Gender Stereotypes: Perception or Rationalization?

Curt Hoffman and Nancy Hurst
University of Alberta, Edmonton, Alberta, Canada

It is proposed that gender stereotypes arise to rationalize the distribution of the sexes into social roles. Ss read descriptions of members of two fictional categories, one having 80% "city workers" and 20% "child raisers," the other with the percentages reversed. They later made personality ratings of each category and of the category subgroups occupying each role. Ss formed role-based category stereotypes that affected their ratings even when targets' roles were specified. Stronger stereotypes arose when the categories were biologically defined or when Ss attempted to explain the category-role correlation. The basic effect was replicated using roles that are not differentially linked to familiar human groups. The findings are interpreted as showing that stereotypes can arise solely in response to a sexual division of labor and serve to rationalize this division by attributing to the sexes intrinsic personality differences.

Of the several interesting questions that can be asked about stereotyping, two are fundamental to an understanding of this phenomenon. The first is, How are stereotypes represented cognitively, and how do they affect perceptions of and actions toward members of stereotyped categories? The second is, How and why do stereotypes acquire their particular contents? The first question has been the topic of extensive theorizing and research in recent years, with the result that we now know a good deal about the structural properties of stereotypes and their cognitive and behavioral consequences (see Hamilton & Trolier, 1986, for a review). The second has received much less attention, and we still know relatively little about the origins of stereotypes or the factors responsible for their contents. We suspect that this is largely because the content origin issue has been a much more difficult problem to conceptualize satisfactorily within the current information-processing paradigm in social psychology. Stereotypes belong to a class of "schemas" that, in general, are thought to have as their goal the representation of external reality (even though they may achieve only a selective or simplified version of this reality). In the case of stereotypes, however, this is tantamount to saying that they contain at least a kernel of truth—a position that social psychology has, understandably, been somewhat reluctant to embrace.

Still, a number of social psychologists have regarded the kernel-of-truth hypothesis as a partial solution to the content origin problem, at least for some stereotypes. Gender stereotypes in particular have often been thought to reflect, in exaggerated form, real differences between the groups in question. Recently, for example, Eagly and colleagues (e.g., Eagly, 1987; Eagly & Steffen, 1984) proposed an account of the origin of gender stereotypes that grants to these stereotypes what amounts to a sizable kernel of truth. Eagly presents evidence that the central core of the stereotypes is the idea that men are more "agentive" (self-assertive and motivated to master) than women and that women are more "communal" (selfless and concerned with others) than men. This belief, it is proposed, stems largely (though not entirely) from the observation of women and men in different social roles. Women are more likely than men to be homemakers and child raisers; men are more likely than women to be breadwinners (i.e., to work outside the home). It is assumed, moreover, that these roles confer on their occupants different traits of behavior and personality—wherein lies most of the kernel of truth in the stereotypes.

Are Gender Stereotypes Based on Observed Sex Differences in Personality?

It seems to us that Eagly and colleagues are almost certainly correct in suggesting that the difference in men's and women's social roles is largely responsible for the gender stereotypes. But in precisely what way is it true that the sexual division of labor gives rise to the stereotypes? Is it true only in the indirect sense that the division of labor leads to a difference in men's and women's personalities, which in turn is the direct (proximate) source of the stereotypes? We argue, to the contrary, that there are certain difficulties with the idea that the gender stereotypes as we know them do in fact arise directly and exclusively from observed sex differences in personality—or even that they could plausibly arise in such a fashion.
In developing this argument, we need to make a sharp distinction between two kinds of sex differences in personality that have been proposed to underlie gender stereotypes. First, there are those differences associated with gender by virtue of women's and men's unequal participation in certain social roles, that is, what we might call "role-based" differences that are (fortuitously) correlated with gender because of the sexual division of labor. It is this kind of difference that Eagly and colleagues regard as primarily responsible for the stereotypes. Second, there are those differences associated directly with gender itself, that is, what we might call "gender-based" differences that exist more or less independently of specific social roles and that distinguish the two sexes generally—not just those women and men occupying the traditional social role for their sex. It is this kind of difference that psychologists ordinarily have in mind when they speak of "sex differences," and which is most commonly at issue in discussions of the veridicality of gender stereotypes. Note that our use of the term gender-based differences carries no implication that such differences stem from biological factors. We use it simply to refer to differences that are unconfounded by extraneous variables such as occupational or familial role. For our purposes, it is largely irrelevant whether one attributes such differences primarily to biology, socialization, or (as does Eagly) highly generalized "gender roles."

Let us first consider the argument that the gender stereotypes stem partly from observation of gender-based differences in personality. In our view, there are two difficulties with this argument. First, research shows that many of the most stereotypical such differences tend to be rather small—too small, we suspect, to allow their discovery in the absence of factors leading people to expect such differences. As Nisbett and Ross (1980) argued, unless a group difference is quite large, "people's covariation detection capacities are far too crude to allow any...purely 'data-based' discovery" (pp. 238–239). According to Eagly's (1987) meta-analytic review of research on the sex difference in influenceability, for example, the mean effect size (d) for this difference is around 0.3. Although there is disagreement on what constitutes a large or small difference, most would agree that a d of 0.3 is small (by comparison, the sex difference in height corresponds to a d of 2.6). Cohen (1977), moreover, argued that a d of 0.5 or so is normally required before a difference is even noticeable at all.

Second, there would seem to be little correlation between the size of a sex difference and the degree to which it is emphasized in the stereotypes. The stereotypes feature prominently a number of dimensions on which men and women differ only slightly (e.g., influenceability) or on which, as far as we now know, they do not differ at all (kindness, patience, ability to make decisions, ability to think clearly, and many others). On the other hand, there are a number of objectively larger differences (e.g., in nonverbal and paralinguistic behavior) that are essentially absent from the stereotypes. For example, Hall (1984) reported an effect size of about 0.7 (reexpressed in terms of d) for frequency of speech errors and for bodily restlessness. It is doubtful, however, that most people are even aware of these differences. At the very least it is clear that women's greater articulateness and men's greater restlessness are, at best, extremely peripheral features of the gender stereotypes. A conservative conclusion from the research on gender-based differences, then, is that it would be difficult to account for the strength, consistency, and patternning of the stereotypes on the basis of such differences as psychologists have been able to measure objectively.

The other candidate for the kernel of truth in the stereotypes is the set of personality differences assumed to be associated with gender by virtue of the sexual division of labor (role-based differences). We are willing to grant, for the sake of discussion, that the personalities of homemakers are, in some real sense, noticeably less agentic and more communal than are the personalities of breadwinners. (In fact, however, we doubt that this is true, but full consideration of the issues involved would take us too far afield.) The major problem with the notion that gender stereotypes are primarily the result of observed differences in the personalities of homemakers and breadwinners is simply that it fails to explain convincingly how or why the stereotypes come to characterize males and females in general. In other words, why do we have gender stereotypes in addition to homemaker and breadwinner stereotypes?

The idea that role-based sex differences, in and of themselves, suffice to explain gender stereotyping would work fine if such stereotyping were limited to cases where targets are known to occupy the traditional role for their sex or cases where the role is unknown but could be assumed (by default) to be the traditional one. But gender stereotyping is very clearly not limited to such cases. A convincing demonstration of this was provided by Taylor, Fiske, Etcoff, and Ruderman (1978, Experiment 3), who obtained strong gender-stereotyping effects while holding rigidly constant not only the target's role, but his or her actual behavior as well. In addition, it has been shown that gender stereotypes are readily applied not only to adult men and women, but also to newborn infants and young children, and that this can occur even if the actual identity of the child is controlled for (e.g., Condry & Condry, 1976; Gurwitz & Dodge, 1975; Rubin, Provenzano, & Luria, 1974). But such examples are probably not even necessary to make the basic point. Anyone who has ever done a study in which college students rate the personalities of other students, for example, knows that gender-stereotyping effects are hard not to obtain, even though this is another case in which subjects are well aware that the targets of their ratings are not occupants of the traditional adult roles.

Eagly and Steffen (1984) reported evidence that is seemingly at variance with our contention that the influence of gender stereotypes extends beyond persons who occupy (or can be assumed to occupy) the traditional role for their sex. Subjects in one of their studies rated "an average man" or "an average woman" whose role was that of homemaker, whose role was that of employee, or whose role was not specified. Targets whose role was unspecified were rated in gender-stereotypic terms, but there was no stereotypic sex-of-target effect when role was specified. Eagly and Steffen readily admitted, however, that in real life even persons in a nontraditional role probably do not escape the gender stereotypes altogether. And although Eagly and Steffen have identified a context in which role occupancy apparently overwhelms gender as a judgmental cue, other research (including the studies cited above and the new studies reported later in this article) clearly indicates that in many contexts stereotypes do have a categorywide influence that is not completely overridden by considerations of role.

It might still be argued that the gender stereotypes result from
a straightforward process of generalization, whereby, for example, traits such as “kind” and “gentle” become attached to the category “female” simply because most (adult) females are homemakers and homemakers tend to have these traits. But is this really a sufficient explanation for the breadth of the stereotypes? Most college students are young, and young people are stereotypically healthier than the general population, yet the stereotype of college students does not include (or does not feature prominently) the trait “healthy.” Generalization from a subcategory (even when it is the majority) to the category as a whole does not seem to occur invariably, automatically, and on just any attribute dimension.

Apart from its logical insufficiency, the mere-generalization hypothesis is also not compelling in a psychological sense. There is something quite implausible, for example, about the idea that people would make direct extrapolations from observation-based beliefs about homemakers and breadwinners in evaluating the personalities of newborns (and it is, of course, even more unlikely that people could have separate “newborn male” and “newborn female” stereotypes based directly on observed sex differences in neonate behavior). We do not deny that generalization is involved in stereotype formation; our point is simply that such generalization, when it occurs, may sometimes itself be a phenomenon in need of further explanation.

The Rationalization Hypothesis

To summarize, we have argued that there are certain difficulties with the idea that gender stereotypes arise directly or exclusively from the observation of sex differences in personality. At best such an account is incomplete, but it would also appear to be somewhat at odds with certain basic facts about sex differences and sex stereotyping. An alternative account, which we believe provides a better fit with empirical knowledge in the area, proposes that the stereotypes are largely an attempt to rationalize, justify, or explain the sexual division of labor. The differential participation by women and men in the roles of homemaker and breadwinner (as well as a few other key roles, such as soldier) is a social fact of such pervasive significance that it would be very odd indeed if cultures and individuals did not feel some need to explain or rationalize that fact. And the most powerful rationale imaginable is probably the simple assumption that there are inherent differences between males and females that make each sex better suited for its role—that each sex has, in other words, a greater capacity for the qualities thought to be necessary to the performance of its traditional function. “Women care for the children, and understandably so—they are by nature kinder, gentler, and more sensitive than men. Men run the businesses and fight the wars, and that is obviously because they are naturally more logical, independent, and competitive than women.” In essence, then, we propose that gender stereotypes be regarded not primarily as summary abstractions of males’ and females’ personalities based directly on observed differences in those personalities, but at least partly as explanatory fictions that rationalize and make sense of the sexual division of labor.

Neither the idea that stereotypes function as rationalizers nor the specific application of that idea to gender stereotypes is original with us (although we believe we are perhaps the first to test experimentally the direct implications of that idea). Allport, for example, embraced this point of view in his book The Nature of Prejudice (1958), in which he concluded that “the rationalizing and justifying function of a stereotype exceeds its function as a reflector of group attributes” (p. 192). More recently, Tajfel (1981) has also written about the justificational function of stereotypes, and O’Leary (1974) and Huici (1984) have looked specifically at gender stereotypes from this perspective. It was Williams and Best (1982), however, who perhaps most explicitly expressed the idea that gender stereotypes serve to rationalize the traditional division of labor between the sexes:

Having found that it is useful, if not essential, for women and men to assume different social roles, a society finds it adaptive to believe that each sex has . . . the characteristics necessary for the successful performance of the assigned functions. If females are to have principal responsibility for the care of the young, it is reassuring to believe that they are—or can become—affectionate, gentle, patient, sympathetic, and so on. If males are to serve as hunters and warriors, it is comforting to believe that they are—or can become—adventurous, aggressive, courageous, energetic, independent, self-confident, and the like. . . . It may be in this context, the “justification of necessity” with regard to different social roles, that many of the sex-trait stereotypes originated. (p. 237)

The present account of gender stereotypes also differs from some others in proposing that people assume an inherent difference in the natural personality tendencies of males and females. “Logically” speaking, of course, assumed gender differences provide a more compelling rationale for the sexual division of labor to the extent that they are regarded as intrinsic to the psychological makeup of the two sexes. We are not suggesting that people are incapable of recognizing exceptions to these assumed natural tendencies or of forming subcategories of women and men who differ from the norm. We are simply suggesting that people in our culture believe that there are inherent male–female differences that predispose the sexes to develop somewhat distinct personalities—even though they may also recognize that upbringing and other experiential factors can strengthen and exaggerate these tendencies as well as suppress or even reverse them in exceptional cases. Our account thus avoids two problems that have proved troublesome for the direct observation theory. First, it easily handles (and in fact predicts) findings that the stereotypes are widely applied to the two sexes in general, including infants, and that their application is not limited to persons who occupy (or can be assumed to occupy) the traditional male and female roles. Second, it explains how the attributes deemed desirable for occupants of the traditional male and female roles become associated with males and females generally (people come to assume that the sexes inherently differ on these attributes), as well as why this takes place (it rationalizes the sexual division of labor).

Overview of the Research

Experiment 1 was designed to test certain predictions from the account of gender stereotypes that we have just outlined. Our most basic proposition is that the stereotypes could, in principle, arise in direct response to the sexual division of labor per se, even in the total absence of any other sex differences. The only additional assumption required is that there exists, prior to formation of the stereotypes, a set of beliefs about
which traits are or are not conducive to successful performance of the kinds of tasks traditionally performed by men and women. The logic of the study was therefore to create an experimental analogue of this hypothesized state of affairs, with the goal of determining whether stereotypes would in fact arise under such conditions. Subjects first read descriptions of individual members of two categories. Although on average the categories had identical personality traits, they did differ with respect to the role most commonly occupied by their members. One consisted primarily of "city workers," the other of "child raisers." In each case, however, there were some individuals who occupied the role more typical of the other category. Subjects later rated each category on a number of personality traits, and also rated the subgroups within each category occupying each of the two roles.

Because the goal of the study was to demonstrate how traits can become stereotypically linked to categories, we obviously could not use the two sexes (or any other familiar human groups) as the categories in question; rather, we required categories about which subjects could have no preexisting impressions. Two fictional groups of intelligent life on another planet were created to serve this purpose. We predicted that

1. Because of the category–role correlation, subjects would form role-based personality stereotypes of the two categories, even though the targets' personalities did not in fact differ as a function of either category or role, and

2. The stereotypes thus formed would affect perceptions of category members even when their social role was also specified.

The first prediction follows from the proposition that stereotypes can arise solely from the distribution of categories into social roles and that they do not, in principle, depend on actual personality differences between the categories. The second, and more interesting, prediction follows from the idea that role-based stereotypes serve to rationalize the category–role correlation and that they do so by attributing to the categories intrinsic psychological differences. To the extent that this is true, the category stereotype should have some effect on all members of the category. More concretely, what this means is that there should be an effect of category membership on perceptions even when social role is specified. This is not to say, obviously, that role is irrelevant, or even that role would have a smaller effect than category; the prediction was simply that, in addition to any main effect of role, there would also be a main effect of category.

The design of the study also included two between-subjects manipulations. After reading the target descriptions but before making any personality ratings, all subjects made estimates of the category–role distribution. Half the subjects were also asked at this time to offer a possible explanation for the observed distribution. To the extent that some form of rationalization or explanation contributes to stereotype formation, this manipulation was expected to affect the strength of the stereotypes. Our third prediction, then, was that

3. The explicit act of explaining (rationalizing) the category–role correlation would result in the formation of stronger stereotypes.

Orthogonally to the manipulation of explanatory set, we also manipulated the nature of the defining distinction between the two categories. For half the subjects, the distinction was biological, whereas for the other half it was nonbiological. Although the hypothesized rationalization mechanism can probably operate whether or not there is an obvious biological difference between the groups in question, one would expect stereotypic perceptions of two contrasting categories to be intensified if (as in the case of gender categories) such a difference does in fact exist. Therefore, our final prediction was that

4. Stronger stereotypes would be formed when the defining distinction between the categories was biological in nature.

**Experiment 1**

**Method**

**Orienting information and manipulation of category type.** A written introduction to the experiment described it as a study of "how people come to understand the structure of social groups," employing "a type of 'make-believe' procedure that is sometimes used in research of this kind." Subjects learned that they would be given information on individual members of two fictional groups, the "Orinthians" and "Ackmians," who inhabit a distant planet. For half of the subjects (the biological condition), the Orinthians and Ackmians were presented as separate species. For the other half (the nonbiological condition), they were presented as distinct subcultures. We shall refer to this variable as category type. The introduction went on to say,

The members of this society live in groups... in the countryside near large cities. Typically,... both species [subcultures] are represented in each of these living groups. Some of the adult members of each group function primarily as child raisers. These adults spend most of their time at the group home, where they take responsibility for the care and teaching of the group's young. Other adult members of each living group work in the nearby cities, where all of the business, industry, technology, and higher education is concentrated. These city workers travel to the city each day and return to the group home in the evening.

Subjects in each category-type condition were also told that there are no male or female sexes on this planet, and that any individual can mate with any other individual, causing both to reproduce. (In the biological condition, it was also stated that mating can only occur between members of the same species.)

**Target descriptions.** Subjects then read individual descriptions of 15 Orinthians and 15 Ackmians, each consisting of a drawing of the target's head and shoulders and a sentence with the target's name, category (Orinhian or Ackmian), social role (city worker or child raiser), and three personality traits (e.g., "Damorin, an Orinhian who works in the city, is resourceful, individualistic, and soft-spoken"; "Dolack, an Ackmian who raises children, is outspoken, compassionate, and reliable"). In the biological condition, the targets were drawn in such a way that there was a distinct family resemblance among the members of each species and a number of obvious differences between the two species. In the nonbiological condition, the only visual difference between the two categories was that the Orinthians were depicted in dark clothing and the Ackmians in light clothing. Each description appeared on a separate page of a booklet, through which subjects were paced at a rate of 1.5 s per page. Instructions for this portion of the study were phrased as neutrally as possible; subjects were simply told to pay close attention to the material presented and that they would later be asked "various questions about the information in these descriptions and the impressions that you formed while reading them." Induction of an explicit memory set was deliberately avoided.

Subjects in each category-type condition received one of two versions of the booklet described above. The versions were identical in every respect save the following: In one, 12 Orinthians were described as city
workers and 3 were described as child raisers; conversely, 12 Ackmians were described as child raisers and 3 were described as city workers. In the other, each of the 30 targets’ roles was simply interchanged with the opposite role. The drawings, names, and traits associated with each of the 30 targets were exactly the same in the two versions, as was the order in which the targets appeared.

The personality traits contained in the descriptions were drawn from a pool of traits that had been pretested for role stereotypicality. We asked 9 men and 9 women to rate 117 traits (on a 9-point scale) for their applicability to breadwinners (persons employed full time outside the home) versus homemakers (persons whose major responsibility is the care of a home and family). Ratings toward the high end of the scale indicated greater breadwinner stereotypicality, and ratings toward the low end indicated greater homemaker stereotypicality. Traits whose mean ratings were 6 or higher and significantly greater \((p < .05)\) than the scale midpoint of 5 were considered agentic traits. Those whose ratings were 4 or lower and significantly less than the midpoint were considered communal traits. Traits whose ratings lay between 4 and 6 and did not differ significantly from the midpoint were considered neutral.

A total of 15 agentic, 15 communal, and 15 neutral traits were used in the descriptions, each of which was attributed once to an Orinthian and once to an Ackmian. Each description included one agentic, one communal, and one neutral trait (the order in which these were listed was counterbalanced within each category/role group). No two traits appearing together in one description were ever paired together in another description; thus, each of the 30 targets was described with a unique set of three traits. Within these constraints, the assignment of traits to targets was random (except that a few inconsistent trait pairings were avoided). The 30 targets appeared within the booklets in a single random order, subject to the constraints that (a) members of the same category did not appear more than three times in succession and (b) none of the targets in the atypical role for their category appeared as one of the first two or last two members of that category (so as not to weaken the perceived category–role correlation).

The net result of these arrangements was that there was no correlation whatsoever between the personalities of the targets and either their category membership or social role. In addition, the use of two counterbalanced booklet versions served as a control for any possible effects of the particular names, drawings, and trait groupings associated with the two categories.

After reading the descriptions, subjects received a questionnaire with the measures and manipulations described below.

**Check on the perceived role distribution.** Page 1 of the questionnaire assessed subjects’ perception of the category–role correlation implicit in the set of target descriptions. Subjects were asked to indicate the approximate percentages of each category who were city workers and child raisers (with a reminder that the two percentages for a given category must sum to 100%). They were considered to have passed this manipulation check if the percentage of city workers indicated for the predominantly city-working category was greater than that indicated for the predominantly child-raising category. By this criterion, 26 subjects (22% of the total) failed and were therefore replaced. The replacement of those subjects who failed to observe the category–role correlation was not considered a threat to external validity, because it can safely be assumed that everyone in this society has observed the correlation between gender and role (the real-life analogue of our category–role correlation).

We did, however, repeat the major analyses with these 26 subjects included. The subjects fell into two groups: those who stated that exactly 50% of both categories occupied each of the two roles \((n = 12)\) and those who actually reversed the true direction of the correlation \((n = 14)\). When only the former group was included, the major results were unchanged. The category-stereotyping effects (although weaker) remained significant at the reported levels, as did the effects of category type and explanatory set. When the latter group was included as well, the category-stereotyping effects were further weakened but remained significant—as did the category-type effects—but the significance of the explanatory-set effects dropped to the \(p < .10\) level.

Ten other subjects were also replaced—3 who heard about the nature of the research in advance, and 7 whose data were incomplete in important ways.

**Manipulation of explanatory set.** For half the subjects (the explanatory condition), page 2 of the questionnaire asked them to think about the category–role relation they reported on the preceding page and asked, “Can you think of a reason why the Orinthians and the Ackmians might tend to occupy the particular roles that they do?” (with instructions to respond in writing). For the other half (the no-explanation condition), this page of the questionnaire was omitted.

**Stereotyping measures.** Two measures of stereotyping appeared on the next pages of the questionnaire. In one, the role-unspecified measure, subjects rated the personal qualities of the “Orinthians in general” and the “Ackmians in general” on six agentic traits (“ambitious,” “assertive,” “competitive,” “independent,” “outspoken,” and “self-confident”) and six communal traits (“affectionate,” “emotional,” “gentle,” “helpful,” “kind,” and “understanding”). Half of these had appeared in the target descriptions, and half were taken from the set of pretested traits that did not appear in any of the descriptions. Three neutral traits (“happy,” “realistic,” and “responsible”) were included as filler items. Ratings were made on a scale ranging from not at all \((0)\) to extremely \((9)\).

The other measure was the role-specified measure. Subjects rated the personal qualities of the Orinthian child raisers, the Ackmian child raisers, the Orinthian city workers, and the Ackmian city workers, respectively, on the same set of traits described above. Half the subjects completed the role-unspecified measure first, and half the role-specified measure.

**Recall measure.** Given the already complex design of the study, it was not feasible to evaluate the possible contribution of memory factors to stereotype formation in a comprehensive fashion, but we did include one memory measure. The next page of the questionnaire asked subjects to list as many traits as they could remember that had appeared in the descriptions of individual Orinthians presented at the start of the experiment; the following page requested a similar listing for the Ackmians. Subjects were told to avoid guessing and to list only those traits they actually remembered. The next page asked subjects to return to the preceding two pages and to indicate, next to each trait they had listed, whether the trait had been used to describe a city worker or a child raiser. Again, subjects were told to avoid guessing and to leave an item blank unless they actually recalled which role the individual in question had occupied.

**Postexperimental inquiry.** The last page of the questionnaire asked subjects to indicate what they believed to be the true purpose of the study.

**Subjects.** The final sample consisted of 48 female and 32 male subjects, who participated in partial fulfillment of a course requirement. There were 5 subjects (3 female and 2 male) in each cell of the 2 (category type) \(\times 2\) (explanatory set) \(\times 2\) (booklet version) \(\times 2\) (order of measures) design. Subjects were randomly assigned to conditions (separately by sex), with the constraint that there be equal numbers of subjects in each of the 16 cells.
Results

The dependent measures were first analyzed for possible effects of the two counterbalancing variables (booklet version and order of measures) in a series of Booklet × Order × Category Type × Explanatory Set analyses of variance. The only significant main effect or interaction associated with either variable was a Booklet × Order × Explanatory Set interaction on subjects' role distribution estimates, \( F(1, 64) = 4.75, p < .05 \). However, given that two of the three variables involved were not manipulated until after these estimates were made, and given the large number of \( F \) ratios tested in these analyses, this single significant effect is best regarded as a fluke. Therefore, all analyses reported below have been collapsed over booklet and order.

Role distribution estimates. Two numbers were extracted from subjects' answers to the questions assessing their perception of the role distribution: the estimated percentage of city workers in the predominantly city-working category, and the estimated percentage of child raisers in the predominantly child-raising category. These percentages were then averaged to form a single estimate for each subject. The overall mean was 69.4% (the correct figure is 80%). No subject stated that 100% of either category occupied the same role; thus, all subjects were aware that each category included some individuals who performed the atypical role for that category. The two category-type conditions did not differ in their estimates (Ms = 69.6% vs. 69.3%, \( F < 1 \)); nor did the two explanatory-set conditions (Ms = 68.9% vs. 70.0%, \( F < 1 \)).

Nature of subjects' explanations. Responses in the explanation condition to the question asking about a possible reason for the role distribution were readily classifiable into two types: those that cited differences in personality between the two categories as the reason, and those that did not. Of the 40 subjects in this condition, 29 (72%) attributed the distribution to personality differences. Here are excerpts from two typical answers:

The Orinthians are on average the kind and sensitive . . . species. The Ackmians are more self-confident and forceful, therefore better suited for working in the city.

The Orinthians and Ackmians might occupy the particular roles that they do, due to their personalities. . . . Patient, kind, understanding, etc., would be characteristics you would want in a child-raiser, and a city-worker you would wish to be determined, energetic, etc.

The other 11 subjects (a) said that they were unable to explain the distribution, (b) attempted (rather unsuccessfully) to explain the distribution in terms of ideas such as sheer tradition or the need for some sort of division of labor, or (c) failed to address the question as stated.

Overall, then, the responses clearly confirmed our expectation that the preferred explanation for the role distribution would be simply that each category's personality causes it to be well suited to its predominant role.

Role-unspecified rating measure. For convenience, the predominantly city-working category shall be called the male-analogue category; the predominantly child-raising category, the female-analogue category. Composite rating scores were computed by subtracting each subject's mean rating of a category on the six communal traits from his or her mean rating of that category on the six agentic traits. Thus, positive values indicate higher ratings on agentic than communal traits; negative values indicate the opposite. Means are shown in Table 1. In order to further simplify the presentation of results, a stereotyping index was then computed by subtracting the score for the female-analogue category from the score for the male-analogue category (see the last column of Table 1). Stereotyping is indicated by a value on this index greater than 0; the larger the value, the greater the stereotyping (i.e., the greater the differentiation of the two categories). A negative value would indicate "reverse" stereotyping. All analyses were performed directly on this index. The use of a male-analogue/female-analogue difference score is, of course, equivalent to treating category (male vs. female-analogue) as a within-subject variable. The difference-score method, however, reduces the dimensionality of the design, so that interactions between category and the two between-subjects factors can be more simply described as main effects of those factors on stereotyping. For the analyses of covariance (ANCOVAS) presented later, on the other hand, it was necessary to include category as a variable.

As Table 1 shows, the male-analogue category (overall \( M = 1.98 \)) was rated quite differently than the female-analogue category (overall \( M = -2.45 \)), mean stereotyping index (\( M_{\text{SD}} = 4.42, F(1, 76) = 139.20, p < .001 \)). This stereotyping effect was not, however, equally strong in all conditions. As predicted, greater stereotyping occurred in the explanation condition (\( M_{\text{SD}} = 5.33 \)) than in the no-explanation condition (\( M_{\text{SD}} = 3.51, t(76) = 2.44, p < .01 \)). The effect of category type was not significant, \( t(76) = 1.58 \). There was no Category Type × Explanatory Set interaction (\( F < 1 \)).

Although generally consistent with predictions, the results of the role-unspecified measure do not provide direct support for the hypothesis that the stereotypes represent or entail beliefs about the intrinsic natures of the categories. The stereotyping evident in subjects' responses to this measure could have resulted entirely from a chain of inference similar to the following.

Table 1

<table>
<thead>
<tr>
<th>Condition</th>
<th>Male-analogue category</th>
<th>Female-analogue category</th>
<th>Stereotyping index*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonbiological</td>
<td>1.04</td>
<td>-1.66</td>
<td>2.70</td>
</tr>
<tr>
<td>Explanation</td>
<td>2.27</td>
<td>-2.69</td>
<td>4.96</td>
</tr>
<tr>
<td>Biological</td>
<td>1.92</td>
<td>-2.40</td>
<td>4.32</td>
</tr>
<tr>
<td>No explanation</td>
<td>2.68</td>
<td>-3.03</td>
<td>5.71</td>
</tr>
<tr>
<td>( M )</td>
<td>1.98</td>
<td>-2.45</td>
<td>4.42</td>
</tr>
</tbody>
</table>

* Male-analogue score minus female-analogue score.

2 Throughout this study, directional tests were used to evaluate predicted effects of explanatory set and category type, whereas nondirectional tests were used to evaluate unpredicted interactions, the category-stereotyping effect (because we did not firmly predict that it would occur in all conditions), and other effects. Instead of noting which type of test was made in each case, we report \( t \) for directional tests and \( F \) for nondirectional tests.
ing: “City workers [child raisers] typically have such and such traits; most members of this category are city workers [child raisers]; therefore, on average, this category has such and such traits.” This would amount to rating the category in general on the basis of the typical category member and need not entail the attribution of categorywide tendencies. The role-specified measure, however, does speak directly to this latter type of attribution.

Role-specified rating measure. Again, composite scores were computed by subtracting the subject’s mean rating of a given group on the six communal traits from the mean rating of that group on the six agentic traits. A stereotyping index was created by subtracting the mean of the two male-analogue scores from the mean of the two male-analogue scores. Results appear in Table 2.

Table 2 shows that even when role was specified, category stereotyping still occurred: The male-analogue category (overall \( M = 1.26 \)) was rated quite differently than the female-analogue category (overall \( M = -1.48 \)). Again, however, the effect was not equally strong in all conditions. Greater stereotyping occurred in the explanation condition (\( M_s = 3.23 \)) than in the no-explanation condition (\( M_s = 2.25 \)), \( t(76) = 1.85, p < .05 \). Greater stereotyping also occurred in the biological condition (\( M_s = 3.32 \)) than in the nonbiological condition (\( M_s = 2.16 \)), \( t(76) = 2.18, p < .025 \). There was no Category Type x Explanatory Set interaction (\( F < 1 \)).

Role, of course, also had a significant effect on ratings, \( F(1, 76) = 101.95, p < .001 \). This simply confirms that the cityworker and child-raiser roles are indeed associated with different personality traits—a necessary precondition for the category-stereotyping effect itself.

In a further analysis of these data, we took into account differences in the schematic relevance of the rated traits. The 12 traits comprising the stereotyping measures vary in their perceived relevance to the city-worker and child-raiser roles, and one might expect corresponding differences in the strength of the stereotyping effect on those traits. To obtain an index of schematic relevance, we asked a new group of subjects (27 female and 19 male) to rate each trait for the extent to which it is conducive to success in the business world, and the extent to which it is conducive to successful child raising. The difference between the two ratings for a given trait was used as an index of its relevance to the city-worker versus child-raiser roles. On the basis of these difference scores, the traits were divided into a high-relevance set (the three agentic and three communal traits with the largest scores) and a medium-relevance set (the three agentic and three communal traits with the smallest scores). Composite trait-rating and stereotyping scores were then computed as before, separately for each set.3

The first result of interest is that, overall, greater stereotyping took place on the high-relevance traits (\( M_s = 2.95 \)) than on the medium-relevance traits (\( M_s = 2.53 \)), \( F(1, 76) = 8.18, p < .01 \). This effect of trait relevance was qualified, however, by a marginally significant Explanatory Set x Trait Relevance interaction, \( F(1, 76) = 3.57, p < .10 \). Greater stereotyping occurred on the high- than on the medium-relevance traits in the explanation condition (\( M_s = 3.58 \) vs. 2.88, respectively, \( p < .01 \)), but not in the no-explanation condition (\( M_s = 2.32 \) vs. 2.18, respectively, \( p > .50 \)).

We also tested the effects of explanatory set and category type separately for the two trait sets. Significantly greater stereotyping occurred in the explanation than in the no-explanation condition on the high-relevance traits (\( M_s = 3.58 \) vs. 2.32, respectively), \( t(76) = 2.07, p < .025 \), but not on the medium-relevance traits (\( M_s = 2.88 \) vs. 2.18, respectively), \( t(76) = 1.44, ns \).

We also tested the effects of explanatory set and category type separately for the two trait sets. Significantly greater stereotyping occurred in the explanation than in the no-explanation condition on the high-relevance traits (\( M_s = 3.58 \) vs. 2.32, respectively), \( t(76) = 2.07, p < .025 \), but not on the medium-relevance traits (\( M_s = 2.88 \) vs. 2.18, respectively), \( t(76) = 1.44, ns \).

3 In the norming study, there were differences in male and female subjects' ratings of the 12 traits. For example, male subjects regarded the trait "independent" as much more conducive to success in business than to success in child raising, whereas female subjects regarded it as more nearly equally conducive to success in both areas. (There were other differences as well.) Therefore, two divisions of the traits into high- and medium-relevance sets were actually made, one of which was applied to the data of the male subjects in the main experiment, and the other to that of the female subjects. For male subjects, the high-relevance traits were "competitive," "affectionate," "gentle," "kind," "independent," and "ambitious"; the medium-relevance traits were "emotional," "outspoken," "understanding," "helpful," "assertive," and "self-confident." For female subjects, the high-relevance traits were "competitive," "affectionate," "gentle," "ambitious," "emotional," and "outspoken"; the medium-relevance traits were "understanding," "kind," "assertive," "helpful," "independent," and "self-confident."
greater stereotyping occurred in the biological than in the non-
biological condition both on the high-relevance traits ($M_b = 3.56$ vs. 2.33, respectively), $t(76) = 2.02, p < .025$, and on the medium-relevance traits ($M_b = 3.08$ vs. 1.98, respectively), $t(76) = 2.24, p < .025$.

Recall measure. The mean number of traits listed in the free-
task recall was 10.56 (both categories combined). Of these, an 
average of 7.51 had in fact appeared in the target descriptions, 
and 3.05 had not. Interestingly, 28.0% of the traits listed, on 
average, were listed for both categories. This indicates that sub-
jects were aware of (and remembered) the fact that there was 
overlap in the descriptions of the two categories (though they 
were not aware that the overlap was complete). The average per-
centage correct on the role identification task (in which subjects 
were not aware that the overlap was complete). The average per-
centage correct on the role identification task (in which subjects 
were told to indicate, next to each trait they had listed, whether 
the target described by that trait was a city worker or a child 
raiser) was 62.7%. Thus, although there was certainly some 
memory for the target descriptions, the extent of this memory 
was modest, as might be expected given the amount and com-
plexity of information and the number of tasks intervening be-
tween exposure and recall.

Several indices were derived from the recall data and then 
correlated with stereotyping. Predictably, subjects were most 
likely to recall agentic traits for the male-analogue category and 
communal traits for the female-analogue category, and the in-
dex having the highest and most consistent relation to stereotype-
ing was one that took this kind of distortion into account. This 
index was based on the proportional score $(A - C)/(A + C + 
N)$, where $A$, $C$, and $N$ refer to the total numbers of agentic, 
communal, and neutral traits listed, respectively. This score was 
first computed for the two categories separately, and the female-
analogue score was then subtracted from the male-analogue 
score, in a manner analogous to the computation of the stereo-
typing index. Within-condition correlations between the recall 
and stereotyping indices ranged from $.62$ to $.79$ for the role-
unspecified measure (mean $r = .70, p < .001$) and from $.42$ to 
$.60$ for the role-specified measure (mean $r = .54, p < .001$).

Analyses of covariance were then used to test the possible 
contribution of recall to stereotyping. For these analyses, the 
recall and stereotyping indices were separated into their male-
analogue and female-analogue components, and category was 
treated as a within-subject factor. The design was thus a Cate-
gory (male vs. female analogue) × Category Type × Explanatory 
Set factorial, with trait ratings as the dependent measure 
and recall scores as the (varying) covariate (Winer, 1971).

The adjusted effect of category on subjects' ratings was, natu-
really enough, considerably smaller, but the effect remained 
highly significant: $F(1, 75) = 60.52, p < .001$, for the role-un-
specified measure, and $F(1, 75) = 38.62, p < .001$, for the role-
specified measure. The obvious conclusion would seem to be 
that stereotype formation in this study was not primarily a function of, or at least did not entirely depend on, memory-
based factors. This conclusion should be regarded as tentative, 
however, because as noted earlier the present study was not set 
up to test the role of memory in a definitive manner. On the 
other hand, there are at least two additional pieces of evidence 
consistent with this conclusion. First, the number of target traits 
correctly recalled was not significantly related to stereotyping: 
Mean within-condition $r_s = -.14$ for the role-unspecified mea-
sure and $-.10$ for the role-specified measure. Second, and per-
haps more important, the extent of overlap in the information 
remembered about the two categories (i.e., the percentage of 
traits listed for both categories) was also not significantly related 
to stereotyping: Mean within-condition $r_s = -.13$ for the role-
unspecified measure and $-.20$ for the role-specified measure. 
Thus, contrary to what might be expected if (lack of) memory 
for the stimulus information were a major factor in stereotype 
formation, subjects who recalled a good deal of the informa-
tion, and even those whose recall correctly emphasized the simi-
larities in the two categories' personalities, did not show sig-
nificantly less stereotyping than did the others. And, as the AN-
COVAs demonstrated, even the distortion in memory did not 
account for all of the stereotyping effect.

Postexperimental inquiry. The possibility that subjects may 
have equated the two categories with the human sexes, in spite 
of our efforts to prevent this from occurring, was of some con-
tern to us. We therefore examined subjects' responses to the 
postexperimental inquiry, in which they were asked to guess the 
true purpose of the study. Only 10 of the 80 subjects (12%) made 
amy mention whatsoever of gender in their remarks (all subjects 
using the word sex or gender or using any sex-linked term such 
as female, male, woman, man, mother, father, etc., were counted 
in this group). Thus, most subjects were apparently unaware 
that the study concerned gender in any way.

Although gender awareness was unrelated to category type (5 
of the 10 subjects were in the biological condition and 5 were 
in the nonbiological condition), it was significantly related to 
explanatory set: Nine of the 10 subjects belonged to the no-
explanation condition, $(p < .025)$. Why this pattern occurred 
is unclear, but it does refute the potential argument that the 
explanatory-set manipulation "worked" only because subjects 
in the explanation condition were more likely to identify the 
categories with the human sexes and therefore were more likely 
to stereotype the categories in gender-related terms. (Similarly, 
the fact that gender awareness was unrelated to category type 
tends to refute the analogous argument pertaining to that ma-
nipulation.) As a further check on the possible effect of gender 
awareness on the results, we repeated the analyses of the stereo-
typing data with the 10 gender-aware subjects deleted. All for-
ermantly significant effects—including those of explanatory set— 
remained significant.

Despite the small number of gender-aware subjects in Experi-
ment 1, it would still be desirable to show that the stereotyping 
effect can be obtained even when the roles in question have no 
differential association with familiar human categories. A sec-
ond study was therefore conducted for this purpose.

Experiment 2

Experiment 2 was a simplified replication of Experiment 1, 
in which the city-worker and child-raiser roles were replaced by 
two new roles. We wanted roles that satisfied two criteria: They 
should be associated with different personality traits, but they 
must not be differentially associated with gender, ethnicity, or 
similar biosocial variables. Two roles that seemed to satisfy 
these criteria were "business persons" versus "academics." The 
former role is associated with extraverted/ambitious traits, and
the latter with introverted/intellectual traits, but they are not differentially linked to familiar human categories. The key word here is differentially: although the roles are clearly associated with men more than women, with Whites more than Blacks, and so forth, the point is that this is equally true of both roles.

Because the only goal of Experiment 2 was to replicate the basic category-stereotyping effect, we did not manipulate category type or explanatory set, but instead ran only a condition corresponding to the biological/no-explanation condition of Experiment 1 (which most closely resembles the real-life situation to which we would like to generalize).

Method

Orienting information. The introduction to the experiment was essentially identical to that used in the biological condition of Experiment 1, except that there was no discussion of gender or reproduction, and the passage describing the city-worker and child-raiser roles was replaced by the following:

In the cities, where most [Orinthians and Ackmians] live, the work force is divided into two major sectors: the "free-enterprise" sector and the "research/educational" sector. To the Ackmian, a "liberal/generous," "friendly/congenial," "good-natured/cheerful," "self-confident/self-assured," "sociable/outgoing," "spirited/lively" and "uninhibited/impulsive," social roles. The extraverted trait pairs were "active/energetic," "domineering/argumentative," "entertaining/talkative," "friendly/congenial," "good-natured/cheerful," "self-confident/self-assured," "sociable/outgoing," "spirited/lively" and "uninhibited/impulsive." The introverted trait pairs were "cautious," "intellectual," "introspective," "introverted," "reserved," "solitary." One fourth of these were "old" extraverted traits, and one fourth were "old" introverted traits; half of each set had originally been attributed to Orinthians, half to Ackmians. The remaining four were new traits.

Postexperimental inquiry. The last page of the questionnaire asked for a detailed statement of the subject's perception of the true purpose of the experiment and of what we were "trying to test or study in this research."

Subjects. The final sample consisted of 22 female and 10 male subjects who participated in partial fulfillment of a course requirement. There were 8 subjects (5 to 6 female and 2 to 3 male) in each cell of the 2 (booklet version) × 2 (order of stereotyping measures) design. Subjects were randomly assigned to conditions (separately by sex) with the constraint that cell sizes be equal.

Results

There were no significant effects associated with either counterbalancing variable (booklet version and order of measures) on any of the dependent measures. All analyses have therefore been collapsed over booklet and order.

Precisely because the roles in question are not differentially linked to familiar groups, there are no convenient labels by which to refer to the target categories. They therefore are called Category A (the one working primarily in the free-enterprise sector) and Category B (the one working primarily in the research/educational sector). Composite scores were derived from the role-unspecified ratings by subtracting each subject's mean rating of a category on the six introverted traits from his or her mean rating of that category on the six extraverted traits; a stereotyping index was then computed by subtracting the Category B score from the Category A score. Similar scores were derived from the role-specified ratings, and a stereotyping index was computed by subtracting the mean of the two Category B scores from the mean of the two Category A scores.

The stereotyping effect, although not as strong as in Experiment 1, was replicated on both measures. On the role-unspeci-
fied measure, Category A received a mean score of 1.98 and Category B a mean score of −0.22; $M_2 = 2.20$, $F(1, 31) = 13.90$, $p < .001$. On the role-specified measure, the mean for Category A was 1.34 and for Category B was 0.06; $M_2 = 1.28$, $F(1, 31) = 11.16$, $p < .01$. The fact that Category B received a nonnegative mean score on the role-specified measure shows that it was not rated higher overall on the introverted traits than on the extraverted traits. This result probably means only that the introverted traits chosen for the rating scales are not as strongly linked to the “research/educational” role as the extraverted traits are to the “free-enterprise” role. The only important result is that Category B was seen as more introverted and less extraverted than Category A.

In their responses to the postexperimental inquiry, not a single subject made any reference whatsoever to a human category based on gender, ethnicity, race, age, or similar variables. We wish to call special attention to this finding, because it confirms our expectation that subjects would not associate the target categories with familiar human groups.

**General Discussion**

This research demonstrates several things relevant to an understanding of gender stereotypes. In Experiment 1, stereotypes arose in direct response to a difference in how two categories of individuals were distributed into social roles; a similar effect was demonstrated in Experiment 2 even though the roles in question were not differentially associated with familiar human groups. The studies thus show that objective sex differences in personality are not necessary to the formation of gender stereotypes; the fact of an unequal role distribution is sufficient. Both studies found that category membership affected ratings of a target group even when social role was specified. This result strongly supports the idea that stereotypes such as those based on gender constitute, in part, beliefs about intrinsic, category-wide tendencies or predispositions, an idea that is central to the present analysis.

In addition, the research provides several pieces of evidence, both direct and indirect, consistent with the proposition that stereotype formation of the kind examined here is at least partly mediated by the attempt to rationalize or explain the category-role correlation. Most directly, inducing subjects to explain the correlation resulted in stronger stereotypes. Moreover, a large percentage of these subjects did in fact offer explanations of the hypothesized type, namely, that the categories tend to occupy different roles because their personalities suit them to those roles. In addition, those traits regarded as most differentially relevant to role performance in the norming study were also more likely to be stereotypically applied to the target categories in the main study, and this was especially true for subjects who attempted to explain the category-role correlation. Although we do not have the data to support the following speculation (and were unable to locate exactly the right kind of data in the stereotype literature), it is our impression that the most differentially role-relevant traits (gentle and affectionate, competitive and ambitious) are also more gender stereotypical, in real life, than are the least differentially relevant traits (understanding and helpful, self-confident and assertive). If this turns out to be true, it would provide further support for the proposition that the gender stereotypes derive in part from a category-role rationalization mechanism.

Furthermore, the ANCOVAs showed that only part of the stereotyping effect disappears when the distortion in memory for the stimulus information is controlled for. Although not evidence for rationalization per se, this finding is consistent with the proposal that stereotype formation of the kind at issue here is not purely an information-processing phenomenon (even taking into account the biases in such processing). Lastly, stronger stereotypes were formed of biologically defined categories, suggesting that stereotype formation is especially likely if it is possible (as in the case of gender) to regard the categories as inherently different on the relevant dimensions. Although also not direct evidence for rationalization, this result is nonetheless exactly what one would expect if, as we have speculated, role-based category stereotypes originate in a rationalization process that operates by positing intrinsic differences between the categories in question.

Several additional points deserve clarification. First, we emphasize that the primary goal of this research was to demonstrate that certain conditions are theoretically sufficient to account for the gender stereotypes—not that such conditions are necessarily the only factor responsible for their existence in the real world. Second, our analysis does not deny the possible existence of important sex differences in personality. It simply asserts that whatever such differences may exist are not exclusively (and probably not even primarily) responsible for the stereotypes. Once formed, however, the stereotypes may well enable perceivers to “detect” small sex differences in personality, and this in turn may be a major factor contributing to the maintenance of the stereotypes. Once in place, the stereotypes may also cause perceivers to view role-based differences in personality (such as any that may exist between male breadwinners and female homemakers) as “evidence” for the stereotypes, even though, logically speaking, the evidential value of such differences is nil.

Third, we remind the reader that the present account is primarily concerned with people’s belief that males are agentic and females are communal. The concepts of agency and communion, however, although highly central to the stereotypes (Eagly, 1987), by no means exhaust the full range of their contents. And it seems likely that at least some of these other features of the stereotypes do in fact have a primary basis in observed sex differences.

Fourth, we do not deny the role of cultural transmission in stereotype formation. The problem of how and why gender stereotypes arise in a historical sense is to some extent separable from the problem of how and why they take hold in the minds of succeeding generations of individuals. The rationalization mechanism, we believe, plays a role in both processes, but as far as the second is concerned, there can be almost no doubt that cultural transmission is also involved. (In addition, the notion that males’ and females’ personalities suit them to their roles may itself be one of the messages that the culture transmits.) Cultural transmission cannot, however, explain how the stereotypes originate in the first place, whereas this is a question to which the rationalization hypothesis has an answer: They originate in an attempt to rationalize the division of labor by
attributing to each sex those qualities deemed necessary for performance of the assigned functions.

What about the origin of the latter beliefs themselves, that is, people's beliefs or schemas concerning the attributes required for the kinds of tasks traditionally performed by women and men? If it is true that gender stereotypes (i.e., their agentic and communal components) are based on these role schemas, it must also be true that the role schemas precede the stereotypes. Given, however, that homemakers and breadwinners are predominantly female and male, respectively, it might seem to follow that the role schemas themselves are just abstractions based on observed differences in the personalities of (certain) women and men. How, then, can we untangle the origin of the gender stereotypes from the origin of the role schemas, given that gender and role are confounded in this way? This seeming circularity is more apparent than real. Even if we grant that the role schemas in question do in fact originate in this manner—which is by no means obvious to us—the interesting and important thing about gender (and some other) stereotypes is not that a subset of the category may indeed resemble the stereotype, but that the stereotype encompasses the category as a whole. The key problem to which our analysis is addressed is the question of how and why certain traits become conceptually linked to biological sex per se. It is these generalized beliefs about the categories of male and female, as well as their breadth of influence, that we have attempted to explain. It is important to note, moreover, that there is at present no evidence to support the assumption that role schemas such as those we have in mind actually do originate in veridical perceptions of roles occupants' personalities. Almost nothing is known with certainty about the origins of people's conceptions of social roles or personality traits. Although space does not permit a discussion here, we think it is possible to construct a coherent account of the origin of the role schemas that does not require the assumption of differences in the actual personalities of the role occupants or the assumption that the schemas are directly grounded in the perception of such differences.

Because the sexual division of labor figures importantly in our account, the question of what gives rise to that particular division of labor deserves mention. The explanation that we find most compelling is that proposed by Williams and Best (1982). The sexual division of labor, according to these authors, derives from the basic fact that, in earlier times at least, it was always necessary for the woman to be available to nurse her infant children. This in turn placed constraints on the woman's mobility, and so it made sense for her to remain at home and assume the other child-care responsibilities as well as most of the homemaking duties in general. The tasks of the man—hunting, fighting, and their contemporary equivalents—fell to him more or less by default; they were tasks that could not be as efficiently performed by the woman because of the constraints imposed on her by the requirements of nursing the young. An additional factor is that it is preferable, from the point of view of group survival, to expose men rather than women to dangerous activities such as hunting and fighting. This is because a group or society can lose most of its men and still reproduce at an acceptable rate, whereas a group that loses too many of its child-bearing women risks extinction. Thus, according to this account, the sexual division of labor has little or nothing to do with differences in males' and females' psychological characteristics at all.

The irony, of course, is that even though in many societies the original reasons for the division of labor ceased to be obvious or compelling quite some time ago, the traditional arrangement has tended to persist. Although gender stereotypes may themselves have some causal role in this persistence, the major factor is undoubtedly just the inherent conservatism of societal structures. Yet in a culture with a value system like our own, mere "tradition" does not constitute an adequate rationale for the conduct of social life. A type of rationale that we generally do find acceptable, however, is the notion that differences in ability and personality justify differences in status and role (cf. Shweder, 1982); it is in this context that the stereotypes serve an important cultural function.

Finally, we should note that although we have concentrated exclusively on gender stereotypes in this article, the rationalization hypothesis is potentially applicable to a variety of other stereotypes as well, including aspects of many ethnic and racial stereotypes. For example, one widely proposed theory (e.g., Allport, 1958) of the origin of the American stereotype of Blacks as lazy, unintelligent, and generally inferior maintains that this stereotype arose originally in order to justify the institution of slavery and that its persistence even today is partly a means of rationalizing Blacks' lower socioeconomic status. We might also point out that there exist other possible forms of rationalization besides the rationalization of role or status differences, some of which may also play a part in stereotyping and impression formation. For instance, as Hoffman (1985) has argued, the attribution of certain personality traits to an individual can serve to rationalize one's decisions about, or conduct toward, that person (e.g., a decision to marry someone or to not admit someone to one's graduate program). The justifications and rationalizations that people seek for the realities of social life may indeed be quite varied, and it is hoped that the present research will prove to be a useful step in the understanding of one such process.

References


Gurwitz, S. B., & Dodge, K. A. (1975). Adults' evaluation of a child as


Received April 29, 1988
Revision received August 17, 1989
Accepted August 18, 1989

---

**Call for Nominations for Psychological Bulletin**

The Publications and Communications Board has opened nominations for a new editor of *Psychological Bulletin*. John C. Masters is the immediate past editor. Candidates must be members of APA and should be prepared to begin receiving manuscripts by September 1990. Please note that the P&C Board encourages participation by members of underrepresented groups in the publications process and would particularly welcome such nominees. To nominate candidates, prepare a statement of one page or less in support of each candidate. Submit nominations to:

Walter Mischel, PhD  
*Psychological Bulletin* Editor Search Committee  
American Psychological Association  
1200 Seventeenth Street, N.W.  
Suite VA-710  
Washington, D.C. 20036

Other members of the search committee are Ann Brown, Lynn Ann Cooper, Alan Kazdin, and Bruce Sales. First review of nominations will begin March 1, 1990, although the nomination deadline is March 31, 1990.