1990

Do Jurors Understand Criminal Jury Instructions? Analyzing the Results of the Michigan Juror Comprehension Project

Geoffrey P. Kramer
Indiana University at Kokomo

Dorean M. Koenig
Thomas M. Cooley Law School

Follow this and additional works at: https://repository.law.umich.edu/mjlr

Part of the Courts Commons, Criminal Law Commons, and the State and Local Government Law Commons

Recommended Citation

This Article is brought to you for free and open access by the University of Michigan Journal of Law Reform at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in University of Michigan Journal of Law Reform by an authorized editor of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.
In our jury system, laypersons rather than legal experts determine the guilt or innocence of persons accused of crime by applying legal standards provided by a judge in the form of jury instructions. The complexity of some legal standards raises concern about the jurors' ability to perform their duties. Their responsibilities can be divided into three parts. First, jurors must determine the facts in the case before them. Second, they must
learn the relevant legal standards from the judge. Finally, they must apply those standards to the facts of the case. Confusion over the meaning of judicial instructions undermines the jurors' ability to fulfill their second function, that of learning the legal standards, and consequently affects their application of the law to the facts. The study of 600 actual jurors discussed in this Article sought to determine whether jurors are able to learn the law from complex instructions. The study's data support the view that juror understanding of judicial instructions is flawed and that changes in the wording and presentation of judicial instructions may be required to improve juror comprehension.

The use of laypersons to convict or acquit accused criminals was drawn from the English system into American law via the sixth amendment to the United States Constitution, which states in part: "In all criminal prosecutions, the accused shall enjoy the right to a speedy and public trial, by an impartial jury of the State and district wherein the crime shall have been committed. . . ."

A jury trial in a criminal case follows a set and orderly pattern. At the end of the trial, after all the evidence and argu-

2. For a thorough description of the study, see infra Part I.
3. For a complete analysis of the results, see infra Part II.
4. U.S. CONST. amend. VI. In 1968, the sixth amendment right to a jury trial was "incorporated" as one of the fourteenth amendment's fundamental liberties and made applicable to any significant state criminal proceeding. Duncan v. Louisiana, 391 U.S. 145, 149 (1968). The standards for state and federal trials are not entirely congruent, however. In Apodaca v. Oregon, 406 U.S. 404, 406 (1972), for example, the Court concluded that a state law allowing a nonunanimous jury to return a verdict was constitutional. In a federal trial, however, absent a waiver, the Constitution requires a unanimous verdict for a criminal conviction. Johnson v. Louisiana, 406 U.S. 356, 369 (1972) (Powell, J., concurring).

6. The petit jury is first impaneled from the larger panel of prospective jurors in a "voir dire" proceeding, at which either the attorneys or the judge questions the jurors to uncover any potential bias. See T. MAUET, FUNDAMENTALS OF TRIAL TECHNIQUES 25 (2d ed. 1988). The attorneys ordinarily make opening statements, although the defense attorney may wait until after the prosecution has presented his entire case. See id. at 51. The trial itself consists routinely of the following: 1) the prosecutor's direct examination of her witnesses; 2) the defense counsel's cross-examination of the prosecution's witnesses;
ments have been presented, the judge instructs the jury on the legal standards that they are to apply. The jurors are told that they are the sole judges of both the facts of the case and the credibility of the witnesses. Most importantly, they are instructed on the elements or constituent parts that the prosecution must prove for each crime with which the defendant is charged. They also are instructed on the concepts of reasonable doubt and the burden of proof.

The jury is then sent to the jury room to deliberate without outside interference. Jury deliberations are not recorded and ordinarily postverdict statements or testimony by jury members may not be used to impeach the verdict except in those few instances where it is shown that outside matters improperly influenced the jury. Judges and attorneys may not contact the jurors during their deliberations, and jurors receive very little help when they are unclear about what is necessary to convict. Thus, juror comprehension of judicial instructions upon recital is crucial to a legally correct verdict.

The function of the jury instructions, however, does not stop with recital of instructions to the jurors and application of in-

---

3) the prosecutor's redirect and the defense counsel's recross; 4) the repetition of the same process with the defense counsel now conducting direct examination on his witnesses, and so on. See id. at 151.

7. See, e.g., Fed. R. Crim. P. 30. Unfortunately, judicial instructions given after the presentation of evidence come too late to assist jurors in analyzing and processing the facts of the case. See, e.g., Kassin & Wrightsman, On the Requirements of Proof: The Timing of Judicial Instruction and Mock Juror Verdicts, 37 J. Personality Soc. Psychology 1877 (1979) (indicating that mock jurors who were given pre-testimony instructions on the requirements of proof demonstrated a reduced tendency to convict).

8. "[T]he factfinder's responsibility at trial, based on evidence adduced by the State, is to find the ultimate facts beyond a reasonable doubt." Francis v. Franklin, 471 U.S. 307, 316 (1985) (emphasis added) (quoting Ulster County Court v. Allen, 442 U.S. 140, 156 (1979)).


10. "Lest there remain any doubt about the constitutional stature of the reasonable-doubt standard, we explicitly hold that the Due Process Clause protects the accused against conviction except upon proof beyond a reasonable doubt of every fact necessary to constitute the crime with which he is charged." In re Winship, 397 U.S. 358, 364 (1970).

11. See infra Appendix II.


13. Id.

14. In fact, judges routinely respond to juror requests to explain instructions by merely rereading verbatim the same instruction. The judge is not obliged to give all of the instructions previously given and need only repeat those specifically asked for by the jury. See People v. Johnson, 167 Mich. App. 168, 175, 421 N.W.2d 617, 620 (1988).
structions to the perceived facts. Jury instructions in criminal cases form a basis for legal arguments raised on appellate review of a criminal conviction.\textsuperscript{15} Thus, although appellate courts generally may not review jury determinations,\textsuperscript{16} the instructions that were given to the jury are duly recorded and counsel may point to errors in the instructions as reason for reversal on appeal. Jury instructions thus form the basis on which the appellate courts shape the substantive criminal law. In order to minimize the use of jury instructions as a tool to obtain appellate reversal, "pattern" or "standard" instructions have been developed.\textsuperscript{17} Perhaps the emphasis placed on technical accuracy explains why the primary role of jury instructions—to teach jurors about the law—has arguably been overlooked and underrepresented in the literature and in the cases.\textsuperscript{18} Highly technical and relatively incomprehensible jury instructions that are legally

\textsuperscript{15} Appellate courts have been reluctant to reverse convictions following jury trials, basing their refusals on court rules as well as case law. They ordinarily will not reverse a conviction even if the jurors did not fully understand the jury instructions. See People v. Lee, 220 Cal. App. 3d 320, 329, 269 Cal. Rptr. 434, 439 (1990). \textit{But see} Hoffert v. Florida, 559 So. 2d 1246 (Fla. Dist. Ct. App. 1990); People v. Tucker, 193 Ill. App. 3d 849, 550 N.E.2d 581 (1990).

One barrier to reversing convictions based on poorly comprehended instructions is that courts will read the instructions "as a whole." City of Minot v. Rubbelke, 456 N.W.2d 511 (N.D. 1990). A concurring judge in People v. Lee commented on this lack of concern over comprehension:

\begin{quote}
One juror then noted: "We have read over several pages again and again and we are having a little bit of difficulty with when specific intent begins." After the judge stated that he could not further clarify the point, another juror asked: "How do you find intentions?" The question was not answered. \ldots It is clear from the above the jurors were having difficulty with the concept of specific intent generally. \ldots I am therefore unpersuaded by that too-oft employed bromide that failure to give a correct instruction may be cured by viewing the instructions as a whole.
\end{quote}

220 Cal. App. 3d at 329, 320, 269 Cal. Rptr. at 434, 439 (White, J., concurring).

A second barrier, derived from court rules, precludes reversal if the defendant did not object adequately to the jury instructions, unless the alleged error in those instructions constitutes "plain error" or "fundamental error." See, e.g., United States v. Jackson, 569 F.2d 1003, 1009 (7th Cir.), cert. denied, 437 U.S. 907 (1978) (finding that even the rejection by the court of twice-tendered valid instructions did not constitute grounds for reversal when the counsel failed to make a distinct objection); State v. Odom, 307 N.C. 655, 661, 300 S.E.2d 375, 379 (1983) (holding that, in applying the "plain error" test, the reviewing court "must examine the entire record and determine if the instructional error had a probable impact on the jury's finding of guilt").


18. Some authors, however, have recognized the importance of comprehensible jury instructions. For a thorough discussion regarding comprehensible jury instructions and
“correct” but less than useful in practice have supported guilty verdicts on appeal.19 No matter how technically accurate an instruction is, it fails in its primary purpose if jurors are unable to use it to deliberate and determine verdicts.

In response to this concern, the Juror Comprehension Project (“the Project”) sought to determine whether jurors understand judicial instructions. This Article reports the results of an empirical study growing out of that Project. The Project investigated how well 600 actual jurors in Michigan understood criminal jury instructions in actual trials. Part I describes the history of the study and explains the procedures and materials used in the study. Part II presents the results of the study, first analyzing juror comprehension of selected concepts, then discussing general factors that influence juror comprehension. Part III concludes that the results show a mixed juror understanding of complex judicial instructions, discusses this mixed understanding, and argues for changes in the current method of jury instruction, including the use of written instructions and simpler language.

I. THE STUDY

The study undertaken by the Juror Comprehension Project is the result of a collaboration between social science researchers from Michigan State University and legal professionals—judges, prosecutors and defense attorneys who were members of the Michigan Criminal Jury Instructions Committee (the “CJI Committee” or the “Committee”).20

an alternative method of evaluation, see A. ELWORK, B. SALES & J. ALFINI, MAKING JURY INSTRUCTIONS UNDERSTANDABLE (1982).

19. Some of the most confusing instructions have been in the terms of art found in homicide law. See, e.g., People v. Kelly, 423 Mich. 261, 292-303, 378 N.W.2d 365, 375-84 (1985) (Levin, J., concurring in part and dissenting in part) (asserting that trial judge gave incorrect felony-murder instruction, telling the jury that it might infer from defendant’s participation in robbery that he had the requisite intent to support a conviction for murder); People v. Borgetto, 99 Mich. 336, 339, 58 N.W. 328, 329 (1894) (holding unobjectionable a jury instruction that stated that malice includes “not only anger, hatred and revenge but every other unlawful and unjustifiable motive”). The arguments against such language are discussed in People v. Morrin, 31 Mich. App. 301, 310-24, 187 N.W.2d 434, 438-46 (1971). See also Koenig, Jury Instructions for Jurors: Proposals for the Simplification of Michigan Instructions on Murder, 21 WAYNE L. REV. 1, 3-4 (1974); Purver, The Language of Murder, 14 UCLA L. REV. 1306, 1308-10 (1967).

For a case other than homicide, see People v. White, 168 Mich. App. 596, 604-06, 425 N.W.2d 193, 197-98 (1988) (upholding jury instruction as to whether defendant was armed with a weapon).

The collaboration began with a 1987 survey of Michigan judges. The 1987 survey helped identify perceived weaknesses in Michigan's standard criminal jury instructions, because judges giving instructions in criminal cases are in the best position to recognize difficulties that jurors might encounter. Based on the results of the 1987 survey, certain instructions were targeted as likely candidates for juror confusion.

A. Overview of the Study

Once the 1987 survey identified potentially problematic instructions, the Project developed a short, easy-to-complete questionnaire. This questionnaire consisted of both true-false and open-ended questions. Actual jurors received this questionnaire immediately after they finished serving on a trial. The Project assessed juror comprehension of instructions by comparing the responses of jurors instructed on a particular standard with jurors who had not received such an instruction. The presumption, of course, was that jurors exposed to instructions would score higher than the jurors not exposed.

This procedure yielded a great deal of data about the comprehension level of ordinary jurors. It also tested comprehension following actual trials and did so without interfering with appellate review of the instructions. Finally, it demonstrated that social science researchers and legal experts can work together productively to test empirically the effectiveness of jury instructions.

B. The Questionnaire

Each true-false question was intended to test comprehension of one jury instruction or one aspect of an instruction. The open-ended questions, which were designed to probe for additional sources of misunderstanding about jury instructions, allowed jurors to report their thoughts and experiences in trying to comprehend and use instructions.

The items contained in the questionnaire were developed in the following manner. First, Dorean Koenig, Geoffrey Kramer and Norbert Kerr wrote approximately seventy-five true-false

21. Id. at 369-71.
22. For a copy of the questionnaire, see infra Appendix I.
questions and submitted them to the CJI Committee for review by committee members. The committee members scrutinized these questions for legal accuracy and clarity. They recommended keeping or discarding each item, suggested rewording, commented about possible misinterpretations, and suggested alternate questions. Committee members also indicated whether the question, as worded, was true or false. Their responses for each question were tabulated and compared. The Committee only retained questions for which there was unanimity about a correct answer and for which there was a clear consensus on the question's relationship to the targeted instruction. The remaining questions were revised by the Committee, and this procedure was repeated with members of the Juror Comprehension Subcommittee of the CJI Committee, as well as a member of the Plain English Committee of the Michigan State Bar. The CJI Committee made every effort to word the questions in simple language. The questionnaire was again submitted to the CJI Committee for feedback and final approval.

The complexity of the items on the questionnaire varies considerably. This variation was intended for two reasons. First, it reflects a reality about jury instructions themselves. Some instructions seem straightforward while others engender considerable debate, even among legal professionals. Second, it reflects the differing abilities of the jurors.

Some instructions are confusing because the underlying legal standards are vague and indefinite. This is indicative of an unwillingness in the law to refine difficult questions that produce substantial differences of opinion among lawyers and judges. It is easier to leave technical and expansive jury instructions in a form that may be incomprehensible to the layperson because legal terminology is not ordinarily a ground for reversal of a conviction. In other words, because juror confusion is not ordinarily grounds for reversal as long as the underlying instructions are technically correct, our legal system provides little incentive to correct errors in juror comprehension. Obviously, if jurors cannot ascertain what the law requires, their decisions will be

23. For an example, see the discussion of reasonable doubt, infra notes 45-56 and accompanying text.

24. See Koenig, Kerr & Van Hoek, supra note 17, at 354-55 (giving an example of judges' disagreement over "reasonable doubt"). Koenig, Kerr, and Van Hoek also discuss the history of the "reasonable doubt" instruction in Michigan, id. at 356-58, and summarize the four different methods used in the United States to explain the concept of reasonable doubt to juries, id. at 372-73.

25. See id. at 362.
tainted. Under our present system, however, there is no method to redress this confusion.

In addition, as a matter of test construction, it is desirable to include items that vary in difficulty. A questionnaire with uniformly easy items probably would do little to assess jurors' comprehension of more difficult aspects of instructions, while uniformly difficult items might differentiate comprehension levels among only the brightest jurors. Variability in the difficulty of items, however, might reveal whether jurors comprehended the basic meaning of the instructions as well as their more subtle distinctions.

C. Procedure

Before using the questionnaire in earnest, we conducted a pilot test of it on undergraduate students in a jury simulation and on a small sample of actual jurors following trials. We hoped to identify questions that might present unforeseen problems. For instance, if virtually all jurors in the pilot sample answered a question correctly, further testing with that question probably would be uninformative. In that case, we would replace the question with another that might better reveal possible sources of juror misunderstanding. In addition, we conducted follow-up interviews with student pilot subjects to reveal unforeseen ambiguities in the wording of the questions. We made minor revisions as a result of this testing.

The first task in administering the questionnaire to actual jurors was to secure the cooperation of a number of courts. The Project contacted several judges from the Detroit Recorder's Court and the Thirtieth Judicial Circuit Court of Lansing, Michigan. These courts were selected because they hear a relatively high number of cases, which facilitated data collection. Specific judges were contacted in order to achieve a balance between judges who had defense-orientated backgrounds and judges who

---

26. The desirability of varying the difficulty of items in a test depends upon the purpose for which the test is given. Variability in item difficulties for our questionnaire was desirable because jurors' abilities were likely to vary broadly. Further, our purpose was to assess whether jurors understood both the elementary and complex aspects of particular instructions. Therefore, questions that varied in difficulty were constructed. It is worth noting that the overall test difficulty was very close to 50% (i.e., half of the items answered correctly), though the range of individual item difficulty varied roughly from 25% to 95% for uninstructed jurors. For a discussion of the relationship between test item difficulty and testing purposes, see A. Anastasi, Psychological Testing 205-06 (4th ed. 1976). See also J. Nunnally, Educational Measurement and Evaluation 133 (1964).
had prosecution-oriented backgrounds. Only those judges who routinely used the Criminal Jury Instructions verbatim participated. A total of twelve different courts eventually provided data, seven from Detroit and five from Lansing.

The Project administered the questionnaire to actual jurors immediately after they finished serving on a trial. Juror participation was entirely voluntary. At the completion of a trial, either the trial judge or the court clerk explained briefly the purpose of the questionnaire and requested that jurors complete it individually and at their own pace. The court exercised its discretion as to whether the judge or a court clerk administered the questionnaire and explained its purpose. Not surprisingly, higher percentages of jurors volunteered in the courts in which the judge administered and explained the questionnaire.\textsuperscript{27} Completed questionnaires were collected and mailed to Michigan State University for analysis. A checklist, filled out by the court clerk, accompanied the questionnaires from each jury, indicating which criminal jury instructions those jurors had heard.

\textit{D. Design}

The study focused on a comparison between instructed and uninstructed jurors. The instructed jurors were those who had been exposed to targeted instructions during the trial. The uninstructed (control) jurors were those who were not exposed to targeted instructions. Because all jurors were requested to answer every question on the questionnaire, each juror could function as an instructed juror for some questions and an uninstructed juror for other questions. For example, jurors who served on a felony murder case heard instructions about felony murder, but not about assault.\textsuperscript{28} They served as instructed jurors for questions dealing with felony murder and uninstructed jurors for questions dealing with assault.\textsuperscript{29} Persons who did not

\textsuperscript{27} The Lansing courts also requested that jurors complete a different, internal questionnaire. This additional request may have reduced participation in these courts.

\textsuperscript{28} Assault is not a felony that will support a conviction for felony murder in Michigan. See Mich. Comp. Laws Ann. § 750.316 (West 1968 & Supp. 1989).

\textsuperscript{29} Unfortunately, we do not have any data comparing control jurors who never served with control jurors who served but were not exposed to particular instructions. It seems reasonable to suppose that the two groups are comparable in their responses to questions for which they did not hear instructions. Jurors who did serve were initially selected from the same pool as those jurors who did not serve, all jurors being selected via a random process, so there should be no systematic differences between the groups. It is also difficult to specify how exposure to some instructions (e.g., reasonable doubt, as-
serve on a jury comprised the control group for questions concerning reasonable doubt and witness credibility because reasonable doubt and witness credibility instructions are given to all jurors who serve in a criminal trial. The rates of comprehension for uninstructed jurors, as revealed by the number of correct and incorrect responses, served as a baseline to evaluate the comprehension of instructed jurors and to make inferences about the effects of the instructions.

E. The Sample of Jurors

In total, 882 jurors, including those who served on a trial and control jurors who did not serve, were asked to complete the questionnaire. Of those requested, 600 completed the questionnaire, resulting in a compliance rate of approximately 68%. For the sample that completed questionnaires, the gender distribution was 50.2% female and 49.8% male. They ranged in age from 19 to 71, with a mean age of 39 and a median age of 37. Their educational level ranged from 7 to 20 or more years of education, with a mean of 13.4 years and a median of 12 years of education.

Testing actual jurors following actual trials has the advantage of realism, but it also presents certain limitations. Primary among these is the lack of control over a variety of factors that might influence juror comprehension. For instance, jurors who volunteer to take a questionnaire may be brighter and generally may have higher rates of comprehension than jurors who do not volunteer. This possibility does not have strong support. Of juries that returned questionnaires, 60% had all twelve members complete a questionnaire, and 75% had at least eleven of twelve members complete the questionnaire. Thus, self-selection within juries was relatively uncommon—either virtually all jurors from a jury completed the questionnaire or few or none did. Also, discussion of jury instructions may occur during deliberation, im-
proving the comprehension of jurors who initially did not understand an instruction. It is therefore difficult to make strong assertions about juror comprehension before deliberation, based on the results of a questionnaire administered after deliberation. One would hope that deliberation improves comprehension of instructions, but postdeliberation testing of juror comprehension may not provide an unambiguously clear picture of how comprehensible certain instructions are when they are given. Some of the limitations of testing actual jurors could be controlled by our procedures, but other limitations could not.

An important question is whether our volunteer sample reflects a truly random sample of jurors. Neither the Detroit nor the Lansing courts keep records of juror demographics, so no direct comparison with their jury pools can be made. A sample of 205 jurors chosen at random over a 16-month period from the Lansing courts does provide a basis for comparison. This random sample had a mean education level of 13.3 years, nearly identical to the 13.4 mean for the volunteer sample. The mean age for that randomly chosen sample was 39.3, again very close to the 39.0 mean age for the volunteer sample. Finally, the proportion of men and women in our volunteer sample does not differ significantly from the proportion in the general population.

We believe our volunteer sample, therefore, is representative of a random sample of jurors drawn from a metropolitan area in Michigan.

The variation in volunteer rate was mostly a function of the courts from which the jurors came. Jurors frequently declined in some courts, while in other courts very few or no jurors declined. Follow-up telephone interviews with the personnel who administered the questionnaires revealed that declining was primarily related to who made the request. There was much greater com-

31. See infra notes 63-64 and accompanying text.
32. A comparison with one relevant population would be ideal. Unfortunately, such a population does not appear to be available. We could find no statewide demographics for jurors (which is not surprising because neither Detroit nor Lansing keep such records). Therefore, we used the sample from a study by Kramer, Kerr, and Carroll for comparison. See Kramer, Kerr & Carroll, Pretrial Publicity, Judicial Remedies, and Jury Bias, 14 LAW & HUM. BEHAV. ___ (forthcoming 1990) (draft on file at the University of Michigan Journal of Law Reform).
33. Id.
34. Id.
36. The educational level of persons from Lansing is slightly higher than the educational level of persons from Detroit. Id. at 668. Thus, our volunteer sample might slightly overestimate the average education level of Detroit's juror pool.
pliance among Detroit Recorder's Court jurors, where judges typically made the request, than among Lansing Circuit Court jurors, where clerks typically made the request. In a few instances, court personnel indicated that declining also may have been related to time of day, with more jurors declining in the late afternoon after long deliberations. No one reported that declining seemed obviously related to juror characteristics such as helpfulness, age, intelligence, or reading ability.

II. RESULTS OF THE STUDY: JUROR COMPREHENSION OF INDIVIDUAL INSTRUCTIONS

To examine the effect of each instruction, the Juror Comprehension Project analyzed every question by comparing average percentage correct obtained by uninstructed jurors with average percentage correct obtained by instructed jurors. Only jurors who had heard the relevant instruction verbatim were included in the instructed group. Next, an analysis was undertaken that subdivided the uninstructed and instructed groups into into high and low education categories. Such an analysis might reveal cases in which jurors at one education level were more (or less) influenced by instructions than jurors at another education level.

A few comments about interpretation should precede the analysis of specific results. First, because the questions are in a true-false format, mere guessing should result in a 50% correct response rate. Scores significantly above that level reflect at least some understanding of the concept tested. Scores significantly below 50% are more difficult to interpret, but probably indicate the presence of misconceptions—in other words, a belief that the wrong answer is correct rather than an absence of knowledge—or a problem with the question itself. This is important because scores for some of the questions are considerably below 50%. Second, differences among groups that have large sample sizes are more reliable than differences among groups that have small sample sizes.

In the results reported below, questions are grouped and reported by topic area; for example, all questions concerned with reasonable doubt are listed under that heading. The Criminal

37. The groups were split at the median, resulting in one group with 12 or fewer years of education and one group with 13 or more years of education.
Jury Instruction number(s) that apply to the relevant topic are in parentheses accompanying each question. Each question is then listed. Underneath each question appear the results for that question. An analysis of variance was performed on each question. Results are presented in the following format:

**Topic Area (Criminal Jury Instruction number(s))**

Uninstructed = A%  Instructed = B%  Δ  p

(number uninstructed)  (number instructed)

A% = the percentage correct for uninstructed jurors, which serves as a baseline of comprehension for that particular question (group sample sizes appear below the percentage, in parentheses);

B% = the percentage correct for instructed jurors (group sample sizes appear below the percentage, in parentheses);

Δ = the change resulting from instruction; and

p = the statistical significance of the difference, which refers to the statistical probability that the observed difference would have occurred by chance alone. This statistic helps interpret the difference (Δ) between groups. As the p value increases, the likelihood that the observed difference has resulted from chance increases. For example, a p value of .3 indicates that there is a 3 in 10 chance that the observed difference is caused by chance or sampling error; a .5 indicates that there is a 5 in 10 chance that the observed difference is a result of chance. Conversely, smaller p values imply that the difference is unlikely to be caused by chance and more likely to be caused by the instructions. Social scientists have traditionally accepted a p value of less than .05 as statistically significant. Values greater than .05 are generally considered nonsignificant. Thus, in judging whether an observed difference may be produced by an instruction rather than by chance, one should consider not only the size of the observed difference, but also the significance level.

A brief discussion of the results follows each series of questions. It should be reiterated that the instructions tested by this procedure were those identified as most problematic. The study’s results, therefore, should not be used to draw overly generalized conclusions regarding juror comprehension of instructions.

---

39. Analysis of variance is a statistical test that allows one to infer whether an observed difference between groups is the result of random chance or the reliable result of some systematic difference in the way that the two groups were treated. In this case, the systematic difference is, of course, exposure versus no exposure to selected instructions.

A. Reasonable Doubt (03:1:04 or 03:1:05)

1. A REASONABLE DOUBT IS BASED ON YOUR COMMON SENSE. [TRUE]
   Uninstructed = 68.8% Instructed = 84.6% Δ = 15.8 p = .005
   (48) (526)

10. A REASONABLE DOUBT MUST BE BASED ONLY ON THE EVIDENCE THAT WAS PRESENTED IN THE COURTROOM, NOT ON ANY CONCLUSION THAT YOU DRAW FROM THE EVIDENCE. [FALSE]
    Uninstructed = 48.0% Instructed = 31.8% Δ = -16.2 p = .020
    (50) (510)

4. YOU HAVE A REASONABLE DOUBT IF YOU CAN SEE ANY POSSIBILITY, NO MATTER HOW SLIGHT, THAT THE DEFENDANT IS INNOCENT. IF SO, YOU SHOULD FIND THE DEFENDANT NOT GUILTY. [FALSE]
   Uninstructed = 24.0% Instructed = 25.2% Δ = 1.2 p = .853
   (48) (514)

22. TO FIND THE DEFENDANT GUILTY BEYOND A REASONABLE DOUBT, YOU MUST BE 100% CERTAIN OF THE DEFENDANT'S GUILT. [FALSE]
   Uninstructed = 26.5% Instructed = 30.9% Δ = 4.4 p = .528
   (49) (505)

   Approximately 69% of uninstructed jurors believe that they are allowed to use their common sense to judge whether a reasonable doubt exists. The reasonable doubt instruction in either of the two forms increases this to approximately 85%. The amount of increase as a result of exposure to the instruction is identical for both high and low education groups.

   Some confusion seems to exist, however, between reasonable doubt and any doubt. A majority of responses from uninstructed jurors revealed a belief that any doubt (or, alternately, anything less than 100% certainty) is equivalent to a reasonable doubt. Reasonable doubt instructions apparently did little to improve jurors' understanding that absolute certainty is not required.41

   Question 10 suggests misunderstanding about whether jurors are allowed to draw inferences or conclusions from the evidence in order to reach a verdict. The comprehension rate actually de-

---

41. The high p values for questions 4 and 22 on reasonable doubt indicate that the observed difference between uninstructed and instructed jurors is likely the result of chance factors, such as random sampling error, and that the comprehension of the two groups of jurors do not differ in reality. Interestingly, uninstructed jurors with high education showed the most misunderstanding of all, responding correctly at a rate of 17%. Receiving instructions raised the high education group's correct response rate to only 31%.
clined as a result of exposure to instruction, though it is difficult to know if this effect comes from exposure to the instruction or from other factors that might be associated with reasonable doubt, such as repetition of the standard or conflict with other instructions.

The responses to questions 4 and 22 reflect a poor understanding of the reasonable doubt standard. Perhaps the questions did not test the information given in the instructions. Despite having been approved both by legal experts and by social scientists, the questions may differ importantly from the corresponding instructions. Questions 4 and 22 attempt to ask jurors about reasonable doubt as a quantifiable concept: question 4 speaks of "any doubt, no matter how slight," and question 22 speaks of "100% certain[ty]." The reasonable doubt instructions, on the other hand, are not cast in quantitative language but instead ask the jurors to examine the evidence qualitatively. Jurors are instructed that a reasonable doubt is one that would cause them to hesitate in making an important decision in their own life, and that the decision is a moral one. An alternative instruction also states that the evidence must convince them to a moral certainty of the defendant's guilt before they vote to convict. Thus, the instructions given may have emphasized one concept while the questions presented may have asked something different—an apples and oranges situation. As stated by Professor Nesson:

As long as the concept [of reasonable doubt] is left ambiguous, members of the observing public may assume that they share with jury members common notions of the kinds and degree of doubt that are unacceptable. . . . The closer reasonable doubt comes to quantification, the more any notion of it being a shared concept will break down.

Indeed, some instructions that have attempted to define reasonable doubt in quantifiable terms have been declared reversible error. Whether these questions test the true meaning of rea-
sonable doubt is certainly debatable. Perhaps questions that addressed more qualitative distinctions would have been more informative, even though these questions are difficult to formulate because of the inherent ambiguity of reasonable doubt. Nevertheless, the minimal and even negative effects of these instructions should raise serious concern about whether instructed jurors understand a crucial element of their charge—the meaning of reasonable doubt.

B. Witness Credibility (3:1:11)

15. **In deciding whether to believe a witness, you may consider not only what the witness said, but how he or she said it.** [True]

Uninstructed = 91.7% Instructed = 84.1% Δ = -7.3 p = .121

18. **If a witness is honest and confident, you can be certain that he or she is accurate in their [sic] perception of events.** [False]

Uninstructed = 77.6% Instructed = 69.0% Δ = -8.6 p = .181

28. **It is the juror’s responsibility to decide if the witnesses are truthful, not the judge’s or the attorney’s.** [True]

Uninstructed = 85.0% Instructed = 92.0% Δ = 7 p = .220

Two things are apparent when one examines the responses to questions dealing with witness credibility. The first is that comprehension levels are generally high, and the second is that comprehension levels appear to decrease as a result of instruction. The latter result is puzzling.

74, 657 P.2d 1157, 1157 (1983) (reversing where the judge instructed “seven and a half, [on a scale of 0-10] if you had to put it on a scale”). But see Kagehiro & Stanton, Legal v. Quantified Definitions of Standards of Proof, 9 LAW & HUM. BEHAV. 159 (1985).

Some famous cases discussing reasonable doubt indicate that some of the best judges throughout history have experienced difficulty defining proof beyond a reasonable doubt. Compare Commonwealth v. Webster, 59 Mass. 295 (1850) (Shaw, C.J.) (providing a technical definition of proof beyond a reasonable doubt) with People v. Brigham, 25 Cal. 3d 283, 292, 599 P.2d 100, 106 (1979) (Mosk, J. concurring) (criticizing Chief Justice Shaw’s definition from *Webster* as confusing, and concluding that proof beyond a reasonable doubt is not susceptible to a technical definition).

47. Questions 28-31 were omitted from the questionnaire given to most jurors after a preliminary analysis because both uninstructed and instructed jurors’ understanding of these concepts was quite high.
One possible explanation is that the reduced comprehension for the instructed jurors is more apparent than real. Generally, social scientists regard significance levels above .05 as suspicious. Because values for both questions 15 and 18 are above that level, it might be that comprehension for the instructed and uninstructed groups are actually the same and appear to differ only because of random sampling variation. If there is an actual difference between the groups such that instructed jurors comprehend less, two explanations may be advanced.

First, judicial instructions regarding witness credibility may actually mislead jurors. For instance, the results of question 15 are consistent with the results of question 22, which concerned reasonable doubt, in that both show instructed jurors endorsing the idea that they are not allowed to draw inferences. It may be that the instructions themselves lead jurors to endorse this idea.

Second, it may be that the repetition of standards stressing the presumption of innocence and proper use of evidence leads jurors to conclude that the law requires them to "stick to the facts" and not draw reasonable inferences from the facts. Why comprehension might decrease for instructed jurors in question 18 is less clear, particularly because in this question jurors seem to believe that they are allowed to draw inferences about credibility from paralinguistic cues.

In any event, the possibility that jurors are misled by the instruction or that the sum total of instructions causes jurors to believe the court requires an overly restrictive standard for drawing inferences is speculative. It cannot be determined conclusively from the present results. The most parsimonious conclusion is that the observed differences resulted from chance rather than that they represent reliable indicia of lower comprehension by the instructed group. Particularly for question 18, for which the p value of nearly .2 is far above what is traditionally regarded as indicating a reliable difference, this conclusion seems warranted.

C. Mixed Direct and Circumstantial Evidence (4:2:02)

5. Facts can be proven through circumstantial evidence. [True]

Uninstructed = 53.9% Instructed = 64.8% Δ = 10.9 p = .008

(290) (219)

48. See supra note 39 and accompanying text.
9. Suppose you see that your neighbor's drapes are closed. When you look a minute later, you see the drapes open and the lights on. The conclusion that someone is at home would be based on circumstantial evidence. [True]

Uninstructed = 81.4%  Instructed = 86.7%  \( \Delta = 5.3 \)  \( p = .113 \)

16. A witness's statement that he or she saw something is an example of direct evidence. [True]

Uninstructed = 74.8%  Instructed = 75.4%  \( \Delta = .6 \)  \( p = .893 \)

23. Suppose that a witness saw the defendant's automobile parked near the scene of a crime. The conclusion that the defendant was in the area would be based upon direct evidence. [False]

Uninstructed = 59.0%  Instructed = 61.2%  \( \Delta = 2.2 \)  \( p = .644 \)

19. The prosecutor may not build a case on circumstantial evidence alone. [False]

Uninstructed = 32.1%  Instructed = 37.6%  \( \Delta = 5.2 \)  \( p = .211 \)

The results from questions 5 and 19 suggest that uninstructed jurors believe that circumstantial evidence may be worthless. The instruction effectively communicates that circumstantial evidence can have value.

Although of marginal statistical significance, responses to question 9 suggest that instruction probably helps jurors distinguish between direct and circumstantial evidence only minimally. The factor that primarily determines a juror's ability to make this distinction is the juror's education level. Averaging across all three questions, uninstructed jurors with less education were 67% correct. Uninstructed jurors with more education averaged 78% correct. The instructions improved the performance of both groups by only about 3%.

The results from question 19, which is less tied to the wording of 4:2:02, suggest a belief by jurors that prosecutors must have more than circumstantial evidence. This is probably consistent with jurors' general belief that evidence must be quite strong as reflected in their responses to questions concerning reasonable doubt. Even for the high-education instructed group, the 54% correct response rate could have resulted from mere guessing.
D. Impeachment of Defendant by Prior Conviction (3:1:08)

7. **YOU CAN NOT CONSIDER A DEFENDANT'S PRIOR CRIMINAL CONVICTIONS FOR ANY PURPOSE.** [FALSE]

Uninstructed = 28.2% Instructed = 56.3% $\Delta = 28.1$ $p = .015$ (539) (16)

25. **YOU CAN CONSIDER A DEFENDANT'S PRIOR CRIMINAL CONVICTIONS IN DECIDING WHETHER TO BELIEVE THE DEFENDANT'S TESTIMONY.** [TRUE]

Uninstructed = 26.2% Instructed = 50.0% $\Delta = 23.8$ $p = .035$ (511) (16)

12. **SUPPOSE YOU LEARN THAT A DEFENDANT HAS SEVERAL PRIOR CRIMINAL CONVICTIONS. BASED ON THIS FACT, YOU CAN ASSUME THAT THERE IS A GREATER CHANCE THAT THE DEFENDANT IS GUILTY OF THE CRIME WHICH HE IS NOW CHARGED WITH.** [FALSE]

Uninstructed = 79.6% Instructed = 75.0% $\Delta = -4.6$ $p = .645$ (539) (16)

Taken together, the results of these questions suggest that uninstructed jurors commonly believe that a defendant's prior record should not be used for any purpose. Roughly three out of four uninstructed jurors endorsed this view. The 23% and 28% increases in correct responses to questions 7 and 25, respectively, suggest that instruction may have reduced this misunderstanding. These increases, however, were only enough to raise correct responses to a level that could have resulted from mere guessing. The results for questions 7 and 25 were independent of educational level.

Question 12 indicates that most uninstructed jurors believe correctly that they should not use a defendant's prior convictions to imply an increased probability of guilt for the current offense. This conforms with the general "should not use" heuristic that most uninstructed jurors seem to employ when considering a defendant's prior record. When education level was accounted for, uninstructed persons with more education were correct more often and benefited more from instruction. Responses improved from 84% to 100% correct. Persons with less education did not benefit from instruction, and in fact showed a decrease that was marginally significant.
E. Homicide (16:2:01, 16:2:02, 16:3:01, 16:3:02, 16:4:01, 16:4:02, 16:4:03)

i. First degree premeditated murder

24. For first degree premeditated murder, it doesn’t matter whether the defendant intended to kill the victim. If the victim was killed by the defendant’s actions, that is enough. [False] (16:2:01)

Uninstructed = 63.6% Instructed = 65.8% \( \Delta = 2.2 \ p = .665 \)

(418) (117)

21. Suppose a defendant was strongly provoked by a victim, but killed the victim later, after he cooled down. The most serious offense you can convict the defendant on is manslaughter. [False] (16:2:01, 16:4:02)

Uninstructed = 75.6% Instructed = 80.3% \( \Delta = 4.7 \ p = .288 \)

(427) (117)

3. Actual intent to kill or create a very high risk of death is required for both first degree premeditated murder and voluntary manslaughter. [False] (16:2:01, 16:4:02)

Uninstructed = N/A Instructed = N/A \( \Delta = N/A \ p = N/A \)

(N/A) (N/A)

Roughly two out of three jurors answered correctly that intent is required for first degree murder. For more-educated jurors, the figure is closer to three out of four. This comprehension level is not impressive, considering that intent is a crucial aspect of first degree premeditated murder. The instructions did not improve substantially the comprehension for either low- or high-education jurors.

Question 21 indicates that three out of four uninstructed jurors answered correctly that if a defendant reflected on the killing, the defendant could be convicted of a crime more serious than manslaughter. Again, uninstructed jurors with more education grasp this concept quite well (85%). Instruction had a minimal impact on the small percentage of people who did not already understand the effect of reflection.

ii. First degree felony murder

6. Suppose that a defendant was robbing a gas station. He had not planned to hurt anyone, but he got scared because he thought that the attendant was reaching for a gun. He shot and killed the attendant. You can still

49. Jurors received two different versions of question 3. We thus cannot draw reliable conclusions from this question.
CONVICT THE DEFENDANT OF FIRST DEGREE FELONY MURDER. [True] (16:2:02)

Uninstructed = 63.4% Instructed = 74.0% \( \Delta = 10.6 \quad p = .037 \) (431) (89)

Question 6 attempted to test comprehension of first degree felony murder by asking jurors to identify an example. Again, roughly two out of three jurors succeeded. Surprisingly, this time the uninstructed jurors with less education were more often correct (73% versus 52%). For the less-educated group, the instruction had no effect; for the more-educated group, the instruction had a large effect (from 52% to 72%). Thus, this instruction raised the comprehension level of the more-educated group to roughly 75%, which corresponds with the comprehension level of the less-educated group.

iii. Second degree murder; second degree murder as a lesser included offense

2. To be convicted of second degree murder, the defendant must have premeditated a killing; that is, he or she had to plan it out beforehand. [False] (16:3:01, 16:3:02)

Uninstructed = 71.5% Instructed = 83.5% \( \Delta = 12 \quad p = .007 \) (425) (127)

Question 2 tests jurors' understanding of a critical difference between first and second degree murder: premeditation. Each instruction produced a substantial improvement in comprehension for both less-educated jurors, who rose to 76% after instruction, and more-educated jurors, who rose to 92% after instruction.

iv. Voluntary manslaughter; voluntary manslaughter as a lesser included offense

26. To be convicted of voluntary manslaughter, a defendant must have killed while influenced by circumstances that normally would produce strong feelings or passion. The killing must have occurred before a normal person would have had time to stop and think about what he was about to do. [True] (16:4:01, 16:4:02)

Uninstructed = 67.6% Instructed = 89.3% \( \Delta = 21.7 \quad p = .001 \) (414) (56)

27. Suppose that a defendant was grossly negligent and careless in his actions. His actions caused someone's death. Though the death was an accident and not at all
INTENDED, YOU CAN FIND THE DEFENDANT GUILTY OF INVOLUNTARY MANSLAUGHTER. [TRUE] (16:4:03)50

Uninstructed = 87.7% Instructed = 75.0% Δ = 12.7 p = N/A (515) (4)

Question 26 tests jurors’ recognition that voluntary manslaughter requires emotional arousal and an absence of premeditation. Though the majority of uninstructed jurors knew this, the instruction appeared highly effective in assisting those who did not. Again, the more-educated group scored higher than the less-educated group, but the instruction improved comprehension for both groups significantly.

F. Attempt as a Lesser Included Offense (9:1:02)

14. EVIDENCE THAT A DEFENDANT MADE DEFINITE PLANS TO COMMIT A CRIME IS ENOUGH TO PROVE THAT HE ATTEMPTED IT. [FALSE]

Uninstructed = 66.6% Instructed = 100% Δ = 33.4 p = N/A (554) (4)

17. SUPPOSE THAT A DEFENDANT HAD PLANNED A ROBBERY OR CRIMINAL SEXUAL CONDUCT (RAPE). HE THEN BEGAN TO COMMIT THE CRIME. HOWEVER, HE WAS INTERRUPTED AND PREVENTED FROM COMPLETING THE CRIME. YOU CAN STILL FIND HIM GUILTY OF ATTEMPTED ROBBERY OR ATTEMPTED CRIMINAL SEXUAL CONDUCT. [TRUE]

Uninstructed = 93.0% Instructed = 100% Δ = 7 p = N/A (558) (4)

Unfortunately, only four persons from one jury were exposed to this instruction and completed a questionnaire. Though the direction of the effect is encouraging and the comprehension of uninstructed jurors appears high, the sampling error from such a small group is likely to render a comparison of instructed versus uninstructed groups statistically insignificant.

G. Specific Intent (3:1:16)

8. SPECIFIC INTENT MEANS THAT A DEFENDANT HAD PLANNED TO DO AT LEAST SOMETHING ILLEGAL, REGARDLESS OF WHETHER HE HAD PLANNED TO COMMIT THE CRIME HE IS CHARGED WITH. [FALSE]

50. The sample size for the instructed group is too small to draw reliable conclusions from this question.
Uninstructed = 25.5% Instructed = 22.3% Δ = -3.2 p = .436
(192) (229)

This question was probably difficult for jurors because it is long and it addresses specific intent as an abstract definition, rather than as an element of a specific crime. More-educated jurors scored significantly higher than less-educated jurors, but even the instructed group with more education responded correctly only at a rate of about 36%. This can be interpreted as indicating that receiving the instruction did not significantly improve comprehension for the abstract meaning of specific intent. The poor rate of correct response merits attention.

H. Assault (17:1:02, 17:6:01, 17:6:02)

11. An assault must include actual physical injury to the victim. [FALSE]
Uninstructed = 49.6% Instructed = 32.3% Δ = -17.3 p = .060
(538) (31)

30. A defendant purposely and violently struck a victim, causing physical injury. This forceful touching is called a battery. [TRUE]
Uninstructed = 82.5% Instructed = 72.2% Δ = -10.3 p = .314
(97) (18)

The results for the two questions on assault appear rather counterintuitive and striking because the instructed group scored lower than the uninstructed group. The significance level for question 11 indicates that the decrease in juror comprehension after instruction on assault is a reliable measurement, by social science standards. For question 30, however, the significance level is well beyond that which indicates a reliable difference. The best conclusion for results from question 30 is that the two groups do not actually differ.

It is clear that the instructions for assault did not improve comprehension for the issues covered in either of these questions, and these instructions very likely worsened comprehension for the issue covered in question 11. One possibility is that the instructions are confusing to jurors. Postquestionnaire interviews with jurors would have helped determine the exact nature of juror misunderstanding; however, we were not able to do this.

An alternative possibility is that the instructions were not poor or misleading, but that unusual factors were present in one or two trials in which jurors were exposed to these assault instructions. Even if only one jury were confused or somehow mis-
led, its responses might be enough to lower significantly the score of the entire instructed group, which had a sample size of only thirty-one.

We were not able to distinguish among these possibilities using our methodology. Nevertheless, it is clear that instructions do little to help jurors distinguish successfully assault from battery.

I. Criminal Sexual Conduct (20:3:01; 20:3:02; 20:5:01; 20:5:02; 20:2:11; 20:2:12)

13. Second degree criminal sexual conduct always involves personal injury and penetration. [False] (20:3:01; 20:3:02)

Uninstructed = 64.4% Instructed = 100% $\Delta = 35.6\%$ $p = .026$

(516) (9)

20. A jury believes that a defendant coerced a victim into touching his groin to get sexual gratification. There was no physical injury to the victim and the touching was only through the clothes. The jury must find the defendant not guilty of any criminal sexual conduct. [False] (20:5:01; 20:5:02)

Uninstructed $= 0.0\%$ No instructed jurors $\Delta = N/A$ $p = N/A$

(413) (0)

31. In criminal sexual conduct cases, you can consider mental anguish as personal injury. [True] (20:2:11; 20:2:12)

Uninstructed $= 77.4\%$ No instructed jurors $\Delta = N/A$ $p = N/A$

(124) (0)

Unfortunately, the questions that were asked about criminal sexual conduct did not correspond to the instructions that were most frequently given in such cases. Question 13 had one jury that received instructions 20:2:01 or 20:2:02, resulting in a very small sample size for the instructed group. Though the results for question 13 are encouraging and statistically significant, a reliable analysis of relevant characteristics among jurors is hindered by the small sample. For the other questions, comparisons between instructed and uninstructed groups can not be made.

J. Presumption of Innocence

29. In a criminal trial, a defendant is considered innocent until proven guilty. [True]
III. General Factors in Juror Comprehension

The results from individual items seem discouraging. Comprehension levels are sometimes quite low for both uninstructed and instructed jurors. Instructions often had little impact on comprehension; sometimes, instruction actually decreased juror comprehension. Yet if one looks at the effects of the instructions across all items, the results are not quite as discouraging as they first appear.

If we examine only those instructions that demonstrate a statistically significant effect (those with a p value of less than 0.05), we note that eight questions showed a positive effect of instructions (questions 1, 2, 5, 6, 7, 13, 25, and 26) and only one showed a negative effect (question 10). Thus, it seems that when instructions do produce an effect they produce a positive effect on jurors' comprehension more often than they produce a negative effect. On the other hand, thirteen questions indicate that instruction had no statistically significant effect on juror comprehension (questions 4, 8, 9, 11, 12, 15, 16, 18, 19, 21, 22, 23, and 24).

A variety of other factors besides exposure to instructions may have affected the number of correct responses to this questionnaire. For instance, more-educated jurors might be expected to score higher than less-educated jurors. This was a difficult issue in constructing the questionnaire because the questions necessarily included some legal language or unfamiliar concepts.

In order to test which factors were associated with correct responses on the questionnaire, a score reflecting the total number of correct responses for each juror was calculated. Then, the total score was used as a dependent measure against which to test the influence of several factors. These included the total number of instructions heard by the juror, the juror's level of education, age, and gender, and whether written or audiotaped instructions
were provided to the juror. Analysis of variance was used to test whether any of these factors, alone or in combination, affected juror comprehension. 51

Not surprisingly, jurors who were exposed to more instructions generally answered more items correctly than jurors who were exposed to fewer instructions. This confirms that juror comprehension increased as a function of exposure to instructions. Though the difference between the means was highly significant, the magnitude of the increase as a result of exposure to more instructions was quite modest. This makes sense because

51. See supra note 39 for a brief description of analysis of variance. The number of instructions was divided at the median, resulting in one group that received 12 or fewer instructions and one group that received 13 or more of the instructions targeted by this study. Education and age were similarly dichotomized around their respective medians (13 years for education; 39 years for age). Gender and distribution of written copies are naturally dichotomized. Analysis of variance was performed on all the factors. An F statistic, which is "the ratio of the variance between categories to the variance within categories," J. Healey, supra note 38, at 163, was also calculated for each factor. The greater the F statistic, the less likely it is that an observation is due to random chance. Id. at 163-64. The following are the means (reported in percentage correct on the questionnaire), F statistics, and significance levels for each factor:

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>Means</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of instructions</td>
<td>&lt;13=47.1; 13 or more=51.7</td>
<td>4.63</td>
<td>.032</td>
</tr>
<tr>
<td>educational level</td>
<td>&lt;13=46.8; 13 or more=53.9</td>
<td>18.18</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>age</td>
<td>&lt;39=50.9; 39 or more=48.3</td>
<td>1.55</td>
<td>.213</td>
</tr>
<tr>
<td>gender</td>
<td>male=52.9; female=46.7</td>
<td>19.77</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>copies provided</td>
<td>no=49.3; yes=57.8</td>
<td>3.90</td>
<td>.049</td>
</tr>
</tbody>
</table>

These means cluster near 50%, which would suggest random guessing. However, nearly all of the means for individual items differ significantly from chance, which contradicts the notion that subjects guessed randomly on the questionnaire. Rather, the total score's closeness to 50% is an artifact of combining individual items that show high (e.g., witness credibility) and low (e.g., reasonable doubt) levels of comprehension.

There were no two- or three-way interactions except between age and educational level. Except for the effect of age, these are all significant main effects. This implies that each of these factors contributes independently to comprehension and no factor depends on levels of the other factors to have an influence.

Four factors affected overall comprehension: number of instructions heard, education, gender, and provision of copies. These factors did not interact with each other. We chose to discuss the effect of education level on the responses to individual questions because the influence of education on comprehension was, for the most part, not difficult to interpret. We did not discuss the effect of gender on responses to the individual questions because we had no sound theoretical reason for this effect. We did not discuss the effect of the provision of written copies on responses to individual questions because too few juries received written copies.
The following graph of the means from Table I is a useful illustration:
jurors who were exposed to many instructions would face more of a memory burden, particularly if the instructions were part of a complex trial, and may show a slower increase in comprehension, or even some decrease, as a result. Additionally, the number of instructions affected the total number of correct responses independently of each of the other factors. This suggests that the benefit of exposure to more instructions was about equal for all groups (e.g., more-educated and less-educated jurors, older and younger jurors, etc.).

Other factors influenced comprehension as well, including the jurors’ educational level and the use of written instructions. Jurors with higher educational levels scored significantly higher than jurors with lower educational levels. Giving more instructions simply raised the scores of both groups equally. The provision of written instructions also affected comprehension levels significantly, with those who received instructions scoring higher than those who did not. The effect of written instructions was statistically significant despite relatively small sample sizes—only nine percent of the instructed jurors received written instructions.

The positive effect of written instructions conflicts with a recent study by Heuer and Penrod. In a field study of Wisconsin jurors, they found that the provision of written instructions did not positively affect jurors’ comprehension in criminal cases, as assessed on a nine-item multiple-choice questionnaire. It is not immediately clear why their results conflict with the current results. In general, tests with more items are more discriminating, provided that the items in the test have some relationship to the dependent criterion (here, comprehension of particular instructions). It is possible that our twenty-six-item measure was more sensitive to changes in comprehension as a result of written instructions than their nine-item measure. In any event, Heuer and Penrod do report that the provision of written instructions helped to resolve disputes among jurors over the meaning of the law, and therefore had some positive effects.

The results of the questionnaire depended on gender. Men scored higher than women. This effect was not a result of differences in educational levels between men and women, so it remains unclear why men scored higher generally on this question-

53. Id. at 420.
54. J. NUNNALLY, supra note 26, at 78.
55. Heuer & Penrod, supra note 52, at 421.
naire. Finally, age was not independently significant in the analysis of variance, although the number of correct answers increased as juror age decreased. Subsequent analysis revealed that younger persons in this study tended to have more education, and it may be that the difference in education rather than age produced this marginal effect.

IV. SUMMARY AND CONCLUSION

This research supports a growing body of literature suggesting that jury instructions are often lost on jurors, and can sometimes even backfire.56 The relatively low rate of comprehension for some concepts, both among more- and less-educated jurors, the apparent ineffectiveness of instructions to improve comprehension, and the negative effect of certain instructions, constitute the most striking findings in the present study. Particularly startling are the results of instructions concerning reasonable doubt, defendant impeachment by prior conviction, and some aspects of mixed direct and circumstantial evidence.

Perhaps the study's results are not as surprising as they seem, given the complexity of the cognitive task confronting jurors, and the problematic nature of these particular instructions. One possible explanation for the ineffectiveness of instruction in these particular areas is that jurors come to court with preexisting beliefs about these issues that are resistant to change. For instance, jurors' beliefs about what the law requires may be derived from exposure to the media or other sources in the popular culture.57 Ideas learned over a lifetime of exposure to these sources may be firmly entrenched and act as rules of thumb in directing jurors' decisions.58 A few such "rules" may have driven how jurors responded to several questions on this questionnaire. For example:

---


57. Hastie, Penrod, and Pennington reported that videotaped mock jurors made inappropriate generalizations from television and film trials, and that "[n]ot even reinstruction on the law from the judge was sufficient to eliminate all of these errors." R. HASTIE, S. PENROD & N. PENNINGTON, INSIDE THE JURY 170 (1983).

1. Jurors decide whom to believe and whom to disbelieve (witness credibility).
2. Jurors should not use circumstantial evidence (mixed direct and circumstantial evidence).
3. Jurors should not use a defendant’s past record against him (defendant impeachment by prior conviction).

When such preconceived notions are congruent with instructions, both uninstructed and instruction-benefitted comprehension levels appear quite high. But when preconceived notions are incongruent with instructions, both uninstructed and instruction-benefitted comprehension levels remain low, well below chance levels for some issues.

Some support for this notion can be found in a recent study by Professor Ellsworth.\(^9\) She videotaped deliberations of mock juries in order to investigate jurors’ understanding and use of facts and instructions. She concluded that:

Much of jurors’ discussion of the law revolved around phrases they were likely to have known before they heard the judge’s instructions. The instructions may have been effective in reminding the jurors of terms they had heard before, but the instructions were not very effective in educating them in new areas, or even in focusing their attention on the meaning of the familiar terms.\(^60\)

Some jury instructions did raise comprehension. When jurors lack preconceptions or rules of thumb, they are perhaps more influenced by instructions, particularly if the instructions are central to the case and find their way into discussion during deliberation. For instance, distinctions between first and second degree murder are probably more obscure and therefore less entrenched and certain in jurors’ minds. The jurors’ open-ended comments support this interpretation: instructions involving homicide were often mentioned as being the most difficult to understand. Even though uninstructed comprehension may have been above chance, jurors’ responses showed that they were influenced positively by these instructions.


\(^{60}\) Id. at 219. Ellsworth also concluded that the jury benefitted little from the ability of its most able member to understand the meaning of instructions, *id.*, a finding consistent with our conclusion that it takes more than one person on a jury to understand an instruction for the instruction to have its intended effects.
A number of factors external to the instructions themselves can influence comprehension. Some factors that influence juror comprehension, such as jurors' educational level, are outside the control of courts. Other factors, such as whether written instructions are provided, are potentially within the court's control. In this sample, only nine percent of instructed juries received written instructions. Several judges commented that they would like to provide jurors with written instructions, but that the court lacked time and resources to do so. This is unfortunate because both the present results and the study by Heuer and Penrod suggest that written copies of judicial instructions can help juries to use the instructions.61 In discussion of this issue within the CJI Committee, it was noted that the increased availability of computers might lead to greater use of written instructions. In particular, if criminal jury instructions were on file, one would need only to retrieve and print the instructions to be used in a given trial. Any modifications by the judge could be incorporated easily.

Clearly the more-educated jurors generally were likelier to grasp the concepts conveyed by an instruction. They generally had better uninstructed comprehension, and probably had better retention of spoken distinctions or concepts. Given the technical/legal nature of the jury instructions, this advantage for more-educated jurors will probably always be a factor. As one attempts to make instructions comprehensible to all jurors, including less-educated jurors, there is likely to be a trade-off between comprehensibility and technical accuracy. Providing written copies may narrow the comprehension gap between more- and less-educated jurors, but the current data suggest that more-educated jurors are more likely to comprehend legal concepts regardless of whether written instructions are provided.

Jury do not individually attempt to apply instructions to a case but rather do so as a group. Therefore, the effects of jury deliberation on comprehension are important and relevant here. In the current study, the comprehension of instructed jurors was tested after jury deliberation. Thus, our measure could not be considered a pure measure of individual comprehension, because postdeliberation responses may also reflect the influence of group discussion.

The most optimistic theory about the influence of group deliberation on comprehension suggests that it takes only one person to understand an instruction at the time it is given. If jury in-

61. Heuer & Penrod, supra note 52, at 429.
structions are discussed later in deliberation, that person can then explain the instruction to the remaining jurors. An alternate theory is that a majority of jurors must understand an instruction at the time it is given. If the instruction is discussed, then the majority viewpoint will prevail in applying the instruction. This research does not directly address how deliberation influenced comprehension, but the current results are suggestive. Because data were collected after deliberation, and because open-ended questions revealed that most juries did discuss some instructions, it appears that the more optimistic theory may be overly optimistic.

This research could not address certain issues. First, we could not ascertain how variations in courtroom atmosphere or judges' behavior and presentation style influence jurors' comprehension of instructions. Presentation factors may be worth investigating in further research on comprehension of instructions. Second, the research could not confront the distinction between what jurors believe the law requires and what standards jurors actually use in making their decisions. We were careful to instruct jurors to respond to the questions based on what the law required, not on what they thought the law should be. Certainly most jurors are inclined to follow the law, if they understand it. But jurors sometimes do not conform their judgments to what the law requires, either because they lack the cognitive control to ignore certain variables, or because they believe that the law conflicts with some fundamental notions of fairness or rationality.

Finally, our procedure was designed to investigate the effectiveness of selected Michigan Criminal Jury Instructions currently in use. The Michigan Criminal Jury Instructions Committee constantly revises instructions, and it is uncertain whether differently worded instructions would result in similar rates of comprehension among actual jurors.

At least two goals—legal accuracy and juror comprehension—underlie the choice of wording in judicial instructions. There

62. Reasonable doubt was mentioned most frequently.

63. See Ellsworth, supra note 59, at 219.

64. See Kramer, Kerr & Carroll, supra note 32.

65. The jury has the power to nullify a judge's instructions. See Y. Kamisar, W. LaFave & J. Israel, supra note 5, at 1295-97. As Justice Holmes stated: "[T]he jury has the power to bring in a verdict in the teeth of both law and facts." Horning v. District of Columbia, 254 U.S. 135, 138 (1920). For a case presenting both sides of this issue, see United States v. Dougherty, 473 F.2d 1113, 1137 (D.C. Cir. 1972) (upholding refusal to instruct jury of its right to acquit defendants without regard to law and evidence) and see id. at 1139 (Bazelon, C.J., dissenting) (asserting that the jury should be told of its nullification power).
seems to exist an inherent tension, however, between communicating to jurors the requirements of the law, and insuring that instructions are technically complete and not reversible on appellate review. This tension has become more strained, for as law becomes more developed and differentiated, it also becomes more difficult for the layperson to understand. Designing instructions to improve comprehension may increase the chance of appeal, but designing technical instructions to guard against appeals may reduce juror comprehension, as indicated in the present study. A misunderstanding of the law precludes jurors from fulfilling their duties to the best of their abilities, as they sit in judgment of their peers.

APPENDIX I

Jury Instructions Study Questionnaire

Today's Date Male or Female
Age Education (highest grade completed)
Did you serve on a jury today? Yes No
Have you ever served on a jury before today? Yes No
If you have served on a jury, list the dates you served (as best you can recall). Also list the crime(s) the defendant was charged with.

Date Charge
Date Charge
Date Charge

Below are some True-False questions based on criminal jury instructions. Please answer each question by circling either T (for True) or F (for False).

EXAMPLE:

T F A defendant has a right to remain silent.

You are not being tested. Rather, we are testing how clear the criminal jury instructions are. When answering these questions, assume that you are a juror. If you have heard jury instructions on the questions below, try to recall what you heard. If you have not heard instructions or are not sure, pick the answer that you think is the correct one.

Please answer every question.

T F 1. A reasonable doubt is based on your common sense.

T F 2. To be convicted of second degree murder, the defendant must have premeditated a killing; that is, he or she had to plan it out beforehand.
T F 3. Actual intent to kill or create a very high risk of death is required for both first degree premeditated murder and voluntary manslaughter.

T F 4. You have a reasonable doubt if you can see any possibility, no matter how slight, that the defendant is innocent. If so, you should find the defendant not guilty.

T F 5. Facts can be proven through circumstantial evidence.

T F 6. Suppose that a defendant was robbing a gas station. He had not planned to hurt anyone, but he got scared because he thought that the attendant was reaching for a gun. He shot and killed the attendant. You can still convict the defendant of first degree felony murder.

T F 7. You can not consider a defendant's prior criminal convictions for any purpose.

T F 8. Specific intent means that a defendant had planned to do at least something illegal, regardless of whether he had planned to commit the crime he is charged with.

T F 9. Suppose you see that your neighbor’s drapes are closed. When you look a minute later you see the drapes open and the lights on. The conclusion that someone is at home would be based on circumstantial evidence.

T F 10. A reasonable doubt must be based only on the evidence that was presented in the courtroom, not on any conclusion that you draw from the evidence.

T F 11. An assault must include actual physical injury to the victim.

T F 12. Suppose you learn that a defendant has several prior criminal convictions. Based on this fact, you can assume that there is a greater chance that the defendant is guilty of the crime which he is now charged with.

T F 13. Second degree criminal sexual conduct always involves personal injury and penetration.

T F 14. Evidence that a defendant made plans to commit a crime is enough to prove that he attempted it.

T F 15. In deciding whether to believe a witness, you may consider not only what the witness said, but how he or she said it.

T F 16. A witness’s statement that he or she saw something is an example of direct evidence.

T F 17. Suppose that a defendant had planned a robbery or criminal sexual conduct (rape). He then began to commit the crime. However, he was interrupted and prevented from completing the crime. You can still find him guilty of attempted robbery or attempted criminal sexual conduct.

T F 18. If a witness is honest and confident, you can be certain that he or she is accurate in their perception of events.

T F 19. The prosecutor may not build a case on circumstantial evidence alone.

T F 20. A jury believes that a defendant coerced a victim into touching his groin to get sexual gratification. There was no physical injury to the
victim and the touching was only through the clothes. The jury must find the defendant not guilty of any criminal sexual conduct.

T F 21. Suppose a defendant was strongly provoked by a victim, but killed the victim later, after he cooled down. The most serious offense you can convict the defendant on is manslaughter.

T F 22. To find the defendant guilty beyond a reasonable doubt, you must be 100% certain of the defendant’s guilt.

T F 23. Suppose that a witness saw the defendant’s automobile parked near the scene of a crime. The conclusion that the defendant was in the area would be based upon direct evidence.

T F 24. For first degree premeditated murder, it doesn’t matter whether the defendant intended to kill the victim. If the victim was killed by the defendant’s actions, that is enough.

T F 25. You can consider a defendant’s prior criminal convictions in deciding whether to believe the defendant’s testimony.

T F 26. To be convicted of voluntary manslaughter, a defendant must have killed while influenced by circumstances that normally would produce strong feelings or passion. The killing must have occurred before a normal person would have had time to stop and think about that he was about to do.

T F 27. Suppose that a defendant was grossly negligent and careless in his actions. His actions caused someone’s death. Though the death was an accident and not at all intended, you can find the defendant guilty of involuntary manslaughter.

Please answer the following questions. Feel free to write on the back if you would like more room. Your thoughts and opinions are important.

Were there any instructions which you or any of the other jurors found confusing? If so, please list what these instructions were about. Also, for each instruction, please tell us what made the instruction confusing.

1. Instruction: ____________________________________________

   1a. What made the instructions confusing or hard to understand?______

2. Instruction: ____________________________________________

   2a. What made the instructions confusing or hard to understand?______

3. Instruction: ____________________________________________

   3a. What made the instructions confusing or hard to understand?______

4. Instruction: ____________________________________________

   4a. What made the instructions confusing or hard to understand?______

If you sat on a jury today, did your jury discuss the meaning of the jury instructions?_____ Yes _____ No If yes, which instructions were discussed?

A. ____________________________________________
We would welcome any comments about the jury instructions or your experience with them.

Comments:

APPENDIX II: COMPOSITE INSTRUCTIONS

CJI 3:1:04

Definition of Reasonable Doubt

A reasonable doubt is a fair, honest doubt growing out of the evidence or lack of evidence in this case or growing out of any reasonable or legitimate inferences drawn from the evidence or the lack of evidence. It is not merely an imaginary doubt or a flimsy, fanciful doubt or a doubt based upon the mere possibility of the innocence of the defendant or a doubt based upon sympathy, but rather it is a fair, honest doubt based upon reason and common sense. It is a state of mind which would cause you to hesitate in making an important decision in your own personal life. By stating that the prosecution must prove guilt beyond a reasonable doubt, I mean there must be such evidence that causes you to have a firm conviction to a moral certainty of the truth of the charge here made against the defendant.

USE NOTE: The jury must be instructed on the concepts of the presumption of innocence, burden of proof and reasonable doubt.

CJI 3:1:05

Definition of Reasonable Doubt (Alternate)

A reasonable doubt is a fair doubt growing out of the testimony, the lack of testimony or the unsatisfactory nature of the
testimony in the case. It is not a mere imaginary or possible doubt, but a fair doubt based on reason and common sense. It is such a doubt as to leave your minds, after a careful examination of all the evidence in the case, in the condition that you cannot say you have an abiding conviction to a moral certainty of the truth of the charge here made against the defendant.