

# HEINONLINE

Citation: 76 Cal. L. Rev. 877 1988

Content downloaded/printed from  
HeinOnline (<http://heinonline.org>)  
Mon Oct 11 13:13:05 2010

- Your use of this HeinOnline PDF indicates your acceptance of HeinOnline's Terms and Conditions of the license agreement available at <http://heinonline.org/HOL/License>
- The search text of this PDF is generated from uncorrected OCR text.
- To obtain permission to use this article beyond the scope of your HeinOnline license, please use:

[https://www.copyright.com/ccc/basicSearch.do?  
&operation=go&searchType=0  
&lastSearch=simple&all=on&titleOrStdNo=0008-1221](https://www.copyright.com/ccc/basicSearch.do?&operation=go&searchType=0&lastSearch=simple&all=on&titleOrStdNo=0008-1221)

# Social Facts: Scientific Methodology as Legal Precedent

Laurens Walker†  
John Monahan‡

*In this Essay, Professors Walker and Monahan examine the use of social science research data to prove contested facts in judicial proceedings. They argue that a court's decision to rely upon social science research should have precedential value only to the extent that it determines whether the social science methodology used is or is not a legally acceptable way to prove a given claim. They further argue that the particular application of a given methodology in a single research effort should not be accorded precedential force.*

*The Authors demonstrate that social science research is commonly used as "social fact" in a broad variety of legal settings and should be subject to the normal rules of evidence law. Applying the Federal Rules of Evidence, they evaluate the precedential force of various aspects of social science research in judicial proceedings. Through this analysis, the Authors draw a distinction between social science methodology and applications of that methodology—concluding that the former is more enduring and general. Accordingly, the Authors contend that only methodologies, and not applications of methodologies, should be accorded precedential force as legally accepted ways to prove certain claims. The Essay concludes that, for the purpose of defining the applicable precedent, methodologies should be narrowly construed at present to produce specific rules of limited generality. This proposal would allow courts to draw upon a variety of social science tools and gradually to develop a set of rules that properly reflect the great breadth of social science research capabilities.*

## INTRODUCTION

In a remarkable decision, the United States Court of Appeals for the District of Columbia Circuit wrestled with some of the most controver-

---

† T. Munford Boyd Professor of Law, University of Virginia

‡ Henry and Grace Doherty Professor of Law, University of Virginia.

We are grateful to our colleagues David Baldus, Richard Bonnie, William Gardner, David Kaye, Edmund Kitch, N. Dickon Reppucci, George Rutherglen, Stephen Saltzburg, Michael Saks, Robert Scott, and Eric Turkheimer for their comments on drafts of this Article, and to our student assistants Randy Broberg, Preston Burton, John Cooper, Don Culkin, Tom Fina, Ella Frederiksen, Sally Groseclose, Robert Hupe, Steve Koldin, Sarah Mills, Pamela O'Hanlon, Roy Oser, Karen Robertson, Fred Wager, and Page Wittkamp for their research, and to Michael G. Colantuono for his editorial assistance.

sial and confusing issues on the border of law and social science. In *Palmer v. Shultz*,<sup>1</sup> a class of female Foreign Service Officers brought suit against George Shultz as Secretary of State alleging sex discrimination in violation of Title VII of the Civil Rights Act of 1964.<sup>2</sup> The plaintiffs offered data showing disparities between the assignments and promotions given male and female Foreign Service Officers, along with statistical analyses in order to demonstrate the improbability that these disparities resulted by chance. The district court disparaged this empirical evidence and concluded that no discrimination had been proven.<sup>3</sup>

The court of appeals, however, ruled that the district court had "committed a number of legal errors"<sup>4</sup> including the application of the wrong test of statistical significance to the plaintiffs' data.<sup>5</sup> The court of appeals noted that not only was the district court bound by "legal principles"<sup>6</sup> derived from several previous court of appeals decisions regarding statistical evidence, but it also was itself constrained by two prior decisions of the United States Supreme Court, *Castaneda v. Partida*<sup>7</sup> and *Hazelwood School District v. United States*.<sup>8</sup> Chief Judge Patricia Wald stated:

[G]iven the language of the Supreme Court in *Castaneda* and *Hazelwood*, we do not believe that we can allow the threshold at which statistical evidence alone raises an inference of discrimination to be lower than 1.96 standard deviations. . . .if plaintiffs in Title VII cases are ever to be allowed to establish a prima facie case by evidence of disparity measuring lower than 1.96 standard deviations, this decision under the current law must be made by the Supreme Court (or Congress).<sup>9</sup>

*Palmer* is just the latest and most striking in a series of cases<sup>10</sup> that

---

1. 616 F. Supp. 1540 (D.D.C. 1985), *rev'd*, 815 F.2d 84 (D.C. Cir. 1987).

2. 42 U.S.C. § 2000e (1982).

3. 616 F. Supp. at 1549.

4. 815 F.2d at 116.

5. *Id.* at 95.

6. *Id.* at 90.

7. 430 U.S. 482 (1977).

8. 433 U.S. 299 (1977).

9. *Palmer*, 815 F.2d at 96 n.9. The court defined "standard deviation" as follows:

The "standard deviation" is a unit of measurement that allows statisticians to measure all types of disparities in common terms. Technically, a "standard deviation" is defined as "a measure of spread, dispersion, or variability of a group of numbers equal to the square root of the variance of that group of numbers." D. BALDUS AND J. COLE, *Statistical Proof of Discrimination* 359 (1980) (emphasis in original). The "variance" of the group of numbers is computed by subtracting the "mean," or average, of all the numbers, "squaring the resulting difference, and computing the mean of these squared differences." *Id.* at 361.

*Id.* at 92 n.7.

10. See, e.g., *Moultrie v. Martin*, 690 F.2d 1078, 1082 (4th Cir. 1982) ("[I]n all cases involving racial discrimination, the courts of this circuit must apply a standard deviation analysis such as that approved by the Supreme Court in *Hazelwood* before drawing conclusions from statistical comparisons"); see also *EEOC v. Western Electric Co.*, 713 F.2d 1011, 1018 (4th Cir. 1983) (reaffirming *Moultrie*); *Hill v. K-Mart Corp.*, 699 F.2d 776, 780 n.7 (5th Cir. 1983) (citing

have treated seemingly technical aspects of social science research as "rules of law" with the force of precedent. Unfortunately, none of these cases has explained why technical aspects of social science research should be accorded precedential status. Nor have commentators made the case for treating aspects of social science research as precedential. Rather, they have criticized the specific rules of law developed by the courts in their treatment of empirical research on various legal questions.<sup>11</sup>

We believe that courts should attribute precedential force to the use of some social science concepts. The controversy that this practice has engendered in the past, we argue, results from the insufficient attention that courts and commentators have given to the *content* of the precedent. In their eagerness to provide guidance through a difficult and uncharted area, courts have overreached by granting precedential value to the specific findings of empirical studies that had previously received judicial approval. To remedy the tendency to overreach, we propose that these precedents be construed narrowly. Thus legal force should be given only to the essential components of relevant social science methodologies, and not to the details of how those methodologies were applied in particular cases.

In Part I, we distinguish the use of social science research data to determine facts at trial from other uses of such data in the legal system. We also illustrate the use of social science data in diverse legal contexts. To determine the precedential force of this form of factual evidence, we must first acknowledge that it is *evidence*. Accordingly, in Part II, we employ the Federal Rules of Evidence to evaluate the utility of various empirical data. In Part III, we clarify the issues raised by that evidentiary analysis by distinguishing between a methodology generic to a class of research and the application of that methodology in any particular study. In Part IV we employ this distinction to discern the precedential component of social science research. We argue that it is a research

---

*Hazelwood* for the proposition that the statistical significance of research evidence has probative value); *United States v. County of Fairfax*, 25 F.E.P. 662, 666 (E.D. Va. 1981) (using the statistical test relied upon in *Hazelwood* to find a prima facie case of discrimination).

11. See, e.g., D. BALDUS & J. COLE, *STATISTICAL PROOF OF DISCRIMINATION* § 9.4, at 188-89 (Supp. 1987) ("An easily avoided problem is the treatment of the test of statistical significance as a rule of law rather than as an aid to interpretation . . . . This approach was never intended by the United States Supreme Court in *Castaneda* and *Hazelwood* and completely misses the point that in discrimination suits, as in all other contexts, tests of statistical significance and confidence intervals do not lay down arbitrary rules for accepting or rejecting data."); Kaye, *Hypothesis Testing in the Courtroom*, in *CONTRIBUTIONS TO THE THEORY AND APPLICATION OF STATISTICS* 331, 352 (A. Gelfand ed. 1987) (Statistical tests such as those used in *Palmer* "are ill suited for courtroom applications"); Rutherglen, *Disparate Impact Under Title VII: An Objective Theory of Discrimination*, 73 VA. L. REV. 1297, 1333-35 (1987) (criticizing lower federal courts for using rigid rules to evaluate statistical evidence).

methodology, and not its application that should have precedential value. In Part V, we consider how broadly methodology should be characterized for the purpose of establishing precedent. We conclude that while the most prudent course for the present is to construe methodology narrowly, the boundaries of what constitutes legally acceptable social science methods will expand over time as the legal system gains experience with those methods. Thus we argue that courts have properly treated the use of certain social science methodologies as precedent to resolve many legal questions, but they have often misunderstood the content of those precedents.

## I

### THE USES OF SOCIAL SCIENCE RESEARCH: SOCIAL FACTS AND SOCIAL AUTHORITY

The use of social science research data as a fact-finding tool, while common in Title VII cases such as *Palmer*,<sup>12</sup> is by no means limited to this context. Additional illustrations can be drawn from a number of areas, including such disparate areas as trademark infringement and obscenity. For example, in *Processed Plastic Co. v. Warner Communications*,<sup>13</sup> Warner, owner of "The Dukes of Hazzard" television series, had licensed several toy companies to produce replicas of a car used prominently in the series. The Processed Plastic Company, which was not licensed by Warner, began to market a toy car that strongly resembled the car used in the television series. Warner introduced evidence that eighty-two percent of children it had surveyed identified a Processed Plastic car as the "Dukes of Hazzard" car.<sup>14</sup> The trial court ruled,<sup>15</sup> and the Seventh Circuit agreed,<sup>16</sup> that this data demonstrated that Warner was likely to prevail on the merits of its claim that Processed Plastic had violated the Lanham Trademark Act<sup>17</sup> by creating consumer confusion.

In another example, evidence based on social science research was offered in *Saliba v. State*,<sup>18</sup> an obscenity case. In that case, a jury convicted the defendant of distributing obscene material after a trial in which the judge excluded the testimony of an expert witness who had

---

12. For a discussion of the use of research data in Title VII cases, see D. BALDUS & J. COLE, *STATISTICAL PROOF OF DISCRIMINATION* (1980 & Supp. 1987); Rutherglen, *Claims of Employment Discrimination Under Title VII of the Civil Rights Act of 1964*, in *STATISTICAL METHODS IN DISCRIMINATION LITIGATION* 33 (D.H. Kaye & M. Aickin eds. 1986).

13. 675 F.2d 852 (7th Cir. 1982).

14. *Id.* at 854-55.

15. *Id.* at 857.

16. *Id.* at 858.

17. 15 U.S.C. § 1051 (1982 & Supp. IV 1986).

18. 475 N.E.2d 1181 (Ind. Ct. App. 1985).

conducted an opinion poll to determine community standards of "patent offensiveness." The Court of Appeals of Indiana overturned the conviction, however, holding that "[e]xpert testimony based on a public opinion poll is uniquely suited to a determination of community standards."<sup>19</sup> The court continued, "[p]erhaps no other form of evidence is more helpful or concise."<sup>20</sup>

To appreciate what is precedential about social science research in cases like *Palmer*, *Processed Plastic*, and *Saliba*, it is first necessary to distinguish between the two generic uses of social science in law: (1) its use as a fact-finding tool, and (2) its use as a tool for developing legal rules.

This distinction was first noted in 1942 by Professor Kenneth C. Davis.<sup>21</sup> Davis distinguished between "adjudicative" and "legislative" fact. He labelled "adjudicative" those facts "concerning the immediate parties—what the parties did, what the circumstances were, [and] what the background conditions were."<sup>22</sup> This application of social science to prove facts particular to a given case is sharply contrasted with the use of social science data to create or modify a rule of law, which Davis called the proof of a "legislative fact,"<sup>23</sup> and which we prefer to call "social authority."<sup>24</sup>

In *Palmer*, *Processed Plastic*, and *Saliba*, social science research was clearly introduced to prove "adjudicative" facts: the research concerned the immediate parties alone and addressed factual issues specific to their litigation. The use of research in this fashion bears little resemblance to the use of research for "legislative" or "authoritative" purposes such as in *Brown v. Board of Education*.<sup>25</sup> The research introduced in *Brown* did not concern only the immediate parties. Rather, the Court relied on the research to establish a new legal rule.

For our purposes, an essential implication of this distinction is that social science evidence used to prove case-specific facts plainly constitutes *evidence*, much like any other factual evidence produced at trial. To discern the precedential value of this form of case-specific "social fact"<sup>26</sup> evidence, we examine the principles and the rules that govern the

---

19. *Id.* at 1185.

20. *Id.*

21. See Davis, *An Approach to Problems of Evidence in the Administrative Process*, 55 HARV. L. REV. 364, 423-25 (1942).

22. *Id.* at 402.

23. *Id.*

24. Monahan & Walker, *Social Authority: Obtaining, Evaluating, and Establishing Social Science in Law*, 134 U. PA. L. REV. 477 (1986) [hereinafter *Social Authority*].

25. 347 U.S. 483 (1954).

26. We use the term "social facts" to refer to those case-specific facts that Davis referred to as "adjudicative facts." Social science research used for the purpose of creating or modifying a rule of law is also sometimes referred to as "social fact." See, e.g., T. MARVELL, *APPELLATE COURTS AND*

introduction of evidence at trial.<sup>27</sup>

## II

### SOCIAL SCIENCE RESEARCH AS EVIDENCE: A RELEVANCE ANALYSIS

Under the Federal Rules of Evidence, the analysis of the admissibility of social science research to prove a case-specific fact begins—and usually ends—<sup>28</sup> with the question of “relevance.” According to Rule 401:

“Relevant evidence” means evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.<sup>29</sup>

There are two components to this definition: To be relevant, evidence must bear on a fact that is “of consequence” to the case, and must make that fact “more probable or less probable” than it would otherwise be.<sup>30</sup> We will consider the precedential aspects of these two components

---

LAWYERS 184 (1978). However, we find this terminology inappropriate since to use research in such a way is a tool for developing legal rules and not for factfinding. See *Social Authority*, *supra* note 24. Finally, we have identified a use of social science research by courts that bears characteristics of both case-specific “social facts” and law-making “social authority.” In this hybrid use, research data are used to construct a frame of reference or context for deciding factual issues crucial to the resolution of specific cases. We call this use of social science research “social frameworks.” See Walker & Monahan, *Social Frameworks: A New Use of Social Science in Law*, 73 VA. L. REV. 559 (1987) [hereinafter *Social Frameworks*].

27. The issue addressed in *Palmer* was whether the “statistical evidence alone” sufficed to make a prima facie case of employment discrimination. The “admissibility” of the plaintiff’s statistical evidence was not challenged. Nevertheless, the Federal Rules of Evidence provide an appropriate and useful vehicle for analyzing the issues in *Palmer* since a concern about probative value characterizes both decisions about admissibility under the Federal Rules and decisions about the existence of a prima facie case under Title VII. See, e.g., Sugrue & Fairley, *A Case of Unexamined Assumptions: The Use and Misuse of the Statistical Analysis of Castaneda/Hazelwood in Discrimination Litigation*, 24 B.C.L. REV. 925, 935 (1983) (the purpose of the “Supreme Court’s reliance on formal statistical techniques in *Castaneda* and *Hazelwood* [was] to assess the probative value of data evidencing possible discrimination”); see also Boardman & Vining, *The Role of Probative Statistics in Employment Discrimination Cases*, 46 LAW & CONTEMP. PROBS., Autumn 1983, at 189, 193-201 (discussing the probative value of various methods of social science research used in employment discrimination cases).

28. Rule 403 states that the issue of relevance need not end the question of admissibility: Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.

FED. R. EVID. 403.

However, we have been unable to locate any case in which social science evidence of a fact in dispute has been held relevant under Rule 401 but nonetheless excluded under Rule 403. See *Social Frameworks*, *supra* note 26, at 575-578.

29. FED. R. EVID. 401. The identical wording is used in Rule 401 of the Revised Uniform Rules of Evidence.

30. See C. MCCORMICK, MCCORMICK ON EVIDENCE § 185, at 541 (3d ed. 1984) (observing

in turn.

### A. Materiality

While it was once mired in the strictures of pleading, the question of whether evidence is “of consequence” or “material” to a fact in issue is now generally treated as a component of relevance.<sup>31</sup> Materiality concerns the relation between the proposition that the evidence is offered to demonstrate and the factual issues in the case. If the evidence is offered to prove a proposition that is not at issue, the evidence is immaterial and therefore irrelevant.

Whether proffered evidence is material depends upon the legal issue before the court: “[w]hat is ‘in issue,’ that is, within the range of the litigated controversy, is . . . controlled by the substantive law.”<sup>32</sup> In the context of social fact evidence, this means that even research flawlessly executed is inadmissible if the substantive law governing the case does not put in issue the fact that the research seeks to establish.

Examples of successful challenges to the materiality of social fact evidence can be found in the three areas of substantive law discussed above. In Title VII cases, studies comparing the percentage of minority teachers in a school district with the percentage of minority students in the district have been held not material to the determination of racial discrimination in hiring. Rather, studies must compare the percentage of minority teachers in the district with the percentage of minority teachers in the labor market.<sup>33</sup> In trademark litigation, research demonstrating that the general public is confused about the source of a product has been held not material to the determination of trademark infringement; only confusion of consumers or potential consumers is germane.<sup>34</sup> In obscen-

---

that there are “two components” to relevance); James, *Relevancy, Probability and the Law*, 29 CALIF. L. REV. 689, 690-691 (1941) (discussing “a double meaning in the requirements of relevancy”).

31. See R. LEMPERT & S. SALTZBURG, *A MODERN APPROACH TO EVIDENCE*, at 150 n.5 (2d ed. 1982) (“The concept of materiality is carried forward in the language ‘of consequence to the determination of the action’ in F.R.E. 401. Although some judges still use the term ‘material,’ in federal courts it should be clear by now that an objection on relevance grounds may include objections to both the logical relevance and the materiality of opposing evidence”); see also *United States v. Carriger*, 592 F.2d 312, 315 (6th Cir. 1979) (observing that the Advisory Committee Note to Rule 401 criticizes the term “material” as “loosely used and ambiguous”).

32. C. McCORMICK, *supra* note 30, § 185, at 541.

33. In *Hazelwood School District v. United States*, 433 U.S. 299 (1977), the Court stated: [T]he District Court’s comparison of Hazelwood’s teacher work force to its student population fundamentally misconceived the role of statistics in employment discrimination cases. The Court of Appeals was correct in the view that a proper comparison was between the racial composition of Hazelwood’s teaching staff and the racial composition of the qualified public school teacher population in the relevant labor market.

*Id.* at 308.

34. Support for this rule can be drawn from Judge Wyzanski’s statement in *American Luggage Works, Inc. v. United States Trunk Co.*, 158 F. Supp. 50, 53 (D. Mass. 1957), *aff’d sub nom.* *Hawley*

ity prosecutions, surveys used to determine which publications or films are patently offensive must address the standards of the appropriate local community in order to be material.<sup>35</sup>

As these examples illustrate, questions of materiality are questions of substantive law, not questions of social science. Even without inquiring into the empirical merits of the evidence, it is proper for a court to exclude social science research that does not purport to demonstrate a fact with which the substantive law is concerned. Since judicial decisions based on the materiality prong of Rule 401 interpret the substantive law governing a case, they have the same precedential value as any other decisions of substantive law. Indeed, decisions on the materiality of social science evidence have uniformly been treated as precedent by subsequent courts.<sup>36</sup> Therefore, the second component of relevance must be examined in order to identify the source of the confusion surrounding the precedential nature of social science evidence.

### B. Probative Value

To be relevant, evidence must make the existence of a fact "more probable or less probable" than it was before the evidence was offered. This component of relevance is usually referred to as "probative value"<sup>37</sup> or "logical relevance."<sup>38</sup> Evidence is logically relevant "only when the probability of finding that evidence given the truth of some hypothesis at issue in the case differs from the probability of finding the evidence given the falsity of the hypothesis at issue."<sup>39</sup> In the context of social science

---

*Products Co. v. United States Trunk Co.*, 259 F.2d 69 (1st Cir. 1958): "[U]nder the substantive law the issue is not whether the goods would be confused by a casual observer (trained or untrained, professional or lay,) but the issue is whether the goods would be confused by a prospective purchaser at the time he considered making the purchase." *Cf. Rolls-Royce Motors, Ltd. v. A & A Fiberglass, Inc.*, 428 F. Supp. 689, 695 (N.D. Ga. 1977).

35. See *Smith v. United States*, 431 U.S. 291, 304 (1977) (observing that evidence of national standards would not be material); *Commonwealth v. Trainor*, 374 Mass. 796, 805, 374 N.E.2d 1216, 1222 (1978) ("The interrogation of residents of only Boston raises the question whether the survey results are competent to prove the standards of the Commonwealth as a whole.")

36. See, e.g., *D. BALDUS & J. COLE*, *supra* note 11, at § 9.4, at 188-89 (Supp. 1987) (materiality in Title VII cases); *S. KANE*, *TRADEMARK LAW* 242 (1987) (trademark cases); *Glassman, Community Standards of Patent Offensiveness: Public Opinion Data and Obscenity Law*, 42 PUB. OPINION Q. 161 (1978) (obscenity cases).

37. Rule 403 uses the phrase "probative value" to refer to the "more probable or less probable" component of relevance. See *supra* note 28; see also *C. McCORMICK*, *supra* note 30, § 185, at 542 (accord); *Friedman, A Close Look at Probative Value*, 66 B.U.L. REV. 733, 734 (1986) (discussing the concept of "probative value" in terms of probability theory).

38. See *United States v. Eckmann*, 656 F.2d 308, 312 (8th Cir. 1981); *R. LEMPERT & S. SALTZBURG*, *supra* note 31, at 153 (evidence is relevant if it "tends logically to prove or disprove some fact in issue"); see also *R. EGGLESTON*, *EVIDENCE, PROOF AND PROBABILITY* 80 (2d ed. 1983) (using the phrase "logical relevancy" and attributing it to James, *supra* note 30); *J. THAYER*, *A PRELIMINARY TREATISE ON EVIDENCE AT THE COMMON LAW* 265 (1898) (using the phrase "logically probative").

39. *Lempert, Modeling Relevance*, 75 MICH. L. REV. 1021, 1026 (1977). The notion of

evidence, this means that even if the research directly addresses a fact of central concern to the substantive law, it will not be admitted unless the research data provides insight into the likelihood that the fact exists.

This is the source of the confusion surrounding the precedential nature of social science evidence. Courts treat prior decisions on the probative value of social science evidence as if they were decisions on questions of law, with the force of precedent. They do so, however, without enunciating what aspect of social science evidence is to be treated like law, and without providing a rationale for such treatment. To clarify these issues, therefore, it is necessary to consider more precisely the nature of the social science research that is offered as evidence.

### III

#### A FOUNDATIONAL DISTINCTION

Our effort to identify what is precedential about the use of social facts turns on a distinction between two levels of abstraction in the use of the term "social science research." The first and more abstract use refers to the *methodology* of the research offered in evidence. The second and more concrete use refers to the *application* of that methodology in a particular study.<sup>40</sup>

##### A. Methodology

"Methodology" is the generic term used to characterize how social scientists go about answering factual questions. It is the overall strategy or plan of attack for generating and analyzing information. A more refined appreciation of the concept requires considering its two principal components: how information is gathered, and how information, once gathered, is interpreted.<sup>41</sup>

Social scientists refer to the methods by which they gather informa-

---

probability in this context is sometimes referred to as a "likelihood ratio." *Id.* at 1025; see also Schnm & Martin, *Formal and Empirical Research on Cascaded Inference in Jurisprudence*, 17 *LAW & SOC'Y REV.* 105, 108 (1982) (noting that in criminal cases "a likelihood ratio expresses the probability of an item of evidence assuming the defendant's guilt, relative to the probability of this same item of evidence assuming defendant's innocence").

40. See Strong, *Questions Affecting the Admissibility of Scientific Evidence*, 1970 *U. ILL. L.F.* 1, 18:

Assuming that the particular device or method used to procure scientific data is a valid one, it remains necessary to consider whether that device or method has been correctly and properly used or applied on the occasion in question. Though this point may appear self-evident, the failure to distinguish between a valid method and a valid use of that method on a particular occasion would appear to have produced more errors in the admission of scientific evidence than any other single cause.

41. See generally J. MONAHAN & L. WALKER, *SOCIAL SCIENCE IN LAW: CASES AND MATERIALS* 33-81 (1985).

tion as research designs.<sup>42</sup> A research design can be thought of as the blueprint of an empirical investigation. Researchers design their methods of collecting data to yield a true (or "valid") estimate of some state of the world—in the present context, a factual issue in contention at trial.<sup>43</sup> In *Saliba*,<sup>44</sup> for example, researchers used a survey of persons answering randomly selected telephone numbers in a given geographic area as the research design to gather information relevant to a community's standards for the depiction of sexual activities.

Once information is gathered the second principal component of methodology is interpreting that data. Typically, social scientists employ statistical methods of analysis in order to transform "raw" data gathered by a research project into a more intelligible form. The statistics necessary to interpret social science information are of two generic types: one describes the data that have been gathered, and the other allows inferences to be drawn from that data.<sup>45</sup> In *Palmer*,<sup>46</sup> for example, social scientists used one statistic, the selection rate, to describe the assignment of men and of women within the State Department, and a statistical technique, hypothesis testing, to make inferences about the plausibility that any differences in assignment were attributable to gender.<sup>47</sup>

Statistics are used initially to describe or summarize the data yielded by the research design. Often, one wants to ascertain how the "typical" subject responded. Numerical techniques that describe the data in this manner are known as measures of central tendency and include the most popularly known statistics, such as the "mean" and "median."<sup>48</sup> In addition, social scientists usually provide an index of just how typical this "typical" score is, i.e., an indication of how similar or dissimilar the scores are to one another. Numerical techniques that do this are known as measures of variability and include the "variance" and the "standard deviation."<sup>49</sup>

After social scientists use descriptive statistics to summarize their

---

42. See, e.g., A. BLALOCK & H. BLALOCK, JR., *INTRODUCTION TO SOCIAL RESEARCH* 59-60 (2d ed. 1982).

43. There are any number of factors that might compromise the accuracy (or "threaten the validity") of these estimates, and social scientists choose their research designs to eliminate, or at least to minimize, these confounding factors. See generally D. CAMPBELL & J. STANLEY, *EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS FOR RESEARCH* (1963).

44. *Saliba v. State*, 475 N.E.2d 1181, 1188-89 (Ind. Ct. App. 1985).

45. See D. BARNES AND J. CONLEY, *STATISTICAL EVIDENCE IN LITIGATION: METHODOLOGY, PROCEDURE, AND PRACTICE* (1986); R. WEHMHOFER, *STATISTICS IN LITIGATION: PRACTICAL APPLICATIONS FOR LAWYERS* (1985).

46. *Palmer v. Shultz*, 616 F. Supp. 1540 (D.D.C. 1985), *rev'd*, 815 F.2d 84 (D.C. Cir. 1987).

47. *Palmer*, 815 F.2d 84, 92 (D.C. Cir. 1987); see also Kaye, *Is Proof of Statistical Significance Relevant?*, 61 WASH. L. REV. 1333 n.2, 1356 n.7 (defining "hypothesis testing").

48. See S. HUCK, W. CORMIER & W. BOUNDS, JR., *READING STATISTICS AND RESEARCH* 22-25 (1974).

49. *Id.* at 25-30. See generally L. KIDDER, SELLTIZ, WRIGHTSMAN, & COOK'S *RESEARCH*

data in an understandable way, they must draw inferences from the data to answer the questions originally posed by the law. Rarely in social science do researchers study all of the people or events about which they would like to draw conclusions. Instead of studying the entire population—say, every member of the jurisdiction in which an obscenity case is being tried—social scientists study a sample drawn from the population. Inferential statistics then allow the scientists to make inferences about populations based upon limited samples of those populations.<sup>50</sup>

### B. Application

The application of a methodology involves a discrete series of events in contrast to the broad principles that characterize a methodology itself. While methodology addresses issues that transcend any one investigation, an application of methodology is confined to the circumstances under which it was conducted.<sup>51</sup> The application of a methodology requires a precise account of how the research was carried out. A survey, for example, if carried out properly, can be dispositive. If carried out improperly, it can be misleading.

How do we know if a survey was carried out properly? We need to ask such questions as: how was the sample chosen? how were the inquiries phrased? who administered the study? and what results were obtained? In *Processed Plastic*,<sup>52</sup> for example, the court precisely described the toy cars made by the contending firms, the age range of the children in the survey sample, the nature of the questions asked, and the percentage of respondents who gave answers to each question indicating consumer confusion.

The distinction between a methodology and its application recalls the distinction between questions of law and questions of fact. In the next section, we develop this similarity as a way to clarify the precedential nature of social science research: methodology should be treated like law, and the application of methodology should be treated like fact.

## IV

### METHODOLOGY AS PRECEDENT

#### A. Methodology: The Analogy to Law

The case for viewing research methodology as precedential rests on

---

METHODS IN SOCIAL RELATIONS 315-41 (4th ed. 1981); *Statistical Inference in Litigation*, 46 LAW & CONTEMP. PROBS., Autumn 1983, at 1.

50. See D. BARNES, STATISTICS AS PROOF: FUNDAMENTALS OF QUANTITATIVE EVIDENCE 31-74 (1983).

51. For a discussion of what should be characterized as methodology and what should be regarded as application, see *infra* Section III B.

52. *Processed Plastic Co. v. Warner Communications*, 675 F.2d 852, 854-55 (7th Cir. 1982).

the observation that methodology has some of the characteristics typically associated with "law."<sup>53</sup> Like law, methodology applies generally and produces results applicable beyond particular instances. Many distinct investigations use the same methodology. Thus, the methodology transcends the particulars of any individual project.

Similarly, while law (particularly common law) is made manifest only by its application to specific facts, it also yields general propositions.<sup>54</sup> This attribute of generality is described as the "precedential effect" of a court decision. A decision assumes the status of precedent in subsequent litigation precisely to the extent that the decision transcends the particular details of the original case. Indeed, the way to deny the status of precedent to a decision is to deny its generality by limiting it to its facts.<sup>55</sup>

### B. *Application: The Analogy to Fact*

The application of a methodology does not bear the same relationship to law possessed by the methodology itself. Unlike methodology and law, a particular study applying a methodology does not yield general propositions beyond the scope of a specific case. Rather, a particular study only produces outcomes applicable to the particular case for which it was commissioned. Thus, it more closely resembles fact.<sup>56</sup> Who administers what instruments to whom on what dates and with what results, for example, are questions which may not transcend the given investigation; rather, they may constitute a simple description of that investigation, with no necessary implications for any other case.<sup>57</sup>

This distinction between methodology and application suggests that social science research is *both* like law and like fact. Methodology is sufficiently general in nature that "law" is a plausible analogy for courts to use when determining the precedential force of the social science research used in previous cases. In contrast, application of methodology is bound to the particular events of the study. Accordingly, courts should treat application as "fact." Thus, the precedential aspect of rulings on social science evidence is restricted to the methodology.<sup>58</sup>

---

53. This observation draws upon the analysis developed in *Social Authority*, *supra* note 24, at 488-95.

54. Monaghan, *Constitutional Fact Review*, 85 COLUM. L. REV. 229, 235 (1985) ("The important point about law is that it yields a proposition that is *general* in character. Fact identification, by contrast, is a case-specific inquiry into *what happened here*.").

55. *See generally*, R. CROSS, PRECEDENT IN ENGLISH LAW (2d ed. 1968).

56. *See* K. DAVIS, *supra* note 21, at 402 (adjudicative facts are those "concerning the immediate parties").

57. *See infra* Section III B.

58. As to the objection that basic methodology lacks the element of "command" integral to the concept of law, see *Social Authority*, *supra* note 24, at 492 (reviewing the modern criticisms of the view that all law emanates from the commands of the sovereign). In addition, positive materials

## V

## CHARACTERIZING METHODOLOGY

Since, under our view, only those aspects of a study that are characterized as methodology constitute precedent binding on succeeding courts, the characterization of "methodology" becomes a pivotal concern.<sup>59</sup> To state, as we have, that methodology is a "more abstract" concept than the description of how a particular research project was conducted will suffice at the extremes. For example, the recognition in *Processed Plastic*,<sup>60</sup> that "surveys" are useful in determining consumer confusion between products, or the view in *Palmer* that "statistical analyses" can assist in demonstrating employment discrimination under Title VII,<sup>61</sup> are statements about methodology. Likewise, the statement in *Saliba* that "[d]uring the first two weeks of February, 1982, five hundred (500) adults from Marion County were questioned by interviewers from Herron Associates of Indianapolis, a field service"<sup>62</sup> is a statement about the application of survey methodology in a specific study.

However, for the assertion that judicial decisions regarding methodologies are binding precedent to be workable, methodology must be defined more precisely than by a blanket judicial approbation of "surveys" and "statistical analyses." What, then, are the contours of the concept of methodology? This question is best addressed inductively, and we take as our grist the *Palmer* court's assertion that "we do not believe that we can allow the threshold at which statistical evidence alone raises an inference of discrimination to be lower than 1.96 standard deviations."<sup>63</sup> This assertion raises the issue of whether the *Palmer* court properly characterized the statistical methodology that the United States Supreme Court approved for use in Title VII cases.

A. *The Analogy to Precedent*

On order to define the contours of methodology, we must first dis-

---

similar to social science have also been analogized to law. Custom is perhaps the earliest example of positive materials treated as law. See, e.g., C. ALLEN, *LAW IN THE MAKING* 121 (6th ed. 1958); 1 W. BLACKSTONE, *COMMENTARIES* \*64. As Professor Merryman wrote: "Authority varies in degree but not in kind, and statutes and cases are more authoritative than other legal and nonlegal writing, but are no more authority." Merryman, *The Authority of Authority: What the California Supreme Court Cited in 1950*, 6 *STAN. L. REV.* 613, 621 (1954).

59. Following our view that methodology is analogous to law, materials used in this characterization process should ordinarily be obtained by courts in briefs rather than by testimony. Courts may also locate these materials through their own efforts. This is a familiar process when prior judicial opinions are reviewed, and should also be used when social science materials are involved. See *Social Authority*, *supra* note 24, at 495-98.

60. *Processed Plastic Co. v. Warner Communications*, 675 F.2d 852, 854-55 (7th Cir. 1987).

61. *Palmer v. Shultz*, 815 F.2d 84, 90 (D.C. Cir. 1987).

62. *Saliba v. State*, 475 N.E.2d 1181, 1188 (Ind. Ct. App. 1985).

63. *Palmer*, 815 F.2d at 96 n.9.

tinguish methodology from its application. This task is similar to the one faced by courts when defining the precedential force of a prior decision. In that task, courts must distinguish the law set down by the decision from the facts of the case. Thus, one way to define the contours of methodology requires an understanding of legal precedent.

How courts define precedent is a profound and perennial issue.<sup>64</sup> Fortunately, Professor Frederick Schauer has recently provided a lucid analysis of the issue.<sup>65</sup> Two alternative approaches to determining the content of precedent emerge: The first is textual; the second, functional.

### 1. Textual Analysis

One way to determine the content of a precedent is to note the signals sent by the original court to other courts. Decisions often contain "canonical language"<sup>66</sup> or an "articulated characterization"<sup>67</sup> that seeks to inform subsequent courts of the range of fact situations that are to be governed under the new rule. Applying such a textual analysis to the *Palmer* court's statements regarding statistical evidence provides a useful starting point for our effort to define methodology.

The *Palmer* court stated<sup>68</sup> that its discretion—and, therefore, the discretion of the district court—in evaluating statistical evidence in Title VII cases was constrained by two precedents of the United States Supreme Court, *Castaneda v. Partida*<sup>69</sup> and *Hazelwood v. United States*.<sup>70</sup> The textual question, therefore, is whether the language of *Castaneda* and *Hazelwood* indicates that the Supreme Court intended to articulate a rule that lower federal courts were to follow in evaluating statistical evidence.

In a footnote to the *Castaneda* decision, the Supreme Court said that a disparity of more than two or three standard deviations in the representation of racial and ethnic minorities in jury pools as compared to their representation in the community "would be suspect to a social scientist."<sup>71</sup> This language does not seem to constitute "canonical lan-

---

64. See generally S. BURTON, AN INTRODUCTION TO LAW AND LEGAL REASONING (1985); K. LLEWELLYN, THE COMMON LAW TRADITION: DECIDING APPEALS (1960); Goldman, *The Force of Precedent in Legal, Moral, and Empirical Reasoning*, 71 SYNTHESE 323 (1987); Schaefer, *Precedent and Policy*, 34 U. CHI. L. REV. 3 (1966).

65. Schauer, *Precedent*, 39 STAN. L. REV. 571 (1987).

66. *Id.* at 579.

67. *Id.*

68. *Palmer v. Shultz*, 815 F.2d 84, 96 n.9 (D. C. Cir. 1987).

69. 430 U.S. 482 (1977).

70. 433 U.S. 299 (1977).

71. 430 U.S. at 497 n.17. Note the minor discrepancy between *Castaneda*, which speaks of a minimum of "two" standard deviations, and *Palmer*, which specifies a more precise "1.96" standard deviations. *Palmer*, 815 F.2d at 92. Given certain assumptions, a disparity of 1.96 standard deviations corresponds to a 5 percent probability that differences as large as those observed could

guage" seeking to lay down a firm rule. Indeed, no such rule of "two or three" standard deviations was needed in *Castaneda* since the uncontroverted empirical data introduced in that case demonstrated that the disparity between the expected and observed number of Mexican-American members of the jury pool was "approximately 29 standard deviations."<sup>72</sup> If the phrase "dicta" has any meaning in current legal discourse, it would seem to apply here.

*Hazelwood*<sup>73</sup> makes it even more evident that the Supreme Court did not intend to set "two or three standard deviations" as the rule for resolving Title VII cases. The *Hazelwood* Court referred to the "two or three" standard deviation test of *Castaneda* as "[a] precise method"<sup>74</sup> for measuring statistical disparities, and stated that "a more precise method" existed, indicating that the *Castaneda* test was only *one* possible precise method.<sup>75</sup> The citation given by the Court for a "more precise method" was a page in an undergraduate statistics textbook<sup>76</sup> which presented a "Bayesian"<sup>77</sup> analysis, found by some commentators<sup>78</sup> to be superior to counting standard deviations as a technique for testing hypotheses. The Court then characterized the analyses given in *Castaneda* and earlier in the *Hazelwood* opinion itself as merely illustrative of the type or class of analyses that may be probative in discrimination cases: "These observations are not intended to suggest that precise calculations of statistical significance are necessary in employing statistical proof, but merely to highlight the importance of the choice [of a comparison group]"<sup>79</sup>—in that case, the choice of the relevant labor market area.

Thus, the language of the Supreme Court in *Castaneda* and *Hazelwood* does not signal a "two or three standard deviations" rule as the intended content of the precedent. Based on a textual analysis of *Cas-*

occur by chance alone, even if there were no real differences between the populations being sampled. The Supreme Court in *Castaneda* evidently rounded 1.96 to the nearest whole number. See D. Baldus and J. Cole, *supra* note 12, at 293-297.

72. 430 U.S. at 917 n.17.

73. 433 U.S. 299 (1977).

74. *Id.* at 309 n.14.

75. *Id.* at 311 n.17. The *Hazelwood* view of *Castaneda* was reaffirmed in *Watson v. Fort Worth Bank and Trust*, 56 U.S.L.W. 4922, 4927 n.3 (June 29, 1988).

76. F. MOSTELLER, R. ROURKE & G. THOMAS, *PROBABILITY WITH STATISTICAL APPLICATIONS* 494 (2d ed. 1970).

77. "Bayes Theorem" allows one to calculate the extent to which new information modifies an existing probability estimate. See D. BALDUS & J. COLE, *supra* note 12, at 304. See generally Finkelstein & Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489 (1970) (proposing the use of Bayesian analysis in criminal trials); Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971) (arguing against the use of inatheoretical evidence in trials).

78. See, e.g., Fienberg & Schervish, *The Relevance of Bayesian Inference for the Presentation of Statistical Evidence and Legal Decisionmaking*, 66 B.U.L. REV. 771 (1986); Kornstein, *A Bayesian Model of Harmless Error*, 5 J. LEGAL STUD. 121 (1976); Lempert, *supra* note 39.

79. *Hazelwood v. United States*, 433 U.S. 299, 311 n.17 (1977).

*taneda* and *Hazelwood*, the *Palmer* court mischaracterized the applicable precedent.

### 3. *Functional Analysis*

The second method proposed by Professor Schauer to determine the scope of a precedent is a functional analysis. He suggests that the characterization of a prior opinion "will vary with the purposes to be served within a decisional domain."<sup>80</sup> In any domain, the characterization adopted will reflect an accommodation between two overarching and conflicting principles: a desire to correctly decide the current case in all of its complexity, and a desire to create simple rules for efficiently deciding future cases.<sup>81</sup>

The first principle seeks to maintain the *flexibility*<sup>82</sup> necessary to tailor a decision to the elaborate contours of the facts before the court. It chafes at the restraints imposed by decisions in other cases, and therefore acknowledges precedent only "in small units."<sup>83</sup> It assimilates only a narrow range of fact situations under the control of any precedent in order to allow courts the discretion to respond to the equities of individual disputes. Where social science evidence is used in Title VII cases, the principle of flexibility would suggest that *Castaneda* and *Hazelwood* be given a narrow reading so that only a limited range of fact situations would be assimilated under their holdings.

The second principle affecting the scope of precedent seeks to achieve the *stability* necessary to provide guidance to other courts.<sup>84</sup> The principle of stability would conserve the resources of judicial decisionmakers, by allowing them to rely upon established rules for deciding cases rather than reasoning through issues anew each time they are presented.<sup>85</sup> Stability, therefore, argues for assimilating a broad range of fact situations under the control of precedent in order to reduce substantially the burden on subsequent courts of reanalyzing issues. Where social science evidence is used in Title VII cases, the principle of stability would suggest that *Castaneda* and *Hazelwood* be given an expansive reading so that a wide range of fact situations would be assimilated under their holdings.

Analyzed in these terms, the *Palmer* court was more concerned with achieving stability in deciding Title VII cases as a class than it was with preserving the flexibility to tailor its decision to the unique facts of a

---

80. Schauer, *supra* note 65, at 602.

81. *Id.* at 601-02.

82. *Id.* at 602.

83. *Id.*

84. *Id.* at 601-02.

85. *Id.* at 599.

particular case. Accordingly, the *Palmer* court used large categories of assimilation in defining the scope of the precedent of *Castaneda* and *Hazelwood*. Based on that broad reading, it concluded that *Castaneda* and *Hazelwood* required it and the district court to apply a “rule” of two or three standard deviations.

### B. A Narrow Characterization

The *Palmer* court’s expansive reading of the text of *Castaneda* and *Hazelwood*, and its tilt toward stability rather than flexibility in functional terms, is understandable. The *Palmer* court admitted, “We are not expert statisticians,”<sup>86</sup> and remarked that “for obvious reasons” it confronted statistical issues “with some trepidation.”<sup>87</sup> Despite this perceived lack of expertise, the *Palmer* court reluctantly noted “the unavailability of embarking upon a journey into the statistical maze.”<sup>88</sup> Given such diffident comments, it is not surprising that the court reached for guidance from decisions of the United States Supreme Court, both to conserve its own resources in working through the “statistical maze,”<sup>89</sup> and to share the responsibility for taking a turn that may later turn out to be a mistake.

Although the *Palmer* court’s reading of the Supreme Court precedent is understandable, our reading of *Castaneda* and *Hazelwood* yields a considerably narrower “holding” on the analysis of social science evidence. We believe that the *Palmer* court undervalued the need for flexibility in dealing with the methodology of social science evidence, particularly at this relatively early stage of its development as a legal tool. As a consequence, the *Palmer* court also erred on functional grounds in characterizing *Castaneda* and *Hazelwood* as precedent for a bright-line rule of “two or three standard deviations” in a broad range of discrimination cases.

We believe that a concern for flexibility should temper a concern for stability in decisions on the methodology of social science evidence since appropriate methodology, especially the statistical aspect, is highly complex and dependent upon the fulfillment of numerous empirical assumptions. For example, in employment discrimination cases differences between selection processes that produce dichotomous outcomes (for example, persons employed and persons not employed) and those that produce continuous outcomes (for example, salary or benefit levels) will call for different statistical methods to test for discrimination.<sup>90</sup> Thus,

---

86. *Palmer v. Smith*, 815 F.2d 84, 93 n.8 (D.C. Cir. 1987).

87. *Id.* at 92.

88. *Id.* at 93.

89. *Id.*

90. See D. BALDUS & J. COLE, *supra* note 12, at 12-13; see also Kaye, *Ruminations on*

methodology will not often be adequately characterized by broad and simple rules like "two or three standard deviations." Indeed a compelling case can be made that the creation of such simple rules has resulted in the use of statistical analyses which were appropriate in their original application, but are not appropriate in the fact situations to which they have been generalized.<sup>91</sup> The stability that comes from simple rules like "two or three standard deviations" should not be purchased if the cost is the loss of the flexibility to accept statistical analyses more carefully tailored to the unique facts of a given case.<sup>92</sup>

Thus, a functional as well as a textual analysis suggests that the *Palmer* court mischaracterized *Castaneda* and *Hazelwood* as establishing a binding rule of "two or three standard deviations." Rather, a functional analysis of those precedents suggests a characterization much like that produced by our textual analysis: One of the statistical tests that may be applied to empirical data in cases with facts like those in *Castaneda* and *Hazelwood* involves calculating standard deviations, and, if such a test is used, a disparity of two or three standard deviations is among those that will suffice to make a prima facie case of employment discrimination. An alternative and even narrower characterization of these precedents relegates all mention of "standard deviations" to the realm of technical application. It asserts that in cases with facts like those in *Castaneda* and *Hazelwood*, a prima facie case of employment discrimination is made if statistical analysis demonstrates that the probability that the observed disparities were produced by chance is five percent or less.<sup>93</sup> More generically, a functional analysis that emphasizes

---

*Jurimetrics: Hypergeometric Confusion in the Fourth Circuit*, 26 JURIMETRICS J. 215 (1986) (criticizing the Fourth Circuit's application of standard deviation analysis).

91. See Kaye, *supra* note 47, at 1335 n.21 ("[O]ne cannot help observing that the apparent infatuation of the courts with this one procedure [standard deviation analysis] is such that, all too often, it is employed to the exclusion of other, more appropriate methods."); see also Meier, Sacks & Zabell, *What Happened in Hazelwood: Statistics, Employment Discrimination, and the 80% Rule*, 1984 AM. B. FOUND. RES. J. 139 (1984) (arguing that standard deviation analysis used alone is unsatisfactory for evaluating allegations of discrimination); Sugrue & Fairley, *supra* note 27 (contending that standard deviation analysis often is misused in discrimination litigation).

92. If the judge finds that the methodology of the research to be offered in evidence has been established by precedent, then, in jury trials, it would seem to follow that the results of the court's characterization of that precedent should be communicated by instruction. See, e.g., *Herron v. Southern Pac.*, 283 U.S. 91, 95 (1931) ("It is the duty of the court to instruct the jury as to the law; and it is the duty of the jury to follow the law, as it is laid down by the court.") (quoting *United States v. Battiste*, 24 F. Cas. 1042, 1043 (C.C.D. Mass. 1835) (No. 14,545) (Story, J.)). Essentially, such instructions would prepare the jury for the testimony of the expert witness who will offer in evidence the application of methodology in a specific research project. See *Social Frameworks*, *supra* note 26, at 592-98.

93. See Kaye, *supra* note 47, at 1362; see also Kaye, *supra* note 90, at 219-20:

[I]t is doubtful that the law should fixate on a single significance level to establish a prima facie case in all situations. But even if one is committed to testing hypotheses by computing P-values and stacking these P-values up against some preordained number [such as .05], it is clear that what counts is the P-value vis-a-vis the arbitrary number that

flexibility suggests that the methodology of social science evidence should be characterized such that only those aspects of the research essential to establishing its probative value are accorded the status of "methodology."<sup>94</sup> Thus, a court cannot prospectively determine whether any particular aspect of a research study that is used as evidence in a prior case, such as the size of the sample or the manner in which it was drawn, should be characterized as methodology or as application. Rather, the proper characterization turns on whether the succeeding court finds that that aspect was essential to the study's probative value in the prior case.

Further, even such a narrow characterization should be both conditional and nonexclusive. It should be conditional in that a ruling that one methodology is probative in a certain fact situation does not necessarily imply that the methodology is probative in other fact situations. For example, while a disparity of two or three standard deviations may be a useful test where the facts of a Title VII case imply that the data are normally distributed, it may yield "substantially misleading and inaccurate" results in other Title VII cases where the data are not normally distributed.<sup>95</sup> Further, the judicial endorsement of two or three standard deviations in Title VII cases has no necessary implications for the use of social science evidence on other legal questions such as the measure of consumer confusion in a trademark case of community standards in an obscenity case.

This narrow characterization of methodology should also be nonexclusive: a ruling that one methodology is probative in a certain situation should not be taken to imply that other methodologies would not be as—or more—probative in that same situation. Recall that in *Hazelwood* the Supreme Court used the "two or three standard deviation" test but gave as a presumably acceptable alternative the more precise "Bayesian" method.<sup>96</sup> The *Palmer* court similarly stated: "[W]e note that there are various methods for deriving a 'test statistic' measured in numbers of 'standard deviations': the z-test, the t-test, etc. We have no opinion on the choice of these methodologies as this case does not call them into question."<sup>97</sup>

---

establishes "significance." How one comes up with the P-value is a technical issue appropriately left to statisticians who must defend the reasonableness of their model and the accuracy of their computations at trial.

94. See *Social Authority*, *supra* note 24, at 498-512 (proposing criteria for courts' evaluation of social science research).

95. D. BALDUS & J. COLE, *supra* note 11, at 172-73 (Supp. 1987); see also Kaye, *supra* note 90.

96. See *supra* text accompanying notes 74-78 and accompanying text.

97. *Palmer v. Shultz*, 815 F.2d 84, 93 n.8 (D.C. Cir. 1987). See generally *Bazemore v. Friday*, 106 S. Ct. 3000 (1986) (discussing the use of less-than-ideal statistics); Fienberg, *comment*, in *STATISTICS AND THE LAW* 41, 42 (M. DeGroot, S. Fienberg & J. Kadane eds. 1986) (noting that "failure of federal judges and legal and statistical commentators to take note of the virtual equivalence of the various procedures available for the comparison of two sample rates has led to

## CONCLUSION

Our effort to understand the precedential nature of social science research introduced as evidence has provided a critique of the rule adopted in *Palmer*. Our concern about the result in that case, however, should not obscure our view that courts may properly treat social science methodology as law and accord it the effect of common law precedent. On this important theoretical point we find the approach of *Palmer*, and of a growing number of other courts, entirely correct.

Furthermore, our proposal that methodological holdings be construed narrowly does not mean that the benefits of stability that would accrue to a broader characterization of methodology will never be achieved in more generous measure. As Holmes once noted in a different context, what begin as questions of fact often become, in time, questions of law.<sup>98</sup> As courts gain experience with the various methodologies underlying social science evidence, a body of case law will develop that more fully reflects the nuances of research design and of the statistical issues involved in different types of cases. The case-specific applications of methodology that we now propose should be argued as matters of fact, will gradually be transformed into more generic questions of methodology to be treated as matters of law. As this common law process of elaborating precedent takes place, the full benefits of stable and uniform rules for treating social science evidence will be realized.

---

great confusion and has contributed to the statistical disarray found in court opinions.”); Kaye, *supra* note 47, at 1362-64 (discussing alternatives to “hypotheses testing”).

98. See O.W. HOLMES, THE COMMON LAW 123-29 (1881); see also *Watson v. Fort Worth Bank and Trust*, *supra* note 75 (“At least at this stage of the law’s development, we believe that such a case-by-case approach properly reflects our recognition that statistics ‘come in infinite variety and . . . their usefulness depends on all of the surrounding facts and circumstances.’ *Teamsters v. United States*, 431 U.S. 324, 340 (1977).”)