

EXECUTIVE SUMMARY

National Institute of Justice Final Report:

REDUCING BIASES IN JOINED CRIMINAL OFFENSES

Steven Penrod                      and                      Sarah Tanford  
University of Wisconsin                      Purdue University

This research was supported by National Institute  
of Justice Grant #80-IJ-CX-0048.

U.S. Department of Justice  
National Institute of Justice

This document has been reproduced exactly as received from the  
person or organization originating it. Points of view or opinions stated  
in this document are those of the authors and do not necessarily  
represent the official position or policies of the National Institute of  
Justice.

Permission to reproduce this copyrighted material has been  
granted by

JOINDER EXECUTIVE SUMMARY

to the National Criminal Justice Reference Service (NCJRS).

Further reproduction outside of the NCJRS system requires permis-  
sion of the copyright owner.

TABLE OF CONTENTS

INTRODUCTION	1
Legal Background	1
Prior Research	4
Limitations of previous research	5
The Present Research	5
METHOD	6
Dependent Measures	8
RESULTS	8
Predeliberation Results	8
Deliberation Results	9
Post-Deliberation Memory Results	10
Relationships Among Variables	11
DISCUSSION	12
Theoretical Implications	12
Confusion	13
Accumulation	14
Criminal inference	14
Instructions	15
Legal Implications	15
Impact of instructions	15
Guidelines for joinder	16
References	19

NCJRS

DEC 8 1984

ACQUISITIONS

95985

## EXECUTIVE SUMMARY

## INTRODUCTION

## LEGAL BACKGROUND

Federal Rules of Criminal Procedure, Rule 8(a) Joinder of Offenses. Two or more offenses may be charged in the same indictment or information in a separate count for each offense if the offenses charged, whether felonies or misdemeanors or both, are of the same or similar character or are based on the same act or transaction or on two or more acts or transactions connected together or constituting part of a common scheme or plan.

Federal Rules of Criminal Procedure, Rule 14. Relief from Prejudicial Joinder. If it appears that a defendant or the government is prejudiced by joinder of offenses . . . in an indictment or information or by such joinder for trial together, the court may order an election or separate trial of counts, grant a severance . . . or provide whatever other relief justice requires.

In layman's terms, Rule 8(a) of Federal Rules of Criminal Procedure states that a single defendant may be tried for more than one related offense in a single trial, even if the offenses occur at different times and places and with different victims. The legal term for trying multiple charges together is "joinder of offenses." Rule 14 states that the courts must at the same time protect the defendant from "prejudice" that may result if multiple charges are joined together, but the rule provides little guidance as to how one goes about (1) determining whether prejudice exists and (2) providing "relief" from prejudice. The legal rules do not provide a clear definition of prejudice, although the nature of prejudice has been addressed by legal commentators. For example, Lempert and Saltzburg (1983) describe prejudice as "harm which results when evidence is inappropriately influential because it appeals to the biases or emotions of the fact finder" (p. 156), and McCormick notes that the problem of prejudice arises from "the danger that the facts offered may unduly arouse the jury's emotions of prejudice, hostility or sympathy" (Cleary, 1972, p. 439).

In practice, the legal solution to prejudicial joinder is in the form of a "severance" of offenses. Prior to the trial of a defendant charged with multiple crimes, the prosecutor may seek a joinder of the offenses, and the defense in turn may file a motion of severance, requesting that the charges be tried separately. If this motion is denied by the trial judge and the defendant is convicted of one or more of the offenses, the convictions may be appealed on the grounds of prejudice resulting from the joinder. In both the initial motion for severance and the appeal following conviction, the courts must decide whether joinder was sufficiently prejudicial to warrant separate trials. Because Rule 14 has never been interpreted by the Supreme Court, there are few authoritative guidelines available to judges who must make such decisions. However, convictions resulting from joined trials are often subject to appeal, and there are a large number of published appellate court opinions available at both Federal and state levels. In examining the reasoning used by judges in these decisions, one gets a flavor of the "intuitive psychology" of the legal profession.

To illustrate the court's psychological reasoning we will examine the case of the United States v. Foutz (1976) which is the leading Federal case on joinder. In this case, the defendant was convicted of two robberies which occurred several months apart, and successfully appealed the convictions on the grounds of prejudice resulting from joinder. In granting the appeal, the judge recognized three possible sources of prejudice that are possible in a joined trial: (1) jurors may confuse the evidence presented in proof of different charges—we will refer to this as the confusion hypothesis, (2) jurors may accumulate or combine evidence across different charges such that identical evidence is greater weight in joined trials than in severed trials—the accumulation hypothesis and (3) jurors may infer that the defendant has a "criminal disposition" based on the fact that he is charged with multiple crimes—the criminal inference hypothesis.

In applying these three theories of prejudice to the Foutz case, the court noted that the evidence for the second crime was strong while the evidence from the first was weak, so that a jury judging the first offense alone might well have acquitted the defendant. The court thought the jury had probably found the defendant guilty of the second robbery and then concluded that if he had robbed the bank once, there was a good chance he had robbed it

before; in other words, they attributed the robbery to the defendant's criminal nature (a criminal inference). In addition, there may have been a "spillover effect of evidence of one crime implicating guilt in other" (accumulation).

At the end of a joined trial, the jury typically receives a special instruction from the judge in addition to the standard jury instructions, the purpose of which is to alleviate potential prejudice resulting from joinder. The exact form of these instructions varies from state to state, but most instructions address at least a portion of the three legal theories of prejudice. The standard Federal joinder instruction reads as follows:

A separate crime or offense is charged in each count of the indictment. Each charge and the evidence pertaining to it should be considered separately. The fact that you may find the accused guilty or not guilty as to one of the offenses should not control your verdict as to any other offense charged (Devitt & Blackmar, Federal Jury Instructions and Practice, 1977, p. 296).

The Federal instruction essentially instructs jurors not to become confused or to accumulate verdicts across charges, but does not instruct jurors to avoid making inferences about the defendant's disposition. The law presumes that instructions will effectively alleviate prejudice, but appellate judges, acting as "intuitive psychologists" do not always agree. In the case of United States v. Foutz (1976) the court did not think the instruction was sufficient:

[W]e cannot presume that the jury adhered to limiting instructions and properly "segregated the evidence into separate intellectual boxes."

In other words, the instruction did not eliminate the possibility of confusion of evidence between charges. As a result of the judge's determination that joinder had been prejudicial in the Foutz case, the convictions were reversed and two new, separate trials for each count of robbery were ordered.

Given that the law recognizes that joinder can be prejudicial, and often is the subject of appeal, the question arises as to the utility of joining charges at all. The main rationale is that it is expedient and saves time and money (Drew v. United States, 1964). However, as the court in Foutz argued, the savings is actually minimal if the evidence for each offense is entirely separate, so that the only real savings is that of choosing only one jury as opposed to more than one.

Thus, the only real convenience served by permitting joint trial of unrelated offenses against the wishes of the defendant may be the convenience of the prosecution in securing a conviction (U.S. v. Foutz, 1976, p. 738).

Another legal precedent for joining charges is the "simple and distinct" test, which holds that joinder will not be prejudicial if the evidence is simple enough that jurors will not become confused (Drew v. United States, 1964). However, even if the assumption that evidential simplicity does reduce confusion is valid, the test does not protect against the other two types of prejudice, accumulation and criminal inference.

Although there is a reasonably large body of case law on joinder, there is little consensus on the criteria that judges ought to apply in their decisions. The issue has not been carefully researched by legal scholars, and there are only a few published legal articles which address the joinder issue. A brief review of the arguments provided in these articles underscores the somewhat conflicting viewpoints among legal scholars concerning the joinder issue.

Remington and Joseph (1961) described some of the conditions under which joinder is generally regarded as appropriate. If several offenses are committed at the same time and place and either damage several victims or do multiple damage to a single victim, it is appropriate to try them together. Joinder is also called for when several offenses occur at different times but are all part of the same scheme or plan. The difficult issue is several unconnected offenses occurring at different times or places with different victims. Remington and Joseph argued that although joinder may be harmful in some circumstances, it may actually be beneficial in others. It may be harassing to the defendant to defend himself in a number of separate trials, and this could outweigh any disadvantage resulting from joining charges. For this reason and for the reason of expedience, they suggest that a single proceeding may be to the advantage of both sides.

Others have (1) emphasized the potential prejudice to the defendant and (2) argued for clear rules governing joinder. For example, Holderman (1977) discussed the effects of joinder under Nebraska law, which allows similar offenses to be joined even if they are not part of a common plan. The decision to grant severance is left to the judge, and the defendant must be able to demonstrate actual prejudice in order to overturn a conviction. Rather than place this burden on the defendant, Holderman recommended a more stringent test of prejudice, such as that used in Foutz (1976), in which the test was whether the evidence purported to be prejudicial would have been admissible if the cases were tried separately.

Like Holderman (1977), Baron (1977) advocated the development of clearer rules to govern joinder decisions. In Tennessee law, the decision to join charges is left to the discretion of the judge, and the test of whether the judge has abused this discretion rests on the element of prejudice. However, as in the Federal rule, prejudice is not defined. It has been characterized in a number of different ways by legal commentators, although none of these definitions is very precise (Cleary, 1972). Baron recommended that multiple charges be joined only if they arise out of a single "criminal episode," and that offenses committed on separate occasions not be joined at all. Baron's recommendation is an even clearer guideline than Holderman's "admissibility test," and it would eliminate the joinder situation that legal scholars consider most problematic (Remington & Joseph, 1961), and the one that is most often subject to appeal.

A similar proposal at the Federal level was offered in the *Yale Law Journal* (Note, 1964-65). This article listed several traditional tests for assessing prejudice arising from joinder, and pointed out the inadequacies of each. The article essentially challenged the intuitive psychological reasoning used by the courts when they conduct a search for absence of prejudice. First, the article questioned whether it is realistic to expect jurors to heed judges' instructions to the jury to confine their decisions to each offense separately. A second common test is "cure by verdict," which assumes that if the defendant is acquitted on any count, the jury must have kept the charges separate, since it was selective in its verdicts. A related test is "cure by concurrent sentencing," which discounts prejudice if the defendant receives one sentence covering multiple counts. The article noted that what both of these "cures" fail to consider is the possibility that the defendant may have been acquitted on all counts if the offenses were tried separately. The final traditional device is that of "overwhelming evidence of guilt" in the record. In other words, if the jury could have reached the same decision on each of the charges tried by itself, then prejudice is not a problem. Using this test, the appellate court in effect becomes the jury, since justices are making judgments about what the jury would have done in a hypothetical situation. The article concluded by stating that the traditional tests of prejudice are simply not adequate, and that the best solution may be to abolish joinder of charges.

Legal scholars may not agree on the solution, but they all agree that joinder is a problem, and that a clear standard is needed to govern joinder decisions. Confusion about the issue among the legal profession no doubt stems from the fact that the conclusions reached by each legal researcher are based on his or her own subjective interpretation of a diverse collection of case law, which is itself a collection of the intuitions of individual judges. From a scientific viewpoint, such an analysis is clearly not an adequate basis for policy formation, and the issue can best be addressed empirically. Of course, non-empirical theorizing dominates the law, and most legal decisions are made in the absence of scientific evidence. Analogous reasoning can be found in the legal responses to problems related to joinder. Therefore, the issues addressed with respect to joinder have additional significance insofar as they suggest other aspects of the trial that could be empirically investigated.

The type of joinder we are concerned with here is joinder of distinct offenses occurring at different times and places. As noted earlier, a defendant may also be charged with multiple offenses arising out of a single act. In addition to joinder of offenses, the law allows for joinder of defendants, i.e., trying more than one defendant in a single trial. A somewhat related situation occurs when a defendant is charged with a single serious offense and the jury is allowed to simultaneously consider conviction on several lesser included offenses. For example, in some states a jury may be asked to consider a defendant's state of mind or intentions with respect to a homicide, and choose from among first degree murder, second degree murder, and manslaughter. All of the above multiple charge situations may have related effects on jurors' decision processes, all have been the subject of a certain amount of legal theorizing, and all lend themselves to empirical investigation.

The issue of prejudicial joinder is just one example of the intuitive psychological

assumptions found in the Rules of Evidence under the more general classification of "prejudicial evidence," and indeed, judges often refer to the Rules of Evidence in their joinder decisions. The psychological implications of prejudicial evidence have been discussed by Penrod and Borgida (1983) who observe that the Rules of Evidence recognize several types of evidence that are potentially prejudicial, e.g., character evidence, evidence of other crimes, prior convictions, and similar happenings. In all instances the issue of admissibility is concerned with the relevance of the evidence to the case at hand, and whether relevance outweighs potential prejudice. Rule 401 of Federal Rules of Evidence defines "relevant evidence" as "evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence," and the amount of relevance ascribed to any given evidence is termed its probative value. Rule 403 of Federal Rules of Evidence provides the classic "balancing test" in the law, which is the fundamental rule used to determine admissibility:

Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice.

Unfortunately, the Rules of Evidence do not define prejudice, nor do they provide guidelines for weighing relevance against prejudice. However, the balancing test has been subject to a good deal of legal scrutiny in evidence handbooks and textbooks, since it is such a basic notion in the Rules of Evidence (Cleary, 1972; Lempert & Saltzburg, 1983; Lilly, 1978).

#### PRIOR RESEARCH

Horowitz, Bordens and Feldman (1980) examined the impact of joining two charges on jurors' assessments of defendant guilt using an audiotaped trial summary. Horowitz et al. employed a 2 (strength of evidence) x 3 (severed or joined with a strong or weak case) x 2 (position) factorial design. All main effects were significant as were several interactions. The significant joinder effect demonstrated that jurors' ratings of defendant guilt were higher when two offenses were joined than when the offenses were tried separately. Bordens and Horowitz (1983) investigated the effects of joining two charges that varied according to case strength, charge similarity (two rapes or a murder and a rape) and case order, again using audiotaped trial summaries. They found that convictions on the first but not the second charge were significantly higher when the charge was joined than when it was severed. Convictions were also more likely when the joined charges were similar than when they were dissimilar. Bordens and Horowitz investigated the processing of trial information by asking subjects to recall evidence from the cases, generate thoughts that related to their verdict preference, and rate the thoughts according to their degree of favorableness to prosecution or defense. Although the reported results are complex, a general pattern was detected. Subjects generated a higher percentage of anti-defendant thoughts when cases were similar than when they were dissimilar. Subjects also made a higher percentage of anti-defendant recall intrusions from the second case to the first when charges were similar. Interestingly, Bordens and Horowitz found that ratings of thoughts against the defendant did not differ in joined and severed conditions. These ratings can be considered an indirect measure of the strength of evidence against the defendant, suggesting that joinder did not affect judgments of evidence strength. However, ratings of thoughts against the defendant bore a strong relationship to verdict ratings ( $r = .76$ ), suggesting that all subjects' judgments were strongly influenced by their assessments of evidence strength.

In two experiments conducted by Greene and Loftus (1981) subjects read excerpts from a trial transcript consisting of a single charge (murder or rape) or two charges (murder and rape). Greene and Loftus found that the defendant was more likely to be convicted of either crime if the two charges were joined than if they were tried separately. Judges' instructions to consider charges separately were ineffective in removing this effect, regardless of whether they came before or after the trial. Greene and Loftus investigated three mechanisms to account for their effects: memory, a change in reasonable doubt standard, and inferences of a criminal disposition. They found that subjects in joined and single conditions were equally accurate on a fact recognition task, so the memory explanation was not supported. There was also no difference between joined and single conditions in the amount of proof needed to convict the defendant. However, subjects in joined conditions rated the defendant more negatively on the dimensions of dangerousness, likeableness, and believability, suggesting that joinder affected inferences about the defendant's character.

In a study designed to investigate each of the three legal theories of prejudice, Tanford

and Penrod (1982) extended this finding to trials with three and four offenses. Subjects read written trial summaries consisting of a single charge or a joined trial of two, three or four charges in one of several combinations. The results indicated that the probability of conviction on a particular charge increased as a function of the number of charges with which it was joined. Tanford and Penrod also obtained support for each of the three legal theories of prejudice resulting from joinder. Subjects in joined trials evidenced more confusion in recall of evidence from a joined trial of three charges than in recall of evidence from three separate charges containing the same information. Subjects judging joined offenses also rated individual items of evidence as more incriminating than subjects who rated the same evidence from a single-offense trial, supporting the process of accumulation of evidence. Subjects also rated the defendant on seven 9-point bipolar scales on a number of trait and behavioral characteristics. On most ratings the defendant fared less favorably in joined than single conditions. Thus supporting the theory of criminal inference.

Tanford and Penrod also examined the relationship between memory, evidence strength and defendant ratings with respect to verdicts and judgments of the defendant's guilt. They found a low, nonsignificant positive relationship between memory intrusions and guilt, suggesting that confusion is not a key mediating factor. They found a strong positive relationship between ratings of the evidence and guilt in both joined and single conditions. What these results suggest is that all subjects were basing their judgments on the strength of the evidence, as they are supposed to legally. In addition, subjects based their judgments on inferences about the defendant, and subjects in joined trials did so to a greater degree than subjects judging single trials.

#### Limitations of Previous Research

The research reviewed on the effects of joinder (and related prejudicial evidence) demonstrates empirically that mock jurors' judgments can be biased by several evidentiary and procedural factors. The empirical studies have essentially confirmed legal intuitions that evidence and procedures can be prejudicial, and have further suggested that the legal remedies may not be adequate. However, these studies suffer from a number of limitations, both in terms of applications to the courtroom and in terms of providing an understanding of the psychological mechanisms underlying judgment biases.

From an applied standpoint, most of the studies reviewed were conducted using procedures which were low in external validity, so that their generalizability to actual trial settings is questionable. All used undergraduate subjects, all used written trial summaries except for the Bordens and Horowitz (1983) and Horowitz et al. (1980) studies which used audiotaped summaries, and none included group deliberation. In order to obtain results that can be applied to the courtroom, it would seem desirable to conduct experiments more closely resembling an actual trial.

#### THE PRESENT RESEARCH

##### Overview of the Present Research

The present research was designed to avoid the limitations of previous research by using procedures that maximized external validity, and by investigating the underlying mechanisms involved in jurors' judgments. The purpose of the research was to examine the effects of multiple charges using a realistic trial simulation in order to obtain results that could be applied to the courtroom, and to provide an understanding of the psychological processes in operation for actual jurors judging a joined trial. The research was designed with several general goals in mind. One goal was to determine whether the results obtained in earlier laboratory experiments could be replicated and extended in a more realistic trial setting. A second goal was to develop a research paradigm that could later be used to investigate other assumptions in the Rules of Evidence, using methods and procedures that would have clear applicability to the courtroom. As a corollary to this, it was intended that the results obtained would provide insight into the general processes involved in judgmental biases in the courtroom.

Specifically, the purpose of the research was to determine under what conditions (and to what extent) jurors will become biased when several charges are tried together in a single trial. It was predicted that joinder would increase the likelihood of conviction, but that the magnitude of these effects would be influenced by three independent variables: (1) The similarity of the offenses charged, (2) the similarity of the evidence contained in the offenses and (3) judges instructions designed to reduce prejudicial effects of joinder.

Similarity was examined for two reasons. First, from an applied perspective, we wanted to provide guidance to judges as a basis for making decisions about when to join charges. The law primarily allows for joinder of similar charges, though in fact joining similar charges might be more prejudicial than joinder of dissimilar charges. Because the courts have looked to similarity as a basis for categorizing charges, the present research was designed to inform the courts about what specific combinations of charge and evidence similarity would be most prejudicial. Second, social psychological research as well as previous research on joinder suggested that similarity would affect the relative contribution of the hypothesized sources of prejudice: confusion, accumulation and criminal inference.

Hypotheses. Based on research and theory in social and cognitive psychology, as well as empirical research on joinder and other prejudicial evidence, a number of predictions were made concerning the effects of multiple charges.

1. It was predicted that (a) a defendant charged with three offenses was more likely to be convicted on any particular charge than a defendant tried for the same crime by itself and (b) it was also predicted that subjects in joined trials would confuse evidence among charges, view the evidence as stronger than subjects in single trials, and make negative inferences about the defendant.

2. The similarity of the joined offenses was predicted to influence the magnitude of the conviction effects by influencing the memory and social inference processes hypothesized to mediate joinder effects. It was predicted that (a) a defendant charged with three similar crimes would be more likely to be convicted than a defendant charged with three dissimilar crimes. (b) Greater confusion between charges was predicted when charges were similar. (c) It was also predicted that subjects would make more inferences about the defendant's criminal character when charges were similar than when charges were dissimilar. A series of similar charges is more likely to evoke a criminal schema. In attribution theory terms, being charged with several similar crimes creates an impression of consistency.

3. The similarity of the evidence contained in the joined offenses was predicted to affect jurors' judgments. (a) Evidence similarity should primarily affect the accumulation of evidence process, since there will be more evidence to accumulate if the evidence for each charge is different. (b) However, more confusion of evidence was predicted for charges containing similar evidence. (c) Although it is not clear that evidence similarity will affect inferences about the defendant, attribution theory suggests that more dispositional attributions might be made when evidence is dissimilar, indicating behavior that is low in distinctiveness. (d) It was predicted that similar and dissimilar evidence might also vary in terms of their probative value, credibility, or informativeness, and these dimensions could also affect jurors' judgments.

#### METHOD

Subjects. A total of 732 subjects participated in the experiment. Of these, 714 were qualified jurors who had been summoned for service in the Dane County, Wisconsin, jury pool for 1981 and 1982. Of these, 492 (69%) subjects had jury experience, while the remaining 31% had been summoned but not seated on a trial. Jurors were first sent a letter describing the study and were followed up with a phone call to schedule them for a session. Jurors were paid \$20 for participation. The remaining 18 subjects were undergraduates at the University of Wisconsin who were registered voters and therefore jury-qualified. Undergraduate subjects received Introductory Psychology course credit for participation, with the exception of one subject who received \$20. Undergraduates were scheduled in order to fill in sessions for which there were not enough jury pool subjects to form six-person groups for deliberation. Twelve groups contained a single undergraduate and three groups contained two undergraduates, and these groups were distributed evenly across the experimental conditions. The sample as a whole was 49% female and 51% male. The mean age was 40 years, with a range from 18 to 82. Subjects represented wide range of socioeconomic status variables such as income, occupation and educational background.

Design. The design of the experiment is presented in Table 1. A partial factorial design with an additional control group was used. The control group judged a trial that consisted of a single burglary charge which we will refer to as the "target" offense. The experimental groups judged a trial that consisted of the same target charge in combination with two other charges that represented the experimental manipulations. The independent variables were (1) charge similarity: identical, similar or dissimilar, (2) evidence similarity: similar or dissimilar and (3) judges' instructions: present or absent. Charge



Table 1. Experimental Design

CONTROL GROUP	EVIDENCE SIMILARITY	CHARGE SIMILARITY			
		IDENTICAL	SIMILAR	DISSIMILAR	
<div style="border: 1px solid black; padding: 5px; display: inline-block;"> <u>B</u><sub>1</sub> [0]         </div>	Similar	<u>B</u> <sub>1</sub> <u>B</u> <sub>1</sub> <u>B</u> <sub>1</sub> [1]	<u>B</u> <sub>1</sub> <u>b</u> <sub>1</sub> <u>b'</u> <sub>1</sub> [3]	<u>B</u> <sub>1</sub> <u>A</u> <sub>1</sub> <u>R</u> <sub>1</sub> [5]	No Instruct- ions
	Dissimilar	<u>B</u> <sub>1</sub> <u>B</u> <sub>2</sub> <u>B</u> <sub>3</sub> [2]	<u>B</u> <sub>1</sub> <u>b</u> <sub>2</sub> <u>b'</u> <sub>3</sub> [4]	<u>B</u> <sub>1</sub> <u>A</u> <sub>2</sub> <u>R</u> <sub>3</sub> [6]	
	Similar	<u>B</u> <sub>1</sub> <u>B</u> <sub>1</sub> <u>B</u> <sub>1</sub> [7]	<u>B</u> <sub>1</sub> <u>b</u> <sub>1</sub> <u>b'</u> <sub>1</sub> [8]	<u>B</u> <sub>1</sub> <u>A</u> <sub>1</sub> <u>R</u> <sub>1</sub> [9]	Instructions

B<sub>1</sub> = target offense

---

Charge      B = burglary (service station), b = burglary (residence),  
 Codes      A = assault      b' = burglary (business)  
             R = robbery

Evidence    1 = circumstantial evidence  
 Codes      2 = eyewitness identification  
             3 = other evidence (fingerprints, informant, or stolen property)

and evidence similarity were crossed factorially and judges' instructions were manipulated for the similar evidence, but not the dissimilar evidence conditions, resulting in 9 experimental groups and one control group. (It was not financially possible to run the complete factorial design using jury pool subjects.

Based on pretesting of the materials, the independent variables were defined as follows: Charge similarity was defined as the type of crime and the circumstances surrounding the crime, where identical charges were three service station burglaries, all committed in the same manner; similar charges were three somewhat similar burglaries committed at different establishments—the target service station burglary, a house burglary, and burglary of a commercial business establishment, and dissimilar charges were burglary, assault and armed robbery charges.

Evidence similarity was defined as the type of evidence brought to trial by the attorneys to prove their case. For similar evidence conditions the evidence for each charge was circumstantial evidence that the defendant was seen driving suspiciously near the scene around the time of the crime with no explanation for his whereabouts. For dissimilar evidence conditions the main evidence was different for each charge. For example, the same circumstantial evidence for the target offense might be combined with a charge containing an eyewitness identification and a charge containing information from an informant.

An independent sample of undergraduates ( $n = 45$ ) rated the evidence from the cases used in the main study in terms of its informativeness, credibility, and probative value (which was explicitly defined). The ratings indicated that dissimilar evidence was rated higher than similar evidence in terms of its informativeness ( $M = 65.9$  versus  $53.5$ ), credibility ( $M = 65.3$  versus  $58.1$ ) and probative value ( $M = 62.4$  versus  $49.4$ ) on scales from 0% to 100%. In our discussion here we will continue to define evidence in terms of its similarity, since it was conceptualized as such, while recognizing that it varied along other dimensions as well.

The judges' instruction manipulation was defined as a special joinder instruction given by the judge along with the standard jury instructions presented at the end of the trial. The instruction was designed by elaborating on sections taken from the Federal and several state joinder instructions. The goal was to create a strong and complete set of instructions containing elements which corresponded to the three legal theories of prejudice from joinder. The instruction manipulation for similar charges read as follows:

1. The defendant is charged with three counts of burglary. These are separate crimes and the prosecutor is charging that the defendant committed all of them. The fact that the defendant is charged with more than one crime is not evidence against him.
2. Each charge and the evidence pertaining to it should be considered separately. You should treat the evidence from each charge as separate and distinct.
3. It is for you to determine whether the defendant is guilty of one, two, three, or none of the offenses charged. The fact that you may find the accused guilty or not guilty as to one of the offenses charged should not control your verdict as to any other offense charged. In deciding the defendant's guilt or innocence on a particular charge, you should consider the evidence pertaining to that charge only, and you should not consider the evidence from the other two charges. Each count charges a separate crime, and you must consider each one separately.

The instructions were much stronger than those used in actual trials, but were realistically patterned after actual instructions and presented in the traditional manner.

Stimulus materials. Subjects viewed videotaped trial re-enactments. The case materials were based on reports of burglary, assault and armed robbery cases tried in Wisconsin. The target offense was adapted from a complete trial transcript. Pretesting of the materials indicated that they met the requirements of the research in terms of case strength and the independent variables. Two experienced trial attorneys were recruited to serve as the attorneys in the trial re-enactments. The trial re-enactments were videotaped at the University of Wisconsin Law School courtroom over the course of a weekend. Each of the thirteen offenses was filmed individually (the joinder manipulations were accomplished through editing). The cases were essentially "tried" spontaneously on camera, resulting in an abbreviated but complete trial lasting from 30 to 45 minutes for each individual offense.

## Joinder Executive Summary

The experimental conditions were created by editing together combinations of three charges each, all of which contained the target offense in combination with two other charges. In fact, it is because the content of the target offense remained constant across charges that it was not necessary to precisely control the content of the trial re-enactments. The edited versions were presented in the form a joined trial is actually conducted. The target offense presented as a single trial constituted the control group for all experimental conditions. Each joined trial lasted from 1 1/2 to 2 hours, and the single trial lasted approximately 50 minutes.

**Procedure.** Subjects participated in evening sessions in groups of six to sixteen per session. Subjects first viewed the videotaped trial in black and white on a 19" television monitor. Immediately following the trial they individually answered a short "pre-deliberation" questionnaire on which they indicated their verdict preference, certainty in verdict and likelihood of defendant guilt. Subjects were then formed into one or two groups of six members for deliberation, and any extra subjects were sent to a separate location to begin answering a "non-deliberating" juror questionnaire (described below). Juries were instructed to deliberate and reach a unanimous verdict, and were given a form on which to record their verdict for each charge. If an experienced juror had previously been foreman of a jury, that juror was appointed foreman, otherwise a foreman was selected randomly by the experimenter. All deliberations were videotaped for later analysis. A time limit of one hour was placed on deliberations, and juries were not allowed to declare themselves hung if they had not deliberated the full hour. Groups were also given warnings by the experimenter after 50 and 55 minutes of deliberation indicating that they had only a few more minutes in which to reach a verdict. Following deliberations, subjects individually completed a "post-deliberation" questionnaire which is described below. Subjects were then debriefed in a group and paid \$20 by the experimenter. The entire session lasted from 2 to 3 1/2 hours.

Dependent Measures

**Pre-deliberation questionnaire.** Prior to deliberation, jurors individually answered a short questionnaire which contained the following measures for each charge: (1) verdict--guilty or not guilty, (2) certainty in verdict on a 9-point scale, (3) probability of guilt of the defendant on a 9-point scale and (4) reasonable doubt standard on a 9-point scale. Due to the failure of a large number of subjects to understand the last question, it was excluded from analysis.

**Post-deliberation questionnaire.** Following deliberations, jurors responded to a longer questionnaire designed to assess the processing of trial information. Non-deliberating subjects completed the same questionnaire without participating in deliberations. The questionnaire covered:

1. Defendant ratings.
2. Memory. Subjects were given two memory tasks designed to assess their degree of confusion between charges. (a) Free recall. For each case subjects were asked to list the evidence that most strongly supported the prosecution's case, and to do the same for defense evidence. (b) Recognition. Subjects were given a multiple choice recognition task. (c) Judges' instructions. Subjects in the instructions conditions were also asked for free recall of the judge's instructions with respect to multiple charges.
3. Evidence ratings. Subjects were asked to rate the strength of the evidence for prosecution and defense. The evidence ratings provided a measure of accumulation of evidence.

## RESULTS

Predeliberation Results

**Verdicts.** Prior to deliberation, subjects provided individual verdict preference (guilty or not guilty), and rated certainty in verdict and probability of the defendant's guilt on 9-point scales. These ratings were analyzed for all subjects (deliberating and non-deliberating), since deliberating and non-deliberating groups were equivalent prior to deliberation. Analyses were performed on the first (target) charge only, since it was the only charge that remained constant across conditions--the other two charges served as the experimental manipulations. Because the design was not a full factorial design, and since

many of the effects of interest involved comparison of different experimental conditions with the single control group, a series of planned comparisons was performed as recommended by Himmelfarb (1975). A modified Bonferroni procedure was used to control error rates (Keppel, 1982). Results with a probability less than .035 are considered significant. Experimental-control comparisons for which the hypotheses predicted higher ratings in joined versus control conditions were tested using one-tailed significance tests, and experimental-control comparisons for which the hypotheses were not directional used two-tailed tests. We will confine our discussion here to the effects of primary interest: joinder (versus the control group), instructions (versus no-instructions), charge similarity, evidence similarity, and interactions among these factors.

The proportion of individual guilty verdicts and probability of guilt ratings obtained in the ten experimental conditions are provided in Table 2. Analysis revealed a significant effect for joinder,  $t(722) = 2.57$ ,  $p < .01$ , Effect Size = .10, with a mean proportion of 39% guilty verdicts in joined conditions, as opposed to 24% guilty verdicts in the control group. There was no effect for instructions; in fact, conviction rates in joined-instructions ( $M = .38$ ) and no-instructions ( $M = .39$ ) conditions were virtually identical. There were no effect for charge similarity or evidence similarity, although there was a tendency of marginal significance,  $F(1,722) = 3.05$ ,  $p = .08$ ,  $ES = .06$  for more convictions in dissimilar evidence ( $M = .43$ ) than similar evidence ( $M = .35$ ) conditions. There were no interactions among any of the variables. On the basis of the verdict results, it can be concluded that joinder significantly increased the likelihood of conviction, and that judges' instructions were totally ineffective in reducing these effects.

**Certainty in verdict.** In analyses of verdict certainty ratings, no significant results were obtained. All subjects were equally confident in their verdicts regardless of condition. Mean certainty ratings ranged from 6.69 to 7.47 on a 9-point scale.

**Probability of guilt.** It was predicted that subjects in joined conditions would judge the defendant as more likely to be guilty than subjects in the single case control group. Although the direction of the means supports this prediction, the joinder effect overall was only marginally significant,  $t(721) = 1.46$ ,  $p = .07$ ,  $ES = .05$ , with means of 5.41 for joined trials and 4.99 for the control group. There was a marginally significant effect for evidence similarity,  $F(1,721) = 3.85$ ,  $p = .05$ , in which subjects in dissimilar evidence conditions rated guilt as higher ( $M = 5.69$ ) than subjects in similar evidence conditions ( $M = 5.21$ ). This result parallels the result obtained for verdict--subjects also returned more guilty verdicts when evidence was dissimilar. There were no instruction effects or interactions on the probability of guilt ratings.

Deliberation Results

**Reliability checks** The deliberation videotapes were coded by two undergraduate assistants. One person coded approximately two-thirds of the juries, the other coded the remaining one-third. In order to assess coder reliability, the two coders independently coded a sample of four juries (approximately 2 hours of deliberation). Correlations between coders were computed on the number of statements coded for each juror under various category headings. The average correlation was .75. Also, close to 90% of all statements were coded into meaningful categories, with only 11.5% coded into the miscellaneous category. (Table 2b).

**Individual and group voting behavior.** Table 3 presents the group verdicts for each of the three charges as a function of the number of jurors who initially voted to convict. It is apparent that all cases were on the weak side, with 63% acquittals on the first charge, 80% acquittals on the second, and 79% on the third. The results indicate that majorities tended to prevail; on the first charge there were only six reversals of initial majorities; two groups with initial 4-2 splits for conviction ultimately acquitted, and four groups with 2-4 splits convicted. For the second and third charges, there were four reversals apiece. Thus, the present data demonstrate the well documented finding that the initial juror vote distribution is a good predictor of the final group outcome (Davis, 1973; Kalven & Zeisel, 1966; Penrod & Hastie, 1979, 1980; Stasser & Davis, 1981).

**Total deliberations.** Separate path analyses of the influence of predeliberation verdict preferences and deliberation contents on final verdicts for each individual charge suggested that for the first charge, the content of jurors' deliberations are strongly influenced by their initial verdict preferences, and these deliberations influence the outcome. For the second and third charges, jurors' final votes are based more directly on their initial vote preferences, and less on their discussions of evidence. We also found that jurors spent less

Table 2 Proportion of guilty verdicts

CONTROL GROUP	EVIDENCE SIMILARITY	CHARGE SIMILARITY			
		IDENTICAL	SIMILAR	DISSIMILAR	
.24 n = (83)	Similar	.32 (77)	.36 (68)	.33 (74)	No Instruc- tions
	Dissimilar	.43 (68)	.41 (72)	.46 (72)	
	Similar	.36 (69)	.43 (76)	.35 (69)	

Table 2b

Proportions and frequencies of statements per juror in each category.

Category	Charge 1		Charge 2		Charge 3		Total	
	%	Freq	%	Freq	%	Freq	%	Freq
CASE FACTS								
Positive	11.7	1.82	10.9	0.73	8.9	0.48	11.7	3.22
Negative	10.7	1.66	4.7	0.31	2.5	0.13	8.4	2.31
Neutral	7.6	1.18	8.4	0.56	8.3	0.45	8.4	2.31
Questions	9.6	1.49	8.6	0.57	9.5	0.51	10.4	2.83
Total	39.6	6.15	32.6	2.18	29.2	1.57	38.9	10.70
ERRORS <sup>1</sup>								
Positive	0.2	0.03						
Negative	0.2	0.03						
Neutral	0.6	0.09						
Corrections	0.9	0.14						
Total	1.9	.29						
VERDICTS								
Guilty	6.5	1.01	0.4	0.03	2.5	0.13	4.6	1.27
Not Guilty	14.3	2.22	33.4	2.23	32.7	1.76	18.6	5.12
Questions	1.2	0.19	0.7	0.05	0.6	0.03	0.9	0.25
Guilty, No Proof	1.5	0.23	1.0	0.07	0.7	0.04	1.2	0.33
Total	23.5	3.65	35.5	2.37	36.5	1.97	25.3	6.96
REASONABLE DOUBT								
Positive	2.0	0.31	0.9	0.06	0.8	0.04	1.6	0.44
Negative	0.5	0.08	0.1	0.01	0.1	0.01	0.3	0.08
Neutral	1.1	0.17	0.3	0.02	0.6	0.03	0.9	0.25
Total	3.6	0.56	1.3	0.09	1.5	0.08	0.8	0.77
EVIDENCE SUFFICIENCY								
Positive	2.6	0.40	3.1	0.21	2.5	0.13	2.7	0.74
Negative	0.9	0.14	0.1	0.01	0.0	0.00	0.6	0.17
Total	3.5	.54	3.2	0.22	2.5	0.13	3.3	0.91
VERDICT ELEMENTS								
Positive	1.6	0.25	1.3	0.09	0.7	0.04	1.4	0.39
Negative	0.0	0.00	0.2	0.01	0.0	0.00	0.1	0.03
Neutral	0.7	0.11	0.7	0.05	0.6	0.03	0.7	0.19
Total	2.3	0.36	2.2	0.15	1.3	0.07	2.2	0.61
MULTIPLE CHARGES								
Confusion	0.8	0.12	1.0	0.07	0.9	0.05	1.0	0.28
Accumulation	0.8	0.12	0.9	0.06	0.5	0.03	1.0	0.28
Inferences	0.9	0.14	1.0	0.07	0.8	0.04	1.0	0.28
Other	0.9	0.14	1.0	0.07	2.3	0.12	1.4	0.39
Total	3.4	0.53	3.9	0.26	4.5	0.24	4.4	1.21
OTHER								
Experiment	4.3	0.67	4.1	0.27	6.5	0.38	4.6	1.27
Directions	6.4	0.99	3.8	0.25	4.4	0.24	5.0	1.38
Outburst	1.0	0.16	0.8	0.05	1.0	0.05	1.0	0.28
Irrelevant, Uncodable	10.6	1.65	9.1	0.61	12.5	0.68	11.5	3.16
Total	22.3	3.46	17.8	1.19	24.4	1.31	22.1	6.08
SUM	100	15.52	100	6.69	100	5.38	100	27.51

<sup>1</sup>Errors were not coded for Charges 2 and 3.

TABLE 3

Jury Verdicts as a Function of Initial Votes

Initial votes for conviction	CHARGE 1 VERDICTS			CHARGE 2 VERDICTS			CHARGE 3 VERDICTS		
	Not Guilty	Guilty	Hung	Not Guilty	Guilty	Hung	Not Guilty	Guilty	Hung
0	8	0	0	15	0	0	30	0	2
1	19	0	3	31	0	1	28	1	1
2	22	4	4	26	1	2	13	1	0
3	12	2	7	5	1	2	6	2	1
4	2	8	4	2	1	2	2	1	1
5	0	4	0	1	0	0	0	1	0
6	0	1	0	0	0	0	0	0	0
Sum	63	19	18	80	3	7	79	6	5

time discussing the second and third charges than they did the first. In addition, the predictive accuracy of the path models for Charges 2 and 3 was far inferior to that of Charge 1. All of the above facts together suggest that jurors may use other information in their judgments of later charges that is not captured in the individual charge models. Specifically, jurors may use information from the charges they judged previously.

In order to investigate this possibility, we incorporated data from each of the three charges into a single path analysis. The path model included the following variables: jurors' predeliberation votes on charges 1, 2, and 3; the total proportion of statements (collapsed across charges) in each of the four main content categories (positive and negative facts, guilty and not guilty statements), and jurors' final votes on Charges 1, 2, and 3. It was assumed that each decision could influence the subsequent one, therefore the variables were entered in the following order: initial vote for Charge 1, initial vote for Charge 2, initial vote for Charge 3, deliberation categories (all entered at the same level) final vote for Charge 1, final vote for Charge 2, final vote for Charge 3.

The path analysis results are presented in Figure 1, with only significant paths entered. In terms of initial verdict preferences, jurors' votes on all three charges were significantly related to each other. With respect to predicting deliberation content, the initial vote on Charge 1 was a significant predictor of all four categories, the initial vote on Charge 3 significantly predicted all categories except not guilty statements, and the only significant effect for Charge 2's initial vote was on negative facts. Thus, as in the analyses of individual charges, vote 1 had a stronger influence on deliberations than votes 2 and 3.

The final vote for Charge 1 was significantly predicted by 3 of 4 content categories (surprisingly, positive facts had no effect). None of jurors' initial votes significantly influenced final votes for Charge 1; as in the individual case analysis the effects of initial votes were indirect rather than direct. For Charge 2's final vote, none of the content categories were significant predictors, although the coefficients for guilty and not guilty statements were marginally significant ( $p < .10$ ). The only significant predictors of Charge 2 final votes were jurors' initial votes on Charge 2, and their final votes on Charge 1. For Charge 3, 3 of 4 content categories significantly predicted final votes, although these effects were weaker than those obtained for Charge 1. Jurors' final votes on the previous two charges were much stronger predictors of final votes on Charge 3 than was deliberation content, and the initial vote on Charge 3 was also a significant predictor of the final vote. The amount of variance accounted for in the final vote (39%) was far superior to the R-squared of .16 that was obtained when Charge 3 was considered alone.

The preceding analysis creates an overall picture of the dynamics of the deliberation process. Jurors spend a good deal of time discussing the first charge, their deliberations are influenced by their initial votes, and they in fact make their decisions on the basis of their deliberations. Thus, it appears that an informal influence process operates to affect decisions on the first charge. For the second and third charges, jurors spend much less time discussing the case, and do not base their decisions as much on deliberations as they do on their previous votes (and presumably, the votes of other jurors). Thus, it appears that a normative social influence process is in operation for decisions on the second and third charges (Deutsch & Gerard, 1955).

Although our analyses suggest that jurors tend to base their subsequent decisions on previous ones, there is a plausible alternative explanation for the results, which can be interpreted using the same theory of social influences. The group verdict results presented in Table 2 indicate that jurors were more strongly pre-disposed towards innocence for Charges 2 and 3 than they were for Charge 1. Thus, pressure to conform to the majority (i.e., normative influence) was probably stronger for Charges 2 and 3, therefore there was less consideration of the evidence.

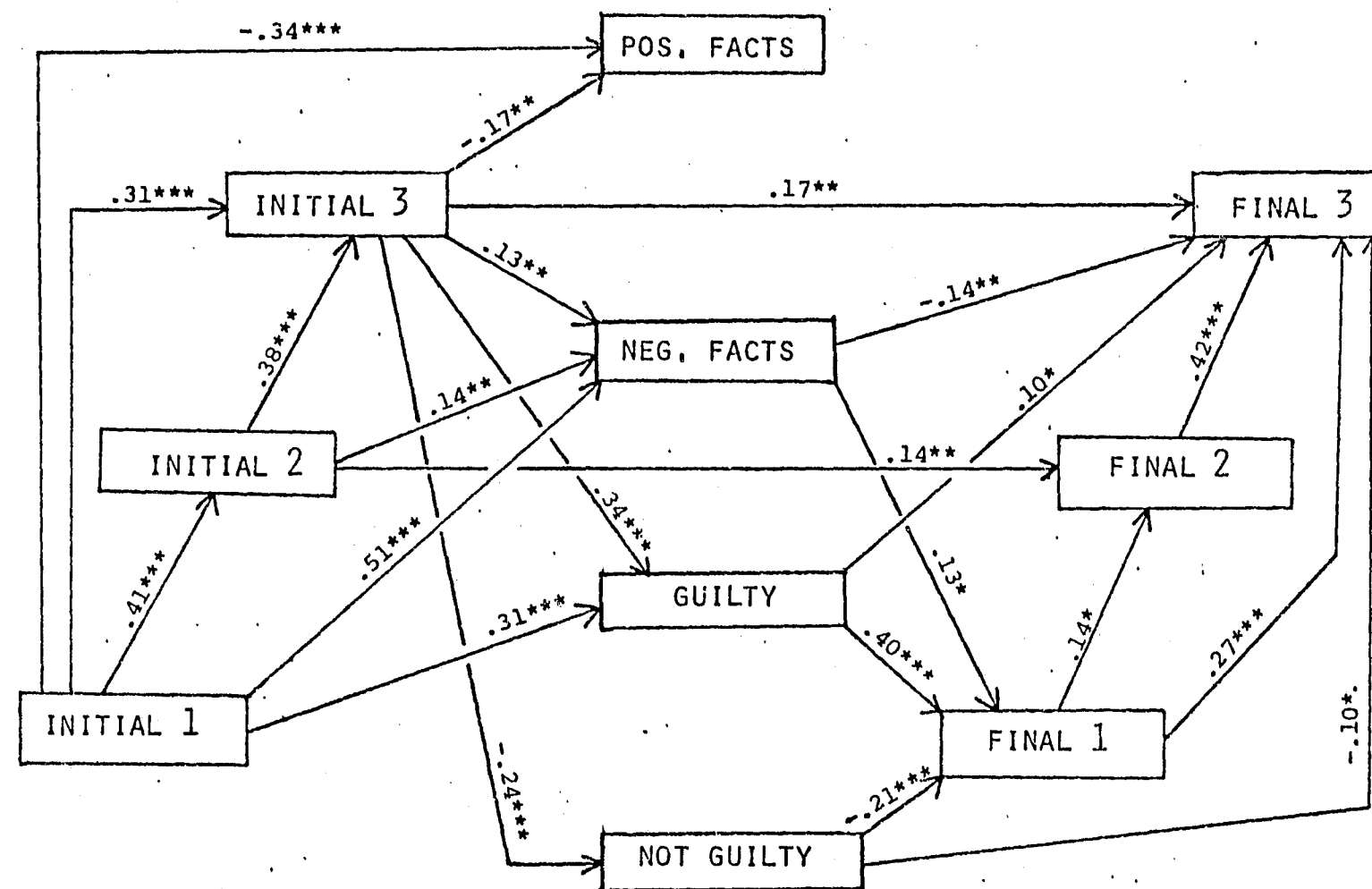
Post-Deliberation Memory Results

Analyses of post-deliberation questionnaire responses were performed on deliberating jurors' responses only, since deliberating and non-deliberating groups were no longer equivalent at the time they completed the questionnaire. In addition, individual subjects were no longer independent, since they had deliberated together in groups, so all analyses employed groups (nested within conditions) as the error term for significant tests.

Recognition task. The recognition task for joined conditions consisted of 16



Figure 1. Percentage of Total Speaking Time for 4 main Categories/Initial and Final Votes For Each Charge



$$F(9,420) = 30.06, p < .001, R^2 = .39$$

# Joinder Executive Summary

multiple-choice items containing 4 correct and 4 incorrect items about the target offense, and 8 "intrusion items" that were facts from the two non-target offenses. Subjects were asked to choose as many or as few of these items as they thought were contained in the target charge testimony. Of course, subjects in the control group had not been exposed to any non-target testimony. They were given the recognition task containing the same correct and error items, along with intrusion items from all joined conditions (which, with few exceptions, were different for each condition), thus subjects in the control group had almost six times as many opportunities to make an "intrusion" error (a total of 45 possible intrusions). Therefore, the control-joined comparison represents a conservative test of the confusion hypothesis.

Overall, subjects were 91% accurate on the correct items, and made 10% factual errors on the incorrect items. There were no joinder effects on these measures. Our primary concern was with intrusion errors. There was a significant joinder effect on the number of intrusions,  $E(1, 90) = 15.40, p < .001, ES = .37$ , with a mean of .32 intrusions in the control group and .87 intrusions in experimental groups. There was a main effect for charge similarity,  $E(2, 90) = 4.76, p < .02$ , which indicated that subjects made more intrusions as charge similarity increased, with means of 1.10, .81, and .69 for identical, similar, and dissimilar charges respectively. There was a marginal Charge similarity x Evidence similarity interaction,  $F(2, 90) = 3.59, p < .05, ES = .23$ , which was rather difficult to interpret. There were fewer instructions in the similar charge-dissimilar evidence condition ( $M = .59$ ) than in any other joined condition, so that this cell tended to "disrupt" the otherwise orderly pattern obtained for charge similarity. There was no effect for instructions, and instructions did not interact with any other variable. The recognition results as a whole indicate that joinder did promote confusion of evidence, but this confusion was not great relative to the total amount possible.

Evidence ratings. Subjects were asked to rate the overall strength of the evidence for prosecution and defense on 9-point scales from weak to strong, and also to rate the incrimination value of two individual evidence items for both prosecution and defense on 9-point scales from innocence to guilt. The responses to the two items were summed to produce four evidence ratings for each subject: (1) prosecution overall, (2) prosecution item sum, (3) defense overall, and (4) defense item sum. None of the analyses on these rating yielded significant effects, although all means were in the predicted direction. Prosecution evidence overall was rated stronger ( $M = 4.23$  vs.  $3.70$ ) and individual items were rated more incriminating ( $M = 11.73$  vs.  $10.98$ ) in joined than single conditions; defense evidence was rated weaker ( $M = 5.26$  vs.  $5.66$ ) and more incriminating ( $M = 8.44$  vs.  $8.03$ ) in joined conditions. However, these differences were small and non-significant, and therefore offer little support for the hypothesis that joinder changes perceptions of evidence strength.

Defendant ratings. Subjects rated the defendant on eleven 9-point trait and behavior scales. Table 4 presents the means and standard deviations of eleven ratings, and the results of a factor analysis that yielded two factors. We have termed the first factor a "criminality-credibility" factor, and the second a "global evaluation" factor. For purposes of analysis, two factor scores were formed and subjected to the same analyses performed on the other dependent measures.

Subjects in joined conditions rated the defendant less favorably in terms of criminality and credibility (Mean = 26.33 overall) than subjects in the control group (Mean = 22.22). Although the results for the charge x evidence similarity ANOVA were not significant, there were patterns for charge and evidence similarity which, although not even of marginal significance ( $p = .08$ ) are worth mentioning. Subjects rated the defendant less favorably as charge similarity increased, with means of 25.78, 26.07, and 27.66 for dissimilar, similar and identical charges. Subjects also rated the defendant less favorably as evidence similarity decreased, with means of 25.84 and 27.15 for similar and dissimilar evidence. These findings are consistent with the pattern that we have been seeing throughout the data—the defendant was judged more harshly on all measures when charges were similar and evidence was dissimilar. As in most of the previous analyses, there was no effect for judges' instructions and no interactions.

## Relationships Among Variables

In order to investigate the process hypothesized to operate in joined trials, the relationships among the variables were examined using path analysis. Table 5 presents the zero-order correlations among the experimental manipulations, ratings of the defendant and

Table 4. Defendant ratings factor analysis

Statistics for each variable			Factor loadings	
Variable	Mean	SD	Factor 1	Factor 2
Sincere	5.09	1.96	.810	.000
Believable	5.10	2.10	.800	.000
Honest	5.07	1.99	.795	.000
Moral	5.06	1.54	.690	.000
Future crime	5.38	2.21	.690	.000
Likeable	4.85	1.69	.632	.000
Typical Criminal	4.89	1.95	.629	.000
Nervous	3.76	1.89	.000	.670
Good	4.90	1.50	.340	.657
Dangerous	3.80	2.19	.000	.626
Attractive	5.00	1.96	.000	.605

loadings less than .25  
have been replaced by 0.

Table 5. Correlations between manipulations, ratings and verdicts

	Joinder	C	E	I	DC	DE	P	D	M	Mean	SD
Joinder										.90	.30
Charge similarity	.55									2.84	1.43
Evidence similarity	.40	.25								.59	.49
Instructions	.21	.12	.53							.29	.45
Defendant Criminality	.17	.17	.01	.00						25.90	7.07
Defendant Evaluation	.19	.13	.15	.07	.41					11.12	2.88
Prosecution Evidence	.07	.10	-.01	.03	.38	.13				4.24	2.50
Defense Evidence	-.07	-.09	-.03	-.04	-.34	-.18	-.48			5.37	2.21
Memory	.16	.18	.08	.04	-.04	-.03	.00	.03		.77	1.10
Verdict	.07	.03	-.02	.03	.33	.12	.55	-.39	-.01	.37	.48

the evidence, memory for evidence, and individual pre-deliberation verdicts. Based on our theoretical predictions, a causal model was devised to specify the hypothesized directional relationships among variables. Hierarchical regression analyses provided path coefficients representing the magnitude of these relationships. Figure 2 graphically depicts the results of the path analysis for the effects of the four manipulated variables and five mediating variables on verdict judgments. Dummy variable coding was employed for the manipulations of joinder (1 = joined, 0 = single), evidence similarity (1 = similar, 0 = dissimilar), and instructions (1 = present, 0 = absent). Charge similarity was scaled to reflect ratings from the manipulation check, resulting in codes of 1, 3, and 4 for dissimilar, similar, and identical charge conditions. Interactions were not coded, since there were no hypothesized interactions and virtually no interactions in the analyses previously reported. The mediating variables consisted of indicators of each of the three hypothesized mediating processes. Defendant criminality and evaluation scores provided measures of criminal inference, overall ratings of prosecution and defense evidence strength served as measures of accumulation of evidence, and the number of recognition intrusions was employed as the measure of confusion of evidence. Regression analysis revealed that these nine predictor variables accounted for 34% of the variance in the verdict data. The most important findings of the analysis are highlighted by the boldfaced lines in Figure 2, which represent all paths with coefficients of .10 or greater. Table 6 presents the path analysis results broken down into direct and indirect causal components.

The model was predicted on the hypothesis that joinder activates a criminal schema which affects jurors' verdicts both directly and indirectly through judgments of the defendant, evidence strength, and memory for evidence. With minor exceptions, the analysis strongly supported this prediction. Both joinder and charge similarity influenced memory directly; however, memory was unrelated to any other variables. Charge similarity was positively related to defendant criminality ratings in addition to memory, but bore little relationship to any other variables. Neither judges' instructions nor evidence similarity had direct or indirect effects on any of the variables.

Our primary concern was with the process whereby joinder influences jurors' decisions. Joinder had a small, positive, direct effect on verdict, while its influence on perceptions of the evidence was negligible. Joinder most strongly influenced perceptions of the defendant's criminality and global evaluations. Defendant criminality ratings influenced verdicts directly, and also strongly affected perceptions of the evidence, having a positive effect on prosecution ratings and a negative effect on defense ratings. However, global evaluations did not significantly influence verdicts or ratings of the evidence. Assessments of the evidence in turn affected verdicts, with strong positive effects for prosecution evidence and weaker, negative effects for defense evidence.

Since the path analysis is based on correlational data, the direction of the effects is not known; however, the results are consistent with the hypothesized pattern of causation. Further support for the hypothesis can be obtained from the decomposition of causal effects in Table 6. The strongest direct effects of verdicts were obtained for prosecution evidence; the strongest indirect effects came from defendant criminality ratings. Joinder exerted its strongest influence on defendant ratings. Thus, the results are consistent with a decisionmaking process whereby joinder leads to inferences about the defendant's criminality, which then influence verdicts both directly (perhaps based on judgments of representativeness) and indirectly (by influencing interpretation and accumulation of incoming evidence).

#### DISCUSSION

The present research has examined a number of issues concerning juror inferencing and judgment processes in multiple-offense trials. In this section, the main findings are discussed in terms of their theoretical and applied significance.

The results indicate that joining multiple charges in a realistic trial situation increases the proportion of individual guilty verdicts obtained on a particular (target) charge relative to the same charge tried by itself. The effects of the manipulations of charge and evidence similarity were relatively subtle compared to the effect of joinder of any sort. Convictions increased regardless of similarity, although there was a marginally significant tendency towards more guilty verdicts when evidence was dissimilar than when it was similar. A very strong set of judges' instructions had no effect on verdict judgments whatsoever. Overall, the results are consistent with previous research using much less realistic methods. Other researchers find that joinder increases conviction rates and

Figure 2. Path Model of representative jurors' judgment process

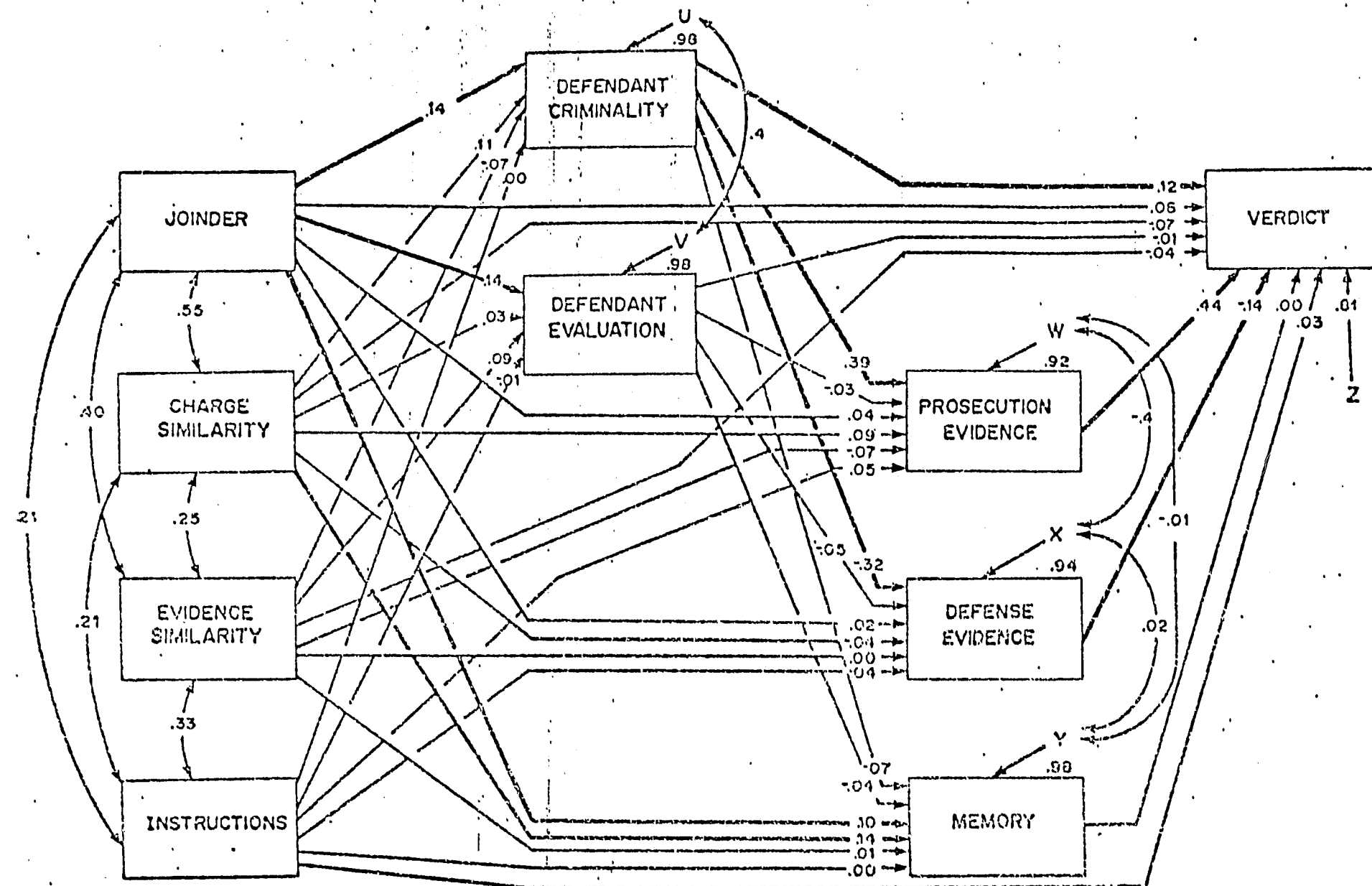




Table 6. Path analysis--Decomposition of causal effects on verdicts

	Total	Direct	Indirect
Joinder	.10	.06	.04
Charge similarity	-.01	-.07	.06
Evidence similarity	-.09	-.04	-.05
Instructions	.06	.03	.02
Defendant criminality	.34	.12	.21
Defendant evaluation	-.02	-.01	-.01
Prosecution Evidence	.44	.44	.00
Defense Evidence	-.14	-.14	.00
Memory	.00	.00	.00

instructions do not significantly reduce convictions (Bordens & Horowitz, 1983; Greene & Loftus, 1981; Horowitz et al., 1980; Kerr & Sawyers, 1979; Tanford & Penrod, 1982). In the present study, joinder also increased the number of guilty and hung group verdicts, relative to the control group. This indicates that the biases induced in jurors prior to deliberation persist through deliberations and affect the final outcome.

Support was obtained for each of the three processes hypothesized to operate in a joined trial. Joinder led to a certain amount of confusion of evidence among charges, particularly when charges were similar, but memory was not related to individual verdicts. Joinder also led to an accumulation of evidence, as measured by ratings of evidence strength, particularly when charges were similar or evidence was dissimilar. Subjects in joined trials rated evidence for the prosecution as stronger than subjects in a single trial, and to a lesser degree rated the evidence for the defense as weaker. This suggests that subjects primarily accumulate evidence against the defendant. Ratings of evidence strength were strongly related to verdicts, more so for prosecution than for defense evidence. Joinder also led to negative inferences about the defendant on dimensions of criminality and global evaluation, and these ratings were significantly related to verdicts. Defendant criminality, but not general evaluation, was strongly related to judgments of evidence strength.

In terms of the three processes postulated to mediate joinder effects, the results are generally consistent with previous research using less realistic stimulus materials. Bordens and Horowitz (1983) and Tanford and Penrod (1982) both found that joinder led to confusion of evidence, although Tanford and Penrod found that confusion was not related to guilt judgments, whereas Bordens and Horowitz found that it was. Greene and Loftus (1981) and Tanford and Penrod (1982) found that joinder led to negative inferences about the defendant, and Tanford and Penrod found that these inferences were strongly related to guilt judgments. Bordens and Penrod found that joinder led to higher ratings of evidence strength, although Bordens and Horowitz found that joinder did not affect ratings of thoughts against the defendant, which could be considered an indirect measure of evidence strength. However, both Bordens and Horowitz and Tanford and Penrod found that ratings were strongly related to verdict and guilt judgments.

The relationships among the variables were integrated into a causal model of judgment processes in joined trials, which is depicted in Figure 2. In the proposed model, joinder leads to negative inferences about the defendant's criminal character. These inferences affect verdicts both directly and indirectly through judgment of evidence strength, which in turn strongly affect verdicts.

#### THEORETICAL IMPLICATIONS

In this section, the main findings of the research are interpreted in terms of three processes hypothesized to operate in joined trials: confusion, accumulation and criminal inference. The path diagram in Figure 2 serves as reference point for this discussion.

##### Confusion

The results indicate that joinder led to a certain amount of confusion among charges on a recognition task, particularly when charges were similar. A strictly cognitive explanation for this result is that recognition intrusions were a result of interference effects in long term memory (Postman & Underwood, 1973). However, the fact that confusion increased as a function of charge similarity suggests a more social psychological explanation for these findings. Similar charges are more easily encoded into a single, coherent representation of the trial than are dissimilar charges. Therefore, it is likely that specific evidence items were recalled in relation to the overall schema, rather than for individual charges. This line of reasoning is supported by research conducted by Hastie and Kumar (1979), which indicated that subjects were more likely to recall schema-incongruent than schema-congruent information. Evidence from dissimilar charges should be less congruent with the overall schema than evidence from similar charges. Further support for a schema-based explanation of the memory results is indicated in the study by Sulín and Dooling (1974), which demonstrated that subjects were more likely to make memory intrusions for a passage that was high in schema-relatedness than one that was not.

Although joinder led to a certain amount of confusion of evidence between charges, confusion was unrelated to verdicts. This finding is consistent with research using other impression formation tasks, which indicates that memory for specific items of information is not strongly related to the overall impression of a stimulus (Anderson & Hubert, 1963;

Dreben, Fiske, & Hastie, 1979; Risky, 1979). The explanation given for these findings is that once the information has been integrated into an abstract representation of the stimulus, the overall impression is independent of the representation of specific items in memory (Dreben et al., 1979). This finding was obtained with "rich behavioral stimuli" (i.e., paragraphs, Dreben et al., 1979, p. 1764), as well as using trait adjectives (Anderson & Hubert, 1963; Risky, 1979). The present study yields a low correlation between memory for specific items and judgments with much richer stimuli than those previously used.

However, we would not want to argue that subjects' judgments were made independent of their memory for any aspect of the trial. Research on the use of the availability heuristic (Tversky & Kahneman, 1973) indicates that people often make judgments on the basis of the most easily remembered information about a stimulus. For example, Reyes et al. (1980) found that manipulating the salience of arguments influenced subjects' judgments of a defendant's guilt. The lack of relationship between memory and verdicts in the present research was likely due to the fact that both recall and recognition tasks asked for memory of brief, discrete case facts, which were probably not the features most available to subjects when making guilt judgments.

On a free recall task, the only difference between joined and single conditions was that subjects in single conditions recalled more total evidence than subjects in joined conditions, both for prosecution and for defense. Joinder did not cause subjects to differentially recall more evidence against the defendant, and recall of evidence was unrelated to verdicts. The recall results underscore the implication of the recognition results that joinder-induced biases are not a result of memory processes.

#### Accumulation

The results indicate that subjects in joined trials rate the evidence for the prosecution as stronger than subjects in single trials, and to a lesser degree rate the evidence for the defense as weaker. The path analysis results further indicate that ratings of prosecution evidence are much more strongly related to verdicts than ratings of defense evidence. From an information integration perspective, this suggests that subjects assign more weight to evidence against the defendant (prosecution evidence) than evidence in favor of the defendant (defense evidence). This is consistent with research demonstrating that negative information is weighted more heavily than positive information in forming impressions (Anderson, 1965; Dreben et al., 1979; Fiske, 1980; Hamilton & Huffman, 1971; Hodges, 1974; Kanouse & Hanson, 1972). In addition, joinder clearly creates a negative overall impression of the defendant, which, if averaged in with judgments of both types of evidence, could make prosecution evidence appear stronger and defense evidence weaker.

If the negative impression of the defendant created by the multiple charge context is represented as a criminal schema, the same evidence rating results can be interpreted in a slightly different manner. Findings from studies reviewed in the introduction indicate that schemas guide the interpretation and organization of incoming information, and that information inconsistent with the schema is often distorted or ignored (Taylor & Crocker, 1981). The causal model in Figure 2 does indicate strong relationships between judgments of defendant criminality and evidence ratings. This suggests that jurors distort the evidence to make it consistent with their criminal schemas, making prosecution evidence appear stronger and defense evidence appear weaker. The paths between evidence ratings and verdicts further suggest that jurors differentially use information to the extent that it is consistent with their schemas. That is, they seem to base their decisions to a greater degree on evidence against the defendant (prosecution evidence) which is consistent with a criminal schema, than evidence in favor of the defendant (defense evidence) which is more difficult to incorporate into a criminal schema.

#### Criminal Inference

The criminal inference hypothesis was tested by asking subjects to rate the defendant on various trait and behavioral characteristics. Factor analysis on these ratings yielded two factors, one representing the defendant's criminality and credibility, and the other representing more global evaluations. Analyses revealed that subjects in joined trials rate the defendant much less favorably on both dimensions than subjects in single trials. These inferences increase further as a function of charge similarity, particularly for undergraduate subjects.

From an attributional perspective, these results suggest that subjects are making

inferences about the causes of the defendant's alleged criminal behavior based on the fact that he is charged with multiple crimes. The multiple charge situation provides information about behavior that is high in consistency, particularly when charges are similar, and thereby is likely to lead to an internal attribution. In terms of distinctiveness, the picture is not quite as clear. It could be argued that being charged with dissimilar crimes indicates behavior that is low in distinctiveness, since it is performed with respect to very different entities. If that were the case, the charge similarity results would not support an attributional interpretation. However, in all joined conditions, the crimes, even though they may have been similar in method, were committed against different victims, on different dates, and in different places, possibly indicating behavior that was low in distinctiveness and therefore more likely to lead to an internal attribution. Finally, in all joined conditions, the defendant's alleged criminal behavior could be considered low in consensus, thus the third component of an internal attribution was present.

Although the defendant rating results can be roughly characterized in attribution terms, they are more consistent with an interpretation that does not assume causal inferences are made in such a scientific manner. It was hypothesized that joinder creates an impression of the defendant as a prototypical criminal. The finding that defendant ratings became less favorable as charge similarity increased supports this hypothesis, since similar charges are more easily incorporated into a criminal schema than dissimilar charges. Defendant ratings were positively related to jurors' verdicts, and this relationship was stronger for ratings of defendant criminality than for global evaluations. Therefore, the defendant rating results suggest that joinder creates a criminal schema, which then influences verdicts to the extent that the defendant appears representative of a typical criminal.

#### Instructions

Judges' instructions had no significant effect on representative jurors' verdicts, and also did not influence jurors' memory, evidence ratings or defendant ratings. The findings are consistent with social psychological research on context effects (Asch, 1946), belief perseverance (Lord, Ross, & Lepper, 1979; Ross, Lepper, & Hubbard, 1975) and schema-based processing (Taylor & Crocker, 1981), all of which indicate that once impressions are formed, they are quite resistant to change.

#### LEGAL IMPLICATIONS

Since the results of the study were obtained using procedures that were high in external validity, they have clear applications to the courts. The study used representative juror subjects, realistic videotaped trials, and included group deliberation. The results indicate that joinder increased the proportion of guilty verdicts on a particular target charge, relative to the same charge tried by itself. This effect was obtained at the level of both individual and group verdicts. At the individual level, 39% guilty votes were obtained in joined conditions overall, compared to 24% guilty verdicts in the control group, so joinder resulted in 15% more guilty verdicts than would otherwise be the case. Statistically, the magnitude of joinder effects was not large, with an effect size ( $r$ ) of .10 for the overall joinder effect. However, the results are of considerable practical significance, if the additional convictions are considered to be conviction errors. Although the absolute magnitude of joinder effects will depend upon numerous factors (type of crime, case strength, etc.), the present results indicate that joinder can substantially increase the chance that an innocent person will be convicted of a crime.

At the group level, joinder increased the number of guilty and hung jury verdicts, relative to the control group. Therefore, pre-deliberation biases persisted through group verdicts, and deliberation did not serve to correct these biases. This finding further emphasizes the applied significance of the results to a degree that would not be possible if deliberation procedures had not been used. The increase in hung juries in joined over severed trials has only tentative implications, since a time limit was placed on deliberations which probably affected the hung jury rate. However, if it is the case that juries are hung more often when deliberating on joined charges, this suggests that some of the supposed expedience of trying multiple charges together may be offset by an increase in hung juries.

#### Impact of Instructions

Although social psychological research as well as previous empirical work on instructions cast some doubt on whether it would be possible to develop effective instructions, we did

want to give the traditional legal remedy for prejudice a fair test. Therefore, a strong set of instructions was devised, patterned after existing instructions yet longer and more complete. The instruction contained elements corresponding to each of the three legal theories of prejudice: confusion, accumulation and criminal inference.

The results of the study indicated that instructions had no effect whatsoever for representative juror subjects. Viewed in light of other failures to develop effective instructions, the present results strongly indicate that the current legal remedy for prejudice resulting from joinder is simply not adequate. In order for instructions to be effective, they would need to disrupt the processes that mediate the effects of joinder on verdicts. The causal model depicted in Figure 2 suggests that joinder effects are mediated through a criminal schema for the defendant. The portion of the instruction manipulation that addressed this process stated that the fact that the defendant was charged with more than one crime should not be used as evidence against him. If, as we have argued throughout, criminal inferences are not a byproduct of a rational, strictly cognitive process, then it is not surprising that simply instructing jurors not to make inferences did not work. It is not clear that any instructions could effectively change these inferences.

#### Guidelines for Joinder

From an applied perspective, the study had two main objectives: (1) to develop guidelines delineating situations in which joinder would and would not be prejudicial, and (2) to design a set of instructions that would effectively reduce prejudice resulting from joinder. With respect to the first goal, the results indicated that the effects of the charge and evidence similarity manipulations were quite subtle compared to the effects of joinder of any sort. Regardless of the experimental condition, there were more convictions on the target charge in the context of a joined trial than on the same charge tried alone. There was a tendency for jurors to convict more often in dissimilar evidence than similar evidence conditions. A likely explanation for this finding is in terms of the probative value of the evidence. Independent ratings of the evidence by a group of undergraduates indicated that evidence defined as "dissimilar" was rated higher than evidence defined as "similar" in terms of its credibility, value, and informativeness. Since ratings on these three measures were highly correlated, together they can be considered a measure of probative value. Although the probative value of evidence for the target charge should have remained the same in all conditions (since it was always the same evidence), subjects apparently used their perceptions of the evidence in non-target charges when making target charge judgments, and therefore convicted more often in dissimilar evidence conditions.

Since joinder significantly increased convictions in all experimental conditions, one possible guideline would be to (1) avoid joining charges at all. This solution can be compared to two legal criteria currently used as a basis for some joinder decisions. The "simple and distinct" test holds that charges can be joined if the evidence from each is simple enough that jurors will not confuse evidence between charges. The present results indicate that this solution is not likely to reduce prejudice. Subjects judging joined offenses did confuse evidence between charges, but confusion was unrelated to verdicts.

The law primarily allows for joinder of similar crimes, and charges are often joined if they can pass the "other crimes" test of admissibility. Rule 404 (b) of Federal Rules of Evidence states:

Evidence of other crimes, wrongs, or acts is not admissible to prove the character of a person to show that he acted in conformity therewith. It may, however, be admissible for other purposes, such as proof of motive, opportunity, intent, preparation, plan, knowledge, identity, or absence of mistake or accident.

Evidence of other crimes may, under the law, be considered probative on the questions enumerated in Rule 404 (b) and therefore be admitted at trial. Under such circumstances the evidence might be termed "legally relevant." Jurors in the present study were asked the degree to which the three charges they judged established a similar motive, intent, a common plan, and the identity of the criminal--these are elements that are legally relevant from one charge to another. Subjects were also asked the extent to which the three charges established a criminal disposition on the part of the defendant--evidence from other crimes is not legally admissible for this purpose. The results indicated that ratings of motive, intent, plan and identity increased significantly as a function of charge similarity, whereas disposition ratings did not. This suggests that jurors' assessments of legal relevance are similar to those of legal professionals. Therefore, if the law wants jurors to use evidence

from other crimes only when it is legally relevant, a possible guideline for joining charges would be to (2) determine that the charges have clear legal relevance for one another.

The law allows joinder of similar charges which would fall into our operationally defined identical and similar charge categories. The charges defined as "identical" clearly met the other crimes requirement, whereas those defined as "similar" fell somewhere in the "gray area" where it was not clear whether they were legally relevant or not. The results indicated that jurors' verdicts were influenced equally in both conditions. However, increased convictions as a result of joined trials that fall into the gray area of legal relevance can be considered more prejudicial, and therefore more likely to be appealed, than convictions for joined charges which are clearly connected. Therefore, an additional guideline for joining charges would be to (3) establish a minimum standard of legal relevance stringent enough to eliminate joinder of charges that fall into the gray area. In fact, for serious crimes the decision on joinder could be made by the grand jury at the time of the indictment, rather than made later at the judge's discretion, since the present results suggest that jurors can assess legal relevance in the manner prescribed by law. Of course, even this solution would not prevent jurors from making inferences about the defendant's criminal character as a result of joinder, which they did in all joined conditions in the present research, and which the law does not want them to do under any circumstances.

The significance of the present results extends beyond the issue of joinder of offenses to other analogous trial situations in which similar processes might operate. As noted in the introduction, there are other forms of joinder in addition to joinder of distinct crimes occurring at different times and places. A defendant may be charged with multiple crimes arising out of a single act (same transaction joinder), and more than one defendant can be tried in a single trial (joinder of defendants). In a related vein, jurors may be allowed to choose from among several verdicts alternatives with respect to a particular crime, and empirical research indicates that the order and seriousness of decision alternatives can affect the verdict reached (Kerr, 1978; McComas & Knoll, 1974; O'Brien et al., 1983; Vidmar, 1972). Although there is little empirical evidence concerning the social inferences processes involved in the above multiple charge situations, the present results suggest that these processes could be meaningfully studied by investigating factors that mediate the effects of the initial phenomena on the final outcome.

The present results have additional significance insofar as they suggest ways in which jurors might become biased as a function of other evidentiary and procedural factors. Central to the Rules of Evidence is the "balancing test," whereby prejudice is weighed against probative value to determine the admissibility of potentially prejudicial evidence. The relevance of evidence of other crimes to the joinder issue has already been discussed. Evidence of prior convictions might produce inferential biases similar to those found in joined trials. As with other crimes evidence, evidence of prior convictions is intended only for purposes of assessing witness credibility, and not to indicate a criminal disposition. However, evidence of prior convictions seems even more likely to be prejudicial than joinder, since the defendant has actually been convicted of previous offenses, rather than just being charged with more than one crime. Therefore, a process model similar to the one proposed for joinder effects might also apply to the effects of prior convictions. If the process is similar, then the law has only partially achieved its purpose with respect to prior convictions. The present results suggest that prior convictions will affect judgments of the defendant's credibility, as the law intends. However, the effects of these inferences are felt primarily through judgments of prosecution evidence, which in turn is used much more than defense evidence in jurors' verdicts. Therefore, the ultimate goal of introducing the evidence, which is to affect verdicts through devaluation of the defendant's testimony, is not realized.

Since the present results suggest that joinder effects are mediated through inferences about the defendant's character, the findings have implications for the prejudicial effects of character evidence, which can also be introduced for the purpose of attacking witness credibility. In fact, introducing damaging character evidence may be sufficient to create criminal inferences of the type induced by joinder, which would then affect assessments of evidence strength and therefore verdicts. Research reviewed by Penrod and Borgida (1983) indicates that character evidence is likely to be prejudicial. Therefore, if character evidence evokes biases similar to those found in joined trials, one way to investigate these processes would be to experimentally manipulate character evidence in order to "activate" various schemas about the defendant.

The effect of judges' instructions with respect to joinder has implications for the use of limiting instructions designed to alleviate the effects of other types of prejudicial evidence. A strong and complete set of instructions had no effect on representative jurors' verdicts, suggesting that traditional legal remedies for alleviating prejudice are not adequate. This finding suggests the need to identify devices other than instructions which might correct the inferential processes that result in biased judgments.

References

- Anderson, N. H. (1965). Primacy effects in personality impression formation using a generalized order effect paradigm. Journal of Personality and Social Psychology, 2, 1-9.
- Anderson, N. H., & Hubert, S. (1963). Effects of concomitant verbal recall on order effects in personality impression formation. Journal of Verbal Learning and Verbal Behavior, 2, 379-391.
- Asch, S. (1946). Forming impressions of personality. Journal of Abnormal and Social Psychology, 48, 258-290.
- Baron, S. A. (1977). A look at the Tennessee multiple offender and the joinder and severance of criminal offenses for trial. Memphis State University Law Review, 7, 457-473.
- Bordens, K. S., & Horowitz, I. A. (1983). Information processing in joined and severed trials. Journal of Applied Social Psychology, 13, 351-370.
- Cleary, E. W. (Ed.). (1972). McCormick's handbook of the law of evidence (2nd Ed.). St. Paul, MN: West.
- Devitt, E. J., & Blackmar, C. B. (1977). Federal Jury Instructions and Practice. St. Paul, MN: West.
- Dreben, E. K., Fiske, S. T., & Hastie, R. (1979). The independence of evaluative and item information: Behavior-based impression formation. Journal of Personality and Social Psychology, 37, 1758-1768.
- Drew v. United States. (1964). Federal Reporter (2d series), 331, 85-94.
- Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. Journal Personality and Social Psychology, 38, 889-906.
- Greene, E., & Loftus, E. (in press). When crimes are joined at trial: Institutionalized prejudice? Law and Human Behavior.
- Hamilton, D. L., & Huffman, L. J. (1971). Generality of impression formation processes for evaluative and nonevaluative judgments. Journal of Personality and Social Psychology, 20, 200-207.
- Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. Journal of Personality and Social Psychology, 37, 25-38.
- Himmelfarb, S. (1975). What do you do when the control group doesn't fit into the factorial design? Psychological Bulletin, 82, 363-368.
- Hodges, B. F. (1974). Effect of valence on relative weighting in impression formation. Journal of Personality and Social Psychology, 30, 378-381.
- Holderman, L. (1977). Joinder of offenses in Nebraska: Judicial discretion v. fair and impartial trial. Nebraska Law Review, 56, 399-409.
- Horowitz, A., Bordens, K. S., & Feldman, M. S. (1980). A comparison of verdicts obtained in severed and joined criminal trials. Journal of Applied Social Psychology, 10, 444-456.
- Kanouse, D. E., & Hanson, L. R. (1972). Negativity in evaluations. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins & B. Weiner (Eds.), Attribution: Perceiving the causes of behavior. Morristown, NJ: General Learning Press.
- Keppel, G. (1982). Design and analysis: A researcher's handbook (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Kerr, N. L. (1978). Severity of prescribed penalty and mock jurors' verdicts. Journal of



- Personality and Social Psychology, 36, 1431-1442.
- Kerr, N. L., & Sawyers, G. W. (1979). Independence of multiple verdicts within a trial by mock jurors. Representative Research in Social Psychology, 10, 16-27.
- Lempert, R. O., & Saltzburg, S. A. (1983). A modern approach to evidence (2nd ed.). St. Paul, MN: West.
- Lilly, G. C. (1978). An introduction to the law of evidence. St. Paul, MN: West.
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. Journal of Personality and Social Psychology, 37, 2098-2109.
- McComas, W. C., & Noll, M. E. (1974). Effects of seriousness of charge and punishment severity on the judgments of simulated jurors. Psychological Record, 24, 545-547.
- Note. (1964-1965). Joint and single trials under Rules 8 and 14 of the Federal Rules of Criminal Procedure. Yale Law Review, 74, 553-566.
- O'Brien, M., Williams, K. D., & Greenberg, J. (1983). Serial position effects of verdict instructions on jury decision making. Presented at the Midwest Psychological Association Meeting, Chicago.
- Penrod, S., & Borgida, E. (in press). Legal rules and lay inference. In L. Wheeler (Ed.), Review of personality and social psychology (Vol. 3). Beverly Hills, CA: Sage.
- Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. Memory and Cognition, 1, 19-40.
- Remington, F. J., & Joseph, A. J. (1961). Charging, convicting, and sentencing the multiple criminal offender. Wisconsin Law Review, 13, 528-565.
- Reyes, R. M., Thompson, W. C., & Bower, G. (1980). Judgmental biases resulting from differing availabilities of arguments. Journal of Personality and Social Psychology, 39, 2-12.
- Riskey, D. R. (1979). Verbal memory processes in impression formation. Journal of Experimental Psychology: Human Learning and Memory, 5, 271-281.
- Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. Journal of Personality and Social Psychology, 32, 880-892.
- Sulin, R. A., & Dooling, D. J. (1974). Intrusion of thematic ideas in retention of prose. Journal of Experimental Psychology, 103, 255-262.
- Tanford, S., & Penrod, S. (1982). Biases in trials involving defendants charged with multiple offenses. Journal of Applied Social Psychology, 12, 453-480.
- Taylor, S. E., & Crocker, J. (1981). Schematic bases of social information processing. In E. T. Higgins, C. P. Herman & M. P. Zanna (Eds.), Social cognition: The Ontario symposium. Hillsdale, NJ: Erlbaum.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 5, 207-232.
- United States v. Foutz. (1976). Federal Reporter (2nd series), 540, 733-740.
- Vidmar, N. (1972). Effects of decision alternatives on the verdicts and social perceptions of simulated jurors. Journal of Personality and Social Psychology, 22, 211-218.

**END**