

Use of Categorical and Individuating Information in Making Inferences About Personality

Joachim Krueger and Myron Rothbart
University of Oregon

In three experiments, we explored the effects of categorical information (stereotypes) and case information (traits or behaviors) on judgments about an individual's characteristics. Subjects judged a target person's aggressiveness on the basis of a description containing both a broad social category and specific case information. In Experiment 1, the description included (a) a category that was either weakly or strongly related to aggressiveness and (b) a behavior that was unrelated, moderately diagnostic, or highly diagnostic of aggressiveness. Trait inferences were a function of both the stereotypic and the behavioral information. A single behavior was not sufficient to override the category effect. In Experiment 2, temporally consistent behaviors were presented as case information; under these conditions, category information had no effect on trait judgments. This finding was extended in Experiment 3 in which subjects predicted behaviors on the basis of the target person's sex and a moderately diagnostic trait.

The neglect of categorical information is a common finding in the literature on human judgment (Kahneman & Tversky, 1973; McArthur, 1972). When human observers make inferences about people's attributes or about the causes of their behavior, they typically rely more on the specific properties of the stimulus person than on the general statistical properties of the class to which the stimulus person belongs. The apparent failure to appropriately incorporate categorical information has attracted much research attention because it violates the assumptions of a number of influential theories. In Kelley's (1967) normative attribution theory, for example, trait ascriptions should depend in part on the magnitude of the perceived social consensus. In particular, the more prevalent a behavior is in a referent population, the lower is the probability of attributing the cause of that behavior to an actor. Research on causal attribution has repeatedly demonstrated insufficient use of consensus or base-rate information (e.g., McArthur, 1972). That is, when inferring the causes for an individual's behavior, subjects tend to underuse information about how others typically behave in the same setting. Similarly, Kahneman and Tversky showed that population base rates are largely ignored when decisions are made about the category membership of a target individual. In a frequently cited study, Kahneman and Tversky described a friend who is shy, likes to write poems, and is small of stature. Asked whether this friend is a professor of Chinese studies or a psychologist, most subjects chose the professor of Chinese stud-

ies. The similarity between the description of the friend and the stereotype of a professor of Chinese studies overrode powerful base-rate differences between the two professional categories.

In social stereotypes, the base-rate fallacy may occur in two ways. First, as Kahneman and Tversky (1973) have shown, observers often ignore the relative size of a social category and estimate the likelihood of category membership solely on the basis of representativeness or similarity. Second, when group membership is clear and individuating information about the target person is available, inferences about personal attributes may be based too heavily on case information. Inasmuch as social stereotypes imply category differences regarding a set of attributes, the judgment literature suggests that such category information will be used when case information is absent but will exert only a minimal influence when case information is present.

Several studies using gender as the categorical variable support this idea (Eagly & Wood, 1982; Locksley, Borgida, Brekke, & Hepburn, 1980). In one experiment (Locksley et al., 1980, Experiment 1) subjects read a five-page description of the behavior of a single male or female stimulus person. Each description contained three episodes in which the person had consistently behaved in an assertive or in a passive way. For example, an assertive target person was reported to have forcefully told an annoying person to go away, to have interrupted a voluble classmate to make a point, and to have broken into a group conversation at a party. The respective passive responses consisted simply of a failure to behave assertively. A second experiment (Locksley et al., 1980) showed that even a single assertive or passive episode is sufficient to override effects of gender stereotypes. Women who were described as having engaged in assertive behavior were rated as assertive as were men who had engaged in the same behavior. Only when no other diagnostic information was available about a target individual did the social stereotypes have an effect on assertiveness ratings. These findings were replicated (Locksley, Hepburn, & Ortiz, 1982,

This research was supported by National Institute of Mental Health Grant MH 40662-02 to Myron Rothbart.

We thank Lewis Goldberg, Susan Green, Oliver John, Robert Mauro, and four anonymous reviewers for their thoughtful comments on an earlier draft of this article.

Correspondence concerning this article should be addressed to Joachim Krueger, Max Planck Institute for Human Development and Education, D-1000 Berlin 33, Lentzeallee 94, Federal Republic of Germany.

Experiment 2), using sets of four behaviors that were diagnostic either for assertiveness or for passivity. When more than half of the behaviors (three or four) were diagnostic of the trait, gender had no effect. Finally, Locksley et al. (1982, Experiment 1) found no category effects when they asked subjects to predict several traits stereotypic of "day people" and "night people" from personal sketches containing both diagnostic trait and behavior information. Again, predictions depended solely on the diagnostic case information.

Although this research (Locksley et al., 1980; 1982) optimistically predicts the neglect of (potentially harmful) stereotypes in the presence of case information, it is not clear that behavioral information relevant to the criterion neutralizes beliefs regarding differences among important social categories. Psychologists have, for example, long lamented the persistence of gender bias in personnel selection tasks. When specific case information of equal quality was available about female and male candidates, selectors tended to favor the men (Dipboye, Fromkin, & Wiback, 1975; Heneman, 1977). Indeed, there is a long research tradition in social psychology that demonstrates the importance of contextual information in social perception. Asch (1946) argued from a Gestalt perspective that behaviors are perceived and interpreted in the context of generalized expectations. Rosenhan (1973) argued that normal behaviors will be perceived as deviant when exhibited by people who have been labeled as schizophrenic in a mental institution. Sagar and Schofield (1980) found identical but ambiguous behaviors to be perceived as more aggressive when engaged in by a Black person than when engaged in by a White person. According to this research, stereotypes can influence the perception of acts, albeit ambiguous acts, by coloring the meaning of those actions in a manner consistent with stereotypic expectations.

Another line of research that may limit the generalizability of the base-rate fallacy focuses on the causal meaning of the base rates. Ajzen (1977), Tversky and Kahneman (1979), and Bar-Hillel (1980) created situations in which the base rates were enriched with causal meaning and found that subjects' responses strongly reflected the base rates. Base rates that were equally strong and predictive, but noncausal, were ignored. Zuckerman (1978) found effects of base rates when the causal meaning of the case information was reduced. He distinguished between two kinds of behavioral events: actions and occurrences. Actions are high in perceived voluntary control or personal causation (e.g., "Jerry attended the Sunday meeting"). Occurrences are perceived to be under personal control to a lesser degree (e.g., "Mary passed the exam in history"). In accord with the predictions, base rates influenced attributions of occurrences to a larger degree than they did actions. This line of research suggests that categorical information is used when plausible causal links can be established between the category and the behavior, and the neglect of base-rate information is most likely when the base rate is noncausal and the individuating information is causally relevant to the behavior.

A variety of techniques have been proposed to increase the magnitude or the salience of the base-rate information. Wells and Harvey (1977), using Nisbett and Borgida's (1975) paradigm, numerically strengthened the consensus (base-rate) manipulation and found that it did affect predictions. Feldman, Higgins, Karlovac, and Ruble (1976) manipulated the percep-

tual salience of base-rate information by presenting videotaped actions of others and found that visually salient base rates were more likely to be used by decision makers.

These studies suggest that general beliefs about differences between social categories are not inevitably neglected in the presence of more specific case information. Rather, the observer's inferences may be determined by the relative strength of the categorical and the individuating information. Reliance on case information may be most likely when the observer is given both weakly diagnostic categorical information and highly diagnostic case information. It is our view that these may have been the conditions under which Locksley et al.'s (1980; 1982) findings were obtained. In their experiments, the categorical information, gender, provided a rather weak basis for predictions of individual attributes. Although an extensive literature confirms the high level of consensus regarding sex stereotypes (Bem, 1974; Spence & Helmreich, 1978), consensus and diagnosticity are not the same. People may agree on perceived differences in assertiveness between the average man and the average woman, but they may also perceive the distributions of male and female assertiveness as highly overlapping, allowing for the likely possibility of assertive women and passive men. Before concluding that individuating information inevitably overrides category information, it would be important to vary the diagnostic strength of the categorical information.

In addition to varying the diagnosticity of the categorical information, one also needs to vary the diagnosticity of the individuating information. Locksley et al. (1980, 1982) provided different versions of case information to their subjects. Typically, the descriptions of the target persons were highly diagnostic of the criterion trait and were consistent across time and setting. Even in its weakest form, the case information still consisted of a one-paragraph description of an unambiguously assertive behavior (Locksley et al., 1980, Experiment 2).

To summarize, Locksley et al.'s (1980; 1982) research established that categorical information that is weakly diagnostic of a trait is neglected when paired with highly diagnostic case information. We conducted three studies in which we varied both the diagnosticity of the category and the individuating information. In our first experiment, subjects were presented with sketches of stimulus persons who belonged to a social category that was either weakly or strongly diagnostic of the trait of aggressiveness. Stimulus persons were described as having engaged in a single behavior that was neutral, moderately diagnostic, or highly diagnostic of that trait. We predicted that when a target individual belonged to a weakly diagnostic category and had engaged in highly diagnostic behaviors, category information would have no effect on ratings of aggressiveness. However, when the categorical information was strong or when weak categorical information was paired with only moderately diagnostic case information, we expected the categorical information to significantly affect judgment.

Experiment 1

Overview

Half of the subjects were presented with information about four target persons who were members of gender categories that

differed strongly in aggressiveness (two construction workers and two housewives), and the other half of the subjects read information about persons who were members of categories that differed only moderately in aggressiveness (two men and two women). In addition to the category information, the personal sketches included information about a single recent behavioral act that was neutral, moderately diagnostic, or highly diagnostic of aggressiveness. For each subject, two of the sketches (one male and one female) included nondiagnostic behavior. The two remaining sketches described either moderately or highly diagnostic behavior. On the basis of this set of categorical and case information, subjects rated the expected frequency and intensity of future aggressive behavior.

Method

Development of stimulus materials. Pretests were conducted to select stimulus materials representing the three components of a prediction task: (a) a trait criterion to be predicted, (b) social categories of varying diagnosticity for the criterion, and (c) behaviors of varying diagnosticity. Eighteen subjects rated men and women, as well as a variety of subordinate gender categories, on 25 scales of the Personal Attributes Questionnaire (Spence & Helmreich, 1978). The trait of aggressiveness differentiated between men and women most strongly. We found that on a 7-point scale men were seen as more aggressive than women (5.21 vs. 3.71). Reviews of research on gender differences confirm that aggressiveness is one of the few stereotypic traits that show a reliable gender difference (e.g., Frodi, McCauley, & Thome, 1977; Maccoby & Jacklin, 1974). Furthermore, aggressiveness was the only trait on which two occupational subcategories of gender (construction worker and housewife) showed a greater difference (5.36 vs. 3.56) than the broad gender categories. Contrary to expectation, the trait of assertiveness did not emerge as a gender-stereotypic trait, confirming our suspicion of the relative weakness of the gender categories in predicting the presence of this trait.

On the basis of diagnosticity ratings, we selected behaviors so that their diagnosticity varied across three levels. Nine subjects rated a sample of behaviors on a 9-point scale of diagnosticity. We selected two behaviors that were rated as neutral ("... recently bought the latest book of a bestselling author" and "... recently went to the barber/hairdresser"). We chose items that were significantly more diagnostic than the neutral items for the moderately diagnostic behaviors ("... complained to a store manager about the quality of a product" and "... yelled at his/her spouse"). Finally, the two highly diagnostic behaviors were rated as significantly more diagnostic than were the moderate items ("... beat his/her child" and "... hit someone who annoyed him/her").

Subjects. In exchange for extra course credit, 53 male and 75 female students from introductory psychology courses at the University of Oregon participated in this study.

Design. Each subject received a booklet containing descriptions of four stimulus persons. Two of these descriptions were of male and two were of female stimulus persons. For half of the subjects, all stimulus persons were simply identified by their first name, to convey male or female gender (weak categories). For the other half of the subjects, the stimulus persons were described as construction workers or housewives (strong categories). For all subjects, one male and one female stimulus person were described by a behavior that was nondiagnostic of aggressiveness. For half of the sample, the other two target individuals were described by moderately diagnostic behavior; for the other half, they were described by highly diagnostic behavior. Because each subject responded to only two of the three levels of behavioral diagnosticity (neutral and low or neutral and high), diagnosticity was both a within- and a between-subjects variable.

Dependent measures. After reading each description, subjects rated the person's aggressiveness on two scales. The first dependent measure was a scale ranging from 0 to 100 on which subjects marked the percentage of future situations in which the target individual would respond aggressively. The second dependent variable was a measure of the expected intensity of future aggressiveness in four aggression-provoking situations. For example: "Michael has been urged by a door-to-door salesperson to subscribe to a number of magazines. Although Michael has expressed little interest in subscribing, the salesperson persists." For each situation, there was a list of seven possible responses, ranging from very aggressive ("Threatens to call police if salesperson does not leave") to passive ("Buys subscriptions although he doesn't want them"). Subjects rated the likelihood of each response on a 7-point scale. We computed a composite index of rated intensity of aggression by summing the top two ratings across the four situations, with higher values reflecting higher perceived aggressiveness.

Results

Because the two measures of predicted aggressiveness (frequency and intensity) correlated highly within subjects and across target persons ($r = .94, p < .001$), the two indexes were combined into a single measure. The frequency measure was transformed to correspond to the metric of the intensity measure, so that the two variables (having the same mean and standard deviation) could be combined with equal weight to provide an overall index of judged aggressiveness.

The results for this composite index are presented in Figure 1. Data were analyzed in a $2 \times 2 \times 2$ (Strength of Category \times Sex of Stimulus Person \times Diagnosticity of Behavior) analysis of variance (ANOVA) with repeated measures on the last two variables. Separate analyses were performed for the subjects who rated the neutral and moderate levels of behavioral diagnosticity and for the subjects who rated the neutral and high levels of diagnosticity.

The first analyses involved the subjects who had been given neutral and moderately diagnostic behaviors. The degree of diagnosticity of the target person's behavior affected trait attributions, thus reassuring us that our moderate-diagnosticity manipulation was successful. Persons who had engaged in moderately aggressive behavior were predicted to be more frequently aggressive in future situations than persons who had engaged in behavior unrelated to aggression, $F(1, 55) = 105.43, p < .001$. Moreover, as expected, men were rated as significantly more aggressive than women, $F(1, 55) = 64.37, p < .001$. More important, there was a significant interaction between sex of stimulus person and strength of category. That is, the difference between the two strong gender categories was significantly greater than the difference between the two weak categories, $F(1, 55) = 11.02, p < .001$. The finding that there was no interaction between category and behavioral diagnosticity supports our hypothesis that the category effects would still be present when paired with moderately diagnostic behavior. Men were rated as more aggressive than women regardless of whether they had engaged in neutral or in moderately diagnostic behavior, $F(1, 102) = 1.14, ns$.

A second ANOVA was performed on subjects who rated neutral and highly diagnostic behaviors. Again, strong effects were found for behavioral diagnosticity, $F(1, 62) = 189.98, p < .001$, and gender category, $F(1, 62) = 71.69, p < .001$. And again, the

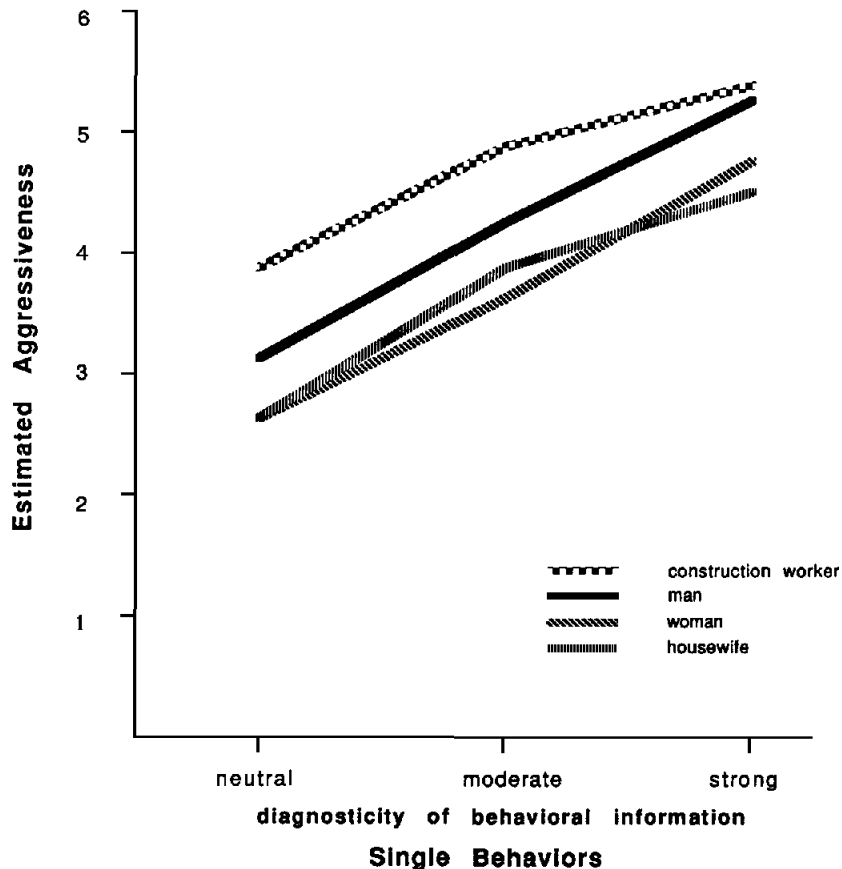


Figure 1. Target person's estimated aggressiveness: Experiment 1.

interaction between gender and strength of category was significant, $F(1, 62) = 5.20, p < .03$. Finally, despite the high diagnosticity of the behavior, the category effects remained equally strong across neutral and highly diagnostic behaviors, as reflected in the finding that there was no interaction between behavioral diagnosticity and category, $F(1, 62) < 1$.

Note that in both of these analyses, support for our prediction came from the absence of an interaction between category and diagnosticity of the behaviors. A more direct test of the hypothesis that category information influences judgment even in the presence of diagnostic case information involves examining only the two conditions in which diagnostic information was presented. Therefore, we performed a three-way ANOVA, treating the two levels of diagnosticity (moderate vs. high) and sex of stimulus person as between-subjects variables. Indeed, the main effect of category was significant. Men were rated as more aggressive than women, $F(1, 102) = 228.41, p < .001$. Moreover, this main effect was not moderated by an interaction between the category and diagnosticity. That is, the category effect persisted even in the presence of highly diagnostic case information, $F(1, 102) = 3.29, ns$. Contrary to prediction, construction workers and housewives were not seen as more different in their intensity of aggressive responses than were men and women, $F(\text{Sex of Stimulus Person} \times \text{Strength of Category}) = 1.14, ns$. Finally, as expected from our pretests, stimulus persons who were described by highly diagnostic behaviors were

rated higher on aggression than persons who had engaged in moderately diagnostic behavior, $F(1, 102) = 3.43, p < .05$.

Discussion

The results of the first experiment provide strong support for the hypothesis that gender categories affect the prediction of the gender-stereotypic trait of aggressiveness even when target individuals are described by behaviors diagnostic of aggressiveness. We had originally expected that categorical information would have a strong effect with irrelevant behavioral information, a somewhat weaker effect with moderately diagnostic behavior information, and a negligible or no effect with highly diagnostic behavioral information. To our surprise, category information had equally strong effects at all levels of behavioral information. These findings differ from research that showed complete neglect of category information (Locksley et al., 1980). The data also partially support the view that highly predictive gender categories affect trait ratings more strongly than weakly predictive gender categories. Construction workers (strong category) were seen as more aggressive than men (weak category) when the target person had engaged in moderately aggressive behavior. The category of housewives, on the other hand, which had been selected to represent the particularly nonaggressive female subcategory, was not rated differently from women in general.

The discrepancies between Locksley's (Locksley et al., 1980;

1982) and our results may be attributable to at least two differences in the prediction tasks. First, the criterion trait of aggressiveness seems to be more stereotypic of gender than is assertiveness. Our failure to find a gender effect of assertiveness in our pretesting suggests that expectations of differences between men and women in assertiveness, if they exist at all, are weak and may thus be neutralized by any other information that is diagnostic of the criterion. Second, the person descriptions in most of Locksley's research (Locksley et al., 1980; 1982) included several behavioral acts, whereas in the present study only a single behavioral act was conveyed.

In an attempt to reconcile the apparent discrepancies between Locksley et al.'s (1980; 1982) and our own results, we conducted a second experiment. In Locksley et al.'s studies the description of the behaviors of the target persons clearly implied temporal stability. Because one difference between the two sets of findings concerns the temporal stability of the target person's behavior, our second experiment pitted the strength of the category against temporally stable behaviors that vary in diagnosticity. Attribution theory (Kelley, 1967) predicts that temporally consistent behaviors should have greater weight in trait inferences because they permit the perceiver to infer correspondent traits in the target person, whereas single behaviors can be attributed to unstable situational conditions.

Experiment 2

The second experiment resembled the first one in its essential features with the exception that the stimulus persons were described as *consistently* engaging in a particular behavior. Using these conditions, we expected the moderate and high diagnostic behaviors to negate the effects of the gender categories.

Method

Subjects were 46 male and 64 female students enrolled in introductory psychology classes at the University of Oregon. Participation in research fulfilled a course requirement. Each subject read descriptions of two male and two female stimulus persons. For each subject, two stimulus persons were described as consistently engaging in a behavior unrelated to aggressiveness, whereas the other two were described as consistently behaving in a way either moderately diagnostic or highly diagnostic of aggressiveness. Subjects were provided with stimulus persons belonging either to highly stereotypic groups (construction workers and housewives) or to weakly stereotypic groups (men and women). Dependent variables again were the expected frequency of future aggressive behavior and the predicted intensity of future aggressive behavior (to be combined into a single, composite measure).

Results and Discussion

Again, the expected frequency and the predicted intensity of aggressive behavior were highly correlated ($r = .93, p < .001$). As in Experiment 1, a single index of estimated aggressiveness was created by transforming the frequency measure into the metric of the intensity measure and then combining the two equally weighted measures.

The results of the single index of aggressiveness are presented in Figure 2. First, data were analyzed for those subjects given irrelevant and moderately diagnostic behavioral information. The effects of behavioral diagnosticity and category were significant, $F(1, 50) =$

$270.33, p < .001$, and $F(1, 50) = 9.06, p < .005$, respectively. Whereas we had predicted that a category effect would obtain under conditions of irrelevant behavioral information but not under conditions of moderately diagnostic behavior, the interaction between category and behavioral diagnosticity was merely marginal, $F(1, 50) = 2.68, p < .10$. No other effects were observed.

For subjects who were given irrelevant and highly diagnostic behavioral information, there was a strong effect of behavioral diagnosticity, $F(1, 52) = 188.22, p < .001$, but no effect of category, $F(1, 52) = 2.08, ns$. The significant interaction between category and diagnosticity, $F(1, 52) = 7.43, p < .007$, supported the prediction that the difference between the gender categories in the neutral condition would disappear when gender was paired with consistent, highly diagnostic aggressive behavior. When both moderately and highly diagnostic behaviors were included in one analysis, the category effect disappeared, $F(1, 52) < 1$, suggesting that temporally consistent diagnostic behavior negated category effects. Finally, persons described by highly diagnostic behavior were rated as more aggressive than were persons who were described by moderately diagnostic behavior, $F(1, 102) = 7.63, p < .007$.

The results of the second experiment supported the prediction that diagnostic and temporally consistent behaviors negate the effect of gender categories. Gender had an effect only when the consistent behaviors were irrelevant in regard to the criterion trait. In Experiment 2, this effect was smaller (.44 mean difference) than in Experiment 1 (.87). A possible explanation for this result may be by the "dilution effect" (Nisbett, Zukier, & Lemley, 1981). Strong individuating information, though nondiagnostic with respect to the criterion trait, may reduce (dilute) perceived differences between predictive categories. This explanation is only a conjecture, however, because a condition with no behavioral information was not included here.

In this experiment, then, temporally stable behavioral information about a target person tended to negate category effects. Male and female target persons who consistently beat their children or yelled at their spouses were rated as equally aggressive. This finding is consistent with the act-frequency approach to personality dispositions (Buss & Craik, 1983). Knowledge about frequently recurring behaviors constitutes a sufficient basis for trait inference. When making judgments about the aggressiveness of a target individual, subjects in this experiment may have been interpreting the behavioral information as trait information, presented in behavioral terms. For example, saying that a person when annoyed consistently hits people may be seen as just another way of saying that the person is violent. Inferences about the degree of a person's aggressiveness then depend on both the extremity of the inferred trait and the perceived correlation of this trait with aggressiveness. For the present example, the trait of violence is both extreme and similar in meaning to aggressiveness. If traits are summary statements about past behavior, they should have the same effect on prediction as consistent behavior. We conducted a third experiment to test the idea that traits (like consistent behaviors) negate the effects of predictive social categories.

Experiment 3

In this study, we compared the predictive category information (gender) to that of diagnostic or counterdiagnostic trait in-

formation. Subjects were provided with both gender and trait information about a stimulus person. Subjects were to predict the likelihood with which the target person would endorse specific behavioral statements. We expected that the category information would have no effect on predictions of individual behaviors and that predictions of behavior would be made solely on the basis of the target person's trait.

Method

Overview. Subjects were presented with male or female stimulus persons who were described by one trait adjective that was either predictive or counterpredictive for a stereotypically masculine or a stereotypically feminine behavior. Subjects then rated the likelihood with which the target person would engage in the behavior.

Subjects and design. The subjects in this experiment were 256 undergraduates at the University of Oregon who received extra credit in introductory psychology courses for their participation. Each subject made a judgment about one male or one female stimulus person who was characterized by traits that were either predictive or counterpredictive of a behavior statement. Half of the statements described stereotypically male behaviors, and the other half described stereotypically female behaviors. Half of the statements were socially desirable, and half were undesirable. There were five between-subjects variables: sex of subject, sex of target, diagnosticity of trait, and favorability and stereotypicality of criterion behavior.

The experimental materials consisted of a set of eight behavior state-

ments and 32 traits. The behavior statements were selected from a pool of statements that had been used by Park and Rothbart (1982) in their research on group stereotypes. Every statement was associated with 1 of 4 traits. As determined by pretests, 2 of the traits were moderately diagnostic of the statement, and 2 traits were fairly counterpredictive. For example, *serious* is fairly diagnostic of endorsing the statement "I enjoy being in leadership positions," whereas the trait *awkward* is counterpredictive of the endorsement of this statement. Furthermore, in the description of the stimulus person, the impact of the trait adjective was mitigated by the qualifier *somewhat*. A complete listing of the pairings of statements and adjectives is given in Appendix A. An example vignette is given in Appendix B.

Results and Discussion

The judgments generated by each of the 256 subjects were analyzed by a five-way ANOVA, with sex of subject, sex of target person, diagnosticity of personality trait, stereotypicality of criterion behavior, and favorability of criterion behavior as the five variables. The dependent variable was the estimate of the likelihood of the endorsement of the criterion behavior by the stimulus person.

No effects were found for sex of subject, sex of stimulus person, and type of statement. More important was the finding that subjects perceived male stimulus persons ($M = 47.81$) as no more likely than female stimulus persons to endorse masculine

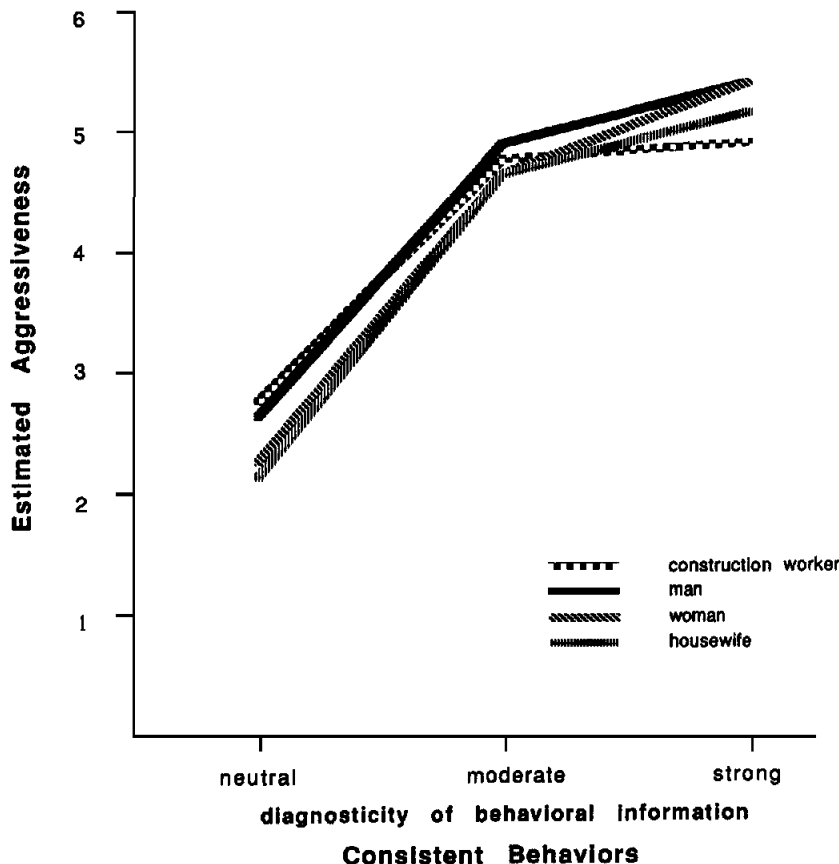


Figure 2. Target person's estimated aggressiveness: Experiment 2.

items ($M = 45.63$) and female stimulus persons ($M = 48.44$) no more likely than male stimulus persons to endorse feminine items ($M = 48.13$). The trait information available for each person overrode the effect of the gender of the behavior and of the stimulus person. As in Experiment 2, only the case information (trait) affected predictions of behavior. When the target person was described with a trait that was predictive of the criterion behavior, the average estimates of the endorsement likelihoods were greater (60.31) than when the trait was counterpredictive (34.88), $F(1, 254) = 80.285, p < .001$. Finally, the main effect of favorability, $F(1, 254) = 10.246, p < .002$, suggests a strong tendency to view desirable behaviors as more probable. In all, it appears that the presence of personality trait information is sufficient to negate effects of gender categories in predictions of a person's preferences.

General Discussion

In our first experiment, we simultaneously varied the strength of the category and the diagnostic strength of single behaviors in influencing judgment about a stimulus person's aggressiveness. The effects of category were present at every level of behavioral diagnosticity. That is, categorical and individuating information had an additive effect on judgment, and individuating information did not negate the effects of category, even when the individuating behaviors were highly diagnostic.

We argued that the nature of the individuating information may have been responsible for the apparent discrepancies between Locksley's (Locksley et al., 1980; 1982) findings and our own results. Locksley et al. used behavioral information connoting passivity or assertiveness, but in our pretests we found no significant perceived differences between men and women on these traits. Because we could not attempt to replicate Locksley's research with a criterion that did not differentiate between the relevant categories, we chose a different dimension, aggressiveness, as the criterion attribute. In our view, the fact that Locksley obtained results while using case information that was only weakly correlated with gender may be a factor of some importance.

Having found a consistent effect of categorical information when it was presented along with single behaviors, we hypothesized that categorical information may play a weaker role when combined with temporally stable behaviors, and we tested that idea in our second experiment. In fact, Experiment 2 showed that stable behaviors, whether highly or only moderately diagnostic of aggressiveness, did override the effects of category.

In the third experiment, we substituted trait adjectives for stable behaviors, paired them with male and female stimulus persons, and asked subjects to predict the likelihood of endorsement of masculine and feminine behaviors. The traits were chosen to be either positively or negatively correlated with the target behaviors. Under these conditions, there was a significant effect due to the traits and no effect due to the gender of the stimulus person or the gender relatedness of the behavior.

These experiments qualify some of the stronger statements that have been made about the underutilization of categorical (or base-rate) information. Indeed, trait terms or traitlike behaviors (i.e., temporal consistency in behavior) did override the effects of category, but single behaviors did not. Clearly, then,

not all case information negates base-rate information, and it remains for future research to explore in more depth when categorical information is dominant over or subordinate to case information. In trying to predict the relative power of stereotypic categories and trait information, it may be necessary to examine more precisely the relation between categories and category attributions. Not all categories are the same, and it does not make sense to think that broad categories, such as gender, each of which includes half the world's population, will be as predictive of people's attributes as more narrow categories. Compare the category *women* to the category *schizophrenic* (Rosenhan, 1973). The latter refers to a small set of individuals with reasonably well-defined attributes. We suspect that the attributes *incoherent* or *autistic* are more prevalent among members of the schizophrenic category than are *timid* or *passive* among members of the women category.

Given such differences between categories, how can one think about the role of categorical and case information in predicting a target behavior or trait? It seems to us that the critical factor is the relative strength of the category versus the relative strength of the case information in predicting the target attribute. In Locksley's (Locksley et al., 1980; 1982) research, it could be argued that gender was only a weak predictor of the target attribute (e.g., assertiveness), whereas the individuating behaviors were far stronger predictors of the criterion. It is not necessarily true, however, that case information is invariably weighted more strongly than category information, as demonstrated in our first experiment. Even with gender categories, which we suspect are inherently weak predictors of behavior, given the behavioral heterogeneity of the category members, our first experiment showed that the presence of a single, even highly diagnostic behavior did not override the effects of the category.

Categories that are associated with more extreme base rates should be even more likely to influence predictions even in the presence of individuating information. Experiment 1 yielded support for this idea: Target persons who were described as housewives and construction workers were rated as more extreme with respect to aggressiveness than were men or women in general. A number of categories in the social world give rise to stereotypic expectations that are far more extreme than those elicited by gender categories (or their subordinates). For these categories, even consistent diagnostic behaviors may not be sufficient to override the category effect. For example, if people are to make a judgment about a schizophrenic's interpersonal sensitivity on the basis of consistently friendly behavior, they may still incorporate category membership into their prediction.

Categories that are associated with extreme base rates are typically fairly narrow. They predict few criteria, but they predict them very well. In psychometric terms, the fidelity of the prediction (or the measurement) increases as bandwidth decreases. Recently, Hampson, John, and Goldberg (1986) discussed this issue in the context of trait attributes, arguing that traits vary in their breadth, or range of applicability, with broader traits being less predictive of behavior in a specific domain. We believe that the bandwidth-fidelity trade-off applies to the issue of categorical and individuating behavior as well. Gender categories, as we have argued, are rather lacking in pre-

dictive power because they are very broad. Popular stereotypes link a plethora of traits to men and women (high bandwidth), but the confidence with which these links are translated into predictions is very modest (low fidelity). As predictors vary in their breadth, so do the criteria. So far, our research has been concerned with rather specific behavioral and trait criteria, and we have learned that people use the individuating information mostly when making their judgments. The reason for this may be that the individuating information is as narrow as the criteria. That is, people rely on a match between the breadth of the predictor and the breadth of the criterion. The categorical information may be neglected only when it is comparatively broad. If, as we suspect, the match between the breadth of the predictor and the breadth of the criterion is an important factor in prediction tasks, then broad gender stereotypes may well be used as predictors if the criteria are of a similarly wide range.

The question of how categorical and case information are combined is complex, and the present research may not be the final word on this process. In our view, however, it seems premature to conclude that stereotypic categories, in the presence of case information, play only a negligible role in judgment. Future research examining the role of category breadth and strength of case information could illuminate this problem.

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.
- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258-290.
- Bar-Hillel, M. (1980). The base rate fallacy in probability judgments. *Acta Psychologica*, 44, 211-233.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42, 155-162.
- Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review*, 90, 105-126.
- Dipboye, R. L., Fromkin, H. L., & Wiback, K. (1975). Relative importance of applicant sex, attractiveness, and scholastic standing in evaluation of job applicant resumes. *Journal of Applied Psychology*, 60, 39-43.
- Eagly, A. E., & Wood, W. (1982). Inferred sex differences in status as a determinant of gender stereotypes about social influence. *Journal of Personality and Social Psychology*, 43, 915-928.
- Feldman, N. S., Higgins, E. T., Karlovac, M., & Ruble, D. N. (1976). Use of consensus information in causal attributions as a function of temporal presentation and availability of direct information. *Journal of Personality and Social Psychology*, 34, 694-698.
- Frodi, A., McCauley, J., & Thome, P. R. (1977). Are women always less aggressive than men? A review of the experimental literature. *Psychological Bulletin*, 84, 634-660.
- Hampson, S. E., John, O. P., & Goldberg, L. R. (1986). Category breadth and hierarchical structure in personality: Studies of asymmetries in judgments of trait implications. *Journal of Personality and Social Psychology*, 51, 37-54.
- Heneman, H. G. (1977). Impact of test information and applicant sex on applicant evaluations in a selection simulation. *Journal of Applied Psychology*, 62, 524-526.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192-238). Lincoln: University of Nebraska Press.
- 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. *Journal of Experimental Social Psychology*, 18, 23-42.
- Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.
- McArthur, L. A. (1972). The how and what of why: Some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology*, 22, 171-193.
- Nisbett, R. E., & Borgida, E. (1975). Attribution and the psychology of prediction. *Journal of Personality and Social Psychology*, 32, 932-943.
- Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: Non-diagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*, 13, 248-277.
- Park, B., & Rothbart, M. (1982). The perception of out-group homogeneity and levels of categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology*, 42, 1051-1068.
- Rosenhan, D. L. (1973). On being sane in insane places. *Science*, 179, 250-258.
- 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.
- Spence, J. T., & Helmreich, R. L. (1978). *Masculinity and femininity: Their psychological dimensions, correlates, and antecedents*. Austin: University of Texas Press.
- Tversky, A., & Kahneman, D. (1979). Causal schemata in judgment under uncertainty. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Wells, G. L., & Harvey, J. H. (1977). Do people use consensus information in making causal attributions? *Journal of Personality and Social Psychology*, 35, 279-293.
- Zuckerman, M. (1978). Use of consensus information in prediction of behavior. *Journal of Experimental Social Psychology*, 14, 163-171.

Appendix A

Favorable and Unfavorable Gender-Stereotypic Behaviors With Their Predictive and Counterpredictive Traits

Male, favorable		Female, favorable	
“I enjoy being in leadership positions.”		“I am patient when working with young children.”	
predictive:	serious, reliable	predictive:	contented, secure
counterpredictive:	detached, awkward	counterpredictive:	careless, moody
“I am not afraid to challenge my superiors if I disagree with their statements.”		“I often tell those close to me how much I care for them.”	
predictive:	dominant, daring	predictive:	helpful, generous
counterpredictive:	passive, submissive	counterpredictive:	irritable, withdrawn
Male, unfavorable		Female, unfavorable	
“If I get the chance to crowd in front of other people waiting in line, I’ll do it.”		“I avoid competition; it scares me.”	
predictive:	stubborn, assertive	predictive:	cautious, timid
counterpredictive:	passive, cautious	counterpredictive:	proud, assertive
“When playing card or board games, I get extremely upset if I don’t win.”		“If I am treated badly or insulted by a salesperson, I am more likely to simply feel bad than to defend myself.”	
predictive:	egocentric, aggressive	predictive:	withdrawn, dependent
counterpredictive:	easygoing, passive	counterpredictive:	stubborn, crafty

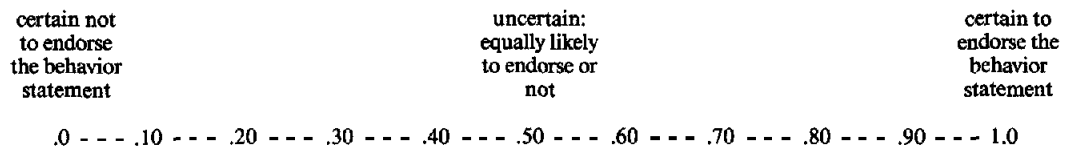
Appendix B

Example of a Description of a Target Person

Michael is a 35-year-old man who was born in the United States. His acquaintances describe him as somewhat serious. How likely do you think Michael is to endorse the following statement?

“I enjoy being in leadership positions.”

Please circle the point on the scale which you feel most closely corresponds to the likelihood of the person to endorse the statement.



Received August 21, 1987
Revision received December 11, 1987
Accepted December 14, 1987 ■