Experimental Psychology and the Courtroom

Gary L. Wells, Ph.D.

ABSTRACT

Experimental social and cognitive psychology cover a range of topics such as attitudes, attribution, decision making, perception, memory, comprehension, and reasoning. The relevance of these topics to the law is apparent in such domains as jury selection, jury behavior, plea bargaining, eyewitness testimony, sentencing, and police investigation strategies. Despite early attempts to apply experimental psychology to law (circa 1900–1920), it was not until the mid 1970s that a robust law-relevant literature in experimental psychology began to unfold. The experimental psychology/law interface continues to experience some problems in (1) generalizing from specific research experiments to real-world conditions, and (2) identifying solutions rather than just problems.

A natural relationship exists between experimental psychology and the law. Both make assumptions about human behavior and cognitive processes such as memory, comprehension, reasoning, and perception, to name just a few. This is not to say that the relationship between experimental psychology and law is a comfortable one, that there is agreement about human behavior or that any significant amount of knowledge is being transmitted between experimental psychology and law. Indeed, the relationship traditionally has been and continues to be a troubling one in which assumptions about human behavior and thought sometimes diverge and communication between experimental psychology and law is often minimal and frequently muddled.

The purpose of this article is to attempt to describe some of the characteristics of an experimental psychology/law interface. Absent from this paper will be the familiar moanings that the legal system neglects the social sciences and that change in the legal system operates at sub-glacial speed. Absent are the cries that a wide gap exists between experimental psychology’s findings and the practices of the legal system. Whatever the merits of those cries and moans, they distract from the

Gary L. Wells, Ph.D., is Professor of Psychology at the University of Alberta. Correspondence and reprint requests should be addressed to: Dr. Gary L. Wells, Department of Psychology, University of Alberta, Edmonton, Alberta T6G 2E9, Canada.
more basic, albeit less glamorous, problems surrounding the interface of experimental psychology and law. Before discussing these problems, a brief history is given.

**SOME HISTORICAL NOTES**

The experimental psychology/law interface has a broken history in which promising attempts at the interface in the early part of this century experienced a 50-year hiatus (circa 1920–1970). There were numerous experimental psychologists dealing with various aspects of law in the early years. Perhaps the best known of these was Hugo Munsterberg whose book, *On the Witness Stand* (Munsterberg, 1908), dealt with perception and memory as they relate to the courtroom. Using an effective mix of findings in experimental psychology, rhetoric, and incredulity regarding the competence of judges and jurors, Munsterberg delivered strong messages. He argued, for example, that neither judge nor juror would know whether a witness saw in late twilight a woman in a red gown or one in a blue gown whereas the experimental psychologist would know that faint light would make red color sensation disappear while blue color sensation would continue to be perceived. Munsterberg’s examples were numerous and persuasive to many people, although some legal scholars (notably Wigmore, 1909) argued with equal persuasiveness that Munsterberg’s claims were mere hyperbole.

French psychologist Alfred Binet (1900), although rarely identified as an experimental psychologist in the modern sense, conducted numerous studies on suggestion effects in memory, not unlike contemporary studies typified in the work of Loftus (1979). German psychologist Louis W. Stern, perhaps the first experimental psychologist to testify as an expert in court (in the year 1903), published widely on the psychology of eyewitness testimony and edited a serial publication (Stern, 1903–1906) featuring studies by Borst, Cramer, Gottschalk, Kosog, Lipmann, Oppenheim, Rodenwalt, Weindriner, and numerous others. Guy Montrose Whipple communicated much of the European work on experimental psychology and law to American researchers in a series of *Psychological Bulletin* articles.

Unfortunately, the 1920s, 1930s, 1940s, 1950s, and 1960s were dry years for the experimental psychology/law interface relative to the 1900–1920 period. Although there were some significant studies during this period (e.g., Cady, 1924; Burtt, 1931; Snee & Lush, 1941), it was not until the 1970s that experimental psychologists began to approach the law again with vigor. Readers need only scan the reference lists in the articles in this special issue to see that the vast majority of experimental research relating to law has been conducted in the last 15 years. There is speculation but no definitive explanation as to why interest in the law by experimental psychologists waned and has now become robust (e.g., see Sporer, 1981; Wells & Loftus, 1984). Whatever the explanation, it seems that the interface is more solid today than it was in Munsterberg’s time. Presently, there are many more researchers sharing the interest, there is support from new journals devoted to the behavioral science/law interface, the American Psychological Association
has created a new division—Division of Psychology and Law, and psychologist’s research hypotheses are sometimes being taken directly from explicit legal decisions (e.g., Wells & Murray, 1983).

THE EXPERIMENTAL PSYCHOLOGY/LAW INTERFACE

What is meant by reference to an experimental psychology/law interface? What can be contributed to law by experimental psychology that is not already intuitive or that could not be contributed by some other discipline? Does experimental psychology have something to offer now or are there only promises of things to come? I’ll try to answer each of these questions. First, experimental psychology, as broadly defined, covers research specialities in such areas as perception, memory, human judgment, decision making, reasoning, and comprehension. Experimental social psychology covers such special topics as attitude measurement, attitude change, persuasion, and attribution. Although other disciplines may research some of these areas (e.g., political science studies in decision making), the experimental methods used by psychologists provide a qualitatively different approach to the issues.

The experimental method, borrowed from physics, allows precise control over variables and, among other things, establishment of cause-effect relationships. The relevance of these research specialties for law is somewhat obvious: Perception and memory are important cognitive functions in witnesses; human judgment and decision making are involved in all aspects of the justice system including police investigations, plea bargaining, voir dire actions of attorneys, jury deliberations, sentencing, parole decisions, and so on; human comprehension is of profound import at many points, including the timing of judge’s instructions and the charging of the defendant; attitude assessment is crucial to voir dire processes; persuasion principles are of concern to attorneys in their case arguments; attribution processes are thought to underlie juror judgments, and so on.

Indeed, the abstract argument for an experimental psychology/law interface is easy to make. Somewhat more difficult is concrete evidence that theory and data in experimental psychology are of practical use to the legal system. In part this difficulty arises from the fact that the preponderance of research in the experimental psychology/law interface has been of the “problem identification” type rather than the “solution to problems” type. An example helps make this point. Recent research has been directed at the question of so called death-qualified or “scrupled” jurors. Witherspoon v. Illinois (1968) set the stage for this issue by allowing some jurisdictions to eliminate persons from jury eligibility who indicate that they would not find a defendant guilty if conviction would lead to execution. Some research suggests that the process of death qualification results in a conviction-prone jury, presumably because it results in a disproportionate number of conservative, authoritarian persons who in turn are more likely to convict a defendant (e.g., see Bronson, 1970; Goldberg, 1970; Jurow, 1971).

Assume for a moment that the conviction-proneness hypothesis is true (some recent data notwithstanding i.e., Moran & Comfort, in press). Note how this
research is oriented only toward problem identification, not problem solution. It is no solution, for example, to allow such excluded venire members to serve on the jury as that would seem to guarantee that juries would not convicit in cases where the death penalty legally is applicable even when the evidence was beyond reasonable doubt. This is not to say that no solution is possible. Research could be directed at finding ways to select jurors who are both death qualified and are not more conviction prone than would be a representative jury. The point here is that most experimental research directed at the law is the type where problems are identified and not the type where solutions are proposed.

It is true, of course, that the clear identification of a problem must precede the identification of solutions. Thus, I am not espousing a criticism of the research strategies used to date. Instead, I am suggesting a description of the current state of research knowledge. In addition, I suggest that the "gap" between psychology and law, usually described as some kind of disagreement between the scientific knowledge of psychology and the fireside inductions of law (Meehl, 1977), is a distraction from the true gap between psychology and law. The true gap, I propose, is between psychological research on the one hand and applicability (or solution readiness) of the research findings on the other hand.

APPLICATION FRAMEWORKS

The gap between identification of problems and identification of solutions is not attributable solely to the natural order of research wherein the former must precede the latter. In part the gap also is due to the lack of "application frameworks" in experimental areas of psychology. What I mean by application framework is a model that classifies research operations according to their applied properties. This may be clearer by stating an example of an application framework.

Suppose a researcher is interested in doing some experiments on eyewitness testimony. The researcher considers four possible variables to explore, namely lineup structure, question wording, cross race vs. same race identifications, and expectations of the witness prior to seeing the witnessed event. Is there any reason to consider any one of these variables more valuable than any other? Do they have equal a priori status as applied variables? According to one application framework these variables do not have equal status. Wells (1978) proposed a distinction between estimator variables and system variables. Estimator variables are defined as variables that have reliable effects on eyewitness accuracy but cannot be controlled by the criminal justice system. Generally, estimator variables include all variables that precede a witnessed event or occur during the witnessed event. The race of the witness and perpetrator are estimator variables as are the expectations of the witness because neither can be manipulated nor controlled by the criminal justice system in an actual case. At best their influence can only be estimated, hence the term estimator variables. The wording of questions from police interrogators as well as the structure of a lineup, however, are system variables because the criminal justice system has control over these variables. If a researcher finds one lineup structure to
be superior to another, the justice system can use this knowledge to improve eyewitness identification accuracy. If cross-race identifications are shown to be inferior to same-race identifications, however, police cannot actually change that fact.

The system variable/estimator variable distinction is, therefore, a useful application framework. Note how the distinction clarifies a priori the research operations (in this case the independent variables) with regard to a relevant dimension (control). The alternative to this application framework is to classify eyewitness variables according to a dominant theoretical conception of human memory involving acquisition, storage, and retrieval mechanisms. Although the acquisition-storage-retrieval conception is a powerful theoretical tool, it fails to classify variables along dimensions of applied relevance.

Application frameworks may themselves become a focus of controversy. The debate surrounding expert testimony by experimental psychologists is one recent example (Loftus, 1983; McCloskey & Egeth, 1983). Generally, experimental psychologists who conduct a law-relevant experiment argue for applicability of the results by direct or indirect reference to expert testimony. A researcher who finds cross-race eyewitness identifications to be less reliable than same-race identifications, for example, can argue that jurors should be warned against holding much faith in cross-race identifications through the use of expert testimony. Unfortunately, expert testimony has served more as a pseudo application framework than it has as a true, defensible application framework. Consider the cross race vs. same race example. It would seem that an argument favoring expert testimony on this issue requires some evidence that jurors do not already understand that cross-race identifications are less reliable than same-race identifications. Additionally, one may need some assurance that expert testimony on the race issue would not make jurors less believing of cross-race identifications than they should be. The problem here is that mere reference to expert testimony is not itself an application framework. Expert testimony can be part of a justifiable application framework only when relevant application dimensions (e.g., probable effects on jurors) have been specified and tested.

Application frameworks need not be complex, as the system variable/estimator variable distinction illustrates. The main requirement merely is that an application framework clarify how certain research procedures, variables, and outcomes translate into applied settings. The concept of application frameworks has been slow to develop in the experimental psychology/law interface and this makes the research literature a less powerful source of information for judges, police, attorneys, and other justice system personnel than might otherwise be the case. Fortunately, experimental psychologists are becoming more knowledgeable of the legal system and are taking better account of the perspectives and needs of the legal system in selecting research hypotheses, procedures, and variables.

This call for increased development of application frameworks is not meant to downplay the importance of psychological theory. As the next section attests, the generalization of research findings depends a great deal on basic research and theory.
GENERALIZATION ISSUES

Perhaps the most difficult problem in applying experimental psychology to the law is how to generalize from one setting to another setting or from one set of people to another set of people. How can it be said that a phenomenon discovered in a university behavioral laboratory applies to jurors in an actual trial? How can it be argued that judges or parole board members in a particular district are making decisions in a way that is representative of those in some other district? There are no simple answers to this issue, but consider the following two examples.

In their research on how judges in San Diego County, California, sentence adult felons, Ebbesen and Konečni (1981) found that the judges thought or spoke about a large and complex number of different factors prior to reaching their decisions. However, their decisions were made on the basis of only a few factors (severity of crime, prior record, offender’s status between arrest and conviction, and probation officer’s recommendation). The research operations, procedurally and statistically, used by Ebbesen and Konečni were precise, justifiable, and reliable over 1400 cases. But how do we know that these results apply to any jurisdiction other than San Diego County?

The answer to this lies with basic research and theory on human judgment and decision making. It is now a well established fact that people rely on simple heuristics even while claiming to use complex rules combining large numbers of factors (see Kahneman, Slovic, & Tversky, 1982). It doesn’t matter whether one is judging livestock (Phelps & Shanteau, 1978) or judging the likelihood of a flood (Slovic, Fischhoff & Lichtenstein, 1982), the mental operations performed by both novice and expert are based on only a few salient factors. Thus, judges in other jurisdictions might claim that San Diego County is an exception and that their own sentencing decisions are based on numerous complex factors. (Indeed, such a claim was made by the San Diego County judges themselves.) But, basic research and theory on human cognitive processing tells us that San Diego County judges probably are representative of other judges because all humans appear to use simplifying heuristics.

Consider another example where generalization is at issue. Using a classroom laboratory setting with university students as subjects, Loftus (1979) presented films and slide sequences followed by questions, some of which are misleading. For example, subjects may be asked whether the car they saw running the stop sign in the film was green or blue (when the sign was actually a yield sign). Her research shows that subjects exposed to this type of misleading question are more likely to report later that they saw a stop sign than is the case for those who were not given the misleading question. How do we know, however, whether this effect is true only for undergraduate students, only if the witnessed event was part of a film, and so on? In part we can generalize on the basis of basic research and theory on human memory. It is a well-accepted conception of human memory that events that occur after encoding can interfere with retrieval of the original event, that memory is not like a videotape, and so on. In addition, direct research attempts to show the influ-
ence of misleading questions with other subject populations, in nonlaboratory settings and with live events have proven successful (e.g., Loftus, 1979, pp. 61–62).

There are two main ways, therefore, to make the case for generalizing from research settings to real-world settings. First, and perhaps most important, basic research and theory can provide the basis for generalization. Extant research evidence and accepted theory regarding the fact that people use simplifying heuristics, for example, provides powerful generalization potential to Ebbesen and Konečni’s (1981) findings on sentencing strategies of judges. The widely accepted concept of interference in human memory gives Loftus’ (1979) research a generalization capacity not otherwise attainable. In addition, generalization can be achieved in part through direct attempts to conduct studies with subject populations, materials and settings that represent the real-world situation to which generalization is desired.

The ability to generalize from research settings, populations, procedures, and variables to “real world” settings, populations, procedures, and variables is dependent on many factors. Undoubtedly there have been some “shaky generalizations” (Meehl, 1977) made by social scientists that are unwarranted. Equally unwarranted, however, are dismissals of research findings by critics who charge that some aspect of the research operations fails to match exactly the real-world conditions. Conscientious reliance on accepted theory, extant basic research, and specific relevant studies can produce responsible generalizations.

DOMAINS OF APPLICATION

Because experimental areas of psychology cover a broad number of social and cognitive phenomena, there are many places in the legal system where such expertise can be used. Nevertheless, there is a tendency for legal system personnel to think of only two or three application domains. Usually these are eyewitness testimony, jury selection, and jury behavior. It is true that these are the most heavily researched domains, but knowledge in experimental psychology can be applied more broadly to legal issues.

Perhaps the least utilized potential application of experimental psychologists is in the area of criminal investigations. Police detectives who “work up” a case are engaged in an activity that greatly depends on reasoning and problem solving. In recent years cognitive psychologists have discovered numerous flaws in human reasoning and problem solving that apply to almost everyone. Consider a simple example. Researching a burglary, a forensic team discovers a footprint at the rear of the store where the robber exited. Failing to match the print to store employees, it is concluded that the print almost certainly was the perpetrator’s. The forensics team consulted charts indicating that fewer than 15% of those in the population whose height is less than six feet six inches would have a foot of that size whereas 85% of those whose height is between six feet six inches and six feet nine inches would have a foot of that approximate size. The detective uses the forensics team’s data to conclude with 85% certainty that the burglar is six feet six inches to six feet
nine inches tall. Is the detective’s reasoning correct? No. The detective has fallen prey to a classic cognitive error of ignoring prior probabilities or base rates (see Kahneman and Tversky, 1973). Specifically, because more than 95% of the population falls outside of the six feet six inch or taller height range, the actual probability that the perpetrator is six feet six inches to six feet nine inches tall is less than 23%. (The answer to this problem is provided by Bayes’ Theorem.)

One may be prone to argue that this base rate problem is an example of what mathematicians have to offer to the legal system, not experimental psychologists. In fact, however, it is experimental psychologists who have discovered these types of systematic errors in human reasoning, albeit mathematicians are the ones who developed the equations and theorems.

Consider another example of faulty cognitive processes involving judgments of likelihood. Recent award-winning research and theory by Daniel Kahneman and Amos Tversky on human judgment indicates that people judge the likelihood of an event by the ease with which relevant instances are recalled (e.g., see Tversky & Kahneman, 1973). The ease with which something is recalled is termed “availability.” Although a seemingly trivial observation, the availability notion helps predict situations where humans err in judging likelihood. For example, most people have experienced the substantial increase in their subjective likelihood of having an automobile accident following having seen an overturned car by the side of the road. Similarly, if an acquaintance who smokes dies of lung cancer, there is a profound increase in people’s judgments of the likelihood that smoking inevitably will lead to lung cancer. Note in these two examples that the salient event (i.e., the overturned car or the cancer death) had robust effects on subjective probability but the actual probabilities remain constant. The general point is that there are many things that determine the availability (ergo subjective likelihood) of an event that are unrelated to actual probability.

How might the availability principle relate to police investigations? Consider a situation where a detective assigned to several burglary cases obtains an unexpected confession in one of the cases. The confession comes from a man who is 53 years old, employed in a white collar position and has strong community ties. The detective then returns his attention to the other, as yet unsolved, burglaries for which there are three suspects. Two suspects are young (18–25 years old), unemployed and have no community ties. The other suspect is 55 years old, employed as an architect and is a member of the Rotary club. Not surprisingly, the detective now considers the architect to be his prime suspect. Why? Because the prior confession of the 53 year old man made it easy to imagine a related instance. Should the confession in the other case have influenced the detective’s subjective likelihood of guilt in this case? No. It remains true that most crimes of this type are committed by young unemployed males with weak community ties. The confession of the 53 year old man does not change the true odds, it changes only the subjective feelings of the detective. It is no safer to buckle one’s seat belt after seeing an accident than it was before seeing the accident; nor is it any wiser to consider a 55 year old a prime suspect after a 53 year old confesses to a similar crime than it was before the confession.
Human judgments of likelihood are notoriously unreliable and police investigators are as human as anyone else. Fortunately, experimental psychologists have conducted sophisticated research that can predict when such judgments are most likely to be in error (e.g., when recent salient events create availability biases, when considering compound events, when statistical regression is operating, when trying to intuit correlations from observations, when prior probabilities vary from .50, when initial estimates are anchored arbitrarily, see Kahneman, Slovic, & Tversky, 1982). Because police investigations are tasks involving judgments of likelihood under conditions of uncertainty, corrective procedures developed by experimental psychologists can be of great use (see corrective procedures research by Dawes, 1979; Fischhoff, 1982).

I used these police investigation examples because they represent a lesser known potential for the experimental psychology/law interface. The better known areas of interface are represented in depth elsewhere in this issue (e.g., eyewitness testimony, jury selection, jury behavior). Other areas of the experimental psychology/law interface include: civil court cases in which, for example, there may be questions about whether a particular slogan or logo was confusable with another company's trademark; whether or not exposure to pornography has a causal relationship to crime; or the probabilistic distribution of errors in polygraph testing.

Although the domains in which experimental psychology can be applied to law are numerous, they also are limited. Experimental psychologists tend to confine their research to problems that can be studied effectively using the experimental method. As a result, numerous problems (e.g., those involving questions of criminal responsibility) are not congruent with the tools of trade used by experimental psychologists. In addition, experimental psychologists tend to select problems that are limited in scope, concrete, specific, and well defined rather than broad, abstract issues that lend themselves to multiple or ambiguous definitions. As a result, experimental psychologists tend not to tackle the "big problems" such as the causes of crime or penal reform and the attention given to some domains such as jury behavior may be disproportionate to its importance in the overall picture of criminal justice (e.g., the vast majority of criminal cases never reach a jury). Nevertheless, the approach of experimental psychologists in dealing with well defined, specific problems may prove more successful in the long run than are broader scope attempts, which tend to exceed bounded rationality and promote dysfunctional frustrations (see Weick, 1984).

**FINAL COMMENTS**

Early writings aside (e.g., Munsterberg, 1908), the experimental psychology/law interface is nascent. In terms of research volume the interface can be described as less than ten years old. As a result it should come as no surprise that the generalization of research findings to real-world settings is often difficult to accomplish. Nor should it be surprising that the identification of problems has been more of a research focus than has the identification of solutions. But these shortcomings must not only be acknowledged and worked on, they must also be kept in perspective.
These are general problems with the experimental psychology/law interface that do not necessarily apply to a given study or set of studies. Indeed, the other articles in this issue of the Journal are good examples of how the shortcomings discussed in the current article are not equally shared across topics and researchers.

The current article’s focus on problems with the experimental psychology/law interface runs the risk of downplaying the importance of the current literature. In fact, however, the current literatures on human perception, memory, comprehension, reasoning, attitudes, and related experimental topics can no longer be ignored by serious legal scholars and practitioners. Experimental psychology is the only empirical counterpart to the fireside inductions of legal participants on many important questions involving human thought and behavior.

REFERENCES

Copyright of Behavioral Sciences & the Law is the property of John Wiley & Sons Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.