WHAT EMPIRICAL RESEARCH TELLS US: STUDYING JUDGES' AND JURIES' BEHAVIOR

Peter David Blanck[1]

Peter David Blanck

TABLE OF CONTENTS
Introduction

---

[1] Associate Professor of Law, University of Iowa College of Law; J.D., 1986 Stanford Law School; Ph. D., 1982 Harvard University; B.A., 1979 University of Rochester. This Article benefitted from my collaborative relationship studying courtroom behavior with Professors Robert Rosenthal and Alan Hart.

INTRODUCTION

In a criminal trial, due process mandates that the trial judge does not show actual bias toward the defendant. Trial judges are not only required to be fair and impartial, they must also "satisfy the appearance of justice."[2] Thus, the trial judge's "appearance" and behavior in a criminal jury trial must never indicate to the jury that the judge believes the defendant is guilty.[3] The appearance of judicial bias alone is grounds for reversal even if the trial judge is, in fact, completely impartial.[4]

The courts, legal scholars, practitioners, and social scientists recognize that trial judges' verbal and nonverbal behavior may have important effects on trial processes and outcomes.[5] Courts caution repeatedly that juries may accord great weight and deference to even the most subtle behaviors of the trial judge.[6] One judge concludes that " juries can be easily influenced by the slightest suggestion from the court, whether it be a nod of the head, a smile, a frown, or a spoken word."[7] Despite the danger of influencing the jury, trial judges in a criminal jury trial, like all human beings, develop certain beliefs about the defendant's guilt or innocence.[8]

The Jury Selection Conference addressed several issues central to the behavior of judges, jurors, litigants, and the press embroiled in the litigation process. These issues included the reactions of judges and juries when prior knowledge exists about the people or the events in the trial before

---

[2] Offutt v. United States, 348 U.S. 11, 14 (1954); see In re Murchinson, 349 U.S. 133, 136 (1955) (noting that due process requires that fair trial be conducted in fair tribunal); Mayberry v. Pennsylvania, 400 U.S. 455, 465-66 (1971) (emphasizing that judge can become so personally embroiled with lawyer in trial that judge loses image of impersonal authority of law); Walker v. Lockhart, 726 F.2d 1238, 1251 (8th Cir. 1984) (Arnold, J., concurring) (stating due process requires trial before unbiased judge and jury); Blanck, Rosenthal & Cordell, The Appearance of Justice: Judges' Verbal and Nonverbal Behavior in Criminal Jury Trials, 38 STAN. L. REV. 89, 89-97 (1985) [hereinafter The Appearance of Justice] (discussing judicial influence and its relation to procedural due process); Blanck, Rosenthal, Hart & Bernieri, The Measure of the Judge: An Empirically-based Framework for Studying Trial Judges' Behavior, 75 IOWA L. REV. 653, 655-57 (1990) [hereinafter The Measure of the Judge] (describing research that studied judges behavior to determine if it "appears" to trial participants to be fair and impartial).

[3] The Appearance of Justice, supra note 1, at 89 & n.3 (discussing basic requirement that judge be impartial and never appear to believe that accused is guilty).

[4] See Bollenbach v. United States, 326 U.S. 607, 611-15 (1946) (holding judge's indication of belief of defendant's guilt constituted reversible error); State v. Larmond, 244 N.W.2d 233, 236 (Iowa 1976) (noting that defendant is not required to show that judge's behavior actually prejudiced jurors, but merely that jurors could have inferred judicial bias); The Appearance of Justice, supra note 1, at 89-90 & nn.4-5 (concluding jurors can become prejudiced from appearance of judicial bias).

[5] See The Measure of the Judge, supra note 1, at 654 & n.12 (discussing books, articles, and studies recognizing importance of judge's behavior). See generally H. KALVEN & H. ZEISEL, THE AMERICAN JURY (1966) [hereinafter THE AMERICAN JURY] (providing classic study of judges and juries); J.P. RYAN, A. ASHMAN, B.D. SALES & S. SHANE-DUBOWN, AMERICAN TRIAL JUDGES: THEIR WORK STYLES AND PERFORMANCE (1980) (encompassing comprehensive analysis of judges' behavior).

[6] See State v. Hamilton, 240 Kan. 539, 544, 731 P.2d 863, 868 (1987) (warning that jury looks to judge for guidance and attributes great weight to his utterances); Marino v. Cocuzza, 14 N.J. Super. 16, 20, 81 A.2d 181, 185 (1951) (commenting that jury is influenced by judge's slightest suggestion).

[7] State v. Wheat, 131 Kan. 562, 569, 292 P. 793, 797 (1930) (Jochems, J., dissenting).

[8] See The Appearance of Justice, supra note 1, at 89 (stating judge may reveal beliefs during trial by directing trial based on judge's own expectations).

them,[9] the methods social scientists employ to study the reactions and the behavior of trial participants during a "live" trial,[10] and the implications of social science research on judges' and juries' behavior regarding actual courtroom behavior.[11]

This Article describes an empirically-based framework and theoretical model for exploring judges' and juries' behavior in actual trials. The research model and program tests the long-standing observation that sometimes subtle and perhaps unintentional nonverbal behavior of judges, as well as other "extra-legal" variables, might alone predict trial processes and trial outcomes.[12] To explore these ideas, Part I of this Article sets forth a research model that employs various legal, extra-legal, and behavioral variables (e.g., the nonverbal behavior of the judge) for describing and documenting trial judges' and juries' behavior.[13] The preliminary findings of the model presented in Part II suggest, inter alia, that extremely prejudicial behavior by judges might sometimes deny defendants their constitutionally protected right to a fair and impartial trial.[14] Part III discusses future uses of our findings. Finally Part IV highlights recent developments regarding what empirical research can tell us about courtroom behavior.

The intention in "modeling" judges' and juries' behavior here is not to suggest that there is a bright line standard for detecting, quantifying, or measuring the legally permissible limits of judges' behavior--for example, for separating a judge's verbal and nonverbal behaviors that are legally appropriate from those that may unduly influence a jury. Nor is it my intention to suggest that trial judges display great stone faces, showing no emotion or reaction to the events in the courtroom. Rather, the more immediate goal of this Article is to highlight a model of study and an empirical method to aid in the description of the behavior of judges and juries. Indeed, appellate courts, in a more legally formal manner, attempt this by describing the effect andpropriety of judges' behavior during the trial.[15] In this regard, appellate courts attempt to balance a number of factors on a "sliding scale" to assess the propriety of judges' behavior. These factors include: (1) the relevance of the behavior; (2) the emphatic or overbearing nature of the verbal or nonverbal behavior; (3) the efficiency of any instruction used to cure the error;

---

[9] See generally Selecting Impartial Juries: Must Ignorance Be a Virtue in Our Search for Justice., Panel Three: The Roles of Juries and the Press in the Modern Judicial System, Annenberg Washington Program Conference, May 11, 1990, 40 AM. U.L. REV. 597 (1990) [hereinafter Panel Three].

[10] See generally Selecting Impartial Juries: Must Ignorance Be a Virtue in Our Search for Justice, Panel One: What Empirical Research Tells Us, and What We Need To Know About Juries and the Quest for Impartiality, Annenberg Washington Program Conference, May 11, 1990, 40 AM. U.L. REV. 547 (1990) [hereinafter Panel One].

[11] See generally id.

[12] See infra notes 79-80 and accompanying text (discussing research results on effects of judges' nonverbal behaviors on trial outcomes).

[13] See infra notes 36-41 and accompanying text (discussing model to describe and document judge behavior). The ideas expressed in the introduction of this Article are developed more fully in Blanck, et al., The Appearance of Justice, and in Blanck, et al., The Measure of the Judge, see supra note 1.

[14] See The Appearance of Justice, supra note 1, at 99-100 (stating that courts usually will not conclude judge's behavior violated defendant's constitutional right to fair trial unless judge's influence is clearly prejudicial to defendant or significantly affects trial outcome).

[15] See United States v. Olgin, 745 F.2d 263, 268 (3d Cir. 1984) (emphasizing that reviewing court is more concerned with matter central to defense than with comment on tangential issue), cert. denied, 471 U.S. 1099 (1985); United States v. Anton, 597 F.2d 371, 374-75 (3d Cir. 1979) (adding judge's comment on defendant's credibility was issue central to defense and one factor in appellate court's reversal of conviction).

and (4) the prejudicial effect of the behavior in light of the entire trial.[16]

Fred Graham, the keynote speaker at the conference[17] highlighted the growing tendency to videotape and televise actual trials to better document and analyze courtroom behavior.[18] As this trend grows, the empirical framework presented in Part I may become increasingly useful to courts and practitioners in systematically assessing the propriety of the behavior of judges and of other trial participants.[19] In part, this is because it is now apparent that verbal and nonverbal behaviors of judges and trial participants, not previously recorded by the written court transcript or trial record, can be preserved and summarized by videotape or by on-line behavioral analyses. Thus, it will be possible for courts and counsel to describe more accurately these behaviors for appellate review.

As suggested elsewhere, legal practitioners may likewise employ the basic framework and model described here to aid in a more systematic and economical assessment of judges' "global" and "micro" behavioral or communicative styles, and of jurors' reactions to those styles.[20] Further, increased sensitivity to a particular judge's behavioral style may help practitioners in the selection and preparation of expert witnesses, at least in terms of maximizing their communicative strengths to judges and juries.

Judges, as a community of professionals, are also interestedin the issues presented by the empirical model of courtroom behavior described in this Article.[21] Judicial training programs across the country teach judges the importance of communication behavior and style in the courtroom.[22] Hopefully, the research framework and model will provide a standard to help judges evaluate the qualities of their behavior and generally to guide future empirical studies of behavior in the courtroom.

---

[16] See The Appearance of Justice, supra note 1, at 95-96 (reviewing appellate courts' factor approach in assessing propriety of judge's behavior); see also Stevens v. United States, 306 F.2d 834, 838 (5th Cir. 1962) (holding as prejudicial trial court's statement certified in record that court would not believe witness on oath).

[17] See generally Selecting Impartial Juries: Must Ignorance Be a Virtue in Our Search for Justice, Keynote Address: The Impact of Television on the Jury System: Ancient Myths and Modern Realism, Annenberg Washington Program Conference, May 11, 1990, 40 AM. U.L. REV. 623 (1990) [hereinafter Keynote Address].

[18] See generally id. (describing media's efforts to make trial proceedings available through television).

[19] But see Grisso, Baldwin, Blanck, Borus-Rotheram, Schooler & Thompson, The Advancement of Scientific Integrity, AM. PSYCHOLOGIST (1991) (forthcoming) [hereinafter Scientific Integrity] (providing reasons why research psychologists should proceed with extreme caution when employing videotaped data); Blanck, The "Process" of Field Research in the Courtroom: A Descriptive Analysis, 11 LAW & HUM. BEHAV. 337, 349-51 (1987) (discussing due process problems and concerns of disruption to the trial process when videotaping trials for research purposes).

[20] See Anderson, Trial by Press?: Pretrial Publicity Doesn't Bias Jurors, Panelists Say, A.B.A. J., Sept. 1990, at 32 (reporting consensus of penalists at Annenberg Washington Program that jurors subject to extensive publicity can put aside preconceptions if judges provide proper instructions and other curative assistance); see also McConahay, Mullin & Frederick, The Uses of Social Science in Trials with Political and Racial Overtones: The Trial of Joan Little, 41 LAW & CONTEMP. PROBS. 205, 213-20 (1977) (recounting juror selection strategies based on empirical model, personality traits, and juror nonverbal behavior).

[21] See Panel One, supra note 9, at 554-57 (expressing concern about judges' behavior and its effect on juries).

[22] The Measure of the Judge, supra note 1, at 676 (discussing program that videotapes and analyzes judge's behavior during trial proceeding).

Legal scholars are similarly interested in the impact of judges' communicative behavior on courtroom fairness. The American Bar Association's recent amendments to the Code of Judicial Conduct include a new canon that emphasizes the need for the appearance of fairness and justice in the courtroom.[23] The commentary to the proposed canon states: A judge must perform judicial duties impartially and fairly. A judge who manifests bias on any basis in the proceeding impairs the fairness of the   proceeding and brings the judiciary into disrepute. Facial expressions and body language, in addition to oral communication, can give to parties or lawyers in the proceeding, jurors, the media and others an appearance of judicial bias. A judge must be alert to avoid behavior that may be perceived as prejudicial.[24]

The program of research described below should facilitate the systematic assessment of this Canon.

Finally, for social scientists, the theoretical model and its preliminary results may help to reveal the richness and complexity of the field study of judges' and juries' behavior. The title of the panel on which I spoke at the conference was "What Empirical Research Tells Us, and What We Need to Know About Juries and the Quest for Impartiality."[25] The program of study described herein suggests that empirical research can tell judges, lawyers, and social scientists a good deal by replacing unsubstantiated myths about courtroom behavior with empirically validated facts.[26] Still, as Part III illustrates, a more cumulative model of judges' and juries' behavior is warranted, and any single social science study, no matter how well conceived and conducted, yields only a limited degree of external or real-world validity.[27] I present our research efforts next as a step toward developing this body of research.

I. THE STUDY OF JUDGES' AND JURIES' BEHAVIOR

This program of empirical study explores how trial judges sometimes intentionally or unintentionally convey their beliefs or biases to juries and the subsequent impact of this on trial processes and outcomes. A trial judge's beliefs or expectations of a defendant's guilt or innocence may be manifested either verbally (by the spoken word) or nonverbally (by facial gestures, body movements, or tone of voice) and can be reflected in a judge's comments on

---

[23] Canon 3(B)(5) states: A judge shall perform judicial duties without bias or prejudice. A judge shall not, in the performance of judicial duties, by words or conduct manifest bias or prejudice based upon race, sex, religion, national origin, disability, age, sexual orientation or socioeconomic status, and shall not permit staff, court officials and others subject to the judge's direction and control to do so.A.B.A. Model Code of Judicial Conduct, 9-10 Standing Committee on Ethics and Professional Responsibility (Aug. 1990).

[24] Id.

[25] Panel One, supra note 9, at 547-72.

[26] See Konecni & Ebbesen, External Validity of Research in Legal Psychology, 3 L. & HUM. BEHAV. 39, 40-42 (1979) (criticizing reliance on simulated legal research for developing practical recommendations); see also Blanck & Turner, Gestalt Research: Clinical-Field-Research Approaches to Studying Organizations, in HANDBOOK OF ORGANIZATIONAL BEHAVIOR 109, 111 (J. Lorsch ed. 1987) (stating clinical field research is more appropriate where goal is to improve practice).

[27] See Rosnow & Rosenthal, Statistical Procedures and the Justification of Knowledge in Psychological Science, 44 AM. PSYCHOLOGIST 1276, 1280 (1989) (suggesting importance of replications of research in social science to aid in more "meta-analytic" or cumulative view of science).

evidence, responses to witness testimony, or rulings on objections.[28]

In exploring the contention that a judge's behavior somehow impermissibly influences the trial process or outcome, the research relies on three main sources which I have developed previously.[29] Briefly, the first source is the vast case law that requires judges to be fair, impartial, and to "satisfy the appearance of justice."[30] The second source involves a survey of judges' and practitioners' views on the importance of verbal and nonverbal communication in the courtroom.[31] The third source draws on the findings of empirical studies in analogous contexts to validate the development of the research model set forth herein.[32] From these sources, it is possible to design an interdisciplinary methodology that explores or models judges' behavior.

A. A Model for the Study of Judges' and Juries' Behavior

The program of study employs a theoretical framework or working model to determine the effects of judges' behavior on jury verdicts and on other trial process variables.[33] The model identifies several types of variables that need to be studied to achieve a systematic understanding of judges' behavior and its potential influence on juries' decisionmaking processes.

The basic elements of the model are as follows:

("A") the background variables of the trial participants, such as the defendant's criminal history;

("B") the judge's attitudes and beliefs about trial processes prior to trial outcome;

("C") the verbal and nonverbal "global" and "micro" behaviors that communicate the judges' attitudes and beliefs to the trial participants; in particular, to their juries;

("D") the outcome of the trial, in terms of the jury's decision;

("E") the extent to which the judge and jury agree as to the trial outcome; and

---

[28] See The Appearance of Justice, supra note 1, at 90-91. Some of the principal ways in which judges can impermissibly influence the criminal trial process include: (1) disparaging remarks toward the defendant; (2) bias in rulings or comments; (3) consideration of matters not in evidence; (4) forming expectations for trial outcome before the defense has presented its case; (5) inappropriate statements of opinion to the jury during the trial; and (6) failing to control the misconduct of counsel. Id.

[29] See Blanck, supra note 18, at 340-42 (providing detailed description of the source material).

[30] See id. at 340 (discussing case law and legal sources used in study).

[31] See id. at 340-41 (discussing results of interviews with local judges, practitioners, and lawyers).

[32] See id. at 342-43 (describing development of working model to link method of study and theoretical interest). The research model is based, in part, on social psychological research in other contexts on how nonverbal behavior might convey an individual's beliefs for social outcomes. See id. at 340-43 (describing social science methodology developed in other applied contexts).

[33] See The Appearance of Justice, supra note 1, at 101-36 (proposing preliminary model and describing its characteristics); see also Rosenthal, Pfungst's Horse and Pygmalion's PONS: Some Models for the Study of Interpersonal Expectancy Effects, in THE CLEVER HANS PHENOMENON 182 (1981) (discussing general model for study of interpersonal expectancy effects).

("F") the relative magnitude of the sentence imposed by the judge.[34]

The model is illustrated in Figure 1 below and the six variables are discussed in turn.

FIGURE 1

Model for the Study of Judges' and Juries' Behavior

Variable Name	Background	Expectancy	Behavioral	Trial Outcome	Judge/Jury Agreement/ Disagreement	Sentence
Relationship Of Variables[†]	A → B → C → D → E → F					
Variables Under Study	Defendant's Criminal History	Judges' Beliefs *Prior* to Jury Verdict About Expected Trial Outcome	Verbal and Nonverbal Global Styles and Micro Behaviors	Jury Verdict	Judges' Beliefs *After* Jury Verdict	Magnitude of Sentence Imposed by Judge

1. "A" background variables

Background variables refer to the more stable attributes of the trial participants, such as gender, race, social status, intellectual ability, and other personal history factors. The model describes, for example, the relationship and impact of the background variables on trial outcome, the "A-D" relationship. Elsewhere, studies show that the background variables of defendants influence significantly judges' and jurors' views of defendants' guilt or innocence.[35] The model also enables a description of the relationship between defendants' criminal histories and judges' subsequent style of verbal and nonverbal behavior in relating to their juries, the "A-C" relationship.

We are particularly interested in the extent to which information about defendants' criminal histories might predict, or be predicted by, the other variables in the model. This is because frequently the defendants' prior criminal history is known to the judge but not to the jury, unless the defendant takes the stand to testify, thus making, for example, an "A-C" relationship even

---

[34] For earlier descriptions of the model, see The Appearance of Justice, supra note 1, at 102 (describing findings for "A-B-C-D" simple relationships only); The Measure of the Judge, supra note 1, at 680-84 (highlighting "A-B-C-D-E-F" model without presentation of preliminary empirical findings).

[†] The simple relatinships are between any two variables in the model. Cumulative relationships(*e.g.,* "A-B-C" predicting "D") involve more than two variables.

[35] See The Appearance of Justice, supra note 1, at 105 & nn.55-57 (reviewing studies demonstrating effects of race of defendant on sentencing in death penalty cases and effects of defendant's criminal history on jury verdicts).

more striking.[36]

2. "B" attitudinal or "expectancy" variables

A judge's attitudes, beliefs, or expectations about the trial outcome can influence the decisionmaking process of the jurors or the actual trial outcome. An "expectation" in this context is the particular belief that a judge has about the trial process or outcome. When judges expect or predict a certain trial outcome, they intentionally or unintentionally behave in a way that indicates what they think the outcome should be. In doing so, they set into motion behaviors and trial processes that may increase the likelihood of the occurrence of a certain trial outcome.[37] This predictive behavior has been called a "self-fulfilling prophecy" or an "interpersonal expectancy effect."[38] The model predicts that trial judges' expectations for a trial outcome may prophesy, become related to, or in some cases, improperly influence the trial outcome. The model explores, therefore, judges' particular expectations for trial outcomes and the strength of those expectations.

3. "C" verbal and nonverbal communicative behaviors

Verbal and nonverbal behaviors communicate judges' beliefs and expectations to the trial participants. In order to maximize the realworld generalization of the findings, portions of actual trials were videotaped to assess judges' verbal and nonverbal behaviors.[39] The initial study in this series analyzed the behavior of five California state court judges who were videotaped delivering final pattern jury instructions to jurors in thirty-four criminal trials.[40] The sampling of

---

[36] Cf. McElhaney, The Rub, A.B.A. J., Dec. 1990, at 80-83 (noting that most important decision for criminal defendant is whether to take the stand and testify on his or her own behalf, and most important factor in making that decision is whether judge is going to let prosecutor cross-examine defendant with prior convictions).

[37] See The Appearance of Justice, supra note 1, at 132-36 (examining results of simple relationships in model for study of judicial behavior and influence shown in Figure 1).

[38] See id. at 91-92 & n.9 (describing how individual's expectations about event can influence person's behavior to make event more likely); see also E. JONES, INTERPERSONAL PERCEPTION 237-59 (1990) (outlining development of expectancy effects).

[39] Independent groups of raters are employed to assess the videotapes' communicative content on different emotional dimensions (e.g., warm, hostile, professional, and etc.). Raters assess several modified versions of the videotapes to isolate the specific verbal and nonverbal "channels" and global behaviors of communication. The altered versions of the tapes included: (1) normal video-and-audio tapes; (2) audio-only tapes (normal speech); (3) visual- only cues (facial expressions and body movements); and (4) tone-of-voice-only tapes (by a filtered audio recording that allowed rhythm, pitch, and tone to be conveyed but not verbal content). Blanck, supra note 18, at 342-53. The study examines also trial judges' "micro" nonverbal behaviors. Micro nonverbal behaviors are seven discretely coded actions regularly employed in the study of nonverbal behavior, including: (1) amount of eye contact with the jury; (2) number of smiles; (3) number of head nods; (4) number of significant hand movements; (5) number of forward leans to and away from the jury; (6) number of significant changes in body position (fewer shifts designated as "postural attention"); and (7) number of self-touching behaviors, such as chin- rubbing. Id. For detailed descriptions of the research method, see id. at 342- 53 (discussing development of working model, data-gathering techniques, and calculation of results); The Appearance of Justice, supra note 1, at 113-36 (describing research strategy and design and providing detailed analysis of results of judicial influence study); The Measure of the Judge, supra note 1, at 667-74 (developing practical framework for describing and assessing judges' behavior).

[40] See The Appearance of Justice, supra note 1, at 113-18 (describing study participants, rating procedures, and research methodology).

final jury instructions in these cases enabled a "control" for the effects of the verbal content of the instructions.[41] In this way, videotaped trials enable the separation and fine-grained comparison of both the verbal (content of speech) and nonverbal (face and body movements and tone of voice) behaviors of the trial judges.[42]

4. "D" outcome variables

Outcome variables refer to the behavior of the expectee (the juror) after interaction with the expecter (the judge). The model measures trial outcome variables in terms of the jury's finding of guilt or innocence. This form of archival data is collected easily because trial outcomes are available as public records.

5. "E" judge/jury agreement/disagreement variables

These refer to the judges' attitudes about the trial process after the jury reaches its verdict. In our model, this variable also refers to the magnitude of the agreement or disagreement between judge and jury in terms of their views of trial outcome. The "E" variable is similar conceptually to the classic study in The American Jury, which explored the sources and explanations of judge/jury disagreement.[43] As described in more detail in Part III, the inclusion of the "E" variable in the

---

[41] See id. at 107 (explaining that reading of pattern jury instructions allows researchers to isolate variables of interest); see also Levi, The Study of Language in the Judicial Process in LANGUAGE IN THE JUDICIAL PROCESS 9 (J. Levi & A. Walker eds. 1990) (highlighting the importance of verbal and nonverbal behavior and "communicative competence").

[42] See The Appearance of Justice, supra note 1, at 109-10 (discussing how separation of verbal, video, and audio channels facilitates studying nonverbal and verbal behaviors). As mentioned above, both the courts and practitioners recognize the particular importance of judges' nonverbal behavior in influencing trial outcomes. A judge's facial expressions or tone of voice alone can influence jury verdicts, sometimes in impermissible ways. For example, in the often cited case, State v. Barron, 465 S.W.2d 523, 527-28 (Mo. Ct. App. 1971), a Missouri appellate court reversed a burglary conviction on the grounds of nonverbal prejudicial error by the judge. When listening to the defendant's brother testify that the defendant was at home watching television when the alleged burglary occurred, the trial judge placed his hands to the sides of his head, shook his head negatively and leaned back, swiveling his chair 180 degrees. Id. More recently, in analogous situations, social scientists demonstrated that teacher, doctor, and psychotherapist nonverbal behaviors significantly influence the course of social interaction. See The Appearance of Justice, supra note 1, at 108-09 & n.67 (providing investigations that study factors affecting expectancy effects and processes of interpersonal communication that transmit effects). The earliest such studies showing that nonverbal cues are systematically involved in the transmission of expectancy effects concerned college students asking other student experimenters to judge whether a person in a photograph had been experiencing success or failure in life. Although all the experimenters read the same pattern verbal instructions, students responded in accordance with the expectations that were induced randomly in the minds of the experimenters. Id. at 109. In other words, the students found the person in the photograph to be more successful if the experimenters were led to believe that the person was more succesful. Because all experimenters read pattern instructions, the results suggested that only the nonverbal component of the interaction led experimenters to make the predictions. To examine this hypothesis here, the research model explores how trial judges' verbal and nonverbal behaviors solely may communicate their beliefs or expectations to their juries. See Blanck & Rosenthal, Nonverbal Behavior in the Courtroom, in APPLICATION OF NONVERBAL BEHAVIORAL THEORIES AND RESEARCH (R. Feldman ed. 1991) (forthcoming) (summarizing analogous studies and their relevance to model).

[43] THE AMERICAN JURY, supra note 4, at 55. Relationships involving the "E" variable are reported here in preliminary form; additional analyses are presently being conducted to explore the variable.

model is an attempt to develop the insight of The American Jury research into a more comprehensive model of judges' and juries' behavior.

6. "F" sentence imposed variable

This variable forms the final link in our model and is assessed in terms of its magnitude relative to the maximum possible sentence under the change. At this point in the research process, the suggestions below with regard to the "F" variable are based on hypothesized predictions. Empirical data on the "F" variable is just now becoming available for analytical purposes.

Nevertheless, a number of empirical studies on sentencing patterns and behavior have been conducted.[44] What is most apparent from this research is that trial judges have great discretion in the sentencing process. In fact, the discretion of the judge provides an important function in the sentencing process since it allows the judge to consider individual and community perceptions of the crime, the background of the criminal, and the circumstances of the particular case.[45] It is precisely these sorts of background and trial process variables that the model of judges' and juries' behavior can assess. Thus, as others have demonstrated empirically,[46] the model will enable the systematic and comprehensive assessment of the impact of judges', defendants', and/or victims' background ("A") and behavioral ("C") variables on the sentencing process.[47]

Along these lines, Professors Ebbesen and Konecni observed more than 400 sentencing hearings in San Diego courts over a two-year period to study empirically the factors that influence judges' sentencing decisions.[48] Similar to the approach of the model presented here, Ebbesen and Konecni sought to isolate the factors that accounted for the systematic variation among judges' sentences. Four factors accounted for the vast majority of variation in sentencing--which may be also employed in subsequent tests of our model-- including: (1) the type of crime, (2) the defendants' criminal history, (3) the status of the defendant between arrest and conviction (e.g., released on bail or held in jail), and (4) the probation officer's sentence recommendation.

Of these four factors, the judges followed the recommended sentence of the probation officer in eighty-four percent of the cases studied. Professors Ebbesen and Konecni concluded that the probation officer's recommendation is, in fact, likely based on the other three variables studied; that is, based on a knowledge of the defendants' criminal histories (the "A" variable in our

---

[44] See L. WRIGHTSMAN, PSYCHOLOGY AND THE LEGAL SYSTEM 372 (2d ed. 1991) (reviewing empirical studies).

[45] Id. at 373. Future study of judges' discretion in the sentencing process is warranted, given the recently revised sentencing guidelines.

[46] Id. at 375-80 (noting impact of judges' age, experience, previous employment as district attorney, political party membership, and philosophy of punishment as predictors of magnitude of sentence imposed).

[47] See Diamond, Exploring Sources of Sentencing Disparity, in THE TRIAL PROCESS 387, 402 (B.D. Sales ed. 1981) (providing empirical analysis of variables predicting judges' sentencing behaviors and patterns).

[48] Ebbesen & Konecni, The Process of Sentencing Adult Felons: A Causal Analysis of Judicial Decisions, in THE TRIAL PROCESS 431-58 (B.D. Sales ed. 1981) (providing empirical analysis of eight judges' sentencing behaviors and patterns).

model) and on perceptions of the seriousness of the crime.[49] In future tests of the model, the sentencing recommendation of the probation officer may be incorporated as an additional post-verdict variable.

To summarize, the theoretical model is designed to aid in the general understanding of judges' and jurors' behavior, and of how judges' behavior may predict and sometimes influence jury verdicts. More importantly, the model provides researchers a framework to study empirically "chains" of variables that together may predict more accurately certain aspects of trial processes and outcome.

II. PRELIMINARY RESULTS OF THE STUDY

This Part presents preliminary results from our ongoing studies of judges' and juries' behavior. Section A summarizes our findings with regard to trial judges' communicative behaviors and styles. Section B presents some preliminary and exploratory results derived from our tests of the model of judges' and juries' behavior.

A. Global and Micro Dimensions of Judges' Behavior

Before examining the simple relationships between any two variables in the model (e.g., the "A-D" chain) and cumulative relationships of the model (e.g., "A-B-C-E" predicting "D"), it is useful to describe briefly the findings regarding what we have named the "global" and "micro" dimensions of trial judges' verbal and nonverbal behavior in relating to their juries.[50] The term "global dimension" describes the general demeanor or mode of judges' communicative and interpersonal behavior that is often conveyed independently of verbal content.[51] Although a particular global behavior or style may reflect a judge's general orientation for relating to others during the trial, judges show different global behaviors at different times, depending on the circumstances of the trial process. For example, when responding to improper attorney behavior, a judge might show more directive or controlling behavior. Conversely, when dealing with a child witness, a judge might show more caring and patient behavior.

Our earlier studies provide a description of trial judges' global behavior. The analyses yielded four basic global dimensions of behavior, namely "judicial," "directive," "confident," and

---

[49] See L. WRIGHTSMAN, supra note 43, at 382 (suggesting this is one reason why, in this study, judges' agreement with probation officers' recommendations were so high).

[50] Detailed statistical analyses and tables of the preliminary findings for the simple relationships of the variables in the model are described in The Appearance of Justice, supra note 1, at 119-136; The Measure of the Judge, supra note 1, at 672-74.

[51] See Blanck, Rosenthal & Vannicelli, Talking to and About Patients: The Therapist's Tone of Voice, in NONVERBAL COMMUNICATION IN THE CLINICAL CONTEXT 99-143 (1986) (providing overview of research program that analyzed five areas of interpersonal communication: descriptive, psychometric, interactional, competence, and transituational factors); Blanck, Rosenthal, Vannicelli & Lee, Therapist's Tone of Voice: Descriptive, Psychometric, Interactional, and Competence Analyses, 4 J. SOC. & CLINICAL PSYCHOLOGY 154, 155-75 (1986) (summarizing research program on therapist's tone of voice).

"warm."[52] These four global dimensions can be delineated further into those that appear more legally or procedurally oriented, as reflected by the judicial and directive dimensions, and into those that appear more emotionally-based, as reflected by the confident and warm global dimensions.[53] Based on our quantitative results and qualitative interviews with judges and practitioners, the following tentative conclusions about the "appearance" of judges' four global dimensions may be drawn:

(1) A judge high on the judicial dimension is rated as more professional, wise, competent, and honest. The judicial dimension is thus focused, perhaps in the broadest sense, on the appearance of judicial propriety and fairness.[54]

(2) A judge high on the directive dimension is rated as more dogmatic and dominant. The directive dimension typifies the qualities of the trial judge as a courtroom leader and as an administrator.[55]

(3) A judge high on the confident dimension is rated as less anxious and less hostile. This dimension may reflect the extent to which the judge appears emotionally comfortable, self-assured, and patient with others during the trial.[56]

(4) A judge high on the warm dimension is rated as warmer, more open-minded, and emphatic. This dimension may reflect the extent to which the judge appears to be supportive of, and courteous toward, the trial participants.[57]

Together, the four global dimensions of behavior provide one of the first empirically-based descriptions of actual judges' behavior. More importantly, the global dimensions of behavior seem useful for predicting trial processes and outcomes in the theoretical model because they reflect practical, interpretable, and externally-valid dimensions of judges' communicative style that are consistent with prior case-oriented and clinically-derived descriptions of judges'

---

[52] See The Measure of the Judge, supra note 1, at 661-66 (discussing principal components of statistical methodology to analyze judges' behavior in more useful and practical ways by reducing number of variables to describe such behavior).

[53] Id.

[54] See Redish & Marshall, Adjudicatory Independence and the Values of Procedural Due Process, 95 YALE L.J. 455, 483-84 (1986) (describing core values of procedural due process to include "appearance" of independent and fair adjudicator).

[55] The "directive" dimension is similar to others case-study descriptions of so-called "managerial judging" techniques. See Resnik, Managerial Judges, 96 HARV. L. REV. 376, 380, 445 (1982) (criticizing more active, managerial stance of judges and advocating return to classical judicial role); cf. Flanders, Blind Umpires -- A Response to Professor Resnik, 35 HASTINGS L.J. 505, 505-22 (1984) (critiquing Resnick's article for overstating extent of judicial activity inconsistent with due process and for employing questionable approaches with established and acceptable practices and models in study). Clearly, future research seems warranted to explore the relationship between judges' behavior and their methods and style of case management.

[56] The Measure of the Judge, supra note 1, at 662-65 (noting that "confident" judge frames commands in form of pleasant requests and is respectful of trial participants).

[57] The "warm" dimension is consistent with the style of positive regard for others advocated by the client-centered therapeutic school first advanced by Carl Rogers. See The Measure of the Judge, supra note 1, at 665-66 (describing analysis of therapeutic interaction and warm dimension).

behavior.[58]

The program of empirical study also explored what we have called the "micro" behaviors of trial judges (for example, eye contact with the jury, head nods, or hand movements).[59] The judges' micro behaviors assessed in the model have been employed regularly in studies of courtroom nonverbal behavior.[60]

Like our findings for the "global dimensions," these findings provide a practical description of two basic constellations of judges' micro behaviors. A first component of "engaged" micro behaviors emerged, with judges high on this component displaying more eye contact, more postural attention, and less self-touching. A second component of "emotional" micro behaviors emerge, with judges high on this component displaying more smiles, head nods, hand movements, and forward leans. The two micro constellations are conceptually similar to the "judicial" and "warm" global dimensions of judges' behavior.

Analysis of the predictive relationship between the "global" and "micro" constellations illustrate the potentially important methodological contribution of the model. The results of this analysis suggest that the "micro" nonverbal behaviors of the trial judge may serve as important predictors of the four global dimensions of judges' behavior.[61] This is an interesting finding given the large body of research demonstrating the important effects of eye contact on social influence in the courtroom.[62] Furthermore, the results show that more engaged micro behaviors, such as eye contact and head nods by the judge directed to the jury, predict judges' judicial and directive global behaviors, while more relaxed micro behaviors, such as less postural attention, predict the judges' warm global behaviors.[63] Together, these results provide preliminary evidence that judges' micro behaviors alone can be used to predict significantly, and with practical benefit, the four global dimensions of judges' behavior.

The finding that micro behaviors of judges predict those same judges' global behaviors suggests methodologically effective and economical shortcuts to researchers and practitioners interested

---

[58] The National Conference of State Trial Judges describes the essential qualities of a good judge as including: graciousness, moral courage, reputation for fairness, mercy, patience, ability to communicate, decisiveness, innovation, open-mindedness, brevity, dignity, honesty, and integrity. Nat'l Conf. St. Trial Judges, A.B.A., THE JUDGES' BOOK 31-38 (1989).

[59] See supra note 38. In our studies, two raters independently code the tapes for the micro behaviors. The median rater reliability for these variables is .71, suggesting the two raters to be very consistent in their rating of the judges' micro behaviors.

[60] See generally The Measure of the Judge, supra note 1. See also Hemsley & Doob, The Effect of Looking Behavior on Perceptions of a Communicator's Credibility, 8 J. APPLIED SOC. PSYCHOLOGY 136 (1978) (rating witness less credible if he or she failed to maintain eye contact with lawyer); Pryor & Leone, Behavioral Stereotypes of Deceptive Communication, TRIAL, June 1981, at 14 (finding less eye contact, more backward leans, trunk swivel, leg movement, self-touching, gesturing, and speech errors associated with witnesses' attempts at deception).

[61] See The Measure of the Judge, supra note 1, at 674; see also supra notes 49-57 and accompanying text (discussing global dimensions of judicial behavior).

[62] Id.

[63] Id. at 673-74.

in studying and assessing judges' behaviors and attitudes during the "live" trial process.[64] Even moderate relationships between the quantifiable micro behaviors and the impressionistic global behaviors could be of value to social scientists, legal researchers, and practitioners.[65] In part, as Fred Graham suggested, this is because of the serious logistical and ethical problems associated with studying and videotaping actual trials to assess judges' behavior.[66] Moreover, researchers could use the easily codable and unobtrusively collected micro behaviors as an index of a judge's global behaviors and style.[67] In this way, the model may also eventually prove to be useful when employed by training or educational programs for judges devoted to the fine-grain analysis of courtroom behavior.[68]

B. Testing the Model

The model generates fifteen "simple relationships" (the correlation between any two variables) and several other "cumulative relationships" (two or more variables predicting a third variable using multiple regression analyses) that are explored below.[69] From a practical point of view, the cumulative relationships in the model, that is, several individual variables employed to predict a single criterion variable, facilitate a more realistic and comprehensive test of the model.

C. Simple Relationships in the Model

1. Background--Expectancy ("A-B") relationships

---

[64] Id. This is true even where our study involves a relatively small sample of judges, all of whom knew that they were being videotaped and assessed by a naive group of raters. Moreover, the findings are particularly encouraging given the brief length of the judges' behaviors that are rated.

[65] See id. at 676 (proposing ways in which model and its results could be used in appellate review or as training tool to modify judges' behavior); see also Blanck, Rosenthal & Vannicelli, supra note 50, at 127-29, 131-33 (detailing analogous line of study showing predictive value of micro behaviors in psychotherapeutic context).

[66] See generally Keynote Address, supra note 16; see also Blanck, supra note 18, at 349-51 (discussing logistical and ethical problems associated with videotaping trials); Blanck & Turner, supra note 25, at 109-23 (stating clinical field research is intervention that inevitably alters behavior and attitudes observed and recorded).

[67] See Blanck & Turner, supra note 25, at 113-15 (delineating micro behaviors studied and coding scheme attached to such behaviors).

[68] The Measure of the Judge, supra note 1, at 674.

[69] In the presentation of our preliminary findings here, statistical significance is indexed by a probability that an observation would have been found if, in the population from which we sampled, the true correlation were zero. We typically present probability values (p) of .10 or smaller because these values are useful in assessing the types of variables under study here. See The Appearance of Justice, supra note 1, at 119-20 n.98. In exploring the simple relationships of the model we employ correlational analyses. The correlation coefficient (r) can take on values between -1.00 and +1.00. A value of -1.00 means that there is a perfect negative relationship, a value of +1.00 means there is a perfect positive relationship and a value of .00 means that there is no linear relationship between the two variables in the model. Correlational analyses describe the predictive relationship between two variables and do not isolate the "causes" and "effects" of that relationship. In exploring the complex relationships of the model we employ multiple regression analyses. The Multiple R (R) represents the relationship between a particular variable in the model (the criterion variable) and the set of predictor variables in the model. R takes on values only between 0 and 1, with the former indicating no relationship and the latter indicating a perfect relationship between the variables. The F and t tests describe the level of confidence that the linear relationship between the criterion and predictor variables is not zero in the population. df refers to the "degrees of freedom" required for statistical significance testing. See The Measure of the Judge, supra note 1, at 669.

This relationship describes how a judge's expectations for trial outcome may be predicted solely from the background variables of the trial participants. The results suggest that judges' beliefs about trial outcomes are related to defendants' criminal histories in predictable ways. For example, judges usually expect a guilty verdict when defendants have serious criminal histories and expect innocent verdicts when defendants do not have serious criminal histories.[70]

2. Background--Behavior ("A-C") relationships

This relationship describes how a defendant's criminal history may be related to a judge's expression of global and micro behaviors during the trial. Preliminary results suggest that information about a defendant's criminal history ("A"), information that the jury is ordinarily not allowed to learn unless the defendant takes the stand to testify, relates to the judge's behavior ("C" variable) when instructing their juries.[71] Specifically, in the overt, verbal channels, judges' behavior seems to be rated as more judicial, directive, and warm when delivering instructions for defendants with more serious criminal histories.[72] The nonverbal channels, however, tell a different story: judges tend to be relatively less judicial, directive, and warm when delivering instructions for defendants with more serious criminal histories.[73] These results suggest that judges may sometimes "leak" or reveal to juries their underlying beliefs about defendants through nonverbal channels alone.[74]

3. Background--Outcome ("A-D") relationships

The background variable of defendants' criminal history also tends to be related predictably to trial outcome. For example, defendants with more serious criminal histories are more likely to be found guilty.[75] Granted, defendants with criminal histories may be more likely to be guilty. Nonetheless, viewed in combination with the "nonverbal" tendency for judges' to reveal criminal history (the "A-C" results), this finding indicates that a defendant's criminal history, a legally irrelevant factor unless the defendant takes the stand to testify, might influence juries more than previously assumed.[76]

4. Background--Judge/jury agreement ("A-E") relationships

This relationship describes how a defendant's criminal history might be related to the judge's views about the trial outcome. Preliminary tests of this relationship suggest a slight trend for

---

[70] See The Appearance of Justice, supra note 1, at 120-21 (illustrating how defendants' prior records may affect judges' expectations for conviction or acquittal).

[71] Id. at 121-24.

[72] See id. at 122 (describing correlation between judges' basic communication style and seriousness of defendants' criminal histories).

[73] See id.

[74] See id. at 122-24 (detailing study results that support theory that judges reveal personal opinions to juries through nonverbal communication).

[75] See id. at 129 (providing statistical correlation between guilty verdicts and criminal history); see also McElhaney, supra note 35 (noting most important decision for defendant with prior convictions is whether to take stand at trial).

[76] See The Appearance of Justice, supra note 1, at 130.

judges' disagreement with their juries' verdicts to be stronger when defendants had more serious criminal histories.[77] We hypothesize further that in cases where the judge knows the defendant's prior criminal history and the jury does not, the judge will view a jury verdict of innocence to be overly lenient. This suggestion, discussed below, is consistent with Kalven and Zeisel's conclusion in The American Jury that in cases of judge/jury disagreement juries tended to be viewed as more lenient than judges.[78]

5. Background--Sentence ("A-F") relationships

This relationship describes how a defendant's criminal history is related to the sentence imposed by the judge. Expectedly, we hypothesize that the magnitude of the sentence imposed by the judge will reflect the relevant prior criminal history of the defendant. This suggestion is consistent with the findings of Professors Ebbesen and Konecni that defendants' criminal histories and perceptions of the seriousness of the crime predict judges' sentences.[79]

Additionally, an interesting subset of cases to study further will be those in which the defendant does not take the stand to testify. For this subset of cases, we might expect that generally defendants receive relatively lighter sentences. This finding should be particularly apparent in cases where the judge's initial expectations ("B" variable) are for an innocent verdict.

6. Expectancy--Behavior ("B-C") relationships

This aspect of the model describes how judges' expectations for trial outcomes relate to their global and micro behavioral styles. The preliminary findings support the suggestion that judges may reveal their beliefs to their juries through nonverbal channels alone. That is, in the purely nonverbal channels, judges expecting a guilty verdict tend to be somewhat less judicial and warm in relating to their juries.[80]

These results imply, but do not prove, that judges' verbal and nonverbal channels in some extreme cases may convey messages concerning the defendant's guilt or innocence to their juries. Intentionally or unintentionally, judges' beliefs may influence their communication styles when relating to their juries, although on the written trial transcript or in response to a pencil-and-paper questionnaire, the judges may "appear" (or actually believe themselves to be) impartial.[81]

7. Expectancy--Outcome ("B-D") relationships

This relationship describes how judges' expectations for trial outcomes relate to the actual trial outcomes. Taken alone, the "B-D" findings for our sample of thirty-four trials suggest that a

---

[77] Simple correlation for this relationship is r = -.16, not significant at p < .10. But this relationship may prove more apparent when assessed in the context of the multiple regression analyses. See infra notes 94-97 and accompanying text.

[78] See infra notes 98-111 and accompanying text.

[79] See Ebbesen & Konecni, supra note 47 and accompanying text.

[80] The Appearance of Justice, supra note 1, at 130-31.

[81] Id. at 133-34; Blanck, supra note 18, at 342.

judge's belief of a trial outcome (as assessed by questionnaires) does not predict accurately the actual trial outcome.[82] The implications of this finding for conclusions drawn in The American Jury research, which relied solely on the questionnaire data of judges, are discussed in Part III. Nevertheless, the preliminary findings suggest that judges that send "expectancy effects" to juries should be held accountable, because biasing messages (at least as assessed by our methods) might not be an inevitable product of courtroom dynamics.[83] As suggested below, empirical testing of the cumulative impact of the model (e.g., "A-B-C-E" predicting the "D" chain) is necessary to understand more completely this relationship.

8. Expectancy--Judge/jury agreement ("B-E") relationships

This relationship describes how judges' expectations prior to trial outcome predict those same judges' agreement or disagreement with their juries' verdicts. Consistent with our discussion above, the initial findings seem to suggest that those judges expecting a guilty verdict prior to trial outcome are more likely to agree with a jury verdict of guilt than with a verdict of innocence.[84] In other words, judges who are convinced early in the trial that the defendant is guilty may likely reflect or "confirm" this attitude in the subsequent magnitude of their agreement with the jury's verdict.

The strength of the "B-E" relationship may be moderated also by the severity of the defendant's criminal history, by the magnitude of the judge's expectations, and by whether or not the defendant testified. This suggestion would parallel Kalven and Zeisel's view in The American Jury that in cases where the judge has some knowledge of the defendant's criminal history that the jury does not have, the judge and jury may in fact be trying two different cases.[85] In other words, had the jury known what the judge knew, it would likely have agreed with the judge.

9. Expectancy--Sentence ("B-F") relationships

We hypothesize that this relationship will show that the sentence imposed by the judge reflects the judge's expectations about the defendant's guilt or innocence formed during the trial. It will be most interesting to explore the types of cases or circumstances in which judges deviate from their "preconceived" notions or biases formed during the trial in sentencing the defendant. We might predict, as Professors Ebbesen and Konecni found, that judges' perceptions and attitudes about a particular crime or defendant help predict those same judges' sentencing behaviors and patterns.[86]

10. Behavior--Outcome ("C-D") relationships

This relationship describes how judges' behaviors alone may predict the verdicts returned by

---

[82] The Appearance of Justice, supra note 1, at 132-33. Simple correlation for this relationship is r = -.20, which is not significant at p < .10, but suggests a trend that expectations may be related to outcomes in predictable ways. See infra notes 95-97 and accompanying text.

[83] Blanck, supra note 18, at 342.

[84] The simple correlation for this relationship is r = -.39, p < .05.

[85] THE AMERICAN JURY, supra note 4, at 121.

[86] See generally Ebbesen & Konecni, supra note 47 and accompanying text.

their juries.[87] The findings suggest a trend for judges' global behavior to be less judicial and directive and their micro behavior to be significantly more engaged when the verdict returned is guilty.[88] Thus, certain communicative channels and/or styles of judges' behavior may predict with greater accuracy their juries' verdicts.[89] Subsequent fine-grain analyses of behavior are needed to understand how the "C-D" relationship varies with the type of the verbal or nonverbal behavior under study.[90]

Perhaps one encouraging general conclusion to be drawn from the examination of the "C-D" relationship is that in most cases where the judges' expectations for trial outcomes are conveyed to juries either verbally or nonverbally, jurors still tend to make their own independent assessments of the evidence.[91] We are examining this suggestion further in marginal or "close" cases. We expect, as obvious as it might sound, that the strength of the evidence will be generally an important predictor of trial outcomes.[92] Alternatively, in "close cases" the judges' beliefs and behavior, the defendants' background variables or other extra-legal factors will play an increasingly important role in predicting trial outcomes and the ultimate sentence imposed by the judge.[93]

11. Behavior--Judge/jury agreement ("C-E") relationship

This relationship describes how a judge's global and micro behaviors at trial may predict (or be predicted by) that judge's views about the trial outcome. Preliminary tests of this relationship suggest a trend for judges' engaged micro behaviors to be related to their views about trial outcome.[94] Although further analysis is required, this potential relationship supports the earlier suggestion that judges' behavior alone may reflect judges' views about the perceived "correctness" of their jury's ultimate conclusion.

12. Behavior--Sentence ("C-F") relationships

This relationship describes how a judge's behavior is related to the sentence imposed by the judge. To the best of my knowledge, this relationship has neither been tested empirically nor is there an empirical framework for assessing such a relationship. For this relationship, we might predict that judges' global and micro behaviors at trial, reflecting relatively less warm or more directive and engaged attitudes, may predict the imposition of more severe sentences. Again, this relationship may provide a promising avenue for future study, given the relative logistical ease

---

[87] See The Appearence of Justice, supra note 1, at 133-34 (documenting study results showing correlation between judges' behavior and jury verdicts).
[88] Id. Further, simple correlation for the relationship between judges' engage micro behaviors (e.g., more eye contact and trial outcome) is $r = .28$, $p < .10$.
[89] Id. at 135.
[90] Id.
[91] See supra notes 36-37 and accompanying text.
[92] See infra notes 102, 106-07 and accompanying text.
[93] Cf. THE AMERICAN JURY, supra note 4, at 134-35 (analyzing evidence in terms of "close" and "clear" cases); see also infra notes 98-106 and accompanying text.
[94] The simple correlation between the "C" and "E" variables was not significant. The partial correlation derived from exploratory multiple regression analyses, however, suggested a stronger relationship--partial $r = .26$, $p < .04$.

with which micro behaviors can be assessed.

13. Outcome--Judge/jury agreement ("D-E") relationship

The initial data exploring this relationship shows that trial outcome is predicted by a knowledge of the magnitude of judges' agreement/disagreement with their juries' verdicts.[95] Not surprisingly, judges are more likely to agree with guilty and to disagree with not guilty verdicts. On its face, this result supports Kalven and Zeisel's general conclusion that judges tend to view jury results as more lenient than their own. Further research is being conducted to isolate this result, given that in our research (unlike The American Jury study) the judge completed the questionnaires both before and after he or she had knowledge of the jury's verdict.[96]

14. Outcome--Sentence ("D-F") relationships

This relationship describes how jury verdicts (that is, guilty verdicts) relate to the sentence imposed by the judge. As suggested above, this simple relationship is likely affected by other variables in the model (e.g., defendants' criminal histories) in meaningful ways. Interestingly, analysis of this simple relationship also may reflect judges' individual disparities in the sentencing process or in their adherence to legislatively-mandated sentencing guidelines.

15. Judge/jury agreement--sentence ("E-F") relationships

This relationship expresses the extent to which judges' views about trial outcome may predict the sentence they impose. We expect that the magnitude of the judge/jury agreement or disagreement will be particularly important in assessing this relationship. In other words, judges may impose relatively less severe sentences when they disagree with the jury's finding of guilt or vice versa. Consistent with the findings of The American Jury described in the next Part, judges' sentencing behavior may be influenced also by evidentiary factors or by other facts known only by the judge (e.g., in some cases defendant's criminal history).

D. Cumulative Relationships in the Model

The basic purpose of the model is to aid in the general understanding of how judges' behavior may influence jury verdicts and trial processes. The model is most powerful or most predictive when examining the chains of variables taken together. One such primary chain is highlighted next.

In exploring the cumulative chains in the model we employ multiple regression statistical

---

[95] Simple correlation for this relationship is r = -.63, p < .01.

[96] Professors Kalven and Zeisel recognize that a possible source of methodological bias existed in their study because they could not be sure that their participating judges completed the questionnaire responses before actually hearing their jury's verdict. See THE AMERICAN JURY, supra note 4, at 52. It will be interesting to explore the extent to which our results parallel those of Kalven and Zeisel's, given that judges in our study made certain responses before and after knowledge of their jury's verdicts. See The Appearance of Justice, supra note 1, at 157-58.

analyses.[97] From a practical point of view, the regression analyses enable a more detailed assessment of the relationship between a set of variables in the model with one other variable in the model. Typically, the analyses will employ several of the variables in the model as predictors of either trial outcome ("D"), judge/jury agreement/disagreement ("E"), or sentence imposed ("F").

A primary chain of this chain that we have pilot-tested is the extent to which trial outcomes ("D") are predicted by the set of variables including: ("A") the defendants' criminal histories, ("B") judges' expectations for trial outcome, ("C") the engaged micro behaviors of the judge at trial, and ("E") the magnitude of judge/jury agreement/disagreement as to verdict.

Initial results for this test of the model suggest that judges' expectations for trial outcome ("A"), judges' engaged micro behavior at trial ("C"), and the magnitude of judge/jury agreement or disagreement ("E") together predict trial outcomes better than any single variable in the model alone.[98]

We are conducting further analyses to explore the direction and magnitude of the "A-B-C-E" predicting "D" chain. The following scenario, however, may be hypothesized: In cases where the evidence is "close", a guilty jury verdict is likely to result where (1) the defendant has a more serious criminal history; (2) the judge (early in the trial) expects a guilty verdict; (3) the judge's micro behaviors at trial are more engaged (or serious); and (4) the judge, at some point in the trial, comes to agree strongly with the jury's ultimate determination of guilt. This is one descriptive example of how analyses of the cumulative chains of variables in the model may explore the long-standing observation that legal and extra-legal factors influence judges' and juries' behavior. As the next Part suggests, the uses and tests of the model in the study of actual courtroom behavior are not confined to these relationships.

III. *THE AMERICAN JURY* REVISITED

The discussion thus far highlights the ongoing effort to explore and refine a model of judges' and juries' behavior. The inclusion of the "E" variable, judge/jury agreement/disagreement, represents an attempt to provide additional insight into the ground-breaking work evidenced in The American Jury.

The American Jury provided knowledge about the operation of judges and juries in actual criminal jury trials. This task was accomplished through the analysis of an extensive survey (questionnaire) of judges' views of the trial process. The basic purpose of the survey was to

---

[97] For a review of multiple regression techniques, see J. COHEN & P. COHEN, APPLIED MULTIPLE REGRESSION/CORRELATIONAL ANALYSES FOR THE BEHAVIORAL SCIENCES 7 (2d ed. 1983) (explaining that multiple regression analyses describe relationships between complex set of predictor variables and single criterion variable).

[98] Multiple R ($R(2)$) for this complex relationship $= .7420$, $F = 20.10$, df (3, 30), $p < .0001$. For this equation, partial r for "B" variable $= .67$, $p < . 001$, for "C" variable $= .39$, $p < .05$, and for "E" variable $= .64$, $p < . 001$. Partial r for "n" variable in this chain $= .25$, $p = .17$ (not significant).

answer the question, "When do trial by judge and trial by jury lead to divergent results?"[99] The data base for The American Jury research consisted of information on some 3,500 trials.

The main focus of The American Jury is an analysis of the frequency of agreement and disagreement between judge and jury by comparing the actual decision of the jury with the survey response from the judge stating how he or she would have decided the case had it been tried before the judge without a jury. In extensive analyses, The American Jury research seeks to understand the possible reasons for judge/jury agreement/disagreement.

Unfortunately, despite the great visibility of The American Jury research, there has been little attempt to replicate and refine that study using both survey methods and the observation of actual trials. Such study is clearly warranted given the dramatic changes that have occurred with regard to the function, composition, and role of the jury over the last several decades.

One long-term purpose of our model of judges' and juries' behavior is to replicate (albeit on a much smaller scale), and to provide a framework for replication by others, the findings of The American Jury research. For example, analyses are presently being conducted of our "E" variable to explore whether, as evidenced in The American Jury, the judges' report of their verdicts wouldhave been the same as their juries roughly seventy-five percent of the time.[100] This particular result is important to replicate, given the debate over exactly what is an "optimal" degree of agreement between judge and jury.[101] That is, while judge/jury agreement 100 percent of the time might be undesirable because it could indicate that the jury was a "rubber stamp" of the judge, agreement only fifty percent of the time could indicate a deviation in the jury's mandate to follow the law or the judge's instructions.[102]

As also mentioned above and discussed at the Conference, the degree of judge/jury agreement/disagreement should not obscure the suggestion that, in most cases, juries base their decisions on the strength of the evidence presented in the particular case.[103] Alternatively, in close or marginal cases, other extra-legal factors, such as those identified by the model, will increasingly impact on the jury decision-making process. In fact, the main focus of The American Jury research is on describing (1) the impact of extra-legal information in the

---

[99] See generally THE AMERICAN JURY, supra note 4 (measuring performance of jury against performance of judge as baseline).

[100] The high level of agreement found in The American Jury study occurred even in very complex cases, leading several legal commentators to speculate that judges may unintentionally communicate their expectations for trial outcomes to their juries through nonverbal channels. See Elwork, Sales & Suggs, The Trial: A Research Review, in THE TRIAL PROCESS 9 (B.D. Sales ed. 1981); Greenbaum, Judges' Nonverbal Behavior in Jury Trials: A Threat to Judicial Impartiality, 61 VA. L. REV. 1266 (1975). Nonetheless, to date this suggestion has not been tested empirically.

[101] See L. WRIGHTSMAN, supra note 43, at 236-37.

[102] Id. at 237.

[103] See Panel One, supra note 9, at 555-56 (Dr. Hans); see also P. Blanck, A. Hart & R. Rosenthal, The Impact of Legal and Extra-legal Factors on Jury Decision-making (1991) (manuscript in preparation) (noting strength of evidence variable could be viewed as another type of background ("A") variable--Professor Kerr suggested this addition to model in informal discussion at Conference); see also Visher, Juror Decision Making: The Importance of Evidence, 11 LAW & HUM. BEHAV. 1-18 (1987) (commenting that in majority of cases studied, jurors base their verdicts more heavily on evidence and law than on extra-legal variables).

twenty-five percent of the cases in which there was judge/jury disagreement, and (2) how such information accounts for their finding that, in the vast majority of these cases where disagreement occurred, the jury was more lenient than the judge.

The central conclusion of The American Jury research, that defendants are generally better off before a jury than before a judge, was questioned at the Conference.[104] The panel members' intuitive responses are not as uniformly clear as The American Jury findings might have led one to believe. In part, this may be, as others have suggested, a recognition that The American Jury conclusions are drawn only from cases in which the defendants chose the jury to hear their case.[105] To make any meaningful empirical conclusions, Kalven and Zeisel would have had to assign randomly defendants to bench or jury trials and then compare whether juries are still more lenient.[106] Of course, such a study is neither ethically nor legally permissible, given the due process rights of criminal defendants.

The program of empirical study set forth and tested by our model builds on the work of The American Jury research by exploring actual trial behavior, using videotape and survey data, from the perspective of the judge, the defendant, counsel, and the jury. In forthcoming analyses of the model, we explore the relationship between other background ("A") variables of the trial participants (e.g., competency of counsel), as rated by different trial participants, with trial outcome ("D" variable) and with judge/jury agreement/disagreement ("E" variable).

In subsequent analyses, we will also explore the complexity of the facts and evidence, as perceived by different trial participants, and its relationship to other variables in the model. In addition, we will closely analyze data on what Kalven and Zeisel have called the importance of "jury sentiments" in close cases.[107] Kalven and Zeisel suggest that jury sentiments about the defendant (e.g., empathy with the defendant) or about the law (e.g., fairness of the law) may provide additional insight into the reasons for judge/jury disagreement in close cases. These analyses may begin to evidence a more comprehensive view, from several vantage points, of what empirical research can tell us about judges' and juries' decision-making processes.[108]

---

[104] See Panel One, supra note 9, at 553-54 (Dr. Blanck), 551-52 (Ms. Gorelick). Jury leniency, however, varies with the type of case; for example, in libel cases jurors are more likely than judges to find a defendant newspaper or television station at fault. See L. WRIGHTSMAN, supra note 43, at 238.

[105] See Elwork, Sales & Suggs, supra note 99, at 10; L. WRIGHTSMAN, supra note 43, at 237.

[106] See L. WRIGHTSMAN, supra note 43, at 237.

[107] THE AMERICAN JURY, supra note 4, at 111-12, 165 (suggesting that closeness of evidence in case makes is possible for jury to respond to sentiment or intuitive feelings "by liberating it from the discipline of the evidence").

[108] The model provides a framework for several other avenues offuture study. First, in terms of the developing theoretical contribution of the model, it will be important to test the model, and to replicate The American Jury research, in the civil jury context. Moreover, comparison of judges' behavior in bench and jury trials is also warranted. Additionally, we have begun research exploring tests of the model of judges' and juries' behavior in different legal cultures and systems. In particular, we plan to explore trial participants' conceptions of the judges' role and the "appearance of justice" in the German and French (Continental) systems of justice. The comparative and cross-cultural research will build on the American studies by exploring how procedural and cultural norms impact trial processes and the participants' ultimate perceptions of fairness. For a recent discussion of these issues, see Lind, Thibaut & Walker, A Cross-Cultural Comparison of the Effect of Adversary and Inquisitorial Processes on Bias in Legal Decisionmaking, 62 VA. L. REV. 271, 282 (1976) (suggesting in experimental study that procedural and cultural difference in degree of judges' bias in decisionmaking may be moderated by mode of

It is worth noting again that there are many things that our empirical research and model may not be able to tell us about judges' and juries' behavior, given the constraints of our sample size or of the inability to assign randomly defendants to bench or to jury trials. Such issues question the generalizability of studies of actual courtroom behavior. With regard to our research, such questions may be raised about the extent to which the results would hold true across the population of trial judges or over different types of trials. The primary answer to these questions lies in replicating and refining with further field and social science experimental research the results of any single study of actual courtroom behavior.

IV. RECENT DEVELOPMENTS REGARDING EMPIRICAL RESEARCH AND THE COURTROOM

At this point, an exact description of what empirical research may or may not be able to tell us about judges' and juries' actual behavior is difficult to provide. A recent Texas federal district court decision, however, may spur new debate and discussion on this issue. In Cimino v. Raymark Industries,[109] the federal court was confronted with resolving mass tort cases arising from asbestos class action litigation. In an attempt to avoid the large transaction costs associated with trying several thousand such cases, the district court structured a "damage only" phase of the trial in which a random sample of plaintiffs was tried. Specifically, the court tried a random sample of 160 of the some 2,300 potential damage cases before it. The 160 cases were drawn, in the court's view, to reflect with a strong degree of statistical confidence that the results of the 160 cases would be comparable to the average result if all 2,300 cases were tried.

The Cimino procedure highlights the growing debate evidenced at the Conference over what exactly empirical research may be able to tell us about actual courtroom behavior and trial outcomes.[110] On the one hand, there exists a core societal interest in supporting individual due process rights and the related values that form the cornerstone of our system of justice. The defendants in Cimino make this point, arguing that due process entitles them to a traditional one-on-one trial in each of the 2,300 cases. Defendants contend that a random sampling of cases, no matter how sophisticated or carefully developed by social science and empirical guidelines, cannot substitute for individual due process rights.

On the other hand, the Cimino court concludes that the reality of mass tort litigation dictates new means for resolving these complex disputes in a timely, equitable, and practical manner. The court concludes that unless its plan or some other empirically-based procedure is employed to permit damages to be adjudicated in the aggregate, the 2,300 cases will simply not be tried and all parties will be denied access to the courts.[111]

---

evidence presentation); Reitz, Why We Probably Cannot Adopt the German Advantage, 75 IOWA L. REV. (1991) (forthcoming) (arguing cultural centrality of judges' role in characterizing legal system).

[109] 751 F.Supp. 649 (E.D. Tex. 1990).

[110] See Panel One, supra note 9, at 561-63 (Dr. Kerr), 563-64 [Dr. Hans]; see also Saks & Blanck, Improving on Justice: The Unrecognized Benefits of Sampling and Aggregation in the Trial of Mass Torts (1991) (unpublished manuscript) (presenting critique of Cimino procedure).

[111] Cimino v. Raymark Indus., 751 F.Supp. 649, 667 (E.D. Tex. 1990).

In the end, the Cimino court "leaves it to the academicians and legal scholars to debate whether our notion of due process has room for balancing [the] competing interests" between core legal rights and the use of social science in the courts.[112] This debate has been furthered at the Conference. Nevertheless, it is fair to say that additional forums are necessary to resolve the issues related to the complex and changing relationship among the law, social science, and empirical study of the courts.

CONCLUSION

In attempting to contribute to the study of judges' and juries' behavior in the "live" courtroom setting, we have worked with judges in a collaborative manner. Through this collaborative venture, we may begin to understand collectively the potential impact of judges' behavior on the people embroiled in the trial process. Indeed, this is a primary focus of the Conference. It is my hope that questions such as how people, the judges, jurors, counsel, parties, and the press, react to the operation of the legal system, and how they manifest their beliefs, attitudes, and biases about that system, can be addressed more systematically through empirically-based models, such as the one presented here. Through these efforts, I hope also to enhance the growing interdisciplinary effort to understand "what empirical research can tell us" about our evolving system of justice.

---

[112] Id. at 666 (emphasizing practical, economical, and societal importance of aggregating procedures in resolving mass tort litigation) (citing Hensler, Resolving Mass Toxic Torts: Myths and Realities, 1989 U. ILL. L. REV. 89, 90 (1989)).