The first Supreme Court case to use psychological research on sex stereotyping was *Price Waterhouse v. Hopkins*. The case was decided in May 1989 and remanded to Judge Gerhard Gesell, who rendered his final decision in May 1990. In this case, a social psychological expert testified to the antecedent conditions, indicators, consequences, and remedies of stereotyping, on the basis of recent cognitive approaches to stereotyping. The testimony was cited in decisions reached in the trial and appeals courts, as well as in the Supreme Court's review. The American Psychological Association filed an amicus curiae brief supporting the validity of the field of stereotyping and the general methods used by the expert. Such legal application provides further lessons for psychological research on stereotyping.

"The firm of Price Waterhouse refused to make Ann Hopkins a partner. Gender-based stereotyping played a role in this decision" (*Hopkins v. Price Waterhouse*, 1990, p. 1). Thus concluded federal district court Judge Gerhard Gesell in his May 1990 opinion that followed the United States Supreme Court landmark decision in the same sex discrimination case (*Price Waterhouse v. Hopkins*, 1989). The case is unique because it represents the first use of psychological evidence about sex stereotyping by the Supreme Court. Previous sex discrimination litigation has used only statistical approaches. Moreover, social science evidence has played a significant role in race discrimination cases in the Supreme Court since *Brown v. Board of Education* (1954), but such evidence had not before been used in sex discrimination cases. Originally filed in 1984, *Hopkins v. Price Waterhouse* was the first suit to be brought under a new ruling that partnership decisions qualify for protection under Title VII of the 1964 Civil Rights Act (*Hishon v. King & Spaulding*, 1984). *Hopkins v. Price Waterhouse* was eventually heard by the Supreme Court, was decided by the Court in May 1989, and was remanded to Judge Gesell, who rendered his final decision in May 1990 with the new rules developed by the nation's highest court for such cases. The federal court of appeals affirmed Judge Gesell's opinion in December 1990. Testimony about the psychology of stereotyping played a crucial role at each stage of the litigation as it made its way through the judicial process. The American Psychological Association (APA) submitted to the Supreme Court an amicus curiae brief (which follows this article). As we will discuss later, the brief had a significant impact on that Court's final decision.

The relevant psychological literature was heavily cited in Judge Gesell's original decision (*Hopkins v. Price Waterhouse*, 1985), and the testimony about the psychology of stereotyping was cited at all levels of the appeal and review process, including the Supreme Court's decision and the subsequent remand. Moreover, the APA brief clearly contributed to the Supreme Court's opinion regarding the credibility of this area of research. Now that the case has been decided, it is appropriate to describe and evaluate both the role of psychology in this process and what the drafters of the amicus brief learned about our field as a result.

Parts of this article were presented as an invited address by Susan T. Fiske at the 97th Annual Convention of the American Psychological Association in New Orleans and as a paper by Donald N. Bersoff at the 1990 meeting of the Association of American Law Schools in San Francisco.

The research material for the accompanying brief was supplied by Eugene Borgida, Kay Deaux, Susan T. Fiske, and Madeline E. Heilman; Donald N. Bersoff wrote the final copy. The authors would like to thank Sarah Burns, the attorney whose idea it was to introduce current sex stereotyping research into the legal system and who commented on an early draft of the amicus brief; Douglas Huron and James Heller, attorneys who had the good sense and courage to try such unorthodox methods as using sex stereotyping research in court; Ann Hopkins, the plaintiff who pursued the case, despite considerable tangible and intangible costs; Ralph Alexander, who drafted a footnote in the brief regarding the use of subjective criteria in personnel decisions; and James Uleman, who commented on an early draft of the brief.

Correspondence concerning this article should be addressed to Susan T. Fiske, Department of Psychology, Tobin Hall, University of Massachusetts, Amherst, MA 01003, or to any of the authors at their respective Departments of Psychology.

---

1 The position of the names of the parties often changes in a case citation as it goes through trial, appeal, and review process. In this case, Ann Hopkins was the original plaintiff in the federal district (trial) court and Price Waterhouse (PW) was the defendant. Hopkins won at trial; PW then appealed to the intermediate federal appellate court (an appellant's name appears first at this level). After it again lost, PW petitioned the U.S. Supreme Court to review the case. As petitioner, PW's name again appears first. Regardless of the sequence of the names of the parties in this article, the authors are referring to the same underlying case.
Factual Background of the Hopkins Case

In 1982, it seemed that Ann Hopkins had established her credentials as a top-notch performer at Price Waterhouse (PW), one of the nation’s big-eight accounting firms. She had more billable hours than any other person proposed for partner that year, she had brought in business worth $25 million, her clients praised her, and her supporters recommended her as driven, hard working, and exacting. Instead of being promoted for her accomplishments, her candidacy was put on hold, and she was not proposed for partner the following year. Hopkins alleged that she was denied partnership because of her gender. Price Waterhouse countered that she was not admitted because she had interpersonal skills problems. According to some evaluators, this “lady partner candidate” was “macho,” she “overcompensated for being a woman,” and she needed a “course at charm school.” A sympathetic colleague advised that she would improve her chances if she would “walk more femininely, talk more femininely, dress more femininely, wear make-up, have her hair styled, and wear jewelry” (Hopkins v. Price Waterhouse, 1985, p. 1117). Instead, Hopkins took the firm to court. In 1982, she filed a complaint in the federal district court of the District of Columbia, alleging a violation of Title VII of the 1964 Civil Rights Act.

She had a strong case in many respects, but her attorneys, Douglas Huron and James Heller, still needed to demonstrate that these stereotypic remarks might account for discriminatory decision making. Huron and Heller had heard about attorney Sarah Burns’s novel construction of a prior sex discrimination case, using current research on sex stereotyping, particularly cognitive categorization theories (e.g., Allport, 1954; Ashmore & DelBoca, 1981; Hamilton, 1979; Tajfel, 1972; Taylor, 1981; for a review of subsequent contributions, see Fiske & Taylor, 1991, chap. 4, 5). Although that case had been settled before trial, the deposition by social psychologist Susan T. Fiske was viewed by the plaintiff’s attorneys as contributing to a favorable outcome for their client. Huron and Heller asked Fiske to testify about the psychology of stereotyping, with particular reference to the facts of their case. When they described their evidence of stereotyping, Fiske agreed, because, in her opinion, the case so closely fit the literature on sex stereotyping. Moreover, it seemed an excellent test of the utility of social psychological theories and research.

Fiske’s testimony in the original Hopkins v. Price Waterhouse case drew on both laboratory and field research to describe antecedent conditions that encourage stereotyping, indicators that reveal stereotyping, consequences of stereotyping for out-groups, and feasible remedies to prevent the intrusion of stereotyping into decision making. Specifically, she testified first that stereotyping is most likely to intrude when the target is an isolated, one- or few-of-a-kind individual in an otherwise homogeneous environment. The person’s solo or near-solo status makes the unusual category more likely to be a salient factor in decision making (e.g., Crocker & McGraw, 1982; Heilman, 1980; Kanter, 1977; McArthur & Post, 1977; Spangler, Gordon, & Pipkin, 1978; S. E. Taylor, 1981; Wolman & Frank, 1975; for recent reviews, see Fiske & Taylor, 1991, chap. 7; Mullen, 1991; Pettigrew & Martin, 1987). Ann Hopkins qualified as a near-solo in the organization. Stereotyping is also likely to intrude when members of a previously omitted group move into a job that is nontraditional for their group. This is certainly true for women who are senior managers and partners in the big-eight accounting firms. The maximum percentage of women partners in a big-eight firm as of 1989 was 5.6%; PW had the lowest with 2% (Berg, 1988). As of May 1990, PW had 27 female partners out of 900 total, or 3% (Lewin, 1990).

Another antecedent condition concerns the perceived lack of fit between the person’s category and occupation (Heilman, 1983). The attributes desirable in a manager—aggressive, competitive, driven, tough, and masterly—are not attributes typically expected of women (e.g., Heilman, Block, Martell, & Simon, 1989; Schein, 1973, 1975). Women who behave in those managerial ways are often disliked (e.g., Brown & Geis, 1984; Costrich, Feinstein, Kidder, Marecek, & Pascale, 1975; Deaux & Lewis, 1983, 1984; Hagen & Kahn, 1975; Heilman et al., 1989; for a recent meta-analytic review, see Eagly, Makhijani, & Klonsky, in press) and create dissatisfaction among their subordinates (Petty & Lee, 1975; Rosell, 1974).

The testimony also addressed antecedents regarding the information environment. Stereotyping is most likely when evaluative criteria are ambiguous (for reviews, see Arvey, 1979; Kanter, 1977; Nieva & Gutke, 1980; as a general principle of schema use, see Fiske & Neuberg, 1990; Fiske & Taylor, 1991, chap. 4, 5; Markus & Zajonc, 1985; Nisbett & Ross, 1980). Moreover, when information about the individual is ambiguous, it is most open to interpretation (e.g., J. M. Darley & Gross, 1985; Heilman, 1984; Heilman, Martell, & Simon, 1988; Lockley, Bongida, Brekke, & Hepburn, 1980, Study 2; Pheterson, Kiesler, & Goldberg, 1971; Rasinski, Crocker, & Hastie, 1985; for reviews, see Arvey, 1979; Nieva & Gutke, 1980; Tosi & Einbender, 1985). Stereotypes provide structure and meaning, and they shape perceptions most when the data themselves are open to multiple interpretations, as suggested by research on the cognitive bases of stereotyping (see Fiske & Taylor, 1991, chap. 4, 5, for a review). For example, a “counting” decision—based on millions of business dollars—is relatively immune to stereotypic biases. However, subjective judgments of interpersonal skills and collegiality are quite vulnerable to stereotypic biases. Cognitive models of stereotyping attest to the effects of well-developed expectancies and stereotypes on the interpretation of ambiguous information. This is not to say that decision makers should not use subjective criteria, merely that one must be alert to the possibility of stereotyping in their application.

The testimony also indicated that the symptoms or indicators of cognitive processes that give rise to stereo-
typing are straightforward: categorical responding, such as unnecessarily labeling and evaluating someone according to gender. In Ann Hopkin's case, people commented that her behavior was evaluated differently "because it's a woman doing it," and they suggested that she "overcompensated for being a woman" (for reviews, see Allport, 1954; Brewer, 1988; Deaux & Kite, in press; Fiske & Neuberg, 1990; Fiske & Taylor, 1991, chap. 4; Tajfel, 1969).

Another symptom of category-based judgment is evaluating people's credentials along dimensions narrowly relevant to their group's stereotype. For example, the sex role appropriateness of Ann Hopkin's social skills, instead of her business-generating abilities, became the primary dimension along which she was evaluated.

Selective perception and interpretation are also indicators of stereotyping. Hopkin's detractors saw her as an aggressive woman and therefore abrasive and difficult, but her supporters and her clients saw her simply as a determined go-getter (for a gender-relevant example, see Taylor, Fiske, Etcoff, & Ruderman, 1978; as a general process in stereotyping, see reviews by Fiske & Taylor, 1991, chap. 4; Hamilton, 1979, 1981; Higgins & Bargh, 1987; Markus & Zajonc, 1985; Nisbett & Ross, 1980, chap. 2, 8).

Finally, stereotyping is also indicated by extreme, polarized evaluations based on limited evidence. For example, some acquaintances claimed that Hopkins was "universally disliked," which was demonstrably not true (e.g., Kanter, 1977; Taylor, 1981; Wolman & Frank, 1975; for a review, see Mullen, 1991). The testimony dealt with other examples, but these examples illustrate the sense of the argument.

The testimony also noted that the consequences of such stereotyping are obvious. Evaluations are based on category membership, not individual merit (for reviews, see Arvey & Campion, 1982; Helman, 1983; Nieva & Gutke, 1980; Olian, Schwab, & Haberfeld, 1988; D. N. Ruble & Ruble, 1982; T. L. Ruble, Cohen, & Ruble, 1984; Terborg, 1977). Negative attributes are exaggerated (Kanter, 1977; Linville, 1982; Linville & Jones, 1980; Taylor, 1981; Taylor et al., 1978), and positive ones can be discounted (Deaux, 1976; Deaux & Emswiller, 1974; Feather & Simon, 1975; Feldman-Summers & Kiesler, 1974; Frieze, Fisher, Hanusa, McHugh, & Valle, 1978; Garland & Price, 1977; Hansen & O'Leary, 1983; Helman & Guzzo, 1978; Hellman & Stoepke, 1985a, 1985b; Nicholls, 1975; Pazy, 1986; Pence, Pendleton, Dobbins, & Sgro, 1982). Furthermore, there are clear constraints on the permissible behavior of the stereotyped person (e.g., Bartol & Butterfield, 1976; Brown & Geis, 1984; Costrich et al., 1975; S. Darley, 1976; Jago & Vroom, 1982; Kristal, Sanders, Spence, & Helreich, 1975; Wiley & Eskilson, 1982).

Regarding possible remedies, Fiske asserted that stereotyping is controllable (Fiske, 1989a; Fiske & Neuberg, 1990), and people are commonly aware that it is inappropriate. Consequently, adequate information undermines stereotyping (e.g., Deaux & Lewis, 1984; Dip-
The Witness: There are general stereotypes of what people particularly expect men to be like and typically expect women to be like. People typically expect women to be strong on the social dimensions. Women are generally expected to be more tender and understanding and concerned about other people, and soft. The Court: You say that people who have dealt with women expect that? People who have dealt with women in the business context expect that or are you talking about people out on the farm?
The Witness: Well, I would—
The Court: I mean we have got to talk about people dealing with people in a business context. Does that lady have an opinion that she is going to offer in this case?
Mr. Huron: Yes, sir.
The Court: Why doesn’t she give me her opinion? and then tell me what she bases it on.
Mr. Huron: Fine. We were trying to—
The Court: And if we did that then I think I would have a better understanding of where you are getting.
Mr. Huron: Dr. Fiske, have you examined whether stereotyping was occurring at Price Waterhouse; I am talking about sex role stereotyping, at the time and in connection with Ann Hopkins' proposal for partnership which began in August '82 until she was placed on hold in March of 1983?
The Witness: I have examined evidence related to that.
Mr. Huron: Have you formed an opinion as to whether or not stereotyping was occurring?
The Witness: Yes, I have.
Mr. Huron: And what is your opinion?
The Witness: I am confident that stereotyping played a role in the decision about Ann Hopkins.
The Court: Well, now, what kind of role and how confident?
Are you able to say that you are confident within a reasonable degree of certainty in your discipline?
The Witness: Yes, I would say so, given—
The Court: All right. That is what I want to know. And then you said it played some part. What part? I don’t know how you would express it in your discipline percentagewise or how you would express it, but minor, major, middle? I don’t know what the terminology is.
The Witness: Well, in lay language I would say it played a major determining role.
The Court: A major determining role, with reasonable certainty?
The Witness: Yes.

Judge Gesell ultimately ruled that an “employer that treats [a] woman with [an] assertive personality in a different manner than if she had been a man is guilty of sex discrimination” (Hopkins v. Price Waterhouse, 1985, p. 1119). He described the firm’s decision-making process as “tainted by sexually biased evaluations” (p. 1120). With regard to Fiske’s testimony, he noted that “a far more subtle process [than the usual discriminatory intent] is involved when one who is in a distinct minority may be viewed differently by the majority because the individual deviates from an artificial standardized profile” (p. 1118). He went on to say that the firm’s “partnership evaluation system permitted negative comments tainted by stereotyping to defeat her candidacy, despite clear indications that the evaluations were tainted by discriminatory stereotyping” (p. 1118). In a footnote, he added, Common sense is confirmed by the literature on the problem of sex stereotyping which suggests that making evaluators aware of the risks of biased evaluations and inquiring as to whether the generalizations are supported by concrete incidents can be effective in eliminating or minimizing stereotyping. (p. 1120, note 15)

The Court of Appeals Decision
Price Waterhouse appealed Judge Gesell’s decision, and the written record of the trial was reviewed by a three-judge panel of the U.S. Court of Appeals for the District of Columbia. Price Waterhouse argued that the social psychology testimony was “sheer speculation” of “no evidentiary value” (Price Waterhouse v. Hopkins, 1987, p. 467).

The majority on the appeals court disagreed, ruling that “partners at Price Waterhouse often evaluated female candidates in terms of their sex . . . the partnership selection process at Price Waterhouse was impermissibly infected by stereotypical attitudes towards female candidates” (Price Waterhouse v. Hopkins, 1987, p. 468). Specifically regarding the testimony, the appeals court stated that “convergent indicators of stereotyping” . . . taken together provided Dr. Fiske a sufficient basis from which to draw her conclusions that Hopkins was a victim of stereotyping. To the extent Price Waterhouse believes Dr. Fiske lacked necessary information, the firm is in fact quarreling with her field of expertise and the methodology it employs. (p. 467)

However, the sole dissenting appeals judge described Fiske as someone “purporting to be an expert” (Price Waterhouse v. Hopkins, 1987, p. 477) in the field and protested “the remarkable intuitions of Dr. Fiske” (p. 478). His dissent implied that the argument was about the accuracy of perceptions of Hopkins’s personality, not the stereotypic manner in which those perceptions were cast. For example, he argued “To an expert of Dr. Fiske’s qualifications, it seems plain that no woman could be overbearing, arrogant, or abrasive: any observations to that effect would necessarily be discounted as the product of stereotyping” (p. 477). He noted that without information about the so-called truth of the matter, “Dr. Fiske’s expertise rose to the occasion. Her arts enabled her to detect sex stereotyping based largely on the ‘intensity of the negative reaction’” (p. 477).

Issues in the Supreme Court
Because they lost at the appellate court level, PW asked the Supreme Court to review the case. Review by the nation’s highest court is discretionary. It agrees to hear only about 150 of the 3,000 cases it is asked to review each year. But because the appellate court decisions in Hopkins and in similar cases had been in conflict, the Supreme Court accepted this case for review. The central legal issues that made this case ripe for review concerned the proper allocation of the burden of proof between the employer and employee, as well as the proper standard of proof required to support allegations of discrimination.
in what are called mixed-motive cases, cases in which the refusal to hire or promote is based on both legitimate and illegitimate reasons. The burden-and-standard-of-proof issues are irrelevant to this article, but some discussion of the nature of mixed-motive cases is necessary to understand the context of social science evidence in this case.

Essentially, the mixed-motive issue asks whether it is permissible to refuse to hire or promote someone on the basis of membership in a protected category (e.g., sex, race, age) if the person also has actual areas of significant incompetence. In effect, one might argue that if people have real performance problems, it does not matter if they lose a job or promotion partly because of their race or gender. One might argue that the ultimate outcome is the same in either case, whether the grounds were stereotyping or performance. However, this argument would imply that one can only identify discrimination when the person is otherwise perfect. Most employees are not perfect, having records that, if closely examined, could indicate flaws. Is it permissible then to discriminate against the vast majority of people who are genuinely flawed in some important way? This question is identified in discrimination cases as the mixed-motive issue; it formed the basis of PW's petition to the Supreme Court.

The mixed-motive issue has important practical consequences. Can one discriminate against someone who has genuine problems, or do the targets of such discrimination have to prove that they were so exemplary that they would have been guaranteed the position if discrimination were absent? The mixed-motive question asks whether all else has to be above reproach for a person to prove discrimination. Alternatively, one could argue that decision makers can never use discriminatory motives in personnel judgments, even in the case of someone with known shortcomings. In this case, the defense was that if Ann Hopkins had actual interpersonal skills problems, she could not also be a victim of discrimination. The Supreme Court thought this a sufficiently important point of employment discrimination law that it ought to be resolved.

The APA Amicus Curiae Brief

The APA decided to enter the case for two reasons. First, the crucial finding of discrimination in this case, as characterized by the court of appeals, was grounded on direct evidence that the employer's selection process "was impossibly infected by stereotypical attitudes toward female candidates" (Price Waterhouse v. Hopkins, 1987, p. 468). As a result, the parties were forced to focus some of their attention on the issue of sex stereotypes. However, they lacked the necessary social science background to present the issue of stereotyping in an empirically based, legally relevant manner. Second, PW consistently disparaged Fiske's testimony by criticizing the methodology and the concepts she used in arriving at her expert opinion that PW discriminated against Hopkins on the basis of sex. APA informed the Court that the methodology and literature Fiske used were consonant with generally ac-
cepted research practice. APA was careful not to support Fiske personally as an expert witness or as a researcher. Thus, as a good amicus should, APA addressed an issue (the psychology of stereotyping, as relevant to this case) that the parties did not have adequate space, time, or expertise to discuss.

Price Waterhouse had conceded that the partners' negative comments "might conceivably be taken as indicating that stereotypical thinking was sometimes present 'in the air' at Price Waterhouse" (Brief for Petitioner PW, 1988, p. 48), but throughout their briefs they disparaged the psychology of stereotyping. They placed the term expert in quotation marks, in a belated effort to discredit the validity of research on stereotyping. In addition, they placed the term sex stereotyping within quotation marks, falsely implying that it is an unaccepted neologism, and they characterized as an amorphous proposition the appeals court finding that the employer discriminated against Ms. Hopkins because of "stereotypical attitudes" (p. 15). They claimed that the finding was derived from "intuitions about unconscious sexism—discernible only through an 'expert' judgment" (p. 17). In addition to labeling Fiske's opinion as "gossamer evidence" and "intuitively divined" (p. 45), PW claimed that Fiske's conclusions were faulty because she never met Hopkins and only reviewed the partners' evaluations of her. Price Waterhouse accused the lower courts of basing a finding of intentional discrimination on a "chain of intuitive hunches about 'unconscious' sexism" which "were, in turn, magically transformed into evidentiary 'facts' by a shift in the burden of persuasion" (p. 44).

The APA's intent in its brief was to disabuse the Court of the notion that sex stereotyping was not an identifiable and legally cognizable source of sex discrimination prohibited by Title VII, and to inform the court of the scientific validity of the methods and literature used in Fiske's testimony. Moreover, the goal of the brief was to represent the research literature as understood by the researchers who provided the material. The implications of the assessment did favor a particular side of the case, and the amicus used the facts of the case to illustrate the relevance of the sex-stereotyping literature, when appropriate. If the implications of the literature had instead favored PW, in the judgment of the researchers, that perspective would have been presented.

The APA's amicus curiae brief was drafted by a panel of social and industrial/organizational psychologists—Eugene Borgida, Kay Deaux, Susan Fiske, and Madeline Heilman—and the final product was compiled, integrated, and rewritten by then-APA-counsel Donald Bersoff. The brief argued for the validity of the field of stereotyping in general, and sex-stereotyping research in particular. The brief noted, first, that empirical research on sex stereotyping has been conducted over many decades and is generally accepted in the scientific community; second, that stereotyping can create discriminatory consequences for stereotyped groups; third, that some conditions that promote stereotyping were present in the firm's work setting; and finally that the firm took
no effective steps to reduce discriminatory stereotyping, although such methods were available.

**Oral Arguments and the Court’s Decision**

On October 31, 1988, the Supreme Court heard oral arguments in *Price Waterhouse v. Hopkins*. Kathryn Oberly argued the case for PW, James Heller for Hopkins. The two lawyers had each presented about a minute of their 30-minute prepared statements when they were interrupted and peppered with questions by the justices for the remainder of their allotted times.

On May 1, 1989, the Court handed down its decision. The vote was six to three. Justice Brennan delivered the opinion, in which Justices Blackmun, Marshall, and Stevens joined; Justices White and O’Connor filed concurring opinions. Justice Kennedy filed a dissenting opinion, in which Justices Rehnquist and Scalia joined. The Court found in favor of Hopkins on the mixed-motive issue and in favor of PW on the standard-of-proof issue. That is, in mixed-motive cases, it is not permissible for employers to use discriminatory criteria, and they (not the plaintiff) must bear the burden of persuading the trier of fact that their decision would have been the same if no impermissible discrimination had taken place. However, the court also said that PW had been held to too high a standard of proof (i.e., clear and convincing evidence) and that Judge Gesell should review the facts to see whether it could win under a less stringent (preponderance of evidence) standard. In all, it was viewed primarily as a victory for Hopkins, although the case was returned to Judge Gesell so that he could decide whether PW was liable under the lower standard of proof.

APA contributed to making explicit, for the first time in any Title VII case, that sex stereotyping was a form of sex discrimination. The Court’s plurality specifically criticized PW’s placement of the phrase sex stereotyping in quotation marks. The plurality said, “such conduct seems to us an insinuation that such stereotyping was not present in this case or that it lacks legal relevance. We reject both possibilities” (*Price Waterhouse v. Hopkins*, 1989, p. 1791). The plurality went on to endorse the importance of sex stereotyping: “In forbidding employers to discriminate against individuals because of their sex, Congress intended to strike at the entire spectrum of disparate treatment of men and women resulting from sex stereotypes” (p. 1791, quoting *Los Angeles Dept. of Water & Power v. Manorhills*, 1978, p. 707). More specifically, APA had argued that sex-stereotypic prescriptive demands to be feminine simultaneous with job-specific demands to be aggressive place women in a double-bind situation; this was apparently convincing to the court in their framing of Hopkins’s dilemma as a “Catch 22.”

The Court did not dismiss Fiske’s expert testimony on grounds cited by PW that she did not personally interview Hopkins. In characterizing this criticism as scientifically naive and irrelevant, APA stated that PW confused the work of research psychologists such as Fiske with that of clinical psychologists who use interviews and other assessment devices to diagnose a patient. The issue in Fiske’s testimony was the presence of discriminatory stereotyping at PW, not the mental status of Hopkins. The proper focus, therefore, was on the conduct of PW partners, reflected in the stereotypic phrasing and content of their evaluations. Fiske brought precisely that focus when she evaluated the conditions at PW that evoke stereotyping and the nature of the comments by its partners in light of the research literature on stereotyping. The plurality agreed and refused to countenance the PW derogation of Fiske’s testimony. The Court said,

> We are not inclined to accept petitioner’s belated and unsubstantiated characterization of Dr. Fiske’s testimony as “gossamer evidence” based only on her “intuitive hunches” and her detection of stereotyping as “intuitively divined.” Nor are we inclined to accept the dissent’s dismissive attitude toward Dr. Fiske’s field of study and toward her own professional integrity. (*Price Waterhouse v. Hopkins*, 1989, p. 1793)

Instead, they said,

> Indeed, we are tempted to say that Dr. Fiske’s expert testimony was merely icing on Hopkins’ cake. It takes no special training to discern sex stereotyping in a description of an aggressive female employee as requiring “a course at charm school.” Nor . . . does it require expertise in psychology to know that, if an employee’s flawed “interpersonal skills” can be corrected by a soft-hued suit or a new shade of lipstick, perhaps it is the employee’s sex and not her interpersonal skills that has drawn the criticism. (*Price Waterhouse v. Hopkins*, 1989, p. 1793)

One can interpret this comment in various ways: as dismissive, saying that the social science testimony was all common sense; as merely taking the social psychological expertise for granted; or as suggesting that one does not necessarily require expert witnesses to identify stereotyping when the evidence is egregious.

Finally, in response to the argument that Ann Hopkins might really be an obnoxious person, as described by some acquaintances, the Court agreed with Judge Gesell that,

> The reactions of at least some of the partners were reactions to her as a woman manager. Where an evaluation is based on a subjective assessment of a person’s strengths and weaknesses, it is simply not true that each evaluator will focus on, or even mention, the same weaknesses. Thus, even if we knew that Hopkins had “personality problems,” this would not tell us that the partners who cast their evaluations of Hopkins in sex-based terms would have criticized her as sharply (or criticized her at all) if she had been a man. It is not our job to review the evidence and decide that negative reactions to Hopkins were based on reality; our perception of Hopkins’ character is irrelevant. We sit not to determine whether Ms. Hopkins is nice, but to decide whether the partners reacted negatively to her personality because she is a woman. (*Price Waterhouse v. Hopkins*, 1989, pp. 1794–1795)

One could not have asked for a better understanding of the psychology of stereotyping. The Supreme Court concluded,

---

2 For a review of the impact of other social science amicus curiae briefs, see Acker (1990).
In the specific context of sex stereotyping, an employer who acts on the basis of a belief that a woman cannot be aggressive, or that she must not be, has acted on the basis of gender.... We are beyond the day when an employer could evaluate employees by assuming or insisting that they matched the stereotype associated with their group. ... An employer who objects to aggressiveness in women but whose positions require this trait places women in an intolerable Catch 22: out of a job if they behave aggressively and out of a job if they don't. Title VII lifts women out of this bind. (pp. 1790-1791)

The State of the Field, as Revealed by the Amicus Process

Strengths

It is important to give credit where credit is due. In considering material for the brief, we were impressed with the advances made and knowledge accumulated during the past decade or so. Indeed, we gave the most weight to that material. Social scientists (and especially psychologists) have contributed significantly to an understanding of the nature and operation of stereotypes. Areas in which increments in knowledge are most pronounced include (a) the nature of categorization processes, including the development of subtypes; (b) conditions that encourage stereotyping; and (c) the influence of stereotypes on performance evaluation and causal explanation (Deaux, 1985; Deaux & Kite, in press). The work on stereotyping is an excellent example of basic research that can have practical impact, both in analyzing the current situation and in suggesting future courses of action.

The cognitive categorization approaches to stereotyping are particularly well suited to courtroom use. That is, the cognitive approach by definition concentrates on people's thoughts, not on their feelings, which means the type of evidence one examines is highly compatible with the kinds of evidence available in the courtroom. We do not have to undertake a speculative analysis of an individual's authoritarian personality or of an entire culture's stance toward diversity; the cognitive analysis is more straightforward.

Moreover, laboratory research has some special virtues for training experts in this kind of endeavor. Beginning with the antecedents of stereotyping, laboratory researchers are well suited to analyze the situation, looking for factors that encourage or discourage stereotyping. Training in experimental methods teaches one to divide a social situation into its critical components. In conducting laboratory research, one learns to isolate and manipulate the significant situational factors. Moreover, the same compatibility holds in examining the indicators of stereotyping. First, psychological researchers, particularly those with a cognitive bent, are used to dealing with the written record, hard evidence of people's judgmental processes. Second, evidence suggesting how people categorize another person and evidence pointing to the attributes considered important in decision making are indicators that are often apparent in the written record. When a person labels another person, the written record or reports of conversations provide evidence that is similar to the responses of subjects in a laboratory experiment. Turning to the consequences of stereotyping, there, too, laboratory researchers are well suited to identify sensitive measures and interpret the meaning of a pattern of response. Finally, in suggesting remedies, many laboratory researchers have built their research programs around finding ways to undercut stereotyping. Experimenters are trained to deepen their understanding of an effect by making it come and go through various interventions. Laboratory scientists often assume that the interventions may be effectively translated into the organizational mold. For example, Fiske's research on interdependence and stereotyping has been heavily influenced by the kinds of structural features that operate in organizations and might usefully generalize to those contexts. This type of laboratory research on basic processes then converges with field research on real-world effects of interventions.

Weaknesses

More work is needed, for both theoretical and practical reasons. There is still a need to test more laboratory findings in field settings and to use such findings to revise current theories. In turn, the key dynamics in field settings need to be incorporated systematically into laboratory experiments.

A critical issue for managers as well as attorneys is how to minimize the use of stereotypes. The psychological literature is strong in showing how pervasive stereotypes are, in demonstrating the myriad conditions that promote them, and in showing how resistant they can be to new information. Relatively few research programs explore the conditions that discourage stereotyping. Psychologists need to work further to discover when stereotypes do and when they do not operate. Such efforts might investigate deterrent strategies, the influence of observers or models on stereotyping, and the broader contexts that may emphasize or minimize stereotypes. These issues, all raised by the Hopkins case, have not been addressed adequately on an empirical basis.

Specific questions of cognitive process are also raised by this case. For example, how are two or more stereotypic categories combined (e.g., manager and woman)? What are the circumstances under which people use alternative subtypes (e.g., aggressive career woman) and more general categories (e.g., women)? The case specifically raises the issue of whether Hopkins was considered an exemplar of "aggressive career woman" or simply as a deviant from both categories, "woman" and "manager." The continuing debate on perceivers' use of category prototypes versus exemplars is relevant here (see Fiske & Taylor, 1991, chap. 4, for a review).

Perhaps the central issue raised by this case is the distinction between descriptive and prescriptive aspects of stereotypes. This issue is particularly relevant to gender stereotyping, but it arises in other contexts as well. Although it is routinely accepted that sex stereotypes have both descriptive and normative components (e.g., Tergborg, 1977), research has been focused primarily on the descriptive component. It has been concentrated on the
following four issues: the attributes that constitute sex stereotypes; the processes by which they are ascribed to men and women; the consequences of such ascriptions on expectations, attributions, and evaluations; and the conditions that regulate their occurrence. Underlying this research is a working assumption that if psychologists can determine how and when stereotypic attributes are used to characterize an individual man or woman, differential treatment can be averted.

The *Hopkins* case forces a reexamination of that assumption. Hopkins was acknowledged as competent, committed, hard-working, and effective. The extent of her accomplishments was not the main point of contention; her talents were not denied (although they may have been underrated), and the attributes considered part of the male stereotyped “competency cluster” were accepted. In short, in terms of work-relevant attributes, she did not appear to suffer the ills of a traditional stereotyped description (e.g., passive, weak, indecisive), and yet she was a victim of sex stereotyping. Why? Despite her work-related competence (or perhaps because of it), she was seen as behaving in ways that are considered inappropriate for women. She met certain expectations, but not others, and she was expected to meet both.

The prescriptive aspects of sex stereotypes (i.e., the behaviors deemed respectively suitable for men and women) are clearly important. Nevertheless, in reviewing the literature for the brief, we found surprisingly little attention focused on this issue. Although there is evidence demonstrating that behaving out of sex role has unfavorable consequences and that women and men who violate the “shoulds” of the sex-stereotypic prescription are found to be objectionable, few attempts have been made to systematically explore this larger phenomenon of prescriptive stereotyping. Even fewer attempts have been made to address the question of how to prevent or remedy it. The interventions may be of a different sort than those targeting descriptive aspects of stereotyping. Here the dynamics are likely to involve motivational and affective factors as well as cognitive ones.

How Complete Does the Database Have to Be?

In her testimony, Fiske based her assertions regarding the role of stereotyping on a well-established and thriving literature. Her assessment was grounded in decades of research. Moreover, the use of such “social framework” testimony is likely to continue in various areas of psychology (Goodman & Croyle, 1989). How broad and deep does the database have to be for psychologists to render expert opinions? A crucial concern emerges about the adequacy of scientific databases as social science is increasingly used in the legal system. The number of APA-sponsored briefs has increased, and their influence has been examined (Acker, 1990; Roesch, Golding, Hans, & Repucci, 1991). The increased use of social science data in law, whether to address legislative or adjudicative matters, has generated important conceptual frameworks and evaluative criteria based on peer review standards for evaluating the scientific adequacy of the database in question (e.g., Monahan & Walker, 1988).

Issues about the quality of the database, which also pertain to the use of social science evidence in the *Hopkins* case, have long been debated with regard to research on eyewitness identification (Kassin, Ellsworth, & Smith, 1989; Loftus, 1983; McCloskey & Egeth, 1983). More recently, this issue has emerged as central to the debate regarding the role of APA’s amicus curiae brief in death penalty cases, most notably *Lockhart v. McCree*, 1986 (see Bersoff, 1987). Again, the adequacy of the scientific database was at the heart of the controversy. Elliott (1991) took issue with APA’s *Lockhart* amicus position that “that stability and convergence of the findings over three decades lends impressive support to their validity” (Bersoff, 1987, p. 68). Elliott was quite critical of the database and argued that the team of experts assembled at APA may have exaggerated the conclusiveness of the database in order to have an impact on legal policy. Ellsworth (1991) strongly rebutted Elliott’s criticisms of the research database and argued that the database was adequate by generally accepted peer review standards, that convergent validity was established in this research domain, and that Elliott’s claim (i.e., scientific authority was overstated) was unfounded. Most important, Ellsworth argued that “to keep silent until our understanding is perfect is to keep silent forever” (p. 277). The team that developed the amicus position in the *Hopkins* case would strongly endorse Ellsworth’s perspective. The amicus experts’ favorable assessment of the sufficiency of the scientific database in the *Hopkins* case, as in *Lockhart*, was determined by the convergent validity of the scientific research literature.

The Interplay Between Research and Testimony

A critic might argue that researchers who have contributed to a particular literature should not testify as experts or contribute to amicus briefs because they have a stake in the validity of that literature, and they might overlook the rough spots or misrepresent the degree of consensus in the state of the research. These ethical issues have been a concern from the outset of the expert testimony, in the drafting of the amicus brief, and in the composition of this article.

Regarding the testimony of involved experts, it is difficult to imagine who is more qualified to testify than someone whose primary activity is to conduct research on the topic at hand. Certainly, perennial expert witnesses who conduct little research are not more credible than occasional expert witnesses whose primary identity lies in science. The main motivations for scientists to be involved in such expert testimony are to get well-established research literatures into the relevant legal settings and to receive feedback about the utility of their work. As researchers apply their expertise to legal issues, further research and greater expertise is inspired. Such a cycle must commence with a well-developed research area before it can be exported, but the point is that intellectual involve-
ment of experts ultimately strengthens both research and its application to law. 3

There are, of course, several safeguards against deliberately or inadvertently misleading legal applications. First, there is the conscience of the individual researcher. This can only be known in the privacy of one's own thoughts, but ultimately there must be some degree of trust in the honor of our colleagues. One's reputation as a scientist is far more important than one's winning any one case. Without peer credibility, a scientist cannot carry on the requisite professional activities that enable his or her continued career.

Another potential criticism of such legal application comes not from misrepresentation of the literature, but in application to a particular setting. The role of expert testimony, whether psychological, medical, statistical, engineering, or whatever, is to discuss a research literature, not to define or address the legal issues involved. The trier of fact (judge or jury) makes the legal decision. The role of the testimony is to be helpful to the trier of fact by providing information about an area of human behavior that might not be familiar. Specifically, expert testimony (a) represents the state of knowledge in the relevant field and (b) applies it to the facts of the particular case. Certainly, there are issues of expert judgment and informed opinion at both stages of the process. But why is the psychology of stereotyping ultimately any different from other areas of admissible expertise? We would argue that there are no fundamental differences that prevent stereotyping researchers from qualifying as experts, representing the database, and applying their knowledge to a particular case.

Within an adversarial system, another safeguard is the alternative view posed by the actual or potential opposing expert, by the judge's neutral expert (a strategy sometimes used for statistical testimony), and by the actual or prospective opposing briefs. The whole premise of an adversarial system is that the most veridical evidence, scientific or otherwise, is the most convincing. Although adversarial safeguards on expert testimony and amicus briefs are not perfect, it is a premise of the legal system that distortion of the data will come out, one way or another.

Turning more specifically to the selection of the amicus drafters, deliberate efforts were made to represent a cross section of expertise and opinion. The APA Legal Affairs Office approved the final composition of the group. Fiske was included because of her work on stereotyping and her intimate knowledge of the case; Deaux was included for her expertise about stereotyping based on gender per se; Heilman was included because of her knowledge of gender stereotyping in work contexts; and Borgida was included for his knowledge of forensic psychology and because his published work has taken a devil's advocate position on some issues in gender stereotyping. All four have published extensively in the relevant areas, have held editorial positions on journals, and have been competitively funded for their research in these areas. Thus, the institutional controls (APA) and the orientations of the individual researchers both worked to assure a balanced perspective on the literature.

One might argue that the amicus drafters, once chosen, were then driven by politics. However, expertise about the effects of stereotyping is not necessarily less neutral than expertise about the effects of a bullet to the brain. Medical experts can have political or self-serving motivations just as easily; they would fail to qualify if it could be shown that they were primarily motivated to push a policy agenda (e.g., gun control) or to become famous (through involvement in a notorious case despite private misgivings). Yet, one might argue, the researchers representing the psychology of stereotyping are somehow less neutral because they may be potential victims of stereotypes themselves and so are not necessarily neutral. By this argument, female scientists studying gender stereotypes have reduced credibility. As this argument implies, it would be like asking someone who lives in a neighborhood with a high crime rate to serve as an expert on the effects of fatal bullet wounds. Clearly, if such a person had scientific expertise in the topic, as established by peer review and reputation, the person's private life would not be relevant. Obviously, this hypothetical expert would not be called in to testify as an expert against a spouse's murderer, but neither is the stereotyping expert called in to testify in a personal grievance. Careful methods, convincing data, reasonable argument, and debate among peers are the bulwark against bias in science. Careful methods, convincing data, reasonable argument, and debate among peers all protect against bias by experts testifying and drafting amicus briefs.

We are not arguing that the database is perfect or that as researchers we are indifferent to the topics we study or their policy implications. Nor are we arguing that expert opinions may not differ. We are arguing that psychologists' considered and self-critical judgments as researchers were reflected in the use of psychological expertise at trial and in the amicus process in this case.

Conclusions

It is an extraordinary opportunity to have psychological research, in an area so well established and thriving, be confronted by some of the most prominent legal minds in the country. We think psychology came out rather well, on the whole. But we also think, as psychologists, that we have our work cut out for us. To enhance our credibility, we must continue to test our theories in a range of laboratory and field settings. We must attend to those factors in the real world that demand research attention: How

---

3 As concrete examples of the benefits to research, we know of one meta-analysis that cited Hopkins as demonstrating some relevant dimensions for analysis (Eagly, Makhijani, & Klonsky, in press) and one archival study of solo status and performance evaluations apparently inspired by Hopkins (Sackett, DuBois, & Nee, 1991). Moreover, in our own work, one research program has been heavily influenced by Hopkins (Fiske, 1989a, 1989b), and one literature review has described the Hopkins case as marking "a rite of passage for research on gender stereotypes" (Deaux & Kite, in press).
can decision makers best guard against the incursions of stereotypic judgments and consequent discrimination? How can people control their stereotypic thinking? What situational and organizational factors influence stereotyping? What are appropriate remedies? Psychology is well equipped to take on these challenges. Meanwhile, the role of expert testimony and amicus briefs is to educate and to bring the pertinent research to the attention of the trier of fact, and then let the courts decide.

REFERENCES


