivational circumstances. The impact of these base rates is particularly impressive when we note that they were not explicitly mentioned by the experimenter (as has typically been the case in past base-rate research) but apparently were spontaneously evoked by the gender of each successive target.

Note also that the base-rate effects we obtained derived in large part from the respondents' everyday contact with men and women, an experience that doubtlessly promotes a systematic association between gender cues (male or female) and height. By contrast, the bulk of the base-rate literature deals with story problems in which the relevant base rates are presented verbally as part of the overall story situation. As Ginossar and Trope (1987) suggested, story problems may evoke a variety of problem-solving approaches, because "some contexts may encourage the use of statistical rules (e.g., reliance on base-rate information), whereas others may encourage the use of nonstatistical rules" (p. 473). The present results suggest that objective (real-life) base-rate effects may be substantially more robust, and the results raise the possibility that base-rate neglect may be less common than earlier work had implied.

Response Scales

The response scale that was used in these studies (feet and inches) is familiar, quantitative, and objective. Scales like this

have a special virtue, in that there is universal agreement that a man of 5'7" and woman of 5'7" are in fact equal in height. Now consider, by contrast, the sort of subjective rating scales that are so commonly used in studies of social judgment, for example, a 1-to-7-point scale with endpoints labeled *very timid* and *very aggressive.* We suspect that scales of this sort may lead to misleading conclusions, if they are assumed to accurately reflect the perceiver's subjective representations of individual men and women. Most important, the endpoints of such a scale may be differentially located when evaluating a man versus a woman. As a consequence, a woman who was rated as *very aggressive* (through implicit comparison with other women) might still be considered less aggressive than a man who received that same rating (because he was thought to be substantially more aggressive than most men, a more extreme standard). This problem is very widespread in social research, and it may contribute to some of the negative stereotyping results that have been reported in the literature (Locksley et al., 1980; Locksley, Hepburn, & Ortiz, 1982; Rasinski, Crocker, & Hastie, 1985).

Cognitive Representation of Base Rates

These studies testify to the robust influence of category knowledge on the respondent's appraisal of individual category members (exemplars). The results are generally consistent with the dictates of Bayes's theorem, in that the respondents were affected both by individuating information about the different targets and by stereotypes pertaining to the heights of men and women. Although we use the terms *base rate* and *stereotype* almost interchangeably, note that the base-rate effects we observed here need not derive from the respondents' reliance on abstract, statistical information; indeed, this seems unlikely. Instead, we believe that the stereotype effects that we report probably reflect the respondents' use of an informal anchoring and adjustment strategy, in which the tentative starting point for a given target may begin in a scale region that seems reasonable for that type of target (e.g., a man). This anchor may then be adjusted through closer inspection of the target figure. Tversky and Kahneman (1974) suggested, however, that the adjustments in such an approach are typically insufficient, leading to judgments that are biased in the direction of the initial anchor (i.e., the sex stereotype in the present case).

The base-rate concept provides a useful operationalization of how people represent the category–trait linkages that make up the content of stereotypes. Researchers have differed in how they account for the varying strengths of association between a stereotypic label and the traits that constitute the stereotype. Fiske and Pavelchak, 1986, for example, proposed a cognitive network model in which a central label node is linked to separate trait nodes by means of associations of varying strength. McCauley and Stitt's (1978) proportional model used a more quantitative representation of the category–trait link. For example, one's stereotype of librarians might hold that virtually all librarians (95%, perhaps) would be characterized by a love of books but that a smaller proportion might be characterized as nearsighted (say, 75%). This probabilistic link between a category and its constituent traits is effectively captured by the concept of the base rate, or the percentage of a given group that is thought to possess a specified trait. The base-rate shorthand not only makes for a useful notation to describe stereotype content, but it also relates easily to regression-based approaches (e.g., path analysis), in which coefficients represent the differing impacts of categories (stereotypes) and individuating information on judgments.

Although our results suggest that respondents were sensitive to category knowledge in producing their height estimates, a skeptical reader might worry about the possibility that the association between sex and judgment was spurious; that is, the sex of the different targets may have been associated with an unspecified "Cue x," which in turn, exerted a direct influence on the respondents' height judgments. For example, the male targets generally may have been more muscular than the female targets, and the more muscular targets may have tended to evoke higher height estimates. Although a model for such a spurious effect is logically possible, it strikes us as unlikely, given the results of our path analyses.

Consider Figure 2, which shows that the standardized path coefficient linking target's sex and the associated height judgment was .624. If this path coefficient represents the product of two constituent paths (from the target's sex to Cue x and from Cue x to the observed height judgments), it places severe and seemingly implausible constraints on the magnitude of these hypothesized relationships. Most important, because the alternative model requires that the product of the constituent paths be equal to .624 and because neither path can exceed a value of 1.00 (representing a perfect relationship between the variables in question), we are inevitably led to the recognition that neither of the constituent paths can fall below a value of .624. Indeed, were Cue x to exert less-than-complete control over the observed height judgments (e.g., a path coefficient, say, of only .80), then the path from target sex to Cue x must have a path coefficient of .78, because $.80 \times .78 = .624$. This line of reasoning leads us to conclude that Cue x (if it exists) must be very closely related to the targets' sex; indeed the two variables would have to be so closely related as to be virtually indistinguishable. On the other hand, this line of reasoning does not exclude the possibility that the relationship between gender and height estimates might be due to the respondents' reliance on several mediating cues, not just one. Although the present results do not preclude this type of multicue mediation, without further explanation, we find this to be an unconvincing interpretation of our results.

Generality

How general are these results? Although we obtained consistent evidence of stereotyped processing, these results were based on a genuine, observable difference between men and women (i.e., the fact that men are normally taller). Would similar results be observed in a case where the respondents' beliefs about a given group derived from common knowledge, or hearsay, rather than from direct, everyday experience?

Many social psychologists would anticipate similar results, whether the respondents' underlying beliefs about a relevant group were anchored in daily living or were derived from hearsay, rumor, myth, or anecdote. They might reason that the

things that one believes have real consequences for his or her subsequent beliefs and actions, regardless of their origins and regardless of their factual accuracy. This assumption receives support from the many studies of stereotypes that are unlikely to have their roots in the respondents' personal experiences (e.g., Darley & Gross, 1983; Devine, 1989; Sagar & Schofield, 1980) and from studies of stereotypes with varying degrees of accuracy (see Mackie, 1973, for a review).

On the other hand, the robust influence of gender (as a cue) in height estimates may be because this association is repeatedly strengthened through innumerable daily contacts. This experience may contribute significantly to the apparent inability of people to disregard the gender of a target when estimating his or her height. This is an issue that warrants further research.

We suggest that even if most stereotypes derive from folk wisdom, a judge might plausibly develop important and influential stereotypes through personal experience. For such cases, our results may be particularly relevant. Here are some examples:

1. Consider the grade school teacher who has noted that his students from middle-class homes seem to show better academic aptitude (as that term is presently understood) than students from economically deprived homes, for whatever reason. Our data suggest that the many teachers who have experienced these group differences may find it difficult to ignore a student's social origins when appraising his or her academic aptitude. That is, empirically based beliefs about the relationship between social class and scholastic aptitude may affect the teacher's appraisal of individual students, even though the teacher is motivated to be fair and even if he has been assured by his principal that the deprived students of the coming term were selected so that their academic ability, on average, is the same as that of the middle-class students.

2. A clinical psychologist may have personally noted that her patients (as is true in society at large) are more likely to suffer from alcoholism if they come from Native American families than if they are from Chinese-American or Jewish-American homes. Suppose this clinician is now serving as an expert witness and is asked to appraise a new patient whose dependence on alcohol is uncertain. Will her judgment concerning possible alcoholism be influenced by the patient's ethnic origins? Our data suggest that she will be influenced, even if the court urges that the clinician base her judgment solely on the patient's thoughts and action, not on his ethnicity. Note, moreover, that Bayes's theorem, along with other formulations, suggests that in trying to arrive at an accurate judgment, our clinician probably should be influenced by ethnic considerations.

Suppose now that in trying to make her judgment, our clinician learns that her patient was an only child. Suppose further that a professional colleague tells her that in his experience, single children show alcoholism rates that are about the same from one ethnic group to the next. According to this view, people from Native American, Chinese-American, and Jewish-American homes are equally likely to suffer from alcoholism if they are only children. Sadly, even if our clinician accepts her colleague's observation, the present data suggest that given her previous personal experience, she may be unable to suppress a reliance on ethnicity (as a cue) when assessing a particular patient's dependence on alcohol.

Although we do not consider these hypothetical scenarios to be unrealistic, they are not intended as substitutes for further empirical inquiry. The height-judgment paradigm was admittedly chosen in part because of its convenience, but we do not feel that the effects we observed are unique to this domain. We believe instead that they may be representative of a large class of stereotypic judgments that, like the biases examined by Fischhoff (1982), are so natural and heuristic as to resist many debiasing efforts.

References

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology, 35,* 303–314.

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

Asher, H. B. (1983). *Causal modeling.* Beverly Hills, CA: Sage.

Ashmore, R. D., & Del Boca, F. K. (1981). Conceptual approaches to stereotypes and stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 1–36). Hillsdale, NJ: Erlbaum.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology, 44,* 20–33.

Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes: Interrelationships among components and gender label. *Journal of Personality and Social Psychology, 46,* 991–1004.

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

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). New York: Cambridge University Press.

Fiske, S. T., Neuberg, S. L., Beattie, A. E., & Milberg, S. J. (1987). Category-based and attributed-based reactions to others: Some informational conditions of stereotyping and individuating processes. *Journal of Experimental Social Psychology, 23,* 399–427.

Fiske, S. T., & Pavelchak, M. A. (1986). Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect. In R. M. Sorrentino & E. T. Higgins (Eds.), *Motivation and cognition: Foundations of social behavior* (pp. 167–203). New York: Guilford Press.

Ginossar, Z., & Trope, Y. (1980). The effects of base rates and individuating information on judgments about another person. *Journal of Experimental Social Psychology, 16,* 228–242.

Ginossar, Z., & Trope, Y. (1987). Problem solving in judgment under uncertainty. *Journal of Personality and Social Psychology, 52,* 464–474.

Hilton, J. L., & Fein, S. (1989). The role of typical diagnosticity in stereotype-based judgments. *Journal of Personality and Social Psychology, 57,* 501–511.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.

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

Krueger, J., & Rothbart, M. (1988). Use of categorical and individuating information in making inferences about personality. *Journal of Personality and Social Psychology, 55,* 187–195.

Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology, 39,* 821–831.

Locksley, A., Hepburn, C., & Ortiz, V. (1982). Social stereotypes and judgments of individuals: An instance of the base rate fallacy. *Journal of Experimental Social Psychology, 18,* 23–42.

Mackie, M. (1973). Arriving at "truth" by definition: The case of stereotype inaccuracy. *Social Problems, 20,* 431–447.

Manis, M., & Ruppe, J. (1969). The carryover phenomenon: The persistence of reinforced behavior despite the absence of a conscious behavioral intention. *Journal of Personality and Social Psychology, 11,* 397–407.

McCauley, C., & Stitt, C. L. (1978). An individual and quantitative measure of stereotypes. *Journal of Personality and Social Psychology, 36,* 929–940.

Neuberg, S. L., & Fiske, S. T. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology, 53,* 431–444.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90,* 339–363.

Rasinski, K. A., Crocker, J., & Hastie, R. (1985). Another look at sex stereotypes and social judgments: An analysis of the social perceiver's use of subjective probabilities. *Journal of Personality and Social Psychology, 49,* 317–326.

Rothbart, M., & Lewis, S. (1988). Inferring category attributes from exemplar attributes: Geometric shapes and social categories. *Journal of Personality and Social Psychology, 55,* 861–872.

Sagar, H. A., & Schofield, J. W. (1980). Racial and behavioral cues in Black and White children's perceptions of ambiguously aggressive acts. *Journal of Personality and Social Psychology, 39,* 590–598.

Shapiro, P. (1986). *The illusory correlation of distinctiveness: A general mechanism of inference with applications to stereotyping.* Unpublished doctoral dissertation, University of Wisconsin—Madison.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Tversky, A. & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 117–128). New York: Cambridge University Press.

Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology, 45,* 961–977.

Received May 25, 1989
Revision received April 30, 1990
Accepted May 1, 1990 ■