1999

Introduction

Dan Bernstein
University of Nebraska-Lincoln

Published in GENDER AND MOTIVATION:
VOLUME 45 OF THE NEBRASKA SYMPOSIUM ON MOTIVATION,
ed. Richard Dienstbier and Dan Bernstein.
Lincoln, NE: University of Nebraska Press, 1999. Pages vii–xxiii.

Introduction

Dan Bernstein
University of Nebraska–Lincoln

This volume is about motivation and gender. The chapters outline recent research and conceptual analysis related to four important motivational constructs—sexuality, emotion, competition, and aggression. In each case the author has examined the relation between the motivational construct and gender; the chapters describe those relations and analyze their origins and implications. There are two primary ideas that connect these accounts of gender and motivation: the authors generally report great diversity within gender groups in the degree to which these motivational characteristics are found, and they note that there is much to be considered in exactly how these motivational constructs are defined and measured. One could easily conclude that there is tremendous overlap in the amount of aggression, sexuality, emotion, and competition shown by males and females, even given conventional conceptions of the measurement of those constructs. When an alternative and thoughtful reconstruction of the motivational variables is added to the analysis, the overlap becomes even greater, and differences disappear or even reverse their order. Faced with data showing substantial overlap in characteristics, one is left to ponder why human perception of gender differences is so richly caricatured and so firmly held. This is an

emergent theme of the volume, a question that is not always directly asked but is suggested by the chapters taken as a whole.

Attribution of Essential Difference

It seems to be a fundamental feature of human perception that complex and changing patterns of sensation are transformed into stable and manageable entities. Shape and size constancy, for example, make it possible for people to navigate sensibly despite the constantly variable patterns of stimulation made by the physical features of the environment on the sense organs. There are parallel phenomena in the perception of people, as the contextually embedded actions of people are formed into coherent and stable accounts of individual difference through the attribution of dispositional traits or personality types. Sometimes these abstracted characteristics leave out so much contextual detail and individual variability that we call them stereotypes, but much of the time our dispositional map allows for successful navigation through the social world with a minimum of damage to self or other.

Sometimes psychological scholarship can inadvertently contribute to the development of generalizations about people, especially to generalizations about groups of people. Our professional task often seems to be the identification of common elements (which we call main effects) among the varied specific individual cases that make up different samples of humanity that are observed. Success in that task can be a function of finding a measurement procedure powerful enough or a sample large enough to reveal an existing main effect (in this case a difference between the groups) despite the wide variability within the groups being compared. Skill in accomplishing this goal is described by the terms *statistical* or *experimental power*, as it may require a high degree of such power to identify an important but small effect in the noisy behavioral contexts in which we work.

It is clear that skilled researchers recognize the limits of what can be said about individuals given such measurement procedures, and there are regular pronouncements warning consumers of the research not to overgeneralize the findings. These warnings are not new to psychology, however, and their long standing is evidence of

the resilience of this form of human perception. In 1910 Robert S. Woodworth published an essay on racial differences in mental traits in the journal *Science;* he made this point extraordinarily clear with regard to both culture and race in two sections. First, while considering the notion of cultural or national "types" (such as the typical German or the typical southerner), Woodworth wrote:

> If we would scientifically determine the facts regarding a group of men, we should, no doubt, proceed to examine all the individuals in the group, or at least a fair and honest representation of them. The first fact that meets us when we proceed in this way is that the individuals differ from each other, so that no one can really be selected as representing the whole number. We do find, indeed, when we measure the stature or any other bodily fact, or when we test any native mental capacity, that the members of a natural group are disposed about an average, many of them lying near the average, and few lying far above or far below it; and we thus have the average as a scientific fact regarding the group. But the average does not generally coincide with the type, as previously conceived, nor do the averages of different groups differ so much as the so-called types differ. Moreover, the average is itself very inadequate, since it does not indicate the amount of variation that exists within the group—and this is one of the most important facts to be borne in mind in understanding any collection of individuals. It is specially important in comparing different groups of men, since the range of variation within either group is usually much greater than the difference between the averages of the groups. The groups overlap to such an extent that the majority of the individuals composing either group might perfectly well belong to the other. (pp. 171–172)

Woodworth then claims that human perception seems naturally to include a tendency to make those erroneous generalizations, in spite of the nature of the actual data. When suggesting that this form of thinking can be misapplied to questions of race he wrote:

> Our inveterate love for types and sharp distinctions is apt to stay with us even after we have become scientific, and vitiate our use of statistics to such an extent that the average becomes a stumbling-block rather than an aid to knowledge. We desire,

> for example, to compare the brain weights of whites and of negroes. We weigh the brains of a sufficient number of each race—or let us at least assume the number to be sufficient. When our measurements are all obtained and spread before us, they convey to the unaided eye no clear idea of a racial difference, so much do they overlap. If they should become jumbled together, we should never be able to separate the negroes from the whites by aid of brain weight. But now we cast up the average of each group, and find them to differ; and though the difference is small, we straightway seize on it as the important result, and announce that the negro has a smaller brain than the white. We go a step further, and class the white as a large-brained race, the negro as a small-brained. Such transforming of difference of degree into differences of kind, and making antitheses between overlapping groups, partakes not a little of the ludicrous. (p. 172)

It seems clear that there are many group average differences to be identified among subpopulations; no one disputes that observation. Woodworth's point has more to do with how people use the data than is it a complaint about the data themselves.

Despite Woodworth's warnings and many others since, reasonable people who are repeatedly exposed to findings reported as significant mean differences or nonchance factors in a multivariate representation sometimes begin to think and talk as if those differences were actually true in most individual cases. Readers begin to use constructs such as culture, personality, gender, clinical diagnosis, and race as if they indicated the existence of distinct categories of people. Perhaps for the same reasons of efficiency that human perception produces constancy of size and shape, the frequent use of culture as a researchable variable results in the production of cultural stereotypes dressed up in statistical significance. Similarly there are many caricatures of gender that arise as an unwanted by-product of a useful reconsideration of the role of gender as a variable in psychological research.

My own interest in this phenomenon was piqued while teaching a course in cross-cultural psychology to advanced undergraduate psychology students. After reading two leading texts in cross-cultural psychology, the students became very fluent in cultural regularities; for example, they were happy to state that "collectivist"

cultures would of course show one pattern of behavior while in an "individualist" culture one would find a different (and implicitly distinct) pattern of action. All of these constructions were in fact drawn from the material presented in the texts, as the authors reviewed study after study that reported statistically significant differences in the behavior of varying cultures. In many cases these differences correlated with abstract dimensions like individualism/collectivism, leading to the students' articulate statements about what the data meant. Ultimately I was distressed because it seemed that students were trafficking in cultural stereotypes, even though the stereotypes were dressed up in academically respectable terms. The fundamental error was being made; a small average difference was being interpreted to imply the existence of truly distinct populations.

The overall theme of this volume is to suggest that data reported in studies examining group variables such as culture, gender, personality, or race should be presented in a complete way that might reduce the tendency of readers to form unwarranted and unwanted caricatures of categories of people. In particular, the presentations in this volume refer to data and inferences based on the variable of gender; this is a crucial variable in human life, because even people who live in homogeneous communities with little cultural diversity generally encounter members of both genders. The importance of gender stereotyping cannot be overemphasized as a phenomenon in human perception, and it is hoped that some of the data presented in this volume will help readers remember that there is great diversity among the members of each gender. Further, it can be useful to compare the literal magnitude of the variability within a group to the mean difference between groups.

Visual and Verbal Representations of Differences

Graphical distributions of the raw data from any compared groups might provide some modest inoculation against unwarranted stereotyping. As an example of how this might be done, consider the following data from a very interesting, competent, and professional cross-cultural study. Gibbons, Richter, Wiley, and Stiles (1996) collected self-report data from adolescents in four different countries. The respondents rank-ordered the importance of ten factors in the

selection of an opposite-sex ideal person, and the authors reported differences among the four countries in the average rankings of seven of the ten characteristics of an opposite-sex ideal. Among the many intriguing findings from the complete study, it was found that American adolescents ranked "being sexy" as very important and "liking children" as less important, whereas Guatemalan adolescents gave those two characteristics the opposite rankings. Gibbons et al. make a coherent argument that the pattern of data is congruent with the notion of collectivist and individualist cultures.

Professor Judith Gibbons of St. Louis University, expressing interest in the results of an additional examination of the data, has graciously provided access to the original data set. The new analysis was very simple; subgroups of the whole data set that might be interestingly compared were identified, and histograms that show the frequency of selection of each rank for two of the ten items in the Gibbons et al. (1996) set of characteristics were produced. Some of these comparisons are of major cultural groups as reported in the original article, but the data set was also divided by gender of the respondent. The graphical representations of the distributions were used to complement the analysis based on differences among means.

Gibbons et al. (1996) reported a significant difference between adolescents in the United States and in Guatemala in the ranked importance of liking children by an opposite-sex ideal person. Figure 1 shows the data represented as a histogram for each country. The difference reported is well represented in the figure, with the modal importance ranking being higher (3rd most important) for Guatemalan adolescents than for U.S. adolescents (10th most important). Visual inspection of the figure also reveals that the range of rankings for both countries was the same; all ten rankings were selected. Knowing an adolescent's home country would not allow one to rule out any of the possible rankings. Within the Guatemalan sample there were only small variations in the percent of participants using each rank, varying from about 5% to 15%. Within the U.S. sample the variation was greater, ranging from around 2% to slightly over 20%. There was, however, substantial overlap in category use; the middle categories of the ranking scale (ranks 3–8) were used by about 60% of the Guatemalan adolescents and by about 55% of the U.S. sample. Knowing only the country of origin for an ado-

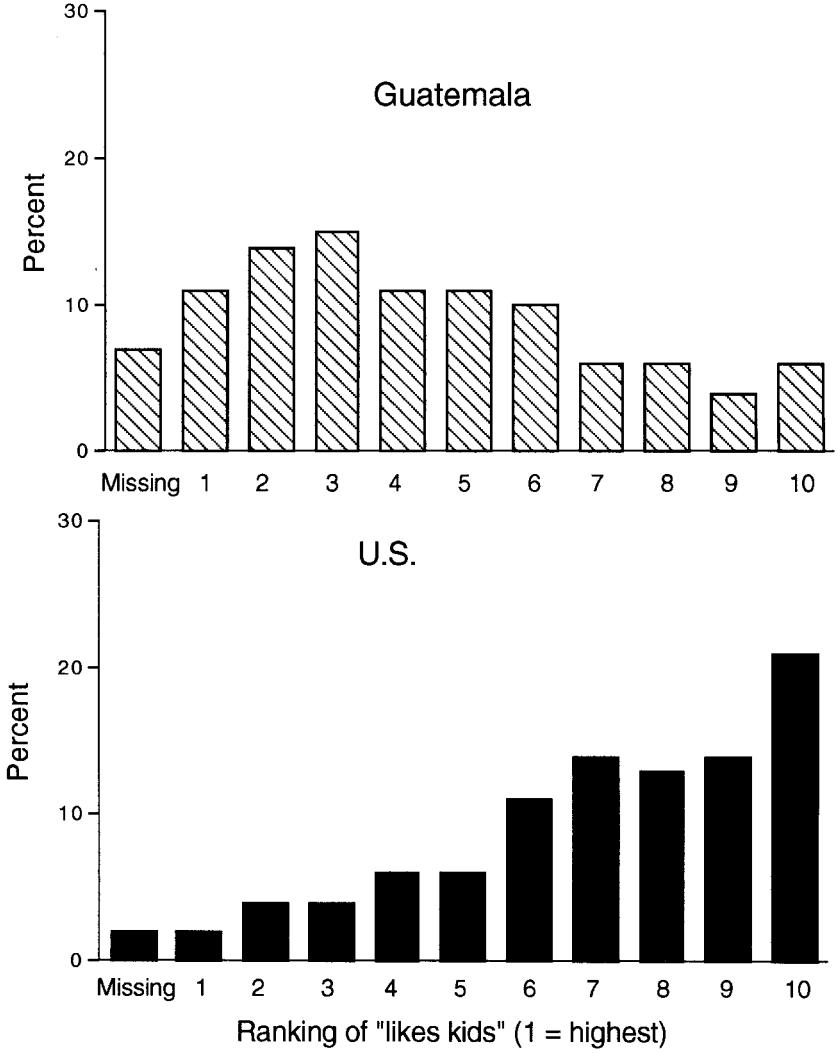

Figure 1. U.S. and Guatemalan adolescent ranking of "likes kids" in selection of opposite-sex ideal person (collapsed across gender).

lescent certainly changes the probability of finding extreme scores on the ranked importance of liking children, but it does not clearly identify a highly likely part of the range of possible ranks.

Similarly, Gibbons et al. (1996) reported a significant difference between adolescents in the United States and in Guatemala in the ranked importance of an opposite-sex ideal person being sexy. Figure 2 shows the data represented as a histogram for each country.

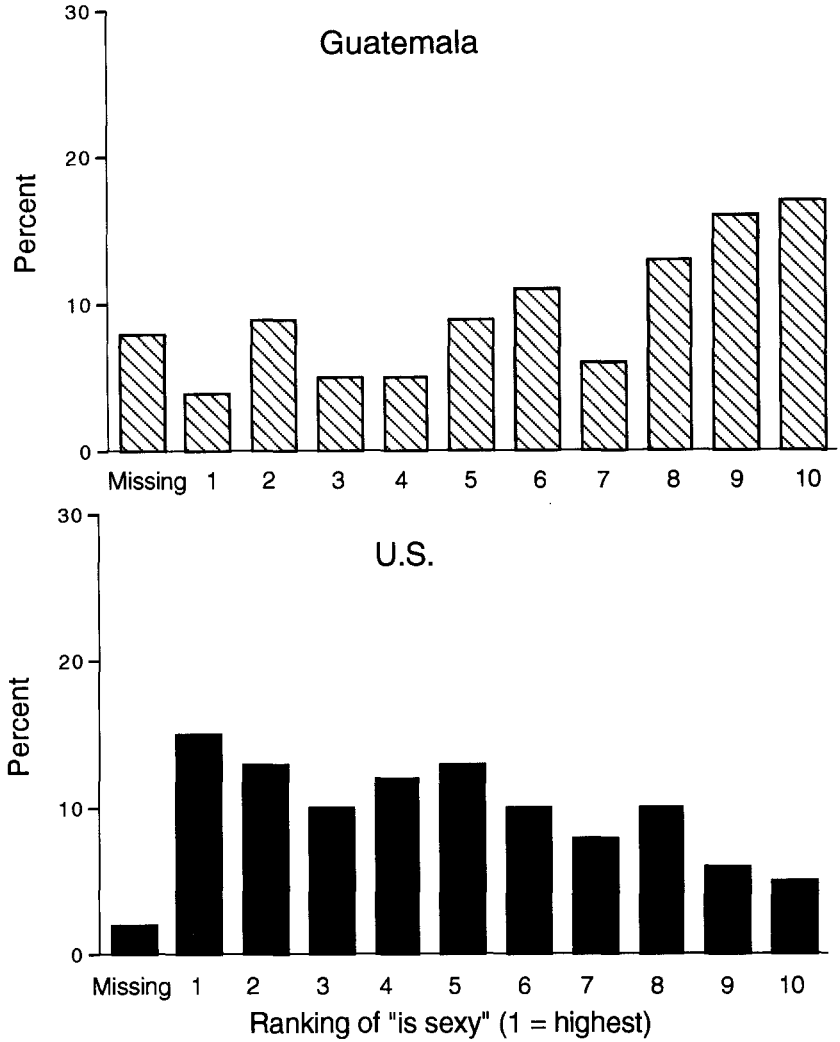

Figure 2. U.S. and Guatemalan adolescent ranking of "is sexy" in selection of opposite-sex ideal person (collapsed across gender).

The difference reported is well represented in the figure; the modal importance ranking of being sexy was higher (single most important) for U.S. adolescents than for Guatemalan adolescents (10th most important). Visual inspection of the figure also reveals that the range of rankings for both countries was the same; all ten rankings were selected. Knowing an adolescent's home country would not al-

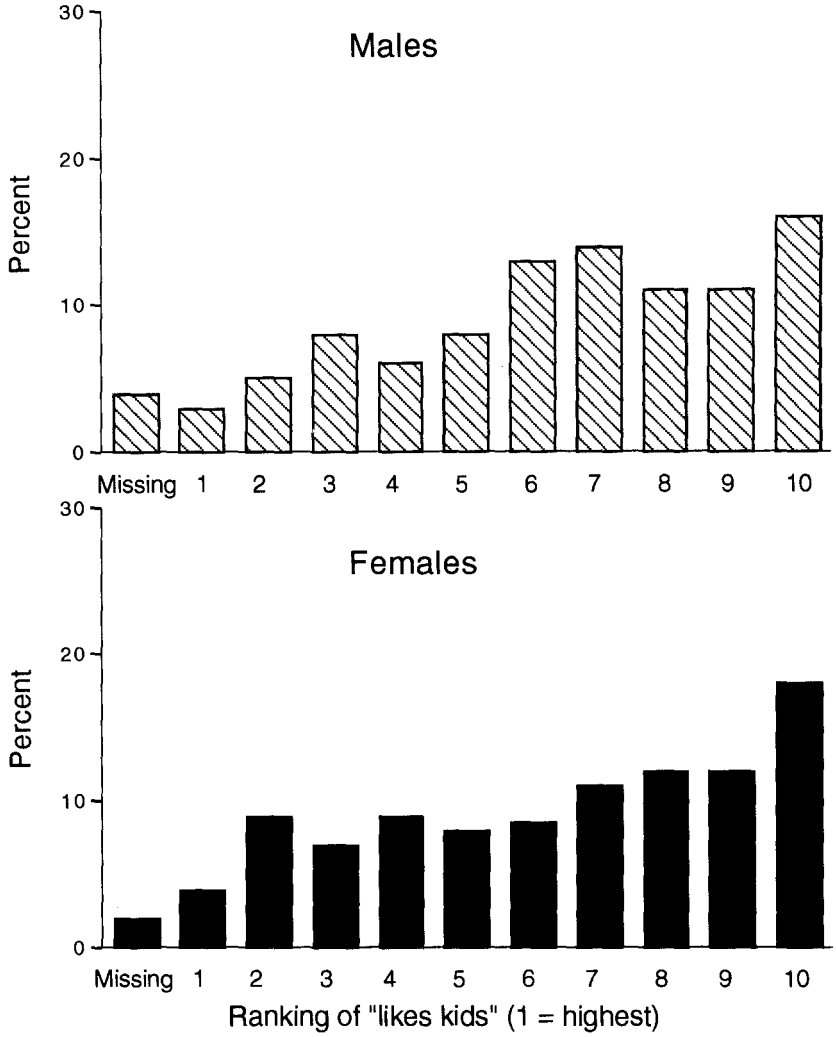

Figure 3. U.S. and Guatemalan adolescent ranking of "likes kids" in selection of opposite-sex ideal person (collapsed across country).

low one to rule out any of the possible rankings. Within the U.S. sample there were only small variations in the percent of participants using each rank, varying from about 5% to 15%. Within the Guatemalan sample the variation was slightly greater, ranging from around 4% to slightly under 20%. Again there was substantial overlap in category use; the middle categories of the ranking scale (ranks

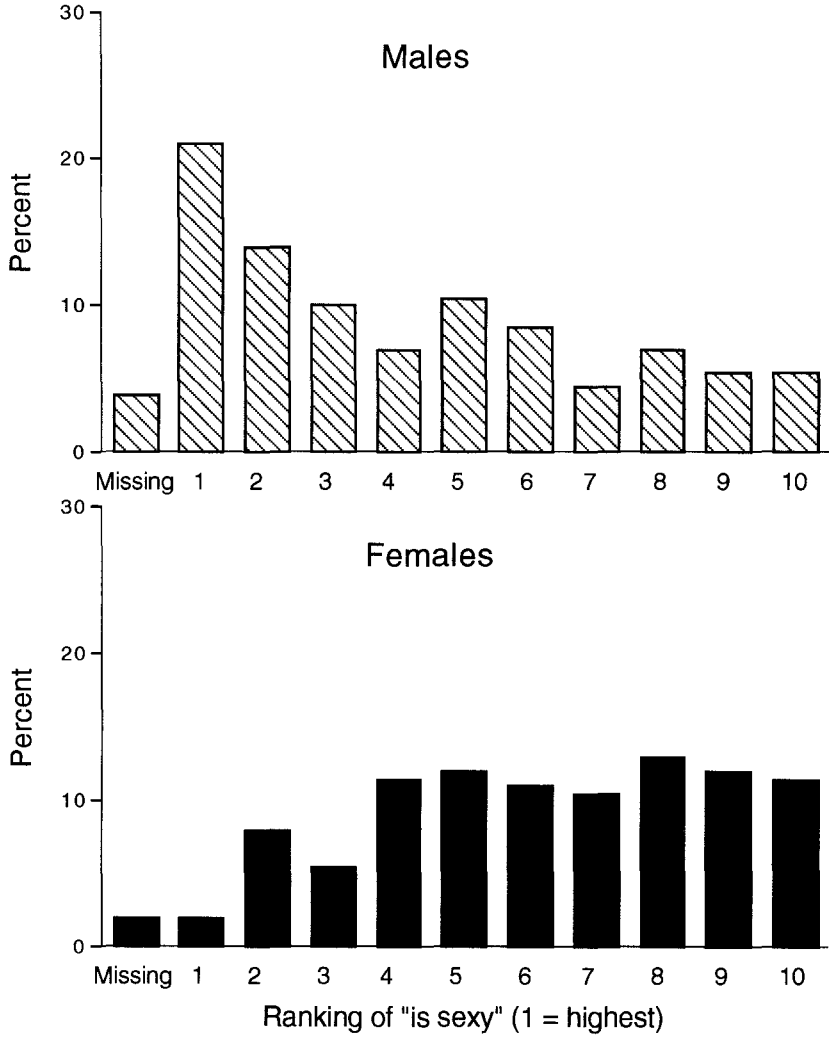

Figure 4. U.S. and Guatemalan adolescent ranking of "is sexy" in selection of opposite-sex ideal person (collapsed across country).

3–8) were used by about 50% of the Guatemalan adolescents and by about 60% of the U.S. sample. Knowing only the country of origin for an adolescent certainly changes the probability of finding extreme scores on the ranked importance of being sexy, but it does not clearly identify a highly likely part of the range of possible ranks.

The same kind of analysis of group effects was conducted by

gender on the data provided by Judith Gibbons; we combined the U.S. and Guatemalan samples and then divided them into male and female participants, collapsed across country. Figure 3 includes two histograms of the rankings of the importance of an opposite-sex ideal person liking children; one shows the rankings by males, and the other shows the rankings by females. Both the modal rank and the general pattern of the rankings are quite similar for both genders. Gibbons et al. (1996) did not test this difference, but these data would likely not show a mean rank difference between the genders. Figure 4 includes two histograms of the rankings of the importance of an opposite-sex ideal being sexy, and there appears to be a difference in the patterns that is comparable to those found with country. For males the modal ranking is the top category, the single most important category, and about 35% of the male participants ranked being sexy as the 1st or 2nd most important characteristic. Among females there was no clear mode, with about 12% selecting all categories from 4th through 10th most important; a little more than 10% of females rated being sexy as the 1st or 2nd most important characteristic.

Given the similarity of pattern in these rankings and the rankings by country analyzed by Gibbons et al. (1996), it is likely that a test of ranks would yield a gender difference on the importance of being sexy. Like the data divided by country of origin, however, there is substantial overlap in the rankings by men and by women. The full range of rankings was used by participants of both genders, and there was substantial overlap in use of the six middle ranks (3–8); roughly 45% of the males ranked being sexy in the middle range, and roughly 65% of females gave this factor ranks in the middle range. As with the analysis by country of origin, knowledge of gender alone does not substantially narrow the range of likely rank of importance of valuing a sexy opposite-sex ideal; there is a probability shift in the distribution, but there is great diversity in both distributions.

As a final analysis, the two variables (gender and country) were combined to examine the ranking of being sexy. Figure 5 shows separate histograms of those rankings for U.S. males, U.S. females, Guatemalan males, and Guatemalan females. An apparent interaction of the two variables would seem to account for the aggregated data. The lowest rankings of the importance of being sexy came

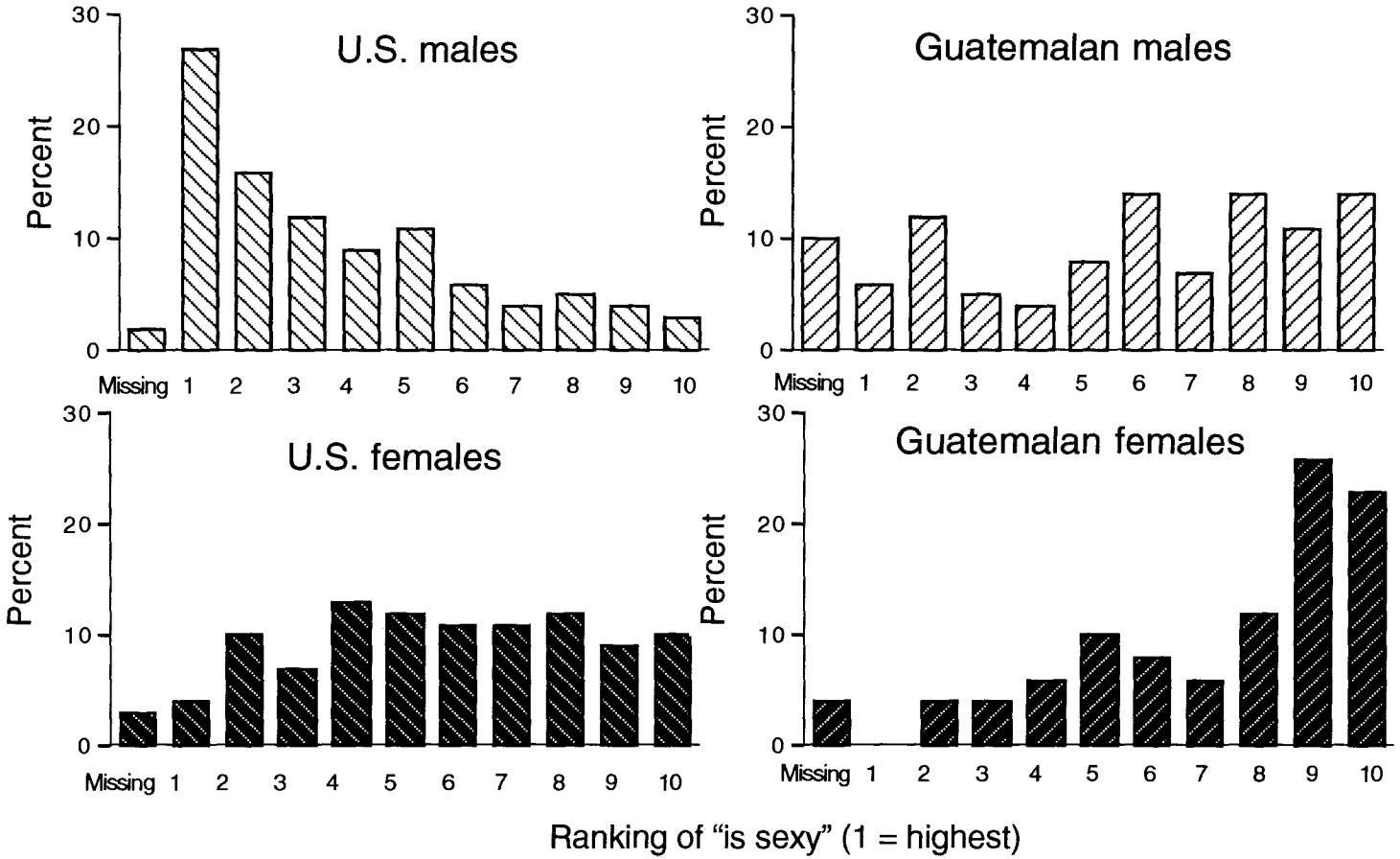

Figure 5. U.S. and Guatemalan adolescent ranking of "is sexy" in selection of opposite-sex ideal person (by gender and by country).

from Guatemalan females, and the highest rankings of importance came from U.S. males. Differences between groups are in fact often quite visible in the comparison of distributions of individual performance. The interaction between gender and country shown in Figure 5 is a good example, and use of this graphical technique may actually strengthen the perception of an important between-group difference.

These separated graphs show the most pronounced variations and yield an important insight into the pattern shown in the aggregate data; yet the figures make clear that there is great diversity in the rankings of the different groups. Three of the four graphs use the full range of ranks (only Guatemalan females do not use the 1st rank), and over the middle range of the ranks there is little systematic variation from an expected flat distribution. Even when a conceptually interesting finding may appear at the group level, there is evidence in the distribution that the group pattern is due to a plurality of individuals whose responses fall in the extremes. The visual representation of the distribution of performance within each group makes it clear that not all individuals show the effect that may be reported as the general finding.

These graphical representations and analyses are neither new nor profound; the suggestion is only that it is useful to present the distribution of data as a complement to any summaries of group differences. This form of data presentation may be useful if it prompts observers to make direct observation of the overlap in the actual performance of the groups being compared. When confronted with the extent of the overlap, including the wide range of scores within each group, readers of the data may be less inclined to create a caricature of a typical group member who is remarkably different from a similarly constructed typical member of the other group(s).

Influencing Readers' Perceptions

One might reasonably ask whether authors who present group data have a responsibility to actively influence how readers construct images of the data. If readers draw inappropriate conclusions from research, whose problem is it? In many cases, responsible authors go to great lengths to warn readers against making the kind of inferences that are described here. For example, the authors of one of the texts

(Smith & Bond, 1993) used in the cross-cultural psychology course described above went to extraordinary lengths to warn readers away from the error of inferring distinct populations. In discussing Hofstede's dimension of individualism and collectivism, they identify the problem specifically: "As applied to the study of cultures, this fallacy would be the mistaken belief that, because two cultures differ, then any two members of those cultures must necessarily also differ in the same manner. For instance, someone might expect that, because America scores higher than Guatemala on individualism, then a particular American is bound to be more independent or individualist than a particular Guatemalan. This is not so" (p. 31). They even provide a diagram of overlapping distributions to highlight the point.

Despite this very clear instruction, however, readers drift into statements that summarize results in ways that ignore the substantial overlap. Perhaps it is in part because the authors report the findings in their field with the shorthand of mean differences found to be statistically significant (presumably for reasons of efficiency and space). When chapter after chapter reports or describes simple differences without showing visually the amount of overlap, the impact of one warning, no matter how clear, is lost. The lesson about a fallacy is overwhelmed by the continuous stream of results presented in language consistent with the notion that differences between groups reflect essential characteristics of the members, rather than slight differences in the modes of very broad distributions.

Another of Woodworth's contributions to psychology was the S-O-R model of learning, in which it is the individual's subjective perception of the world that influences responding. That model presupposes an active person (O for organism) constructing a personal representation of the stimuli (S) perceived, and this model anticipated the contemporary cognitive view of learning a new response (R). Knowing that there are different perceptions of the same stimuli, authors should be interested in any presentation procedures that are more likely to generate subjective representations that are more correct (by the author's standards). I would particularly draw the reader's attention to the visual presentation of data provided by Reed Larson and by Nicki Crick in their chapters on emotion and aggression in this volume; these figures allow the reader to see how the differences reported play out in the context of the diversity of emotionality both between and within genders.

Authors may wish to examine empirically whether a particular visual representation of the performance gives rise to caricatures and stereotypes more or less often than does a numerical representation of central tendency. If there is such an effect, it would also be interesting to ask why it occurs. In principle there could be sufficient information in the numerical data provided to make the same judgment; with means, standard deviations, and measures of skew and kurtosis, an expert judge could draw the same conclusions from either numbers or a graph of the data. Recent work by John Flowers and colleagues (Flowers & Hauer, 1995; Flowers, Buhman, & Turnage, 1997) has demonstrated that observers make very good estimates of correlation from temporally distributed auditory pitch, given sufficient experience with the task. Similar research could examine the tendency of experienced and novice judges to form group caricatures, comparing performance based on visual and numeric representation of the range of individual performance.

Contents of This Volume

The chapters in this book represent analysis by leading scholars whose work is widely appreciated by a variety of audiences. Carol Tavris has done integrative work in many areas of psychology, including gender, health, anger, and critical thinking in higher education. Her textbooks are excellent examples of challenging writing for students, and her writing has earned her both awards and a large readership. Reed Larson does research on adolescence, studying crime, anorexia, mood, leisure, and solitude. His present research on emotionality is especially valuable for its identification of forms of strong emotional experience that are common in men and the ecological approach to observation of emotion in everyday life. Nicki Crick studies aggression that manifests itself in ways other than physical violence, and generates an important account of relational aggression—interpersonal aggression that damages relations rather than body parts. This perspective gives a very different look at the distribution of fundamentally aggressive motivation across genders. Leonore Tiefer is a practicing sexologist whose writing highlights those ways in which gender roles in sexuality are socially constructed. Her perspective is an essential complement to the

generally held view that sexual activity is dominated by biological influences. Diane Gill has studied competition and gender in a wide variety of contexts, including identifying the advantages and disadvantages of being competitive. Her work has often highlighted the intensely competitive behavior of elite female athletes.

The authors invited to present papers at the 1997 Nebraska Symposium on Motivation were identified because their scholarship has demonstrated (at least to some readers) that there is overlap in the distributions of male and female performance. It is not the general position of the authors or the editor that gender is an unimportant variable, nor would most of these authors reject the notion that essential differences in the biological nature of men and women may have contributed to whatever distributions of performance emerge. These authors mainly leave that question unanswered or even unaddressed. There is no shortage of writing available that articulates the utility of a biological account of the origin of gender-related patterns of behavior. At issue in the present volume is exactly what can be learned from the magnitude and nature of the differences that are found.

The claim is not being made that biology is totally irrelevant to human life, particularly in the identification of behavior patterns in gender groups. Instead, these authors report those ways in which there is (and should be) remarkable diversity within each gender group. Most psychologists assume that the behavior of an individual is a joint product of nature and nurture, and that point is not at dispute. Regardless of the particular mix of biological and social influences that generated the behavior, the data presented by these authors argue that knowing a person's gender does not allow one to readily or narrowly identify where an individual will fall on a dimension of sexuality, aggression, competition, or emotion. Put another way, knowing how aggressive, competitive, or emotional a person is will not often allow one to state with any certainty the person's gender. Except at very extreme levels, the distributions show remarkable overlap, and that fact is too often lost in conversations dominated by simple reports of measures of central tendency.

In Appreciation

It should be noted that gender studies is not a special professional interest of mine, and I approached the development of this sympo-

sium as an opportunity to explore some practical ideas about gender by hearing from people who specialize in that topic. I am deeply grateful to Marcela Raffaelli, Deb Hope, Renee Michael, and Lynn Marcus for their excellent conversations and for suggestions of readings and speakers. I am privileged to have such excellent colleagues who encourage me to continue my education with their assistance. I am also happy to have worked closely with Karen Smith and Kris Veit in the development of the visual analysis of group data. It was a great pleasure to work with interesting and creative people who enjoy solving puzzles and exploring data. Finally, I would like to thank Carol Tavris for her contribution to the development of this symposium. The quality of her conversations about psychology in general, and gender studies in particular, is a model to which we should all aspire. While no single edited volume can be a complete account of any important topic, I believe that the chapters included here provide an interesting and thought-provoking introduction to an important perspective on the relation between gender and motivation.

REFERENCES

Flowers, J. H., & Hauer, T. A. (1995). Musical versus visual graphs: Cross-modal equivalence in perception of time series data. *Human Factors, 37,* 553–569.

Flowers, J. H., Buhman, D. C., & Turnage, K. D. (1997). Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. *Human Factors, 39*(3), 341–351.

Gibbons, J. L., Richter, R. R., Wiley, D. C., & Stiles, D. A. (1996). Adolescents' opposite-sex ideal in four countries. *Journal of Social Psychology, 136*(4), 531–537.

Smith, P. B., & Bond, M. H. (1993). *Social psychology across cultures: Analysis and perspectives.* New York: Harvester Wheatsheaf.

Woodworth, R. S. (1910). Racial differences in mental traits. *Science, 31*(788), 171–186.