
10 The Unbearable Accuracy of Stereotypes

Lee Jussim, Thomas R. Cain, Jarret T. Crawford
Rutgers University–New Brunswick

Kent Harber
Rutgers University–Newark

Florette Cohen
Rutgers University–New Brunswick

Sixty years of empirical research has taught us much about stereotypes. Stereotypes can arise from, and sustain, intergroup hostility. They are sometimes linked to prejudices based on race, religion, gender, sexual orientation, nationality, and just about any other social category. They can serve to maintain and justify hegemonic and exploitative hierarchies of power and status. They can corrupt interpersonal relations, warp public policy, and play a role in the worst social abuses, such as mass murder and genocide. For all these reasons, social scientists—and especially social psychologists—have understandably approached stereotypes as a kind of social toxin.

Perhaps equally understandable, but scientifically untenable, is the corresponding belief that because stereotypes contribute to these many malignant outcomes, that they must also be—in the main—inaccurate. The tacit equation is, if stereotypes are associated with social wrongs, they must be factually wrong. However, the accuracy of stereotypes is an empirical question, not an ideological one. For those of us who care deeply about stereotypes, prejudice, and social harmony, getting to the truth of these collective cognitions should guide inquiry about them.

Unfortunately, this has not always been our experience. Because of his inquiries into stereotype accuracy, the first author has been accused by prominent social psychologists of purveying “nonsense,” of living “in a world where stereotypes are all accurate and no one ever relies on them anyway,” of calling for research with titles like “Are Jews really cheap?” and “Are Blacks really lazy?,” of disagreeing with civil rights laws, and of providing intellectual cover for bigots.¹

These reactions are understandable, if one remembers that social psychology has a long intellectual history of emphasizing the role of error and bias in social perception, and nowhere has this emphasis been stronger than in the area of stereotypes. To enter this zeitgeist and to argue for the need to take seriously the possibility that sometimes, some aspects of some stereotypes may have some degree of accuracy, therefore, is to risk making claims that are unbearable to some social scientists. However, science is about validity, not “bearability.” It is about logic and evidence.

In this chapter we review conceptual issues and empirical evidence regarding the accuracy of stereotypes. By doing so we hope to correct some long-held beliefs about stereotypes, and to thereby remove some of the obstacles to the systematic investigation of stereotype accuracy and inaccuracy. The chapter has three main objectives: providing a logically coherent, defensible, and practical definition of “stereotype”; reviewing empirical research on stereotype accuracy; and considering the role of stereotypes in increasing or reducing accuracy in person perception.

ARE STEREOTYPES INACCURATE BY DEFINITION?

Given the frequency with which stereotypes are assumed to be inaccurate, both in popular culture and the social scientific literature, the first order of business is defining stereotype. The accuracy issue becomes “settled” if stereotypes are defined as inaccurate. In this section we explain why a more agnostic approach is needed.

We begin by seriously considering the consequences of defining stereotypes as inaccurate, as have so many researchers before us. When researchers define stereotypes as inherently inaccurate, or assume that stereotypes are inaccurate, there are only two logical possibilities regarding what they might mean: (a) All beliefs about groups are stereotypes and all are inaccurate; or (b) not all beliefs about groups are stereotypes, but stereotypes are the subset of all beliefs about groups that are inaccurate. We next consider the implications of each of these possibilities.

ALL BELIEFS ABOUT GROUPS CANNOT POSSIBLY BE INACCURATE

No social scientist has ever explicitly claimed that all beliefs about all groups are inaccurate. Thus, toppling the assertion that all stereotypes are inaccurate might appear to be refuting a straw assertion. Unfortunately, however, this straw assertion, even if it is merely an implicit rather than explicit assertion, appears to have an ardent scientific following. For decades, stereotypes were predominantly defined as inaccurate, with virtually no evidence demonstrating inaccuracy (e.g., Allport, 1954/1979; Aronson, 1999; Campbell, 1967; Schultz & Oskamp, 2000; see reviews by Ashmore & Del Boca, 1981; Brigham, 1971). Furthermore, among those who define stereotypes as inaccurate, statements regarding what sort of beliefs about groups are accurate (and, therefore, not stereotypes), almost never appear (for concrete examples, see, e.g., Aronson, 1999; Campbell, 1967; Devine, 1995; Jones, 1986; Schultz & Oskamp, 2000; Allport, 1954/1979, remains a lone exception). Accurate beliefs about groups, therefore, appeared to be an empty set.

Furthermore, in their empirical studies, the social sciences have considered people’s beliefs about almost any attribute (personality, behavior, attitudes, criminality, competence, demographics) regarding almost any type of group (in addition to race, sex, class, occupation, dorm residence, sorority membership, college major, and many more) to be a stereotype (see, e.g., reviews by APA, 1991; Fiske & Neuberg, 1990; Kunda & Thagard, 1996; or the meta-analyses reviewed in Jussim, Harber, Crawford, Cain, & Cohen, 2005). It seems, then, that for all practical purposes, the social sciences consider any and all claims and beliefs about groups to be stereotypes.

Putting these points together: Stereotypes are defined (by some) as inaccurate. All beliefs about groups are stereotypes. Therefore, regardless of whether any researcher has explicitly made this claim, any perspective assuming that all beliefs about groups are stereotypes, and defining stereotypes as inaccurate, is logically compelled to reach the conclusion that all beliefs about groups are inaccurate.

This conclusion is untenable on purely logical grounds. It would mean that (a) believing that two groups differ is inaccurate and (b) believing two groups do not differ is inaccurate. Both these conditions are not simultaneously possible, and logical coherence is a minimum condition for considering a belief to be scientific. On logical grounds, alone, therefore, we can reject any claim stating or implying that all beliefs about groups are inaccurate.

Many researchers do not hold such an extreme view. Next, therefore, we consider the less extreme position. The only logical alternative (if one defines stereotypes as inaccurate) to claiming that all beliefs about groups are inaccurate is the following: Not all beliefs about groups are inaccurate, but stereotypes are the subset of beliefs about groups that are inaccurate.

IF STEREOTYPES ARE THE SUBSET OF BELIEFS ABOUT GROUPS THAT ARE INACCURATE, THERE IS NO “STEREOTYPE” RESEARCH

If all stereotypes are inaccurate by definition, then only inaccurate beliefs about groups can be considered stereotypes. *Accurate* beliefs about groups, then, must constitute a different phenomenon altogether. This is not a logical problem as long as those who subscribe to this view stick to their definition and live with its implications.

IF STEREOTYPES ARE DEFINED AS INACCURATE BELIEFS ABOUT GROUPS THEN ONLY EMPIRICALLY INVALIDATED BELIEFS CONSTITUTE STEREOTYPES

Accurate beliefs about groups are not stereotypes. Beliefs about groups of unknown validity cannot be known to be stereotypes. This perspective has a major drawback: It invalidates nearly all existing research on “stereotypes.” This is because so little of the stereotype research has assessed the accuracy of the beliefs under investigation. Without such an assessment, beliefs cannot be known to be stereotypes. No research on “stereotypes” has ever been framed as follows:

Is this belief about that group a stereotype? We are going to figure out whether this belief about that group is a stereotype by assessing whether that belief is inaccurate. If this belief is inaccurate, we will conclude that it is a stereotype. If this belief accurately described that group, we will conclude that it is not a stereotype.

This, however, is precisely how research *must* be framed before one can know one is studying a stereotype, if stereotypes are the subset of beliefs about groups that are inaccurate. If the accuracy of a specific belief being researched has not been first determined, then it is impossible to know whether that belief is a stereotype. The nature, causes, and effects of beliefs of unknown accuracy cannot contribute to knowledge of stereotypes if only inaccurate beliefs are stereotypes.

Holding social psychology to this restrictive definition would mean discarding decades of research purportedly addressing stereotypes. Why? Because almost none of it has empirically established that the beliefs about groups being studied are in fact erroneous. There would be nothing left—no studies of the role of “stereotypes” in expectancy effects, self-fulfilling prophecies, person perception, subtyping, memory, and so on. Poof! We would have to throw out the baby, the bathwater, the tub, the bathroom, and indeed tear down the entire scientific and empirical house in which all our current understanding of “stereotypes” exists.

In the future, those researchers who define stereotypes as inaccurate, or even emphasize their inaccuracy, must provide clear answers to each of the following definitional questions: Do they consider all beliefs about groups to be stereotypes? Do they define all beliefs about groups as inaccurate? Or do they define stereotypes as the subset of beliefs about groups that are inaccurate? If the latter, how do they distinguish between accurate beliefs about groups that are not stereotypes and inaccurate beliefs about groups that are stereotypes?

A NEUTRAL DEFINITION OF STEREOTYPE

Fortunately, many modern definitions of stereotypes do not define stereotypes as inherently inaccurate, and are instead agnostic in terms of stereotype accuracy. One of the simplest of these definitions, and the one we use throughout this chapter, was provided by Ashmore and Del Boca (1981), who stated that “a stereotype is a set of beliefs about the personal attributes of a social group” (p. 21). Stereotypes, as defined by Ashmore and Del Boca, may or may not be accurate and rational, widely shared, conscious, rigid, exaggerations of group differences, positive or negative, or based on essentialist or biological rationales. Stereotypes may or may not be the cause or the effect of prejudice, or the cause of biases and self-fulfilling prophecies.

It is good that Ashmore and Del Boca's definition does not specify these things—that leaves these aspects of stereotypes open to the kinds of empirical investigation this topic deserves. One of the great values of truly believing in the neutral definition is that it does not presume that any time a person holds or uses a stereotype, something inherently bad (or good) is happening. Instead, it opens the door for understanding when stereotypes wreak damage, when they simply reflect social reality, and, possibly, when they actually perform a social good.

Our rejection of defining stereotypes as inaccurate is not equivalent to defining them as accurate. Accuracy is an empirical issue, which naturally raises a question: How (in)accurate are people's beliefs about groups?

ARE STEREOTYPES EMPIRICALLY INACCURATE?

This section reviews empirical investigations of stereotype accuracy. It includes a discussion of an important level of analysis issue with respect to understanding stereotype (in)accuracy, a brief review of common methods for assessing stereotype accuracy (and their limitations), and a discussion of the complexities and richness involved in assessing accuracy. After presenting an overview of those conceptual issues, this section then reviews the research that has assessed the accuracy of people's stereotypes.

STEREOTYPE ACCURACY AND LEVELS OF ANALYSIS

The following statement summarizes a class of criticisms of stereotype accuracy that has periodically appeared in the social psychological literature (e.g., APA, 1991; Fiske, 1998; Nelson, 2002; Schneider, 2004; Stangor, 1995):

Even if it can be successfully shown that perceivers accurately judge two groups to differ on some attribute: (a) Perceivers should not assume that their stereotypes of the group automatically fit all members of the group; (b) perceivers cannot apply their beliefs about the group when judging individuals; and (c) if perceivers do apply their beliefs about the group when judging individuals, they are likely to be wrong much of the time because few members perfectly fit the stereotype.

According to this type of analysis, all stereotypes are already known to be largely inaccurate so there is no need to assess their accuracy.

There is merit to these points. Few, if any, members of a group are fully defined by stereotypes. Assessments of any individual based solely on stereotypes will generally be lacking. However, this logic implies nothing about stereotype accuracy. Instead, it is a claim about the accuracy of applying stereotypes of groups to specific group members.

Stereotype accuracy issues occur, therefore, at two different levels of analysis, each captured by a different question. First, how accurate are people's beliefs about groups? Just as a person might not accurately remember how many games Roger Clemens won in 2000 (inaccuracy in person perception) and still remember that the Yankees won the World Series that year (accurate belief about Clemens's group), inappropriate application of a stereotype does not mean that the stereotype is itself inaccurate. A person may correctly know that, on average, women earn about 70% of what men earn, but have no accurate knowledge whatsoever about how much Nancy earns.

Second, does people's use or disuse of stereotypes in judging individuals increase or reduce the accuracy with which they perceive differences between small groups of individuals with whom they have personally come into contact? This is the accuracy version of the "stereotypes and person perception" question. Do, for example, general stereotypes of male superiority in athletics lead the coach of a soccer team to erroneously view the particular boys on the team as better than the particular girls on the team, when they really have equal skill?

SOME PRELIMINARY CAVEATS

We next briefly summarize some key points made in previous research on accuracy that we need to draw on here, although space considerations preclude an extended discussion. First, stereotypes undoubtedly sometimes lead to errors, biases, self-fulfilling prophecies, and a variety of unfair and unjustified outcomes. The research on these topics, however, typically has provided little information about their accuracy (Funder, 1987, 1995; Jussim, 1991, 2005).

Second, methodological difficulties once plagued accuracy research. Those difficulties, however, have been resolved by statistical, methodological, and conceptual advances within the field of accuracy research over the last 20 years (e.g., Funder, 1987, 1995; Jussim, 1991, 2005; Kenny, 1994; Ryan, 2002). Accuracy is now a thriving area of research within social psychology.

Third, this chapter does not address the issue of how group differences originate. Why groups differ is a fundamentally different scientific question than whether people perceive those differences accurately. Whether group differences result from genetics, childhood environment, socialization, culture, or roles is beyond the scope of this chapter.

Fourth, the genesis of stereotypes is also irrelevant with respect to evaluating their accuracy. A belief's accuracy must be assessed on its merits, not on its sources. Assessing accuracy of beliefs is a different endeavor than assessing processes leading to those beliefs (Jussim, 2005).

DIFFERENT ASPECTS OF STEREOTYPE (IN)ACCURACY

TYPES OF STEREOTYPE ACCURACY

Stereotype accuracy has been commonly assessed in either of two ways in the scientific literature. *Discrepancy scores* assess how close to perfection people's beliefs come. The stereotype belief (e.g., "how tall [rich, aggressive, etc.] is the average woman in the United States") is compared to criteria (e.g., the average height [wealth, aggressiveness, etc.] of the average woman). The difference indicates how far people are from perfection. Smaller discrepancies equal greater accuracy.

Research on stereotype accuracy has also frequently used *correlations* to assess how well people's beliefs about groups correspond to what those groups are like. Stereotype beliefs are correlated with criteria (e.g., people's ratings of women's average height, wealth, and aggressiveness, could be correlated with criteria for women's height, wealth, and aggressiveness). Higher correlations indicate greater correspondence of the stereotype with criteria—that is, higher accuracy.

Discrepancy scores and correlations have been used to assess two types of stereotypes: cultural and personal stereotypes. *Cultural stereotypes* refer to the extent to which a stereotype is shared by the members of a culture, or a particular sample, and are usually assessed by sample means (e.g., the mean belief about women's height in a sample is the best estimate of the cultural stereotype for women's height for the group sampled). *Personal stereotypes* are simply any individual's beliefs about a group, regardless of whether that belief is shared by others.

WHAT IS A REASONABLE STANDARD FOR CHARACTERIZING A STEREOTYPIC BELIEF AS "ACCURATE"?

There is no objective gold standard with which to answer this question. Perfect or near perfect accuracy is reserved for a very select set of endeavors (e.g., moon landings, measuring atomic weights, etc.). Even when social scientists generate hypotheses that predict differences on some outcome between groups (whether experimental or demographic), they do not require a correlation of 1.0 between group and outcome to consider their hypotheses confirmed. Indeed, social psychologists are often quite satisfied with correlations of .2 or less (Richard, Bond, & Stokes-Zoota, 2003). So, the issue is, what is a reasonable standard for lay accuracy in stereotypes? Because there are two

broad types of accuracy, discrepancy from perfection and correspondence with real differences, there must be two separate standards. Each is discussed next.

DISCREPANCIES

The Bull's-Eye

A bull's-eye is as good as it gets in target practice. Bull's-eyes are not microscopic geometric points. They usually have perceptible width, which means one can legitimately hit a bull's-eye without being Robin Hood, who could hit the target dead center, then split his own arrow on the next shot. Our standard is that, for the type of social perceptual phenomena usually studied by social psychologists, a bull's-eye is within 10% of dead center. There is nothing magic about 10%. Reasonable people may disagree about this standard, and it might not be always appropriate, but when judging proportions and probabilities, as is common in the study of stereotype accuracy, getting within 10% is doing pretty well.

Some studies, however, do not report their results as percentages. Most that do not, however, report their results as effect sizes or can be readily translated into effect sizes—real and perceived differences between groups in standard deviation units, which can be translated into percentages. If Kay perceives Group A as $.25 SD$ higher on some attribute than Group B, this means that Kay perceives the average person in Group A to score higher on that variable than 60% of the people in Group B. Bingo! That is a 10% difference (we assume a normal distribution, so the average person in Group B scores at the 50th percentile, so the difference is 10%). Therefore, for studies assessing stereotype accuracy using effect sizes, we characterize a perceived difference as accurate if it is within $.25 SD$ of the real difference.

Our standards often do not correspond to those used by the original authors. McCauley's research (see Tables 10.1 and 10.2) often used "less than 10% off" as his criteria for accuracy; we differ by a single percent, because we characterize 10% off as accurate. Others used statistical significance as their standard (e.g., if the perceived difference statistically exceeded or underestimated the real difference, they concluded it was not accurate). Although these standards have their own advantages and disadvantages, discussing them is beyond the scope of this chapter.

Near Misses

Accuracy is a matter of degree—it is not all or none (Jussim, 2005). Therefore, it does not seem reasonable to characterize a belief that is 10% off as "accurate" and one that is 10.1% off as "inaccurate." So how should we characterize near misses? As near misses. A near miss is not accurate, but it is not too far off. Continuing with the archery metaphor, one can still rack up some points if one hits the target, even if one does not hit the bull's-eye; not as many points as when one hits the bull's-eye, but more than if one misses the target completely.

What, then, is a reasonable standard for a near miss? We use more than 10% off, but no more than 20% off. Within 20% is certainly not a bull's-eye, but it is not completely out of touch with reality, either. It is certainly far more accurate, say, than being 40% off or more. Here, too, people can disagree about what reasonably constitutes a near miss.

Following the same rationale as for accuracy, when results are only reported in standard deviations, we use "more than $.25 SD$ discrepancy (between belief and criterion) but no more than $.50 SD$ discrepancy" as our standard for near misses. If Tom's belief is that Group A's mean is $.5 SD$ higher than Group B's mean, when there really is no difference, he erroneously believes that the mean of Group A group exceeds the scores of about 70% of the members of Group B, when, in fact, it only exceeds the scores of 50% of Group B; again, this is a 20% difference.²

Types of Discrepancies

The literature has focused on two broad types of discrepancies. By far, the most interesting and important discrepancy involves perceiving differences between groups. Do people perceive a larger or smaller difference between groups than really exists, or do they perceive the difference accurately? These types of discrepancies directly test the exaggeration hypothesis that has been so long emphasized in the scholarly literature on stereotypes. It is also important for practical reasons. These discrepancies, when they show that people exaggerate real differences on socially desirable attributes, indicate whether people unjustifiably perceive one group as “better” than another (more intelligent, more athletic, etc.). When they show that people underestimate real differences on socially desirable attributes, they indicate that people unjustifiably see groups as more similar to one another than they really are.

There is a second type of discrepancy reported in the literature that is still relevant as “inaccuracy,” but has considerably less theoretical or practical importance with respect to stereotypes. Independent of perceiving how two (or more) groups mutually differ on a given attribute (e.g., height), sometimes people have a general tendency to overestimate or underestimate the level of some attribute for all groups. For example, let’s say men and women in the United States average 72 and 66 in. in height, respectively. Fred, however, believes that men and women average 74 and 68 in., respectively. He consistently overestimates height by 2 in. (this is a fairly meaningless “elevation” effect; see, e.g., Judd & Park, 1993; Jussim, 2005), but he does not exaggerate sex differences in height.

CORRESPONDENCE WITH REAL DIFFERENCES: HIGH ACCURACY

How much correspondence should be considered “accurate”? Again, this is a judgment call. Nonetheless, we advocate holding people to a high standard—the same standards to which social scientists hold themselves.

J. Cohen (1988), in his classic statistical treatise imploring social scientists to examine the size of the effects they obtained in their studies and not just the “statistical significance” of the results, suggested that effect sizes above .8 could be considered “large.” Such an effect size translates into a correlation of about .4 (in the remainder of this article, effect sizes are discussed exclusively as correlations). By this standard, correlations of .4 and higher could be considered accurate because they represent a “large” correspondence between stereotype and reality.

This standard has been supported by two recent studies that have examined the typical effect sizes found in clinical and social psychological research. One recent review of more than 300 meta-analyses—which included more than 25,000 studies and over 8 million human participants—found that mean and median effect sizes in social psychological research were both about .2 (Richard et al., 2003). Only 24% of social psychological effects exceeded .3. A similar pattern has been found for the phenomena studied by clinical psychologists (Hemphill, 2003). Psychological research rarely obtains effect sizes exceeding correlations of .3. Effect sizes of .4 and higher, therefore, constitute a strong standard for accuracy. Last, according to Rosenthal’s (1991) binomial effect size display, a correlation of at least .4 roughly translates into people being right at least 70% of the time. This means they are right more than twice as often as they are wrong. That seems like an appropriate cutoff for considering a stereotype reasonably accurate.

CORRESPONDENCE WITH REAL DIFFERENCES: MODERATE ACCURACY

Moderate correspondence, of course, is less than high correspondence. It reflects a mix of accuracy and inaccuracy. Following the same standards as science (J. Cohen, 1988; Richard et al., 2003), we characterize correlations between people’s beliefs and reality ranging from .25 to .40 as moderately accurate. Such correlations do not reflect perfect accuracy, nor do they reflect complete inaccuracy.

TABLE 10.1
Are Racial/Ethnic Stereotypes Inaccurate?

Study and Stereotype	Perceivers and Criteria	Personal Stereotype Accuracy Discrepancies	Consensual Stereotype Accuracy Discrepancies	Personal Stereotype Accuracy Correlations	Consensual Stereotype Accuracy Correlations
Ryan (1996); beliefs about differences in the personal characteristics (achievement, personality, athletic, intelligence, etc.) of African American and White University of Colorado students	Perceivers: Random samples of 50 African American and 50 White University of Colorado students Criteria: Self-reports from the same samples	Not available	African-Americans beliefs about: African-Americans: 5, 5, 7 White Americans: 5, 5, 7 Differences: 5, 5, 7 (errors tended to exaggerate real differences) White Americans' beliefs about: African-Americans: 5, 5, 7 White Americans: 5, 2, 10 Differences: 9, 4, 4 (no consistent pattern of exaggeration or underestimation)	African American perceivers: .42 ^a White perceivers: .36 ^a	African American perceivers: .73, .53, .77 ^{b,c} White perceivers: .77, .68, .72 ^{b,c}
McCauley and Stitt (1978); beliefs about demographic differences (high school degree, on welfare, etc.) between African Americans and other Americans	Perceivers: Five haphazard samples (church choir, union members, students, etc.), <i>N</i> = 62 Criteria: U.S. Census data	Not available	White American targets: 17, 13, 5 African-American targets: 17, 14, 4 Differences between African-Americans and Other Americans: 27, 8, 0 (all discrepancies underestimated the real differences)	Not available	Beliefs about: ^{a,b} African-Americans: .60 Americans: .93 Differences between African Americans and other Americans: .88
Ashton and Esses (1999); beliefs about the achievement of nine Canadian ethnic groups	Perceivers: 94 University of Western Ontario students Criteria: Toronto Board of Education data	Accuracy: 36 of 94 Exaggeration: 33 of 94 Underestimation: 25 of 94 ^d	Not Available	.69	Not available
Wolsko, Park, Judd, and Wittenbrink (2000) ; beliefs about differences between African Americans and White Americans on a variety of attributes (work ethic, intelligence, criminality, etc.)	83 White University of Colorado undergraduates Criteria: U.S. Census data and other reports	Not available	Accuracy in judging stereotypical attributes and near misses in judging counterstereotypical attributes. Perceivers overestimated counterstereotypical attributes more than stereotypical attributes, which means they underestimated how stereotypical each group was. They also underestimated race differences, but by less than 10%, i.e., they were accurate	Not available	Not available

Note: Bold numbers represent accurate judgments, italicized numbers represent inaccuracies, and regular typeface numbers represent near misses. Accuracy means within 10% of the real percentage or within .25 SD. For discrepancy scores that were near misses and inaccuracies, parenthetical statements indicate the number of errors that exaggerated real differences, underestimated real differences, or were in the completely wrong direction (reversals). Not available means both that it was not reported and that we were unable to compute it from the data that were reported. See text for explanation of personal and consensual stereotype accuracy discrepancies and correlations. Ryan's (1996) results refer to her stereotypicality results, not her dispersion results.

- ^a For simplicity, if the study reported more than one individual level (average) correlation, we simply averaged all their correlations together to give an overall sense of the degree of accuracy.
- ^b These correlations do not appear in the original article, but are computable from data that were reported.
- ^c For each group of perceivers, the first correlation is the correspondence between their judgments and the self-reports of their own groups; the second correlation is the correspondence between their judgments and the self-reports of the other group; and the third correlation is the correspondence between the perceived difference between the groups and the difference in the self-reports of the two groups.
- ^d Ashton and Esses (1999) computed a personal discrepancy score for each perceiver, and then reported the number of perceivers who were within .2 SD of the criteria, the number that exaggerated real differences (saw a difference greater than .2 SD larger than the real difference) or underestimated real differences (saw a difference more than .2 SD smaller than the real difference). Ashton and Esses (1999) examined beliefs about nine different Canadian ethnic groups; and discrepancy results refer to the number of participants showing each pattern. The results reported here refer to their Table 2, which reports the number of perceivers within .2 SD of the real difference. They did not report results from which the number of perceivers within .25 SD of the real difference could be identified.

TABLE 10.2
Are Gender Stereotypes Inaccurate?

Study and Stereotype	Perceivers and Criteria	Personal Stereotype Accuracy Discrepancies		Personal Stereotype Accuracy Correlations	
		Personal Stereotype Accuracy Discrepancies	Consensual Stereotype Accuracy Discrepancies	Personal Stereotype Accuracy Correlations	Consensual Stereotype Accuracy Correlations
McCauley, Thangavelu, and Rozin (1988); McCauley and Thangavelu (1991); beliefs about the sex distribution into different occupations ^a	Perceivers: College students, high school students, rail commuters (N = 521 over the five studies) Criteria: U.S. Census	Not available	56, 28 (26 underestimates), 6 (all underestimations)	Not Available	.94-.98 ^b across five studies
Swim (1994) ^c ; beliefs about sex differences on 17 characteristics	Perceivers: Introductory psychology students (N = 293 over two studies) Criteria: Meta-analyses of sex differences	Not available	18, 7 (6 underestimations, 1 reversal), 7 (5 exaggerations, 2 underestimations)	Not Available	Study 1: .78 Study 2: .79
Briton and Hall (1995); beliefs about sex differences in nonverbal behavior	Perceivers: 441 introductory psychology students Criteria: Meta-analyses of sex differences in nonverbal behavior	Not available	Female perceivers: 9, 0, 8 (5 exaggerations, 2 underestimations, 1 reversal) Male perceivers: 11, 6 (3 underestimations, 1 exaggeration, 2 reversals), 0	Not Available	Female perceivers : .74 Male perceivers : .68
Hall and Carter (1999); beliefs about sex differences on 77 characteristics	Perceivers: 708 introductory psychology students Criteria: Meta-analyses of sex differences	Not available	Not available	.43	.70
Cejka and Eagly (1999); beliefs about the sex distribution into different occupations	Perceivers: 189 introductory psychology students Criteria: U.S. Census	Not available	Male-dominated occupations: Accuracy Female-dominated occupations: Near miss (underestimation) ^d	Not available	.91
Diekmann, Eagly, and Kulesa (2002) beliefs about the attitudes of men and women ^f	Perceivers: 617 college students over three studies Criteria: Actual attitudes, as indicated in the General Social Survey, which is a recurring nationally representative survey of Americans	Near miss (on average, over all attitudes; they did not report results separately for each attitude)	Accuracy ^e	Male targets: .45 ^e Female targets: .54 ^e When judging sex differences: .60	Male targets: .66 ^e Female targets: .77 ^e When judging sex differences: .80

<p>Beyer (1999)^f; beliefs about the sex distribution into different majors and mean grade point average of men and women in those majors</p>	<p>Perceivers: 265 college students</p> <p>Criteria: College records regarding the majors and grade point averages of its attending students</p>	<p>Not Available</p>	<p>Majors Male perceivers: 6, 5 (3 underestimations, 2 reversals), 1 (1 <i>underestimation</i>) Female perceivers: 7, 4 (1 underestimation, 1 exaggeration, 2 reversals), 1 (1 <i>underestimation</i>) Sex differences in grade point average Male perceivers: 4, 5 (3 underestimations, 2 reversals), 3 (2 <i>underestimations</i>, 1 reversal) Female perceivers: 1, 7 (2 underestimations, 1 exaggeration, 4 reversals), 4 (1 <i>underestimation</i>, 3 reversals)</p>	<p>Proportion: Male perceivers: .48 Female perceivers: .52 Grade point average Male targets: .35 Female targets: .34</p>	<p>Proportion: Male perceivers: .80 Female perceivers: .79 Grade point average Male targets: .35 Female targets: .34</p>
---	--	----------------------	---	--	--

Note: Bold numbers represent accurate judgments, italicized numbers represent inaccuracies, and regular typeface numbers represent near misses. Accuracy means within 10% of the real percentage or within .25 SD. For discrepancy scores that were near misses and inaccuracies, parenthetical statements indicate the number of errors that exaggerated real differences, underestimated real differences, or were in the completely wrong direction (reversals). Accuracy data are separated according to perceiver groups or target groups only when these data are reported in the original studies. "Not available," therefore, means both that it was not reported and that we were unable to compute it from the data that were reported. See text for explanation of personal and consensual stereotype accuracy discrepancies and correlations.

- ^a These studies are grouped together because they were so similar.
- ^b These correlations do not appear in the original article, but are computable from data that were reported.
- ^c Swim (1994) sometimes reported more than one meta-analysis as a criterion for a perceived difference. In that case, we simply averaged together the real differences indicated by the meta-analyses to have a single criterion against which to evaluate the accuracy of the perceived difference.
- ^d Cejka and Eagly (1999) examined beliefs about 80 occupations, but did not report results separately for each occupation. Instead, they simply reported overall discrepancies within male-dominated occupations and within female-dominated occupations. There was a slight overall tendency to underestimate the sex difference in distribution into male-dominated occupations (9.3%), which is accurate by our standards, and a somewhat stronger tendency to underestimate the sex difference in distribution into female-dominated occupations (17.9%), which qualifies as a near miss by our standards.
- ^e For Diekmann et al. (2002) we have averaged several correlations together to give an overall assessment of accuracy. Unfortunately, Diekmann et al. did not report their results separately for each attitude. Instead, they reported results averaged over all attitudes. They did, however, report results separately for men and women targets. Therefore, there were a total of six consensual stereotype accuracy discrepancies reported—three studies by men and women targets. All six were underestimates; that is, there was a general tendency to underestimate how much people supported the various positions. However, this tendency was not very large. All six consensual stereotype discrepancy results underestimated support by only 2% to 8%.
- ^f For Beyer (1999), all results are reported separately for men and women perceivers, except the individual correlations for grade point average (GPA). Because there was no significant sex of perceiver difference in these correlations, Beyer reported the results separately for male and female targets. GPA is not given in percentages and Beyer did not report standard deviations. To determine whether a result was accurate, therefore, we used the conservative criteria of within .10 grade points for accuracy, and within .20 grade points for a near miss. There was an overall tendency to inaccurately overestimate students' GPA. This table, therefore, only reports results regarding the accuracy of the perceived differences in GPA (which is most relevant to stereotyping anyway). There are many reversals in Beyer's data because, although the stereotype is, apparently, that men perform better in school than do women, women's GPAs were actually higher in every major, including the masculine ones (math, chemistry, etc.).

Using Rosenthal's (1991) binomial effect size display, a correlation of .3, for example, means that people are right almost two thirds of the time. Now, this also means they are wrong a little over one third of the time, but two out of three isn't bad.

CAVEATS AND CLARIFICATIONS

Systematic Errors

In social science research, "errors" are usually random. In contrast, in stereotypes, even if the stereotype has considerable accuracy, a major source of concern is the potentially nonrandom nature of the errors. In the preceding discussion on moderate accuracy, for example, a perceiver could be right two thirds of the time. There is, however, a big difference if the one third worth of errors are random versus systematically biased against one group. Fortunately, discrepancy score analyses were intentionally designed to assess systematic errors (Judd & Park, 1993; Jussim, 2005), and we report such errors in our review.

We Only Review Stereotype Accuracy Data

Many of the studies reviewed herein addressed many issues other than accuracy. However, we confine our discussion to those aspects that involve the accuracy of stereotypes. Other issues are beyond the scope of this chapter.

Differences in Terms

None of the studies described in this chapter use our exact terminology of personal and consensual stereotypes that can be evaluated using either discrepancies or correspondence with real differences (or both). Often, they simply discuss "stereotypes." Regardless, we do make that distinction and describe their results accordingly, regardless of whether they described their results this way.

Some researchers have distinguished between personal and consensual stereotypes, although they generally use somewhat different terminology than we do. For example, consensual stereotypes are sometimes discussed as "aggregated" results or stereotypes (because they aggregate across all perceivers). Personal stereotypes are sometimes discussed as "individual" stereotypes; and the Judd, Park, and Ryan group uses the phrase "within-subject sensitivity correlations" to refer to what we call personal stereotype correlations.

CRITERIA FOR INCLUSION

To be included here, the empirical studies assessing the accuracy of stereotypes needed to meet two major criteria. First, they had to compare perceivers' beliefs about one or more target groups with measures of what that group was actually like. Studies assessing social cognitive processes, even when those processes are widely presumed to be flawed and invalid, are not included here, because such studies provide no direct information about accuracy (Funder, 1987; Jussim, 2005).

Second, studies needed to use an appropriate target group. Sometimes, researchers have assessed people's beliefs about a group, and used as criteria the characteristics of a haphazard sample of members of the target group (e.g., Allen, 1995; Martin, 1987). These studies have an important disconnect between the stereotype they are assessing and the criteria they use. Consider, for example, a study in which perceivers provide their beliefs regarding men and women in general, and the criterion sample is a convenient but haphazard sample of college students (Allen, 1995). In this case, even if perceivers' stereotypes corresponded perfectly with men and women in general, they might not correspond to the characteristics of this criterion sample, if the criterion sample's characteristics differ from those of men and women in general. Consequently, such studies were not included in our review.

ACCURACY OF ETHNIC AND RACIAL STEREOTYPES

Table 10.1 summarizes the results of all studies assessing the accuracy of racial and ethnic stereotypes that met our criteria for inclusion. We review the most noteworthy of their results here. First, consensual stereotype discrepancies are a mix of accurate and inaccurate beliefs. Nonetheless, most judgments were either accurate or near misses. Only a minority were more than 20% off.

Second, people's consensual stereotype discrepancies for between group differences are consistently more accurate than are their consensual stereotype discrepancies for characteristics within groups. For example, in the Ryan (1996) study, Whites' consensual stereotypes regarding Whites and regarding African Americans each were accurate only 5 of 17 times (10 of 34, total). However, their judgments of differences between Whites and African Americans were accurate 9 times out of 17. A similar pattern occurred in the McCauley and Stitt (1978) study (see Table 10.1).

Third, these results provide little support for the idea that stereotypes typically exaggerate real differences. Exaggeration occurred, but it occurred no more often than did underestimation, with one exception. The only study to assess the accuracy of personal discrepancies found that a plurality of people was accurate, and that there was a slightly greater tendency to exaggerate real differences than to underestimate real differences (Ashton & Esses, 1999, summarized in Table 10.1).

Fourth, the extent to which people's stereotypes corresponded with reality was strikingly high. Consensual stereotype accuracy correlations ranged from .53 to .93. Personal stereotype accuracy correlations were somewhat lower, but still quite high by any standard, ranging from .36 to .69.

ACCURACY OF GENDER STEREOTYPES

Table 10.2 summarizes the results of all studies of gender stereotypes that met our criteria for inclusion. Results are broadly consistent with those for ethnic and racial stereotypes. In most cases, at least a plurality of judgments was accurate, and accurate plus near miss judgments predominate in every study. Inaccuracy constituted a minority of results. Again, some results showed that people exaggerated real differences. There was, however, no support for the hypothesis that stereotypes generally lead people to exaggerate real differences. As with race, underestimations counterbalanced exaggerations.

Again, consensual stereotype accuracy correlations were quite high, ranging from .34 to .98, with most falling between .66 and .80. The results for personal stereotypes were more variable. Once they were inaccurate, with a near-zero correlation with criteria (Beyer, 1999, perceptions of female targets). In general, though, they were at least moderately, and sometimes highly accurate (most correlations ranged from .40–.60; see Table 10.2).

STRENGTHS AND WEAKNESSES OF RESEARCH ON THE ACCURACY OF RACIAL, ETHNIC, AND GENDER STEREOTYPES

Several methodological aspects of these studies are worth noting because they bear on the generalizability of the results. First, although most of the studies only assessed the accuracy of undergraduates' stereotypes, several assessed the accuracy of samples of adults (McCauley & Stitt, 1978; McCauley & Thangavelu, 1991; McCauley, Thangavelu, & Rozin, 1988). Some of the highest levels of accuracy occurred with these adult samples, suggesting that the levels of accuracy obtained do not represent some artifact resulting from the disproportionate study of undergraduate samples. Nonetheless, additional research on the accuracy of noncollege samples is still needed.

Second, the studies used a wide variety of criteria: U.S. Census data, self-reports, Board of Education data, nationally representative surveys, locally representative surveys, U.S. government reports, and so on. The consistency of the results across studies, therefore, does not reflect some artifact resulting from use of any particular criteria.

Third, the studies examined a wide range of stereotypes: beliefs about demographic characteristics (McCauley & Stitt, 1978; Wolsko, Park, Judd, & Wittenbrink, 2000), academic achievement (Ashton & Esses, 1999; McCauley & Stitt, 1978; Wolsko et al., 2000), and personality and behavior (Ryan, 1996; Wolsko et al., 2000). The consistency of the results across studies, therefore, does not reflect some artifact resulting from the study of a particular type of stereotype.

Fourth, personal discrepancies were the least studied of the four types of accuracy. Thus, the studies do not provide much information about the extent to which individual people's stereotypes deviate from perfection.

INACCURATE STEREOTYPES

Despite the impressive and surprising evidence of the accuracy of stereotypes, there is some consistent evidence of inaccuracy in stereotypes. In the United States, political stereotypes tend to have little accuracy (e.g., Judd & Park, 1993). Many people in the United States seem to have little knowledge or understanding of the beliefs, attitudes, and policy positions of Democrats and Republicans.

A recent large-scale study conducted in scores of countries found that there is also little evidence of accuracy in national stereotypes regarding personality (Terracciano et al., 2005). It is probably not surprising that people on different continents have little accurate knowledge about one another's personality (e.g., that Indonesians do not know much about, say, Canadians, is not very surprising). However, somewhat more surprising is that people from cultures with a great deal of contact (various Western European countries; Britain and the United States) also have highly inaccurate beliefs about one another's personality characteristics.

Although the Terracciano et al. (2005) study was impressive in scope and innovative in topic, it suffers from one of the limitations that excluded several studies from this review. Specifically, the criteria samples were haphazard samples of convenience, rather than random samples obtained from target populations. The extent to which this explains their low level of accuracy is unknown until research is conducted on the same topic that obtains criteria from random samples. In general, why some stereotypes have such high levels of accuracy and others such low levels is currently unclear and is an important area of future research.

THE ROLE OF STEREOTYPES IN ENHANCING OR REDUCING THE ACCURACY OF PERSON PERCEPTION

WHAT SHOULD PEOPLE DO TO BE ACCURATE?

On the Use of Inaccurate Versus Accurate Stereotypes in Judging Individuals

Relying on an inaccurate belief to judge an individual will not increase accuracy, as can easily be seen with a nonsocial example. If Armand believes that Anchorage, Alaska is warmer than New York City, and he relies on that belief for making guesses about where it is going to be warmer, today, tomorrow, the next day, and so on, he will be wrong most of the time. Even though he may pick up an occasional hit on the rare days that Anchorage really is warmer than New York, he will be wrong far more often than he is right.

Stereotypes are no different. If Celeste believes that professional (American) football players are unusually tiny, and if she relies on that stereotype to guess their sizes, she will be very wrong. Relying on an inaccurate stereotype to judge an individual decreases the accuracy of that judgment.

But what happens when people rely on a largely accurate stereotype to judge an individual? Given that the prior section demonstrated moderate to high accuracy in many stereotype beliefs, this becomes a pressing question. It turns out that there are some conditions under which relying on an accurate stereotype can increase accuracy in judging an individual, and there are some

conditions under which relying even on an accurate stereotype will not increase accuracy. Those conditions are the major focus of the next sections of this chapter.

Understanding the role of stereotypes in increasing or reducing accuracy involves understanding three different person perception situations. How perceivers should go about being as accurate as possible will be different in each of these three situations (because this discussion involves understanding the role of stereotypes in person perception, in all situations, perceivers know the target person's group membership). In each case, we first present an example involving nonsocial perception, in which the issues may, perhaps, be easier to understand, and which will certainly be less complicated by political correctness concerns.

Definitive Individuating Information

The first situation involves having vividly clear and relevant individuating information about a particular target. We refer to such individuating as “definitive” because it provides a clear, valid, sufficient answer to whatever question one has about a target. For example, when judging academic accomplishments, we might have standardized test scores and class rank and grade point average for a college applicant; when judging sales success, we might have 10 years of sales records for a salesperson; and when judging personality, we might have multiple expert judges' observations of, and well-validated personality test scores for, a particular individual. When we have this quality and quantity of information, how much should we rely on stereotypes?

If one discovers from a credible source (say, the Weather Channel) that it is 80 degrees today in much of Alaska, but only 60 in New York, what should one conclude? The fact that it is usually colder in Alaska is not relevant. Today, it is warmer in Alaska.

Professional basketball players tend to be tall—very tall. It is very rare to find one shorter than 6'4.” It is, therefore, reasonable to expect all basketball players to be very tall.

Once in a while, though, a short player makes it into the National Basketball Association (NBA). Spud Webb was a starting player in the 1990s, and he was about 5'7.” Once one knows his height, should one allow one's stereotype to influence one's judgment of his height? Of course not. His height is his height, and his membership in a generally very tall group—NBA players—is completely irrelevant.

In situations where one has abundant, vividly clear, relevant individuating information about a member of a group, the stereotype—its content, accuracy, and so on—becomes completely irrelevant. One should rely entirely on the individuating information.

Useful but Not Definitive Individuating Information

In many other situations, people may have some useful information, but not the definitive information presented in the first situation. Sometimes information is ambiguous, or limited in quality or degree.

Small Amounts of Information

When we meet a person for the first time, we might have only physical appearance cues (which will usually reveal sex, but which may or may not clue us in on race or ethnicity, attractiveness, neatness, wealth, concern with fashion, etc.). Or, although we might not be following the election for Town Council closely, we just happen to hear on the radio a 10-second sound bite from a candidate in which she claims that property taxes are too high.

Ambiguous Information

Some information is inherently ambiguous—its meaning and interpretation are unclear. Is a shove playful horsing around or assault? Is that a warm, friendly smile, or a superior sneer? Is that extreme compliment flattery or sarcasm? In these cases people have information, but its meaning or interpretation is unclear.

Inferences Versus Observations

Behavior can be observed directly. Most other aspects of psychology—beliefs, attitudes, motivations, personality, intentions, and so on—are not directly observable; they must be inferred on the basis of behavior. Whereas it is possible to definitively know (most of the time) whether David smiled, without lots of other information, it is not so easy to figure out whether David is a “happy” person. Whereas it is relatively easy to grade a student’s test, without lots of other information, it is quite hard to know whether that high test score reflects the student’s brilliance or the simplicity of the test. There is an inherent ambiguity in going from behavior to inferences about underlying attributes.

Predicting the Future Versus Evaluating the Past

The future is inherently ambiguous. It is not possible to know exactly what will happen in the future (history is littered with the inaccurate predictions of the holy, the greedy, the political, and the superstitious). Nonetheless, we must make predictions about the future all the time. Whenever we select people for admission to college, graduate school, or jobs, we are, essentially, making a prediction that the chosen person is the best for that position, or, at minimum, that he or she is likely to succeed reasonably well. Because the future is inherently unknowable, however, we can almost never have enough information to render such predictions definitive. So, with respect to making predictions about the future, nearly all information has some degree of ambiguity.

What Should People Do With Useful but Not Definitive Individuating Information?*Alaska Versus New York*

You get one piece of information about each location. You learn that Jane, a lifelong resident of Anchorage, considers it “cold” today and Jan, a lifelong resident of New York, considers it “cold” today. Note that the “information” that you have is identical regarding the two places. Should you, therefore, predict that they have identical temperatures?

That would be silly. It ignores the wealth of information you already bring to bear on the situation: (a) It is usually much colder in Anchorage; (b) “cold” can mean lots of different things in different contexts; and (c) people usually adapt to their conditions, so, if it is usually 40 degrees in your neighborhood, you would probably judge 20 degrees as cold; but if it is usually 60 degrees in your neighborhood, 40 might be seen as quite cold. To ignore all this would be foolish, and, most of the time, doing so will lead you to an inaccurate conclusion about the weather in the two places.

In other words, in this situation, to the extent that your beliefs about the general characteristics of Alaska, Alaskans, New York, and New Yorkers are reasonably accurate, they should influence your interpretation of “cold” and your prediction regarding the weather in each place.

Stereotypes and Person Perception

The logic here is identical. Consider stereotypes of peace activists and al Qaeda members. You hear the same thing about an individual from each group: They have “attacked” the United States. Should you interpret this to mean that they engaged in identical behaviors? Not likely. The “attack” perpetrated by the peace activist is most likely a verbal “attack” on U.S. war policies; the al Qaeda attack is probably something far more lethal.

The same principles hold regardless of whether the stereotypes involve groups for whom stereotypes are deemed acceptable (e.g., peace activists or al Qaeda) or groups for whom stereotypes are deemed socially unacceptable (e.g., genders, nationalities, races, social classes, religions, ethnicities, etc.). For example, if we learn both Bob and Barb are regarded as “tall,” should we conclude that they are exactly equal in height? Of course not. Undoubtedly, Bob is tall for a man, and Barb is tall for a woman, and, because men are, on average, taller than women, tall means different objective heights for men and women (implicit acceptance of these “shifting standards” has been thoroughly demonstrated; e.g., Biernat, 1995).

What about judgments about more socially charged attributes, such as intelligence, motivation, assertiveness, social skill, hostility, and so on? The same principles apply. If the stereotype is accurate and one only has a small bit of ambiguous information about an individual, using the stereotype as a basis for judging the person will likely enhance accuracy. For the statistically inclined, this is a very basic application of Bayes's theorem (e.g., McCauley, Stitt, & Segal, 1980) and principles of regression (Jussim, 1991). Let's assume for a moment that 30% of motorcycle gang members are arrested for violent behavior at some point in their lives, and 0.3% of ballerinas are arrested for violent behavior at some point in their lives. People who know this are being completely reasonable and rational if, on dark streets or at lonely train stations, they avoid the bikers more than ballerinas, in the absence of much other individuating information about them.

In all of these cases, the stereotype "biases" the subsequent judgments. At least, that is how such influences have nearly always been interpreted in empirical social psychological research on stereotypes (see, e.g., Devine, 1995; Fiske & Neuberg, 1990; Gilbert, 1995; Jones, 1986). It is probably more appropriate, however, to characterize such phenomena as stereotypes "influencing" or "informing" judgments. Such effects mean that people are appropriately using their knowledge about groups to reach as informed a judgment as possible under difficult and information-poor circumstances. If their knowledge is reasonably accurate, relying on the stereotype will usually increase, rather than decrease, the accuracy of those judgments (see also Jussim, 1991, 2005).

No Individuating Information

Alaska and New York

If you are given absolutely no information, and are asked to predict today's high temperature in Anchorage and New York, what should you do? If you know anything about the climate in the two places, you will predict that it will be warmer in New York. Indeed, you should predict this every time you are asked to do so. Would this mean your beliefs about climate are somehow irrationally and rigidly resistant to change? Of course not. All it means is that you recognize that, when two regions systematically differ and you are asked to predict the day's temperature, and are given no other information, it will always be better to guess that the place with the higher average temperature is warmer than the place with the lower average temperature.

Stereotypes and Person Perception

If you are given no information other than race, and you are asked to predict the income of Bill, who is African American, and George, who is White, what should you do? If you know about the average incomes of African Americans and Whites in the United States, you will predict that George is richer. Indeed, you should predict this every time you are asked to make a prediction about the income of an African American and White target about whom you have no other information. Would this mean your beliefs about racial differences in income are somehow irrationally and rigidly resistant to change? Of course not. All it means is that you recognize that, when the average income of two racial groups differs and you are asked to predict the income of an individual from those groups, and are given no other information, it will always be better to guess that the person from the group with the higher average income has more income.

WHAT DO PEOPLE DO WHEN THEY JUDGE INDIVIDUALS?

Process

People *should* primarily use individuating information, when it is available, rather than stereotypes when judging others. Do they? This area of research has been highly controversial, with many researchers emphasizing the power of stereotypes to bias judgments (Devine, 1995; Fiske &

Neuberg, 1990; Fiske & Taylor, 1991; Jones, 1986; Jost & Kruglanski, 2002) and others emphasizing the relatively modest influence of stereotypes and the relatively large role of individuating information (Jussim, Eccles, & Madon, 1996; Kunda & Thagard, 1996).

Fortunately, literally hundreds of studies have now been performed that address this issue, and, even more fortunately, multiple meta-analyses have been performed summarizing their results. Table 10.3 presents the results from meta-analyses of studies assessing stereotype bias in many contexts. It shows that the effects of stereotypes on person judgments, averaged over hundreds of experiments, range from 0 to .25. The simple arithmetic mean of the effect sizes is .10, which is an overestimate, because the meta-analyses with more studies yielded systematically lower effect sizes ($r = -.43$ between effect size and number of studies). The few naturalistic studies of the role of stereotypes in biasing person perception have yielded similarly small effects (e.g., Clarke & Campbell, 1955; Jussim et al., 1996; Madon et al., 1998).

How small is an effect of $r = .10$? It is small according to J. Cohen's (1988) heuristic categorization of effect sizes. It is among the smallest effects found in social psychology (Richard et al., 2003). An overall effect of .10 means that expectancies substantially influence social perceptions about 5% of the time (as per Rosenthal's [1991] binomial effect size display). This means that stereotypes do not influence perceptions 95% of the time.

In general, therefore, based on more than 300 experimental studies and a smaller number of naturalistic studies, stereotypes have only very modest influences on person perception. Of course,

TABLE 10.3
Meta-Analyses of the Role of Stereotypes in Person Perception

Meta-Analysis	Topic/Research Question	Number of Studies	Average Stereotype Effect
Swim, Borgida, Maruyama, and Myers (1989)	Do sex stereotypes bias evaluations of men's and women's work?	119	-.04 ^a
Stangor and McMillan (1992) ^b	Do expectations bias memory?	65	.03
Sweeney and Haney (1992)	Does race bias sentencing?	19	.09
Mazella and Feingold (1994)	Does defendant social category affect mock juror's verdicts?		
	Defendants':		
	Attractiveness	25	.10
	Race (African American or White)	29	.01
	Social class	4	.08
	Sex	21	.04 ^a
Kunda and Thagard (1996)	Do stereotypes bias judgments of targets in the absence of <i>any</i> individuating information?	7	.25
Kunda and Thagard (1996)	Do stereotypes bias judgments of targets in the presence of individuating information?	40	.19

Note: Effect size is presented in terms of the correlation coefficient, r , between stereotype (or expectation) and outcome.

All meta-analyses presented here focused exclusively on experimental research. Individuating information refers to information about the personal characteristics, behaviors, or accomplishments of individual targets. The effect size shown in the last column for each of the meta-analyses represents the overall average effect size obtained in that study. Effect sizes often varied for subsets of experiments included in the meta-analysis. Only meta-analyses of outcomes, not of moderators or mediators, are displayed.

^a A negative coefficient indicates favoring men; a positive coefficient indicates favoring women.

^b This meta-analysis is included here because many of the studies involved stereotypes.

there is always the possibility that researchers have not searched in the right places or in the right way for powerful stereotype biases in person perception. At minimum, however, the burden of proof (for the existence of widespread, powerful stereotype biases in person perception) has shifted to those emphasizing such powerful biases.

The existence of small stereotype effects, however, does not necessarily mean that people do generally rely heavily on individuating information. However, the empirical evidence shows, in fact, that they do. The one meta-analysis that has addressed this issue found that the effect of individuating information on person perception was among the largest effects found in social psychology, $r = .71$ (Kunda & Thagard, 1996). In other words, people seem to be generally doing the right thing—relying on individuating information far more than stereotypes.

But what about that .10 effect of stereotypes? Doesn't that demonstrate inaccuracy? It might, at least when the stereotype itself is clearly inaccurate. However, it does not necessarily demonstrate inaccuracy for two reasons: (a) Most of the studies examining these issues have examined experimentally created fictitious targets who had no "real" attributes, so that there was no criteria with which to assess accuracy; and (b) an influence of an accurate stereotype (typically characterized as "bias" in the literature) does not necessarily translate into inaccuracy. Indeed, some "biases" mean that people are being as accurate as possible under the circumstances (Jussim, 1991, 2005). Therefore, the next section reviews the very small handful of stereotype and person perception studies that have actually addressed the accuracy issue.

Accuracy

Accuracy in Perception of Small Group Differences

Madon et al. (1998) examined the accuracy of seventh-grade teachers' perceptions of their students' performance, talent, and effort at math about 1 month into the school year. Madon et al. assessed accuracy in the following manner. First they identified the teachers' perceptions of group differences by correlating teachers' perceptions of individual students with the students' race, sex, and social class. This correlation indicated the extent to which teachers systematically evaluated individuals from one group more favorably than individuals from another group. Next, Madon et al. assessed actual group differences in performance, talent, and effort by correlating individual students' final grades the prior year (before teachers knew the students), standardized test scores, and self-reported motivation and effort with students' race, sex, and social class. The teachers' accuracy was assessed by correlating the teachers' perceived differences between groups with the groups' actual differences.

Madon et al. (1998) found that teachers were mostly accurate. The correlation between teachers' perceived group differences and actual group differences was $r = .71$. The teachers' perceptions of sex differences in effort, however, were highly inaccurate—they believed girls exerted more effort than boys, but there was no sex difference in self-reported motivation and effort. When this outlier was removed, the correlation between perceived and actual group differences increased to $r = .96$.

We are aware of only two other studies that have addressed whether people systematically and unjustifiably favor or disparage individuals belonging to certain groups (Clarke & Campbell, 1955; Jussim et al., 1996). Both yielded evidence of accuracy accompanied by small bias.

All three studies (including Madon et al., 1998), however, were conducted in educational contexts—Jussim et al. (1996) addressed teachers' perceptions of students, and Clarke and Campbell (1955) addressed students' perceptions of one another. It remains an empirical question whether this pattern of accuracy and small bias in perceptions of demographic differences between individuals with whom one has extended contact is unique to classrooms, or characterizes social perception more broadly.

This pattern of moderate to high accuracy in perceptions of differences between small groups can occur for either of two reasons. First, perceivers might have jettisoned their stereotypes completely, and judged targets primarily on the basis of relevant individuating information. Second, perceivers

might not have jettisoned their stereotypes. If their stereotypes (e.g., “girls perform slightly higher in math classes than do boys”) were accurate (girls actually did perform slightly higher than boys), teachers could also have reached accurate perceptions of differences between boys and girls by applying their stereotype.

The research described thus far does not distinguish between these explanations. Regardless of the explanation, however, this research does lead to one clear conclusion: In the few studies that have examined stereotypes and person perception under naturalistic conditions, there is no evidence of stereotypes powerfully and pervasively distorting perception. There was some evidence of bias and distortion, but the far stronger pattern has been accuracy. The next section, therefore, reviews studies that have not only assessed accuracy, but have also assessed the sources of both accuracy and bias in person perception.

Does Relying on a Stereotype Increase or Reduce Accuracy in Person Perception?

What does empirical research indicate about whether people’s reliance on stereotypes increases or reduces the accuracy of their judgments? Only a handful of studies provide data capable of addressing this issue, and they are discussed next.

Occupational Stereotypes: C. E. Cohen (1981)

C. E. Cohen (1981) examined whether people more easily remember behaviors and attributes that are consistent with a stereotype than those that are inconsistent with that stereotype. Perceivers in her study viewed a videotape of a dinner conversation between a husband and wife (they were actually husband and wife, but they were also experimental confederates trained by Cohen). Half of the time, this conversation led perceivers to believe the woman was a waitress; half of the time, the conversation led perceivers to believe the woman was a librarian. The remainder of the conversation conveyed an equal mix of librarian-like and waitress-like attributes and behaviors.

Perceivers were then given a series of choices regarding objective aspects of the woman in the videotape (e.g., wore glasses . . . did not wear glasses). Their task was to select the correct description. Perceivers consistently remembered 5% to 10% more behaviors or features that were consistent with the woman’s supposed occupation than behaviors or features that were inconsistent with her supposed occupation. For example, they were more likely to accurately remember that the “librarian” wore glasses and liked classical music, whereas they were more likely to accurately remember that the “waitress” had a beer and no artwork in her house (even though the tape was identical, showing the woman wearing glasses, liking classical music, having a beer, and not having artwork in her apartment). This pattern occurred across two studies and regardless of whether the memory test occurred immediately after the videotape or up to 7 days later. Thus, it appeared that people selectively remembered stereotype-consistent information better than they remembered stereotype-inconsistent information.

C. E. Cohen (1981) also reported results regarding the accuracy of her perceivers’ memories. Across the two studies, accuracy levels were quite high—ranging from a low of 57% to a high of 88% and averaging about 75% in the first study and about 66% in the second study. Overall, therefore, she found high (about 70%) accuracy and small (about 5%–10%) but real bias.

The results from her second study were particularly relevant with respect to understanding whether the stereotype increased or reduced accuracy. In this study, half of the perceivers learned of the woman’s supposed occupation before viewing the tape (so the stereotype was activated prior to viewing); half learned of it after viewing the tape. In comparison to receiving the label after viewing the tape, when people received the label first, they more accurately remembered both stereotype-consistent and stereotype-inconsistent information. On average they correctly remembered 70% of the target’s attributes (regardless of their degree of stereotype consistency) when they received the label first; they correctly remembered only about 63% of the target’s attributes when they received the label last. The upshot here, therefore, is that, although the label biased memory in such a manner as to favor stereotype-consistent information, having the label up front also increased overall accuracy.

Why? Most likely, the label provided some sort of organizing scheme for perceivers, which facilitated their understanding and interpretation of both stereotype-consistent and stereotype-inconsistent attributes. Stereotypes may “bias” perception and, simultaneously, increase accuracy.

Residence Hall Stereotypes: Brodt and Ross (1998)

The utility of an accurate stereotype was also demonstrated by Brodt and Ross (1998). College students made predictions about the behaviors and preferences of other college students who lived in one of two dormitories. The students in the “preppie” dorm were widely seen as politically conservative, wealthy, and conventional. The students in the “hippie” dorm were widely seen as politically left wing with unconventional practices and preferences. Perceivers (other students who did not live in either dorm) viewed photographs of individual targets, were informed of each target’s dorm, and then made predictions about each target’s behaviors and attitudes. Perceivers’ predictions were then compared to the targets’ self-reports on these same preferences and attitudes.

When perceivers predicted targets to be consistent with their dorm (for a preppie dorm resident to have preppie attributes or for a hippie dorm resident to have hippie attributes), 66% of their predictions were correct (they matched the targets’ self-reports). When perceivers jettisoned their dorm stereotypes, and predicted targets to be inconsistent with their dorm, 43% of their predictions were correct. Relying on the preppie–hippie dorm stereotypes enhanced the accuracy of person perception predictions.

Sex Stereotypes: Jussim et al. (1996) and Madon et al. (1998)

Both Jussim et al. (1996) and Madon et al. (1998) examined the accuracy of teacher expectations. (Madon et al., 1998, was described previously; Jussim et al., 1996, was similar, except that it was conducted in sixth grade rather than seventh grade, and it did not examine the accuracy of perceived differences between students from different demographic groups.) Both found that, when controlling for individuating information (motivation, achievement, etc.), student social class and race or ethnicity had little or no effect on teacher expectations. Thus, teachers essentially jettisoned their social class and ethnic stereotypes when judging differences between children from different social class and ethnic backgrounds. Although this finding is in many ways laudable, teachers relying entirely on individuating information does not help address the question of whether relying on a stereotype increases or reduces accuracy.

Both studies, however, found that sex stereotypes biased teachers’ perceptions of boys’ and girls’ performance (standardized regression coefficients of .09 and .10 for performance, and .16 and .19 for effort, for Madon et al. and Jussim et al., respectively). In both studies, teachers perceived girls as performing higher and exerting more effort than boys. Because these effects occurred in the context of models controlling for individuating information, they are best interpreted as stereotypes influencing teacher perceptions—bias effects, in traditional social psychological parlance.

Did these sex stereotyping bias effects increase or reduce the accuracy of teachers’ perceptions? They did both. In the case of performance, the sex stereotype effect increased teacher accuracy. The real performance difference, as indicated by final grades the prior year, was $r = .08$ and $r = .10$ (for the 1996 and 1998 studies, respectively, girls received slightly higher grades). The regression model producing the “biasing” effect of stereotypes yielded a “bias” that was virtually identical to the real difference. In other words:

The small independent effect of student sex on teacher perceptions (of performance) accounted for most of the small correlation between sex and teacher perceptions (of performance). This means that teachers apparently stereotyped girls as performing slightly higher than boys, independent of the actual slight difference in performance. However, the extent to which teachers did so corresponded reasonably well with the small sex difference in performance. In other words, teachers’ perceptions of differences between boys and girls were accurate because teachers relied on an accurate stereotype. (Jussim et al., 1996, p. 348)

The same conclusion, of course, also characterizes the results for the 1998 study.

On the other hand, the results regarding effort provided evidence of bias that reduced accuracy. There was no evidence that girls exerted more effort than boys. Therefore, the influence of student sex on teacher perceptions of effort (i.e., teachers' reliance on a sex stereotype to arrive at judgments of effort) led teachers to perceive a difference where none existed. This is an empirical demonstration of something that, logically, has to be true. Relying on an *inaccurate* stereotype when judging individuals can only harm one's accuracy.

SUMMARY AND CRITICAL EVALUATION

Our review has shown that it is logically incoherent to define stereotypes as inaccurate, that it is unusual (but not unheard of) for stereotypes to be highly discrepant from reality, that the correlations of stereotypes with criteria are among the largest effects in all of social psychology, that people rarely rely on stereotypes when judging individuals, and that, sometimes, even when they do rely on stereotypes, it increases rather than reduces their accuracy. Many scholars, scientists, and people of good will, we do not doubt, will find these conclusions unbearable.

Therefore, the next order of business is to identify important limitations and qualifications to these conclusions. We are going to (a) clearly state many of the things the stereotype literature does not show; (b) state what it does show; and (c) describe many of the limitations to existing research on stereotype accuracy. We hope that doing so reduces the extent to which some readers may misinterpret our claims about what the stereotype research does show, and what lessons can be learned from it.

WHAT THE STEREOTYPE RESEARCH DOES NOT SHOW

1. It does not show that all stereotypes are always perfectly 100% accurate. We know of no researcher who has ever made this claim.
2. It does not show that prejudice and discrimination do not exist, or are trivial and unimportant. Prejudice and discrimination are terribly important, and can be terribly destructive. The research reviewed in this chapter has not addressed prejudice and discrimination.
3. It does not show that people correctly explain why group differences exist. Inasmuch as social scientists do not agree as to why group differences exist, it is probably not possible to assess the accuracy of most lay explanations for group differences.
4. It does not show how people arrive at their stereotypes. There is very little research on where stereotypes come from. Much speculative discussion emphasizes hearsay, family socialization, and the media (e.g., Allport, 1954/1979; Katz & Braly, 1933; Pickering, 2001). The extraordinary levels of accuracy shown in many of the studies reviewed in this chapter, however, do suggest another source is the primary basis of stereotypes—social reality.
5. The amount of research that has addressed the accuracy of people's perceptions of differences between small groups of individuals they know personally (stereotypes and person perception) is quite modest, and does not yet provide a sufficiently broad foundation on which to reach any general conclusions. It appears as if relying on accurate stereotypes seems to mostly enhance accuracy, but that conclusion should be held tentatively, pending further studies.

WHAT THIS RESEARCH DOES SHOW

1. The claim that stereotypes, as beliefs about groups, are inherently inaccurate has been falsified.

2. A more modest claim, one that does not define stereotypes as inherently inaccurate, is that they are generally or frequently inaccurate. This also has been falsified. The scientific evidence provides more evidence of accuracy than of inaccuracy in social stereotypes. The most appropriate generalization based on the evidence is that people's beliefs about groups are usually moderately to highly accurate, and are occasionally highly inaccurate.
3. This pattern of empirical support for moderate to high stereotype accuracy is not unique to any particular target or perceiver group. Accuracy has been found with racial and ethnic groups, gender, occupations, and college groups.
4. The pattern of moderate to high stereotype accuracy is not unique to any particular research team or methodology. It has been found by a wide variety of American and Canadian researchers; by those using Judd and Park's (1993) componential methodology; by those using noncomponential methodologies; and regardless of whether the criteria are obtained through official government reports, meta-analyses, or the self-reports of members of the target group.
5. This pattern of moderate to high stereotype accuracy is not unique to the substance of the stereotype belief. It occurs for stereotypes regarding personality traits, demographic characteristics, achievement, attitudes, and behavior.
6. The strong form of the exaggeration hypothesis—either defining stereotypes as exaggerations or as claiming that stereotypes usually lead to exaggeration—is not supported by data. Exaggeration does sometimes occur, but it does not appear to occur much more frequently than does accuracy or underestimation, and may even occur less frequently.
7. The exaggeration hypothesis—as a *hypothesis*—can still be retained. Exaggeration sometimes does occur. Understanding when stereotypes are more likely to exaggerate real differences, more likely to underestimate real differences, and more likely to be accurate is an important question for future research.
8. In contrast to their reputation as false cultural myths perpetrated by exploitative hierarchies against the disenfranchised, consensual stereotypes were not only the most accurate aspect of stereotypes, not only more valid than nearly all social psychological hypotheses, but they were stunningly accurate by any standard. Correlations of $r = .70$ and higher are almost never repeatedly obtained in any area of social or psychological research. Using Rosenthal's (1991) binomial effect size display to translate correlations into intuitively meaningful relationships shows that correlations of .6 to .9 mean that consensual stereotypes are about 80% to 90% accurate.

Table 10.4 compares the frequency with which social psychological research produces effects exceeding correlations of $r = .30$ and $r = .50$, with the frequency with which the correlations reflecting the extent to which people's stereotypes correspond to criteria exceed $r = .30$ and $r = .50$. Only 24% of social psychological effects exceed correlations of $r = .30$ and only 5% exceed $r = .50$. In contrast, all 18 of the aggregate and consensual stereotype accuracy correlations shown in Table 10.1 and Table 10.2 exceed $r = .30$, and all but two exceed $r = .50$. Furthermore, 9 of 11 personal stereotype accuracy correlations exceeded $r = .30$, and 4 of 11 exceeded $r = .50$.

This is doubly important. First, it is yet another way to convey the impressive level of accuracy in laypeople's stereotypes. Second, it is surprising that so many scholars in psychology and the social sciences are either unaware of this state of affairs, unjustifiably dismissive of the evidence, or choose to ignore it (see reviews by Funder, 1987, 1995; Jussim, 1991, 2005; Ryan, 2002). When introductory texts teach about social psychology, they typically teach about phenomena such as the mere exposure effect (people like novel stimuli more after repeated exposure to it, $r = .26$), the weapons effect (they become more aggressive after exposure to a weapon, $r = .16$), more credible speakers are more persuasive ($r = .10$), and self-serving attributions (people take more responsibility for successes than failures, $r = .19$; correlations all obtained from Richard et al., 2003). How much time and space is typically spent in such texts reviewing and documenting the much stronger

TABLE 10.4**Ethnic and Gender Stereotypes Are More Valid Than Most Social Psychological Hypotheses**

	Proportion of Social Psychological Effects Obtained in Research^a	Proportion of Consensual Stereotype Accuracy Correlations^b	Proportion of Personal Stereotype Accuracy Correlations^b
Exceeding .30	24%	100% (18/18)	81% (9/11)
Exceeding .50	5%	89% (16/18)	36% (4/11)

^a Data obtained from Richard, Bond, and Stokes-Zoota's (2003) review of meta-analyses including thousands of studies. Effects are in terms of the correlation coefficient, r .

^b From Tables 10.1 and 10.2. Within parentheses, the numerator is the number of stereotype accuracy correlations meeting the criteria for that row (exceeding .30 or .50) and the denominator is the total number of stereotype accuracy correlations reported in Tables 10.1 and 10.2. Because Table 10.1 summarizes the results for five studies for McCauley, Thangavelu, and Rozin (1988), the .94–.98 figure is counted five times. These numbers probably *underestimate* the degree of stereotype accuracy, because all single entries in Tables 10.1 and 10.2 only count once, even though they often constitute averages of several correlations found in the original articles.

evidence of the accuracy of people's stereotypes? Typically, none at all. For a field that aspires to be scientific, this is a troubling state of affairs. Some might even say unbearable.

IMPORTANT LIMITATIONS

There are, of course, many important limitations to the existing work on the accuracy of stereotypes. First, the accuracy of two of the other major types of stereotypes—religion and social class—have, as far as we know, never been examined. Although we can think of no reason why patterns of accuracy should differ for these types of groups, we will never know until the research is actually conducted.

Second, the existing research has overwhelmingly examined the stereotypes held by college students, largely because those samples are convenient. Is this important? Maybe. Suggesting it may not be that important has been the research by McCauley and colleagues (see Tables 10.1 and 10.2), and by Clabaugh and Morling (2004) showing that the accuracy of noncollege groups is nearly identical to that of college students. Nonetheless, more research with noncollege samples is needed.

Third, there are many different types and aspects of accuracy, and few studies report results addressing all of them. Ideally, more research in the future will provide more comprehensive assessments of the various types of stereotype accuracy.

Fourth, most of the research on stereotype accuracy to date has been conducted in the United States and Canada. Perhaps stereotypes in other countries are less (or more) accurate.

ARE STEREOTYPES EVER HIGHLY INACCURATE?

The Evidence Reviewed in This Chapter

Evidence of major inaccuracy is rare but it is not entirely absent. First, even the studies that we have reviewed have shown that people are better at judging differences between groups, and at judging the rank order of attributes within a group, than they are at judging the exact level of particular attributes within a group. In other words, the analyses assessing correspondence, which correlated people's beliefs with group attributes or group differences, consistently found strong evidence of accuracy, whereas the analyses assessing discrepancies provided a more mixed picture, including a fair amount of bull's-eyes, a fair amount of near misses, and a fair amount of inaccuracy. Even when people do not exaggerate or underestimate real differences, the evidence we reviewed showed that, often, they either consistently over- or underestimate the level of an attribute in a group.

Second, on average, personal stereotypes corresponded well with groups' attributes (i.e., individual beliefs about groups correlated moderately to highly with criteria). Nonetheless, some personal stereotypes were highly inaccurate. Nearly all of the studies reporting personal stereotype accuracy correlations found at least some people with very low—near zero—correlations. Whether these are simply more or less random fluctuations and measurement error, or whether some people are systematically more accurate than others, is an important question for future research. Possible candidates for individual differences that would predict systematic variations in accuracy would be intelligence (are smarter people more accurate?), education (are more highly educated people more accurate?), exposure to and experience with groups (the “contact hypothesis”; e.g., Allport, 1954/1979, has long suggested that contact with a group reduces prejudice, in part, by disconfirming erroneous stereotypes), nonverbal sensitivity (actually, Hall & Carter, 1999, already showed that people lower in nonverbal sensitivity hold less accurate sex stereotypes, but it would be useful to see if this pattern replicates), and ideology/motivated egalitarianism/universalism (which, despite the intended benevolence of an egalitarian ideology, seems to lead people to hold less accurate stereotypes; Wolsko et al., 2000). Despite the existing evidence showing only weak relations between prejudice and stereotyping (Park & Judd, 2005), perhaps under the right (or wrong) conditions, deeply held prejudices and hostilities can sometimes lead to highly distorted stereotypes.

Speculations on Other Conditions of Inaccuracy

Studies that examined people's beliefs about groups and then used as criteria the self-reports of haphazard samples of members of the target group (Allen, 1995; Martin, 1987; Terracciano et al, 2005) consistently find more evidence of what those researchers interpret as “inaccuracy” than do researchers who used whole populations or random samples of targets (the research summarized in Tables 10.1–10.3). The disconnect between the stereotype and criteria, however, renders such results difficult to interpret.

The existence of so few clear and strong demonstrations of widespread stereotype inaccuracy does justify the conclusion that research on the accuracy of stereotypes usually finds evidence of moderate to high accuracy, and only rarely finds evidence of low accuracy. It does not, however, necessarily justify concluding that stereotypes are hardly ever inaccurate. Perhaps researchers have just not yet looked in the right places or in the right ways for stereotype inaccuracy.

For example, education and mass communication levels are so high in the United States and Canada, where most of the stereotype accuracy research has been conducted, that, perhaps, in general, people are more exposed to social reality in these places (and, probably, in other Western democracies) than in many other places around the world. Perhaps poverty and ignorance help breed stronger inaccurate stereotypes. Perhaps the propaganda of demagogues in authoritarian regimes helps perpetuate inaccurate stereotypes. The Jim Crow American South; South Africa under apartheid; the Indian caste system; the Nazis' racial beliefs; and beliefs about Christians, Jews, and Hindus held by 21st-century Islamists are a few examples where it seems plausible to speculate that stereotypes might be more inaccurate than found in the research reviewed here.

Unfortunately, because the powers that be under such systems are not likely to be open to challenges to their authority, it will probably be very difficult to perform studies of stereotype (in) accuracy in such contexts. If it is difficult to perform research in the contexts most likely to produce stereotype inaccuracy, the scientific literature will be skewed toward providing more evidence of stereotype accuracy than may be actually true of people in general, around the world.

THE SCIENTIFIC AND SOCIAL VALUE OF STEREOTYPE ACCURACY RESEARCH

Stereotypes can be accurate. Some scholars and lay people resist this conclusion, believing that crediting any accuracy to stereotypes is tantamount to endorsing prejudice. We argue that the opposite is more likely true—that acknowledging the accuracy of some stereotypes provides the

epistemological clarity needed to more effectively address prejudice and bigotry, and to more effectively investigate the nature, causes, and moderators of stereotypes.

Distinguishing Accurate From Inaccurate Stereotypes

Not all stereotypes are accurate, and those that are inaccurate may be the most damaging. A special and important case is that of manufactured stereotypes, which are intentionally designed to despoil the reputation of particular social groups. A few notorious examples include 19th-century American stereotypes of indigenous peoples as uncivilized savages, stereotypes of civil rights workers as Communist fifth columnists, and the perpetual stereotype of Jews as seeking world domination. All these manufactured stereotypes served nefarious agendas, and all were (and are) patently false.

However, exposing the fallacious nature of these libelous stereotypes requires criteria and tools for assessing stereotype accuracy. These tools must be calibrated against a standard of authenticity, just as do the tools for demonstrating counterfeit and fraud in art and business. Whereas Jews do not seek world domination, it is not always absurd to believe that certain groups seek domination over, if not quite the world, at least large parts of it (consider, e.g., Rome, Nazis, Communists, Imperial Japan, the Mongolian Khans, and, possibly, some modern Islamic extremists, etc.). Without standards and methods for assessing (in)accuracy, it becomes impossible to reliably sort out valid from bogus beliefs.

Investigating the Dynamics of Stereotypes

Stereotypes are not static phenomenon, but shift with circumstance, policy, social contact, and other forces. To what degree do stereotypes map these changes? How responsive are they to social shifts, or to targeted interventions? Why do some stereotypes shift rapidly and others remain entrenched? Perhaps not surprisingly, if one makes the common assumption that stereotypes are inaccurate, and answers these questions by assumption, one is not likely to even consider such questions, let alone provide answers to them (e.g., see discussions of stereotypes in Aronson, 1999; Devine, 1995; Fiske, 1998; Gilbert, 1995; Jones, 1986). However, answers to some of these questions have indeed begun to be provided by researchers who make the alternative assumption, that stereotypes might be influenced by social reality (e.g., Eagly & Diekmann, 2003; Oakes, Haslam, & Turner, 1994).

Generating a Coherent Understanding of Both Past and Future Research

The decades of research on the role of stereotypes in expectancy effects, self-fulfilling prophecies, person perception, subtyping, and memory, are jeopardized if all stereotypes are regarded as wholly inaccurate. This past research will be haunted by a definitional tautology; that is, that people who believe in stereotypes are in error because stereotypes are erroneous beliefs. On the other hand, accepting that stereotypes range in accuracy makes this past research coherent, and allows for more edifying interpretations of past and future research, such as “people in X condition, or of Y disposition, are more likely to believe in, subscribe to, and maintain false stereotypes, whereas people in A condition, or of B disposition are more likely to believe in, subscribe to, and maintain accurate stereotypes.”

In sum, accepting that stereotypes can sometimes be accurate provides the means to distinguish innocent errors from motivated bigotry, assess the efficacy of efforts to correct inaccurate stereotypes, and reach a more coherent scientific understanding of stereotypes. We believe that this proposition can advance the depth, scope, and validity of scientific research on stereotypes, and thereby help improve intergroup relations.

NOTES

- 1 Some of these appear in print, some have occurred at a conference, and one was in a review of a manuscript submitted for publication. At the May 2004 American Psychological Society (now Association for Psychological Science) conference panel on Stereotyping, Discrimination, and the Law, “Nonsense” was Lee Ross’s characterization of my description of Brodt and Ross (1998) as showing that relying on an

accurate stereotype can increase accuracy of person perception (he is the Ross on that study, which is discussed in some detail later in this chapter and is readily available to the general scholarly public because it was published in a widely circulated journal). Living “in a world where all stereotypes are accurate” was Susan Fiske’s introductory comment as she began her talk at the same conference. “Disagreeing with civil rights” is also from Fiske (1998, p. 381), and refers specifically to McCauley, Jussim, and Lee’s (1995) concluding chapter to their book, *Stereotype Accuracy* (in that chapter they argued that, in the absence of perfectly diagnostic individuating information, people would make more accurate person perception judgments if they relied on rather than ignored accurate stereotypes—exactly the result empirically found by Brodt and Ross, 1998). Stangor (1995) did not specifically accuse any particular person of “supporting bigots”; instead, he indicted the entire scientific attempt to assess the accuracy of stereotypes as potentially supporting bigotry. “Are Blacks really lazy?” and “Are Jews really cheap?” were a reviewer’s comments on a draft of the article eventually published by *Psychological Review* (Jussim, 1991), which argued that, if social psychologists wanted to make claims about the inaccuracy of stereotypes (which, given the frequency of such claims they apparently wanted to do very much), it behooved them to perform research that actually empirically assessed the accuracy of stereotypes. Although this call was removed from that particular article, it appeared in many others, and, in fact, has been answered by many researchers over the last 15 years. This chapter reviews that evidence. It is, perhaps, worth noting that, of the scores of empirical studies and meta-analyses reviewed, not a single one is titled anything like “Are Blacks really lazy?” or “Are Jews really cheap?”

2. Standard deviations are not related to percentiles in a linear manner. Therefore, .52 SD comes closer to capturing a 20% difference than does .50 SD. But .50, is a round number, is easier to use and remember, and ease of use has its own value. A difference of .50 SD actually means the mean of one group is higher than the mean of 69.15% of the members of the other group, which is close enough to 70% for this chapter.

REFERENCES

- Allen, B. P. (1995). Gender stereotypes are not accurate: A replication of Martin (1987) using diagnostic vs. self-report and behavioral criteria, *Sex Roles, 32*, 583–600.
- Allport, G. W. (1979). *The nature of prejudice* (2nd ed.). (Original work published 1954). Reading, MA: Addison-Wesley.
- APA Brief (1991). In the Supreme Court of the United States: Price-Waterhouse versus Hopkins. *American Psychologist, 46*, 1061–1070
- Aronson, E. (1999). *The social animal* (8th ed.). New York: Worth.
- Ashmore, R. D., & Del Boca, F. K. (1981). Conceptual approaches to stereotypes and stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 1–35). Hillsdale, NJ: Erlbaum.
- Ashton, M. C., & Esses, V. M. (1999). Stereotype accuracy: Estimating the academic performance of ethnic groups. *Personality and Social Psychology Bulletin, 25*, 225–236.
- Beyer, S. (1999). The accuracy of academic gender stereotypes. *Sex Roles, 40*, 787–813.
- Biernat, M. (1995). The shifting standards model: Implications of stereotype accuracy for social judgment. In Y. T. Lee, L. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy* (pp. 87–114). Washington, DC: American Psychological Association.
- Brigham, J. C. (1971). Ethnic stereotypes. *Psychological Bulletin, 76*, 15–38.
- Briton, N. J., & Hall, J. A. (1995). Beliefs about female and male nonverbal communication. *Sex Roles, 36*, 79–90.
- Brodt, S. E., & Ross, L. D. (1998). The role of stereotyping in overconfident social prediction. *Social Cognition, 16*, 225–252.
- Campbell, D. T. (1967). Stereotypes and the perception of group differences. *American Psychologist, 22*, 817–829.
- Cejka, M. A., & Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin, 25*, 413–423.
- Clabaugh, A., & Morling, B. (2004). Stereotype accuracy of ballet and modern dancers. *Journal of Social Psychology, 144*, 31–48.
- Clarke, R. B., & Campbell, D. T. (1955). A demonstration of bias in estimates of Negro of ability. *Journal of Abnormal and Social Psychology, 51*, 585–588.
- Cohen, C. E. (1981). Personal categories and social perception: Testing some boundaries of the processing effects of prior knowledge. *Journal of Personality and Social Psychology, 40*, 441–452.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Academic.
- Devine, P. (1995). Prejudice and outgroup perception. In A. Tesser (Ed.), *Advanced social psychology* (pp. 467–524). New York: McGraw-Hill.
- Diekmann, A. B., Eagly, A. H., & Kulesa, P. (2002). Accuracy and bias in stereotypes about the social and political attitudes of women and men. *Journal of Experimental Social Psychology*, 38, 268–282.
- Eagly, A. H., & Diekmann, A. B. (2003). The malleability of sex differences in response to changing social roles. In L. G. Aspinwall & U. M. Staudinger (Eds.), *A psychology of human strengths* (pp. 103–115). Washington, DC: American Psychological Association.
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 357–411). New York: McGraw-Hill.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, 23, 1–74.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.
- Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin*, 101, 75–90.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 102, 652–670.
- Gilbert, D. T. (1995). Attribution and interpersonal perception. In A. Tesser (Ed.), *Advanced social psychology* (pp. 99–147). New York: McGraw-Hill.
- Hall, J. A., & Carter, J. D. (1999). Gender-stereotype accuracy as an individual difference. *Journal of Personality and Social Psychology*, 77, 350–359.
- Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, 58, 78–79.
- Jones, E. E. (1986). Interpreting interpersonal behavior: The effects of expectancies. *Science*, 234, 41–46.
- Jost, J. T., & Kruglanski, A. W. (2002). The estrangement of social constructionism and experimental social psychology: History of the rift and prospects for reconciliation. *Personality and Social Psychology Review*, 6, 168–187.
- Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, 100, 109–128.
- Jussim, L. (1991). Social perception and social reality: A reflection-construction model. *Psychological Review*, 98, 54–73.
- Jussim, L. (2005). Accuracy in social perception: Criticisms, controversies, criteria, components, and cognitive processes. *Advances in Experimental Social Psychology*, 37, 1–93.
- Jussim, L., Eccles, J., & Madon, S. J. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. *Advances in Experimental Social Psychology*, 29, 281–388.
- Jussim, L., Harber, K. D., Crawford, J. T., Cain, T. R., & Cohen, F. (2005). Social reality makes the social mind. *Interaction Studies*, 6, 85–102.
- Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *Journal of Abnormal and Social Psychology*, 28, 280–290.
- Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: Guilford.
- Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, 103, 284–308.
- Madon, S. J., Jussim, L., Keiper, S., Eccles, J., Smith, A., & Palumbo, P. (1998). The accuracy and power of sex, social class and ethnic stereotypes: Naturalistic studies in person perception. *Personality and Social Psychology Bulletin*, 24, 1304–1318.
- Martin, C. L. (1987). A ratio measure of sex stereotyping. *Journal of Personality and Social Psychology*, 52, 489–499.
- Mazella, R., & Feingold, A. (1994). The effects of physical attractiveness, race, socioeconomic status, and gender of defendants and victims on judgments of mock jurors: A meta-analysis. *Journal of Applied Social Psychology*, 24, 131–135.
- McCauley, C., Jussim, L., & Lee, Y. T. (1995). Stereotype accuracy: Toward appreciating group differences. In Y. T. Lee, L. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy: Toward appreciating group differences* (pp. 293–312). Washington, DC: American Psychological Association.
- McCauley, C., & Stitt, C. L. (1978). An individual and quantitative measure of stereotypes. *Journal of Personality and Social Psychology*, 36, 929–940.

- McCauley, C., Stitt, C. L., & Segal, M. (1980). Stereotyping: From prejudice to prediction. *Psychological Bulletin*, *87*, 195–208.
- McCauley, C., & Thangavelu, K. (1991). Individual differences in sex stereotyping of occupations and personality traits. *Social Psychology Quarterly*, *54*, 267–279.
- McCauley, C., Thangavelu, K., & Rozin, P. (1988). Sex stereotyping of occupations in relation to television representations and census facts. *Basic and Applied Social Psychology*, *9*, 197–212.
- Nelson, T. (2002). *The psychology of prejudice*. Boston: Allyn & Bacon.
- Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotyping and social reality*. Cambridge, MA: Blackwell.
- Park, B., & Judd, C. M. (2005). Rethinking the link between categorization and prejudice within the social cognition perspective. *Personality and Social Psychology Review*, *9*, 108–130.
- Pickering, M. (2001). *Stereotyping*. New York: Palgrave.
- Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*, 331–363.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (2nd ed.). Newbury Park, CA: Sage.
- Ryan, C. (1996). Accuracy of Black and White college students' in-group and out-group stereotypes. *Personality and Social Psychology Bulletin*, *22*, 1114–1127.
- Ryan, C. S. (2002). Stereotype accuracy. *European Review of Social Psychology*, *13*, 75–109.
- Schneider, D. J. (2004). *The psychology of stereotyping*. New York: Guilford.
- Schultz, P. W., & Oskamp, S. (2000). *Social psychology: An applied perspective*. Upper Saddle River, NJ: Prentice-Hall.
- Stangor, C. (1995). Content and application inaccuracy in social stereotyping. In Y. T. Lee, L. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy* (pp. 275–292). Washington, DC: American Psychological Association.
- Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*, *111*, 42–61.
- Sweeney, L. T., & Haney, C. (1992). The influence of race on sentencing: A meta-analytic review of experimental studies. *Behavioral Sciences and the Law*, *10*, 179–195.
- Swim, J. K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *Journal of Personality and Social Psychology*, *66*, 21–36.
- Swim, J., Borgida, E., Maruyama, G., & Myers, D. G. (1989). Joan McKay vs. John McKay: Do gender stereotypes bias evaluations? *Psychological Bulletin*, *105*, 409–429.
- Terracciano, A., Abdel-Khalek, A. M., Ádám, N., Adamovová, L., Ahn, C.-K., Ahn, M., et al. (2005). National character does not reflect mean personality trait levels in 49 cultures. *Science*, *310*, 96–100.
- Wolsko, C., Park, B., Judd, C. M., & Wittenbrink, B. (2000). Framing interethnic ideology: Effects of multicultural and color-blind perspectives on judgments of groups and individuals. *Journal of Personality and Social Psychology*, *78*, 635–654.