

BYU Law Review

Volume 1975 | Issue 2

Article 5

6-1-1975

Critique - Data in Search of Theory in Search of Policy: Behavioral Responses to Videotape in the Courtroom

Gordon Bermant

Follow this and additional works at: <https://digitalcommons.law.byu.edu/lawreview>

 Part of the [Courts Commons](#), and the [Litigation Commons](#)

Recommended Citation

Gordon Bermant, *Critique - Data in Search of Theory in Search of Policy: Behavioral Responses to Videotape in the Courtroom*, 1975 BYU L. Rev. 467 (1975).

Available at: <https://digitalcommons.law.byu.edu/lawreview/vol1975/iss2/5>

This Article is brought to you for free and open access by the Brigham Young University Law Review at BYU Law Digital Commons. It has been accepted for inclusion in BYU Law Review by an authorized editor of BYU Law Digital Commons. For more information, please contact hunterlawlibrary@byu.edu.

Critique — Data in Search of Theory in Search of Policy: Behavioral Responses to Videotape in the Courtroom

Gordon Bermant*

*Do not trust data unsupported by theory*¹

This symposium marks a significant place in the history of the relationship between behavioral scientists and jurists in their attempts to understand the consequences of adopting the videotape medium for use in trial presentations and recording. The early literature on legal uses of videotape was concerned more with presenting the medium's law-related potential and with generating enthusiasm for the technology than with providing critical analysis.² Subsequent reviews dealt with accumulated court experience with the medium and outlined areas of social and legal concern.³ Then, surprisingly rapidly, data appeared from experiments with, and systematic assessments of, juror responses to videotaped material.⁴ Thus in the 4 years since Judge McCrystal arranged the first prerecorded videotape trial presentation (PRVTT),⁵ a reasonable body of legal and applied scientific literature has developed. By bringing together the most active participants on the behavioral science side of the issue, this symposium serves the important function of providing the legal community with a single source for review and analysis of the most recent findings. And perhaps more importantly, the symposium's articulation of both scientific and legal methods and concerns in the area of videotape technology serves as a model for the treatment of

*Coordinator, Behavioral and Social Sciences, Battelle Memorial Institute; Affiliate Professor of Psychology, University of Washington, Seattle, Washington.

¹Attributed to Lord Adrian.

²See, e.g., Morrill, *Enter—The Video Tape Trial*, 3 JOHN MARSHALL J. PRAC. & PROC. 237 (1970).

³See, e.g., Doret, *Trial by Videotape—Can Justice Be Seen to Be Done?*, 47 TEMP. L.Q. 228 (1974) [hereinafter cited as Doret]; Comment, *Videotape Trials: Legal and Practical Implications*, 9 COLUM. J.L. & SOC. PROB. 363 (1973); Comment, *Video-Tape Trials: A Practical Evaluation and a Legal Analysis*, 26 STAN. L. REV. 619 (1974).

⁴See, e.g., Bermant, Chappell, Crockett, Jacobovitch & McGuire, *Juror Response to Prerecorded Videotape Trial Presentations in California and Ohio*, 26 HASTINGS L.J. 975 (1975) [hereinafter cited as Bermant]; Bermant & Jacobovitch, *Fish Out of Water: A Brief Overview of Social and Psychological Concerns about Videotaped Trials*, 26 HASTINGS L.J. 999 (1975) [hereinafter cited as Bermant & Jacobovitch]; Miller, Bender, Florence & Nicholson, *Real vs. Reel: What's the Verdict?*, 24 J. COMMUNICATION, Summer 1974, at 99.

⁵McCall v. Clemens, No. 39, 301 (C. P. Erie County, Ohio, Nov. 18, 1971).

other issues in which the two perspectives need joining.⁶

Of course, scientists have a special obligation when they bring their methods to bear on issues of such practical importance as the operation of the courts. That obligation is to present their findings in ways that allow nonspecialists to separate the substantive ore from the methodological dross and establish priorities in regard to the likely practical significance of the results reported. However, this clarity must not be bought at the price of obscuring from view the scientist's methods and forms of reasoning. The relationship between results and conclusions must be clearly shown. The role of the critic in this sort of symposium, therefore, is to provide a separate, clarifying opinion on the important relations among the methods, results, and conclusions presented by the other participants.

I. IDENTIFYING THE RELEVANT ISSUES

Increased court-related use of videotape apparently offers a substantial number of potential benefits.⁷ When evaluating the benefits of videotape to the legal system, however, *two* classes of use must be distinguished, with the benefits of each class measured against their respective financial and social costs.

A. *The Use of Videotape as the Official Court Record: Issues and Problems*

First is the use of videotape as the official court record. While in principle many of the social and psychological issues associated with videotape trial presentations also arise in the context of trial records, in fact the problems are not so pressing because of the specialized and sophisticated audience for these records: appellate court judges. This is not to say that videotaped records provide no new problems, but only that they are not the same as those presented when juries, not judges, are the intended audience. As Short, Florence, and Marsh

⁶A recent example of the attempted joining of social science research and legal analysis, and one in which some controversy has arisen, concerns the relative merits of twelve- and six-person juries. The Supreme Court, in *Williams v. Florida*, 399 U.S. 78 (1970), and *Colgrove v. Battin*, 413 U.S. 149 (1973), relied partially on several empirical studies of jury size: INSTITUTE OF JUDICIAL ADMINISTRATION, *A COMPARISON OF SIX- AND TWELVE-MEMBER JURIES IN NEW JERSEY SUPERIOR AND COUNTY COURTS* (1972); Bermant & Coppock, *Outcomes of Six- and Twelve-Member Jury Trials: An Analysis of 128 Cases in the State of Washington*, 48 WASH. L. REV. 593 (1973) [hereinafter cited as Bermant & Coppock]; Note, *Six-Member and Twelve-Member Juries: An Empirical Study of Trial Results*, 6 U. MICH. J.L. REFORM 671 (1973); Note, *An Empirical Study of Six- and Twelve-Member Jury Decision-Making Processes*, 6 U. MICH. J.L. REFORM 712 (1973). The interpretations offered in these studies, and the Court's reliance on them, have been criticized by Zeisel & Diamond, "Convincing Empirical Evidence" on the Six Member Jury, 41 U. CHI. L. REV. 281 (1974).

The debate was renewed in a group of articles appearing in *Trial*. Thompson, *Six Will Do!*, TRIAL, Nov./Dec. 1974, at 12; Zeisel, *Twelve Is Just*, 10 TRIAL, Nov./Dec. 1974, at 13; and Saks, *Ignorance of Science Is No Excuse*, 10 TRIAL, Nov./Dec. 1974, at 18.

⁷McCrystal has listed 26 potential advantages of videotaped trials in McCrystal, *Videotape Trials: Relief for Our Congested Courts*, 49 DENVER L.J. 463 (1973) [hereinafter cited as McCrystal].

point out in their symposium article,⁸ there is some concern that appellate court review of videotaped records will overextend the appellate process. However, it seems reasonable to expect that official opposition to the use of video recordkeeping, based initially on traditions against cameras in the court and decisions like *Estes v. Texas*,⁹ will fade if it can be shown that the advantages outweigh the increased costs to the system. Questions here seem primarily legal and technical, not psychological or social.

B. Prerecorded Videotape Trial Presentations: Issues and Problems

The second major use of the medium is as a substitute for material that would otherwise be presented in a different form during a traditional trial. Partial uses include reports of sobriety tests, confessions, depositions, physical evidence, and so on. When the substitution is complete, so that all testimony and, perhaps, even the opening and closing arguments and judge's instructions are presented via videotape, then the triers of fact are faced with a complete prerecorded videotape trial presentation (PRVTT). It should be remembered that PRVTT is *not* the record of a live trial in a typical courtroom setting. Many advantages of the prerecorded trial are due to the range of time and place settings within which testimony can be taken. For example, in the 14 prerecorded land-condemnation cases heard in Judge McCrystal's court in 1974,¹⁰ the State of Ohio needed to transport its appraisers from Columbus, the state capitol, to Sandusky, the trial site, only once, at which time their testimony on all 14 cases was taken. Had the trials been conducted live, they would certainly have stretched out over a period of time that would have required the appraisers to make many trips. And in Judge McCrystal's second PRVTT, *Swain v. Norfolk & Western Ry.*,¹¹ involving an industrial accident, the plaintiff's testimony was taken on the accident site. Thus, for PRVTT and partial uses of the medium, there will generally be both spatial and temporal discontinuities in the material presented to the triers of fact. The editorial discretions to be allowed in the out-of-court taping sessions and in the final juxtapositions of separate testimonies constitute at least a theoretical problem for the legal community, because, as Eisenstein proved by example long ago, the juxtaposition of any two scenes on film produces a unique and sometimes unpredictable effect.¹²

⁸Short *et al.*, *An Assessment of Videotape in the Criminal Courts*, section IV, *A, supra* this issue [hereinafter cited as Short].

[Editor's note: Because of the time constraints imposed upon the publication of this issue, references by Mr. Bermant to the other articles of this symposium are cited to the appropriate section, part, and subpart of the articles and *not* to the number of the page containing the relevant material.]

⁹381 U.S. 532 (1965).

¹⁰The responses of the jurors to these PRVTT are reported in Bermant, *supra* note 4, at 988-92.

¹¹No. 39, 494 (C. P. Erie County, Ohio, Jan. 24, 1973).

¹²Eisenstein asserted that "*two film pieces of any kind, placed together, inevitably*

On the assumption that a jury is the intended audience for this videotaped material, psychological concerns arise on three levels. First, there are questions about the jurors' *cognitive or intellectual responsiveness* to the medium: how much and how well can they remember and sift through what is presented to them, and how does this compare with their memories and analytic capabilities when the same facts are presented to them in a live trial format? Second, there are questions about the jurors' *emotional responsiveness* to the medium: how does the presentation of legal material on videotape affect the level and quality of emotional arousal, particularly when compared with the effects of presenting the same material during a live trial? Third, and most important, there are questions about the jurors' *legal accuracy* in response to the medium: do the combined cognitive-emotional changes (if any) associated with the presentation of videotape material lead to changes in the validity of the jurors' judgments? As Doret put it in the title of his review of the videotape issue, "Can justice be seen to be done?"¹³

These three questions are easier to ask than to answer with certainty. This is not because the techniques currently available within psychology are incapable of providing reliable measures of cognitive or emotional responses under controlled laboratory conditions. Rather, it is the existence of four deeper problems that makes the "live v. videotape" question, intuitively simple as it may seem, particularly thorny. The problems have to do with the logic of experimentation.

1. The videotape medium: measuring differences between modes of presentation

The first problem involves the generality of comparisons. As discussed below in more detail, the research results so far available suggest that juror behavior may be at least as responsive to differences in mode of presentation within the video medium as to some live v. videotape comparisons. If this is true, then experiments that attempt to establish a definitive "the same" or "different" verdict in regard to the live v. videotape question are simply misguided, because, unless the particular video technique used in the experiment is the only legally permissible one, changes in the video technique (such as angles, lenses, switching ratings on views, and split-screen use) in other experiments or actual applications could lead to the opposite outcome. Put in somewhat more technical terms, the problem is to locate the relevant sources of variance and manipulate them simultaneously over significant enough portions of their ranges to determine what the important interactions are. It would be naive to imagine that one experiment, or one set of experiments, could establish definitive results for all the potential interactions among types of

combine into a new concept, a new quality, arising out of that juxtaposition." S. EISENSTEIN, *THE FILM SENSE* 4 (1942).

¹³Doret, *supra* note 3.

case material, characteristics of attorneys and witnesses, and characteristics of the medium of presentation.

2. *The null hypothesis and the problem of establishing similarities*

The second problem, related to the first, concerns the form of quantitative reasoning reported in the symposium. The inferential statistics employed by all three groups of investigators start with the arbitrary assumption that there are no differences to be found among the conditions being compared, for example, between videotape and live presentations. This assumption, known as the null hypothesis, comes under statistical test when an observed difference between the conditions is analyzed to determine the probability of a difference of that size arising by chance alone. The statistical test allows the investigator to arrive at a figure representing the probability that the null hypothesis is true. When that probability becomes acceptably small (typically, this means 5 or fewer chances in 100 that the observed difference occurred by chance), then the investigator rejects the null hypothesis and claims, with his confidence level quantitatively established, that there *is* a nonrandom difference between the conditions.

Clearly this form of quantitative reasoning is better suited for establishing differences than for establishing identities. One can never prove the null hypothesis, or the assumption of "no differences"; one can only establish the probability of its truth. Moreover, the null hypothesis is not derived from a theory or prior set of observations. It is rather a logically and statistically convenient starting place for comparative measurements and inductions about the likelihood that two samples of observations were drawn from the same underlying population. Hence, when the policy question is one of establishing a condition of "no difference," as in the question "Is a videotaped trial inferior in any legally significant way to a live trial?" these statistical techniques cannot provide the kind of definitive answer that would allow policymakers to act solely on the basis of the experimental outcomes.

3. *The absence of relevant theory*

The third problem has to do with the roles of normative and scientific theories in establishing conclusions relevant to policy. In regard to the advent of videotape in the courts, we are faced with an almost total lack of relevant theory at two levels. First, we have no substantive theory with which to predict or understand changes in legal decision making as a function of the medium of presentation. Second, we have no normative theory to inform us which among a set of observed decisions is legally the most accurate or appropriate. As a consequence of theoretical poverty, investigators are effectively forced to go fishing for experimental outcomes in the sea of legal, social, and psychological variables. And, because of the nature of a fishing approach to research, the investigators will invariably

find statistically significant differences. But unless we know both how to evaluate these differences in terms of their impact on practical legal outcomes, and, given that evaluation, whether the status quo *ought* to be changed in light of it, our experimental results just flop around like strange fish on the bottom of the boat: we really don't know if they are safe to eat. The symposium has landed a few such fish.

4. *The problem of determining levels of analysis*

Fourth and finally, there is the problem of determining the various levels of analysis at which the effects of videotape implementation are to be sought and measured. We have argued elsewhere that apparently disadvantageous features of communication during a live trial such as inadmissible testimony and unprofessional lawyer conduct may be advantageous in terms of the more general social instructions or messages these features transmit.¹⁴ This is a debatable matter. But there can be no doubt that full implementation of videotape (PRVTT) on a national basis would change the public's image of courtroom justice. Arguments regarding the nature and desirability of this change will revolve around concepts and values that transcend technical and first-order psychological considerations about videotape; they will deal with the social significance of importing additional technology into currently "technology-free" settings like the courts.¹⁵ However, this problem lies beyond the scope of the current symposium and will not be further discussed.

Let us now pay attention more explicitly to three of the symposium contributions. The studies conducted at Michigan State University (the MSU study)¹⁶ and Brigham Young University (the BYU study)¹⁷ need to be analyzed in parallel because of the partial overlapping of their methods and concerns. The study conducted by Ernest H. Short & Associates (the California study)¹⁸ is sufficiently different to be treated separately; it will be discussed first.

II. THE CALIFORNIA STUDY

The California study is divided into five sections, of which the third, "Psychological and Behavioral Impacts of Videotape," and the fifth, "Recommendations and Conclusions," will be of primary concern here. There are, however, two parts of the study's second

¹⁴See Bermant & Jacobovitch, *supra* note 4.

¹⁵See generally THE AMERICAN BAR ASSOCIATION AND THE AMERICAN INSTITUTE OF ARCHITECTS, *THE AMERICAN COURTHOUSE — PLANNING AND DESIGN FOR THE JUDICIAL PROCESS* (1973).

¹⁶Miller *et al.*, *The Effects of Videotape Testimony in Jury Trials: Studies on Juror Decision Making, Information Retention, and Emotional Arousal*, *supra* this issue [hereinafter cited as Miller].

¹⁷Williams *et al.*, *Juror Perceptions of Trial Testimony as a Function of the Method of Presentation: A Comparison of Live, Color Video, Black-and-White Video, Audio, and Transcript Presentations*, *supra* this issue [hereinafter cited as Williams].

¹⁸Short, *supra* note 8.

section that warrant comment. In the discussion of picture composition the authors recommend the use of zooms to close-up; justification for the recommendation is in terms of technical need, that is, to provide a more detailed view of a participant than the basic establishing perspective can provide. What is not discussed, however, is the extent to which the close-up view should be carried.¹⁹ That such a guideline is required is suggested by research results from our laboratory indicating that facial close-ups significantly alter the impressions formed by observers.²⁰ The California study makes a similar argument for split-screen technique in the discussion of special effects, and indeed, the justification for the use of this special effect is sensible. However, it is worthwhile to note that in the collection of results reported by the Michigan State group, one apparent exception to the generally null outcomes was the difference between split-screen and full-screen presentations on ratings of attorney credibility.²¹ Although this difference was only marginally reliable and is difficult to interpret in any case, the fact remains that even such cautious recommendations as those in the California paper may lead to unique sets of juror responses. This comment is not a criticism of the recommendation; it is rather an illustration of the point, made above, that without a prior commitment to what is desirable in juror responsiveness, rational decisions about techniques of various sorts are difficult to make. Thus, given the assumption that a certain videotape technique such as split-screen enhances estimates of attorney credibility, does one argue *for* split-screen on the basis that attorney credibility is a social desideratum or *against* it on the basis that jurors should be encouraged to develop a healthy scepticism of the advocates on both sides of the issue?

The primary behavior focus of the California study is on the responses of participants to the videotaping of approximately 75 preliminary hearings. Presumably, although it is nowhere stated in the paper, approximately one-half of the hearings were conducted under conditions of videotape recording. A total of 100 witnesses, 44 attorneys, and 14 judges were either interviewed or observed. The concern in the study is with the possible disruption of proper courtroom demeanor and responsiveness produced by videotape recording. The data deal only with the recording of preliminary hearings; however, the authors' assumption that their findings are generalizable to other in-court recording applications seems plausible. Unfortunately, no data are reported on juror responsiveness to recording or playback.

¹⁹Such a guideline is provided in the generally excellent manual *Guidelines for Pre-Recording Testimony on Videotape Prior to Trial*, prepared by the Federal Judicial Center in November 1974 and available from the Federal Judicial Center, Washington, D.C. 20234. There is a good deal of intuitive psychology in the prescription on page 22 that "[w]hen focusing on a witness to cover his verbal testimony, the operator should always maintain the witness' face in the picture, and should never go in closer than the entire face."

²⁰The research results are reported in a manuscript by Jacobovitch, Bermant & Crockett currently in preparation.

²¹Miller section II, c, 3, b.

Hence, the behavioral aspect of the study is primarily relevant to questions about videotape as a trial record and secondarily relevant to questions about videotape as a substitute for live trial presentations. The purely technical recommendations and cost estimates should be equally useful for both classes of videotape applications.

It is not possible to provide a detailed critique of the methods employed in the study because the authors do not supply the details. The repeated references to the use of analysis of variance techniques are particularly interesting given the heavy reliance on verbal reports that needed to be coded and placed in an acceptable metric before the analysis could be carried out. It would be of value to other investigators in the field if the California group would provide a detailed description of its coding procedures and the particular forms of variance analysis used.

The study does provide one example of the form in which judges and attorneys were asked to comment on witness conduct during videotaping sessions. The jurists were virtually unanimous in their disagreement with the assertion that "[w]itnesses are unresponsive to questioning when being videotaped." Taken on its own, the unanimity is uninformative because the form of the question effectively biases the outcome. Only perceptions of substantial video-induced witness taciturnity could lead one to agree with the assertion as worded. The authors report that these responses are supportive of self-report and observational measures. One would have liked the opportunity to assess the data generated by these other measures in more detail.

Although the scientific reader is left uneasy by the lack of relevant detail reported in the study, there seems little serious reason to doubt the fundamental conclusion that the addition of videotape recording equipment (as specified) to the setting of a pretrial hearing had little if any lasting behavioral impact on the principal participants. The study is overwhelmingly negative in regard to the finding of differences. If the assumption be valid that the pretrial hearing is a suitable model of a full trial, then the same conclusion would hold for trials. However, two cautionary points should be made. First, as the authors emphasize, the study does not report on the conduct of court principals after they have had the opportunity to view themselves on videotape playback. The self-confrontational aspect of exposure to the medium may have a measureable effect on the behavior of lawyers and judges.²² While this may be of more psychological than legal interest, it is a factor that should be investigated longitudinally, for it may provide insight into one of videotape's long-range consequences.

Second, conclusions about the minimal behavioral impact of the videotape recording procedure in pretrial hearings do not necessarily

²²For an overview of the effectiveness of videotape as a self-confrontational medium, see M. BERGER, *VIDEOTAPE TECHNIQUES IN PSYCHIATRIC TRAINING AND TREATMENT* (1970).

generalize to out-of-court settings. However, it seems plausible at this stage of our understanding to conjecture that the behavioral demand characteristics²³ of most legal settings are sufficiently intense and explicit that the addition of unobtrusive videotape recording equipment to them will not have large behavioral consequences for the participants.

The conclusions and directions for future research suggested by the California group are well chosen. But it is also important to emphasize that increased court-related videotape use will, inevitably, lead to variations in local practice that may produce abuses of the medium to the detriment of certain parties. For example, a source of potential abuse is the method by which pretrial videotaped materials are financed. One of our early concerns about prerecorded testimony was that it might differentially benefit those best able to pay the recording costs.²⁴ Cost-benefit analysis as practiced in the California study is from the perspective of the state. The significance of these figures is not the same in criminal and civil settings. The emphasis on criminal courts in the California study should not cause the reader to overlook the reality that the bulk of prerecorded trial experience has been and may well continue to be in civil settings. Research and analysis should be conducted to determine the fairest means of financing videotape use for the sorts of civil cases that currently produce the largest trial backlogs. Just rules for videotape cost-bearing combined with sound, strict technical guidelines for videotape use, such as those proposed by the Federal Judicial Center,²⁵ will go a long way to prevent abuse of the medium.

III. THE MICHIGAN STATE UNIVERSITY AND BRIGHAM YOUNG UNIVERSITY STUDIES

In contrast to the California study, the MSU and BYU studies deal with the influence of various forms and features of videotaped testimony on juror decision making. Both studies attack the "live v. videotape" question head-on and come up with generally similar results but different conclusions. Thus the MSU group ends up with a generally sanguine attitude about the widespread introduction of PRVTT: "There exist no strong grounds for arguing that videotape will exercise a negative impact on juror decision making."²⁶ The BYU group, on the other hand, argues for a more conservative policy

²³A behavioral demand characteristic is a feature of experimental or other social setting that tends to control the forms of behavior in that setting, independent of other presumably controlling factors. See Orne, *On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and their Implications*, 17 AM. PSYCH. 776 (1962).

²⁴Bermant, Chappell & McGuire, *Videotaped Trials: Advantages and Disadvantages*, June 1973 (unpublished paper presented at the First National Symposium on Crime and the Media, John Jay College of Criminal Justice, New York City, N.Y.).

²⁵FEDERAL JUDICIAL CENTER, *GUIDELINES FOR PRE-RECORDING TESTIMONY ON VIDEOTAPE PRIOR TO TRIAL* (1974).

²⁶Miller section V.

position. They are more concerned than the MSU group about the validity of current experimental procedures, and they would like to see more research done before implementation becomes widespread: "These and other findings of differences between live and media trials should stand as a caution to those proposing the immediate and widespread implementation of videotape trials."²⁷ The critique of these studies will include discussions of features that separate them as well as those they hold in common. It begins with some general comments that apply to both studies.

A. Features Common to Both Studies

1. Decision making: juror v. jury

First it must be emphasized that all the results reported in these studies refer to decisions by individual jurors (or mock jurors), not decisions by juries. Neither MSU nor BYU has investigated the impact of videotape presentations on the group decision-making process of juries. As a practical matter, the decision not to study the group process might be defensible on grounds of financial and temporal economy. Many more data can be collected in a given period of time with a given number of dollars if the experiments are terminated without actual jury formation and group discussion. But these are true economies only if the results obtained from individuals are predictive of the results that would have been obtained from groups. Put more technically, we need to be sure that there is not a significant interaction between the unit of decision making (individual v. group) and the medium of presentation (live v. videotape) before we can be confident that the differences (or lack of them) reported to arise in individual judgments as a function of the medium adequately predict the results for group judgments. What are the grounds, if any, for having this confidence? Neither the MSU nor the BYU group brings any evidence to bear on the issue. The MSU group announces its interest in the individual rather than the group and is careful to talk only about jurors and never about juries. The BYU group does not deal explicitly with the issue but mentions in a footnote²⁸ a result by Kalven to the effect that the verdicts of individual jurors on their first ballots is highly predictive of the eventual group decision. While this result is partially relevant, it does not get at the question of interaction directly.

On the other side of the issue are arrayed numerous results from experimental analysis of individual and group decision making.²⁹ For example, there is a large literature dealing with systematic shifts in decision outcome from the average of a set of individual decisions to the consensus decision made by a group composed of those individ-

²⁷Williams section IV, B.

²⁸Williams n.67.

²⁹See, e.g., Moscovici & Doise, *Decision Making in Groups*, in *SOCIAL PSYCHOLOGY — CLASSIC AND CONTEMPORARY INTEGRATIONS* 230-88. (Nemeth ed. 1974).

uals.³⁰ There is a host of both obvious and subtle influences that can work in group settings to move the consensus or final decision away from the decision that would be predicted on the basis of the pre-deliberation decisions of all the individuals involved. For example, in a very recent study designed to assess the influence of sexism on jury deliberations, we found that the pre-deliberation verdicts of groups of mock jurors did not vary as a function of the sex of the defense attorney in a simulated murder trial.³¹ However, during jury deliberations, issues arose which led to the rendition of significantly more guilty verdicts by juries exposed to a female defense attorney than by juries exposed to a male defense attorney. For present purposes, the point to be made is that legally significant changes can occur in individual decisions as a result of jury deliberations.

The most conservative assumptions under these circumstances are that the existence of differential media effects, measured in jurors, predict effects that would occur in juries, but that the absence of effects measured in jurors does not predict their absence in juries. Adoption of these assumptions in the present circumstance would force those who argue for a lack of practical difference between live and videotape presentations to prove their point by studying group deliberations. However, this approach is probably unrealistically pristine. The final section of the critique readdresses this issue in the context of policy recommendations.

2. *The use of actual jurors*

A second, briefer methodological point common to both studies concerns the experimental participants. Both the MSU and the BYU groups went to considerable lengths to maximize the validity of their findings by choosing research participants who were as much under the contextual sway of the courts as possible at the time of their participation. The cooperation of the Flint, Michigan, court in the initial *Nugent v. Clark* experiment is certainly to be applauded, as is the cooperation afforded the BYU team by the 4th Judicial District Court of Utah.³² The importance of basing conclusions about the

³⁰Research in this area was undertaken in order to comprehend the dynamics of what was claimed to be a decisional conservatism induced by group pressure on corporate boards of directors. Research revealed that for certain kinds of decisions, group influences led to the assumption of higher degrees of risk than would be predicted knowing the pre-deliberation preferences of the individuals. Subsequent research led to the more general hypothesis that the decisional shift induced by group processes was importantly related to the nature of the material under consideration. For a critical review of the relevant literature, see Cartwright, *Determinants of Scientific Progress — The Case of Research on the Risky Shift*, 28 AM. PSYCH. 222 (1973).

³¹Bermant & McGuire, *Effects of Sex of Attorney and Race of Defendant on Jury Verdicts in a Simulated Murder Trial* (manuscript currently in preparation).

³²There is, however, a disquieting feature of the initial MSU study that should not go unnoticed. The research participants were intentionally misled by the court during the conduct of the research. Of course, the rationale for the deception is clear enough, and the participants were subsequently debriefed concerning the true nature of the enterprise. Nevertheless, we need to question the advisability of trading, even slightly, the rule of honesty in the court against the need for experimental verisimilitude. This

behavior of jurors and juries on experiments involving actual jurors cannot be overemphasized. In general, one's confidence in the validity of the reported findings varies substantially with the authenticity of the participant population. If currently serving jurors are recruited as research participants, then at least one does not have to evaluate the likelihood that the observed results will generalize to the population of practical interest, for the jurors are that population.

3. *Media effects on the judgment of jurors*

Finally, the following basic point about the results of both studies needs to be emphasized: neither the MSU nor the BYU study has demonstrated that the medium of trial presentation has a statistically significant effect on the primary legal judgment of jurors.³³ Thus any disagreements between the two groups about the advisability of wide-scale videotape implementation, or the most desirable form thereof, are based on different interpretations of the influence of rather general psychological processes on primary legal judgments. As already mentioned, there exists no substantive theory for the unequivocal prediction of differential legal judgments as a function of variation in these psychological processes. Hence, at this stage, disagreements between the two groups are best construed as disagreements in trans-scientific judgment.³⁴

In turning now to a somewhat more detailed analysis of the separate reports, the critique will attempt to treat purely behavioral or technical matters only to the extent that they bear on legally relevant conclusions.

B. *The MSU Experiments*

1. *Critique of the Nugent v. Clark studies*

The *Nugent v. Clark* dramatization is the most complex and realistic trial simulation yet accomplished. The investigators were extremely sensitive to the demands of structural verisimilitude in simulation; the research sets a new standard for this work.³⁵ However, by following good experimental design principles and creating a videotape trial as similar to the live trial as technically possible, the investigators created a document different in many important par-

is a very serious question that should be debated by forums of jurists and behavioral scientists.

³³By the term "primary legal judgment," I mean assessment of negligence, guilt, size of appropriate award, etc., in other words, the class of judgments jurors are in fact called upon to make for legal purposes.

³⁴Weinberg, *Science and Trans-Science*, 10 *MINERVA* 209 (1972), provides a lucid exposition of the distinction between scientific and trans-scientific issues. Some important questions can be asked but not answered in the language of science. It is at this point that adversarial processes become most useful.

³⁵The concepts of structural and functional verisimilitude in trial simulation are spelled out in Bermant, McGuire, McKinley & Salo, *The Logic of Simulation in Jury Research*, 1 *CRIM. JUSTICE & BEHAVIOR* 224 (1974) [hereinafter cited as Bermant, McGuire, McKinley & Salo].

ticulars from a PRVTT. For example, a PRVTT as generally envisioned is composed of a series of segments—testimony of different witnesses recorded at different times and places. Thus a PRVTT necessarily lacks the continuity and perhaps cohesiveness which inheres in a live trial by reason of its presentation at one time in one place. It is of course unfair to criticize the MSU study for not doing what it did not set out to do; such criticism is not the intention here. The intention is rather to reemphasize the unique and basically open-ended characteristics of PRVTT that need to be examined in order to assess their behavioral impact. Thus, while the first MSU experiment was well designed and conducted, its negative results do not provide the kind of assurance we should like to have. When the authors conclude, “On the basis of this study and the impressions we gleaned while conducting the research, we find the videotaped trial format not guilty of any of the charges of detrimental effects on jury responses,”³⁶ they are speaking about a videotape trial format which is unlikely to be definitive of the model PRVTT of the future.

Nugent v. Clark Study 2, contrasting full-screen with split-screen videotape, landed one of the curious fish alluded to in the introduction. A statistically significant change was found in estimates of the credibility of one of the attorneys, Mr. Simmons. Faced with the finding, the investigators attempt to interpret it. The interpretation is fundamentally ad hoc. Nevertheless, as already mentioned, if the finding be taken seriously³⁷ it suggests that changes within the video medium can be more influential in determining certain aspects of juror behavior than are changes between live and video conditions. This interpretation highlights the need for explicit guidelines on videotape formats and makes the “live v. videotape” question seem less pressing.

The *Nugent v. Clark* Studies 3 and 4, dealing with the effects of deleting inadmissible testimony, are particularly interesting given the value placed on this feature of PRVTT by the video medium’s advocates.³⁸ The deletion of inadmissible testimony has been held as the major legal advantage of videotaped testimony. If, over a reasonably wide range of deletions, no appreciable changes in juror behavior can be found, then one of the major presumptive reasons for change loses some of its force. How much force it loses is a question of policy, not of science.

2. Critique of the information retention studies

The second section of the MSU study, dealing with information retention under live, color video, and black-and-white video condi-

³⁶Miller section II, B, 4.

³⁷It is in the nature of the statistical tests employed that statistically significant findings will appear accidentally at a rate directly related to the criterion point for defining significance. Thus, if the Miller group claims as significant results with a significance level of .05, then they are running a 5 in 100 chance of making a mistake about the nonrandom character of every significant difference they report. The more comparisons they make, the more likely they are to report false significances.

³⁸See, e.g., McCrystal, *supra* note 7.

tions, presents both behavioral scientist and lawyer with some formidable problems of interpretation. So many possibly differentiating variables were introduced between the first and second experiments of the session that it becomes difficult to determine with precision which of the several effects reported are to be considered practically significant. There was a provocative finding in the first experiment of the session that was not followed up, namely a faster decline, absolutely slight but relatively stable, in information retention in live as opposed to videotape conditions. No matter how small the magnitude of the effect (and it appears to be very small indeed), it would be worthwhile to know if it is replicable. A simple replication experiment to pin down the effect more securely would have been welcome. But instead we are left with the finding flopping around on the bottom of our intellectual boat, while the investigators move on to Study 2 in which almost every relevant variable has been changed.

Study 2 in this series provides information about estimates of witness credibility as well as information retention scores. For this experiment the information retention scores are not broken down over time so that the declining retention scores discovered in Study 1 could not be checked for replicability. Indeed, Study 2 is introduced as if Study 1 had shown reliably greater information retention for videotape than for live presentations. But that assertion is an over-interpretation of the experimental results. The observed significant interaction may have been due as much to higher retention of information during the first 13 minutes of the live presentation as to lower retention scores for the last 13 minutes. An overall information retention advantage for videotape cannot be asserted on the basis of the significant interaction between information retention and medium of presentation, over time, in the absence of a significant main effect for medium of presentation. Only a careful replication and extension of Study 1 can clarify this issue.

The observed interaction between apparent confidence of the witness and mode of video presentation in the determination of witness credibility scores is psychologically interesting, but its place in policy discussions about videotape implementation is unclear. However, the discussion of the "modal" personality is of practical interest. In the State of Washington, for example, there is a class of Superior Court trials in which all the testimony is read to the jury by counsel.³⁹ Thus the entire trial is conducted through the mediation of a "modal" witness. The results presented here, as well as the discussion of read transcript material in the BYU study, could be used to make an effective argument that the reading of the transcript

³⁹These are workmen's compensation cases in which the plaintiff, dissatisfied after a series of increasingly elaborate administrative procedures, takes the state's Department of Labor and Industries (DLI) to court. Plaintiff's counsel and the state's attorney cooperate in reading the testimony from the final administrative hearing to the jury; no new evidence is presented. A sample of 128 of these cases formed the basis for the empirical study of the effects of jury size on trial outcome reported in Bermant & Coppock, *supra* note 6.

should be replaced by the display of a videotape of the administrative hearing on which the transcript is based.

The findings of Study 3, in which the modal witness condition was dropped and an extended information retention test was administered for both black-and-white and color videotape, provide the basis for the MSU group's final recommendation that the less costly black-and-white format could be the medium of practical choice. These findings will be discussed later in a direct comparison with the findings and conclusions of the BYU group.

3. *Critique of the emotional arousal studies*

The application of galvanic skin response measurements to the black-and-white v. color videotape question raises so many psychophysiological and psychological issues that much more space would be required for adequate exposition than the results warrant. Even granting without analysis the authors' conclusion that black-and-white videotape produces greater emotional arousal than color videotape, there is no way to determine how that difference in arousal level would be influential in altering the practical reasoning jurors are called upon to perform.⁴⁰

C. *The BYU Experiment*

The material on which the BYU trial simulation was based was a straightforward land condemnation action involving one witness on either side of the suit. The question at issue was technical and, it seems fair to say, relatively dull given a juror's expectation for the kinds of material that might be presented in court. In fact, the ordinariness of the case material recommends it as the basis for simulation, for it is just such cases that can benefit from the timesaving and other administrative advantages claimed for PRVTT. Unfortunately, as the BYU authors emphasize, the strength in the case was primarily on one side, perhaps thereby preventing differences due to medium of presentation to be expressed freely.

The BYU investigators subjected their data to a complex process of statistical manipulations. Figures 1 through 7 of the BYU paper represent high-order abstractions from the original raw data provided by the mock jurors. One needs to consider the extent to which these numbers and the labels associated with them are accurate reflections of psychological realities with policy implications.

To understand what is at issue, consider Appendix 2 of the BYU article, in which are listed factor names associated with sets of pairs of polar adjectives. While there is a certain intuitive reasonableness in the groupings under the several labels, intuition was not the basis for the groupings nor should it be trusted on its own. The polar adjective

⁴⁰For one account of the likely realities of behavior in a jury room, see H. GARFINKEL, *STUDIES IN ETHNOMETHODOLOGY* 104-15 (1967). For an overview of the galvanic skin response literature, see W. PROKASY & D. RASKIN, *ELECTRODERMAL ACTIVITY IN PSYCHOLOGICAL RESEARCH* (1975).

pairs are listed together because of their relative affinity in a multi-dimensional space constructed out of a technique called factor analysis. What the technique does is to go repeatedly through the matrix of intercorrelations between polar adjective pairs and extract information (variance) from it in regard to which polar adjective pairs are related to each other. The results of the process are several sets of numbers. Each set is called a *factor*. The number in each set, one for every polar adjective pair, are the strengths or *loadings* of that polar adjective pair on that factor. The technique is intentionally constructed to insure that the various factors are uncorrelated (orthogonal); it is a mathematical Procrustean bed.⁴¹

Faced with sets of orthogonal factor loadings, the investigator undertakes difficult interpretive tasks. Each factor must be interpreted for its psychological significance, but there are no guarantees that such significance will be present. The results of interpretation are the *factor labels*. In the current instance, these are competency, honesty, friendliness, appearance, and objectivity. Factor labeling is a matter of art, and it is an open question whether the labels are properly denotative of psychological reality. Moreover, a judgment needs to be made before the labeled factor is brought to the discussion of the substantive problem under consideration. That judgment is based on the amount of information (variance) about the original correlation matrix brought together (accounted for) by each of the factors. It is in the nature of the technique that each successive factor brings together less information. Eventually the meaningfulness of the original matrix is exhausted, and the technique simply reiterates through the random remainders. The hazard facing the investigator at this point is overinterpretation of the matrix: a factor will be labeled that deserves no label. Again, this is a matter of art and judgment. But as a rule of thumb, it has been suggested that factors with eigenvalues less than 1.0 ought not to be interpreted.⁴²

Given this background, how are the BYU factors to be evaluated in regard to their usefulness for making intermedium comparisons? All the factors are, by the eigenvalue rule, technically interpretable. However, in my opinion, the interpretations of appearance, objectivity, and handsomeness—calmness are stretched beyond the likely significance of the findings. Although technically proper, the analyses seem psychologically overextended.

It is interesting to compare the BYU findings for competency and honesty with the MSU findings on witness credibility from Study 3 of the information retention studies. If these various labels be taken as valid denoters of psychological constructs, then it is reasonable to

⁴¹For a useful source of technical information, consult H. HARMAN, *MODERN FACTOR ANALYSIS* (2d ed. 1967). Factor analysis was originally used in psychology in respect to studies of intelligence, and has subsequently found favor in several research areas. Unfortunately, it is also often used in attempts to create order out of chaos algorithmically; the attempts seldom succeed.

⁴²*Id.* I thank Professor Allen Edwards, Dr. Carl Bennett, and Dr. Michael Lindell for advice on this point.

equate credibility with competency and honesty. MSU found that estimates of witness credibility were lower for black-and-white than for color videotape. However, BYU did not find lower competency or honesty ratings for the monochromatic medium. The lack of congruence on this issue highlights one of the difficulties of comparing research results from different laboratories when there are no agreed upon standards and dimensions of measurement. The critique returns to this point later with the recommendation that greater effort be exerted to standardize methods in this area of research.

Parts *B* through *E* of section III (the research results) of the BYU article present a welter of analyses that are difficult to place together in a coherent framework. Take for example two questions of general policy relevance: (1) Are there reasons in these data to suggest a slow down or prohibition of the use of PRVTT? (2) Are there reasons to prefer color videotape to black-and-white? A comparison of the answers to the two questions supplied by the separate analyses of data appears in the table below.

Analysis	Live v. Video	Black-and-White v. Color
Landowner compensation (absolute amounts and distributions)	No significant difference reported	No significant difference reported
Dollar awards \times juror ratings	Landowner predicts outcome in video, expert in live	No significant difference reported
Preferences for trial participants	Live differs from black-and-white, not from color	No significant difference reported
Juror reactions to trials	Live differs from color in ease of attention, not at all from black-and-white	No significant difference reported

One of the observed differences between live and videotape presentations—in the relationship of the dollar awards to juror ratings of trial participants—must be interpreted in light of the particular circumstances of this trial. The other two differences, in which the live presentation produced different outcomes from one of the videotape presentations (color or black-and-white) but not the other, are difficult to interpret because the two videotape presentations are not reported to be different. Thus, the data are equivocal on the major policy issues. There are, however, other relevant issues addressed by the data. In particular, the study provides good reasons for replacing

the reading of transcripts by audio or video communication whenever possible.

IV. CONCLUSIONS AND RECOMMENDATIONS

First, in regard to the general policy of videotape implementation: there are *not* sufficient grounds in any of the studies reported to warrant holding back careful, on-site evaluation of PRVTT in a wide variety of cases and settings. This conclusion is not in disagreement with our earlier conclusion that other considerations may warrant a conservative timetable for the increased utilization of PRVTT.⁴³ While the BYU investigators may be correct in their concern about videotape implementation, their data do not, in my opinion, support their concern.

This conclusion does not imply that no additional research on PRVTT or partial videotape uses is required, but rather that the research be extended to include actual trials. We need a carefully designed evaluation of videotape usage in courts of sufficient case volume that random assignment of cases to videotape or live presentation would, in a brief period, generate a sufficient data base. There are probably numerous legal and administrative issues to be addressed and resolved before this program of evaluation could be undertaken. All concerned parties should participate in the development of the evaluation program. Conclusions from standard research methodologies need to be supplemented with the seasoned opinions of judges, lawyers, court reporters, and administrators. Bringing an increased technological load into the courtroom process may alienate some. Any possible alienating effects could be reduced, however, by soliciting the active assistance of all who will be affected by the innovation in constructing details of its operation.

Second, in regard to the black-and-white v. color videotape question: the data presented by the MSU and BYU groups suggest that the decision need not be forced by first-order psychological considerations. In my opinion, concern about biasing effects of black-and-white videotape is not supported by the data presented in the symposium. This does not mean, however, that courts should opt for black-and-white video on the grounds of short-run economy. Whatever the case be now, there is little doubt that in a relatively few years black-and-white videotape will have all but disappeared from American life. The greater cost of color videotape equipment at present needs to be weighed against its eventual ubiquity.

Third, experimental research of the sort presented in the symposium should be continued with greater standardization of materials and methods. There is great scientific value in straightforward replication of experiments in different laboratories. Scientific understanding of media effects on legal decision making will grow more rapidly

⁴³Bermant & Jacobovitch, *supra* note 4.

and surely if investigators can agree in advance on the materials and measuring instruments they will use.

Fourth, there is a need for sound testable theory relating the rational processes of legal decision making to the nonrational effects of changes in media of communication. As a minor contribution to such theoretical development, the following hypothesis is offered for experimental test: *the more evenly balanced or ambiguous the legal issues on the two sides of a case, the more influential will be the extralegal factors in the case, including the medium through which the case is presented to the jury.* This simple idea, if properly refined and operationalized, could serve as the theoretical foundation for a number of interesting and practical experiments.

The offered hypothesis will perhaps render an additional service. The attempted operationalization of the concept "even balance" in regard to the legal issues in a case should provide the beginning of the kind of normative theory of legal decision making that is required in order to properly evaluate the influence of technological changes in courtroom practice. The pioneering work of Kalven and Zeisel in comparing decisions of judges and juries is an example of the kind of data base required to develop the relevant normative theory.⁴⁴ Cases of varying degrees of balance could be devised by reference to the evaluations of expert panels on the appropriateness of decisions made by juries in large numbers of cases.⁴⁵ Alternatively, simulated cases could be created wherein the degrees of legal balance or ambiguity were systematically varied.⁴⁶ In either case, trials of the cases could be presented through different media, and the impact of the medium on trial outcome could be determined. Other extralegal factors could be investigated in the same experiments, in much the same way as has already been done by the MSU and BYU groups. The advantage of the method suggested here is that one would have some idea about how the results *ought* to turn out because one would then have some standards of justice and equity against which to measure results. This calibrating procedure would be of substantial value in moving the issue of videotape trials from the abstractions of laboratory research to the concrete realities of courtroom technology.

⁴⁴H. KALVEN & H. ZEISEL, *THE AMERICAN JURY* (1966).

⁴⁵This procedure would expand the concept of functional verisimilitude of trial simulation as described by Bermant, McGuire, McKinley & Salo, *supra* note 35.

⁴⁶Consider for example the typical automobile-accident-at-an-intersection case wherein both parties claim to have had the green light. Over a large number of such cases juries find for the plaintiff approximately half the time, *i.e.*, as if flipping a coin. *JURY VERDICT RESEARCH INC., 3 PERSONAL INJURY VALUATION HANDBOOKS, LIABILITY RECOVERY PROBABILITIES* 26a-27 (1970). It may be conjectured that when a decision hinges on who is telling the truth and who is lying, one expects extralegal factors to be maximally effective in determining outcomes. Apparently the case presented by the BYU group represents a substantially greater degree of imbalance.