In Praise of Pedantic Eclecticism:
Pitfalls and Opportunities in the Psychology of Judging
(in David E. Klein & Gregory Mitchell, eds.,
THE PSYCHOLOGY OF JUDICIAL DECISION MAKING,
Oxford University Press (forthcoming))

Dan Simon

USC Legal Studies Research Paper No. 08-27

LEGAL STUDIES RESEARCH PAPER SERIES

University of Southern California Law School
Los Angeles, CA 90089-0071

In Praise of Pedantic Eclecticism: Pitfalls and Opportunities in the Psychology of Judging

Dan Simon
USC Gould School of Law
USC Department of Psychology

In David E. Klein & Gregory Mitchell, eds., *The Psychology of Judicial Decision Making*. Oxford University Press (forthcoming)

Though not always acknowledged as such, the subject matter of psychology—namely, human cognition and behavior—has long played a role in important jurisprudential debates. How do judges make decisions (*e.g.*, Cardozo, 1921; Holmes, 1881; Posner, 2008)? Are judicial decisions determined by the law or are they driven by judges' predispositions (*e.g.*, Edwards, 1998; Frank, 1930)? Are assertions of judicial constraint genuine (*e.g.*, Altman, 1990; Kennedy, 1986)? Are judges better fact finders than jurors (*e.g.*, Kalven & Zeisel, 1966)? How do judges weigh the numerous incommensurable and conflicting considerations involved in their decisions (*e.g.*, Cardozo, 1921; Llewellyn, 1960)? Yet, as a discipline, psychology is rarely recognized in the debate. Though hardly a nascent field (see, Frank, 1930; Schroeder, 1918), the psychology of judging remains an under-developed body of research. Hopefully, this collection of essays will pique legal scholars' interest in this line of inquiry.

The reluctance to apply psychological research to the domain of judicial decision making can be understood to be based on at least two concerns.[1] First, unlike other disciplines that have successfully impacted legal discourse—most notably, philosophy and economics—psychological insights tend not to be deduced from overarching concepts or axiomatic-like characterizations of human behavior. The multi-determined nature of human behavior defies parsimony and makes experimental psychology a rather messy field. Psychological findings are made piecemeal, one finding at a time. Properly presented, they are limited to certain classes of circumstances, and are qualified by counter influences.

A second, and more salient, concern stems from the fact that psychological findings are generated mostly in the highly controlled environment of the laboratory, which are starkly different from worldly human practices in real life. The concern, then, is about the *external*

1 In this Chapter, the term psychological research refers to research based on experimental studies.

validity of experimental findings, which speaks to their generalizability to settings outside the laboratory. Psychologists, who are ever so sensitive to situational effects on human behavior (e.g., Lewin, 1935; Ross & Nisbett, 1991), are the first to acknowledge that results obtained in any given study could have come out quite differently under different variations of their experimental design. It is not hard to see why critics question the relevance of findings obtained by testing a relatively small group of lay subjects (say, sixty psychology undergraduates at a mid-western university), performing hypothetical tasks, with limited knowledge, and under specific instructions. Psychological studies have been criticized for the non-representativeness of the subject samples, the artificiality of the experimental setting, the disconnectedness from institutional contexts, the glossing over of individual differences, the lack of appropriate incentives, the inconsequentiality of the tasks, and more (*e.g.*, Konecni & Ebbesen, 1986; Mitchell, 2002; Sears, 1986; Yuille & Cutshall, 1986).

These reservations warrant a healthy skepticism towards even cautious applications of experimental findings to real world situations (see Mitchell, 2003). This concern seems doubly warranted when applying the findings to specialized domains such as judicial decision making. One tempting and seemingly compelling way to overcome this obstacle in the context of judging is to run the experiments on judges.

This Chapter examines some key aspects relating to the experimental design for informing the psychology of judging. In particular, it addresses the prevailing concern with external validity. Two central points will be made here. First, it is important to acknowledge that all types of validity are intricately intertwined. Attempting to fix one aspect of validity can entail a relaxation in other aspects of the experimental design. In other words, compensating for one weakness can be overwhelmed by greater compromises elsewhere, thus resulting in a net loss in validity. To prevent this, one needs to be pedantic about the design of the entire study.

At the same time, the concern with external validity should not be exaggerated. The discrepancy between the experimental environment and real world settings does not automatically bar all applications of findings from the former to the latter. It does, however, require cautious work and oftentimes also more data. When experimental findings meet the rigorous demands of external validity, they can be applied safely to real world domains, including specialized ones. This possibility opens up the field of judging to a wide range of methodological approaches and thus offers the benefit of insights originating from diverse perspectives. To demonstrate the possibilities in this regard, the discussion will center on the application of a particular body of research—coherence based reasoning—to judging.

I. In Praise of Pedantry

Fred Schauer's discerning and provocative essay *Is There a Psychology of Judging?* (this volume) provides a good basis to examine some important aspects of experimental design. The essay pursues the idea of a unique and genuine psychology of judging. Schauer starts with an appealing proposition that a psychology of judging ought to be what it is claims to be—an examination *of judging*, as performed *by judges*. Explanations that rely on cognitive processes that are performed also by non-judges and in non-judicial domains might make for an interesting application of psychological research, but they do not "constitute a discrete area of inquiry." Merely applying basic findings to the work of dentists and plumbers is "a long way from saying that there is a psychology *of* dentistry or plumbing." (p. 6).

Schauer proposes to distinguish decision making by judges from decision making by other people, even lawyers, for the same reason that it would be a mistake to equate the mathematical reasoning of a Harvard mathematics professor with that of ordinary folks balancing their checkbooks. He contends also that the field ought to concentrate on decision making in domains that are unique to judging, such as finding, interpreting, and making law. Even if auto mechanics and psychiatrists approach the finances of their small business in similar fashions, it does not follow that they do the same when they fix cars or provide psychiatric care. The underlying intuition is that judicial training, acculturation, experience and role may lead to deep cognitive effects that "generate process- and not just content-based differences between the cognitive mechanisms of judges and those of non-judge humanity."(p. 5).

Schauer's proposal is constructive in that it actually suggests an experimental project to test his proposition. The central hypothesis focuses on the role of *second-order reasoning* in judicial decision making. Second-order reasoning stands for higher order decision rules that are supposed to trump reasons that would otherwise provide a sufficient basis for a decision. Second-order reasoning is deemed central to the judicial function in that it forces judges to abide by a hierarchy of reasons, and specifically, to yield to higher order considerations even when they feel that doing so leads to suboptimal or unwise outcomes for the case at hand. Schauer intuits that judges engage in second-order reasoning differently, more effectively, and with greater fidelity than non-judges. Specifically, the studies would test whether judges are more inclined to follow a precedent or rule with which they disagree, and to limit their attention to a narrow range of permissible sources even when other sources lead to what they consider to be superior results. The studies are designed to compare judicial decisions with decisions made by lawyers and law students, and to compare decisions made by different classes of judges, such as elected and appointed judges, trial and appellate judges, and judges with different personal backgrounds. At

bottom, Schauer's endeavor is driven by the aspiration to bridge two strands of the external validity gap.[2] Undoubtedly cognizant of the complexity of the issues involved, Schauer emphasizes the tentative and exploratory nature of his proposal, characterizing it as a set of hypotheses, an invitation to consider a possible line of research. This Chapter follows up on that invitation.

As a general matter, Schauer should be lauded for promoting experimental research of judicial decision making, and for the attempt to solve the nagging problem emanating from the external validity gap. Still, before expending the limited experimental resources (there is no abundance of judge-subjects), one ought to think hard about the proposed studies to ensure that they are capable of providing the insights they aim to discern. More generally, Schauer's proposal offers a precious opportunity to think critically and constructively about how experimental psychology can be best employed in examining legal questions.

To evaluate the issues at hand, a brief methodological detour would be helpful.[3] It is imperative to acknowledge that external validity does exhaust the validity challenges facing experimental research. The threshold criterion of the validity of any psychological experiment is that it be *internally valid*, which stands for the degree to which the experimental treatment explains the observed results. Studies are said to be internally valid when they demonstrate that the variables that were set up or manipulated by the experimenter ("independent variables") were indeed the cause of variation in the focal point of the hypothesis (the "dependent variables"). Internal validity rests heavily on the researcher's ability to design the study so that it provides maximal control over the experimental environment. Control is necessary for the restraining of unintended factors that might affect the dependent variable.[4] There seems no reason to believe

---

[2] Note that the proposal bridges two ingredients of external validity, namely, the population of participants and the type of decision tested. It does not bridge other aspects, such as realism, consequentialism, and the like.

[3] For useful discussions on methodological aspects of experimental psychology, see Aronson, Wilson, & Brewer, 1998; Mitchell & Jolley, 2007).

[4] In the following examples, imagine a jury simulation that is intended to test the effect of gruesome photographs from the crime scene on verdicts in the murder trial. The hypothesis in this example is that the exposure of the fact finders to the photographs will result in an arousal of negative emotion, which will increase their tendency to convict.

A typical violation of internal validity is the failure to control for alternative potential causal effects, also known as confounds. In this example, assume a comparison between one group that receives a case that contains gory photographs and another group that receives a different case that does not contain photographs. Given the discrepant stimuli, different rates of conviction (if obtained) could not be explained as driven necessarily by the exposure to the gory photographs. They could readily be explained by the fact that the two groups received different cases. The evidence in the former case might have been more incriminating.

that Schauer's proposal would be lacking in internal validity. A finding of differences in decisions made by judges and non-judges in a well controlled environment could be fairly understood to be related to differences in the manner in which the two groups make decisions.

Detecting differences, however, can be a far cry from understanding their underlying mechanisms at work. To bridge this explanatory gap, studies need to be shown to have *construct validity*, which stands for the degree to which one can correctly identify and explain the operative psychological constructs. This important and somewhat neglected facet of validity stands for the degree to which studies accurately operationalize their theoretical constructs, which is essential for the correct explanation of the relationship between the experimental treatment and the observed phenomena. Absent this validity, one cannot draw reliable inferences from the observed results.[5] Construct validity is needed, first, to overcome the fact that human cognition is notoriously multi-determined. Thus, to reliably interpret experimental findings, studies need to be able to isolate the hypothesized mechanism and rule out alternative explanations. Studies designed to have this capability are said to have *discriminant validity*. Second, the experimental design needs also to overcome the fact that psychological phenomena are generally not directly observable. To attain construct validity, the design needs to be able to identify the psychological mechanisms that drive the observations. Studies designed to be capable of identifying the correct construct are said to have *content validity*. Content validity is enhanced also by the extent to which the construct fits into a broader underlying theory. Theories are particularly useful when they are capable of explaining other related constructs.[6]

It should be acknowledged that the various forms of validity are often in tension with one another. Notably, the high levels of experimental control that are essential for ensuring internal and construct validity, cut against the generalizability of the findings. Likewise, tests that are designed to have a large degree of external validity, such as archival research and field studies, typically lack important aspects of control, most notably, random assignment of subjects to the various conditions. Herein lies the tension in Schauer's proposal. Recall that the proposed research seeks to discern differences in the underlying cognitive processes between judges and

[5] In the abovementioned example of the study testing the effect of gruesome photographs on jury verdicts, observing the heightened conviction rates in the presence of gruesome evidence does not in itself provide a satisfactory understanding of the effect. While it is possible that the effect was driven by the arousal of emotion (the hypothesized cause), it is also possible that it was driven by the fact that the photographs contained incriminating information that tied the defendant to the crime.

[6] The content validity in the study testing the effect of gruesome photographs could be enhanced by showing that the finding can be explained by a theory that explains other effects of emotional arousal, such as in judgments of tort liability (Lerner, Goldberg, & Tetlock, 1998) and aggressive behavior (Bushman, 1995).

non-judges. This aspiration can be problematic, especially since such deep constructs are most difficult to study. Even if the proposed studies yielded the expected results—namely, that judges' decisions were better aligned with second-order principles than decisions by non-judges—the underlying operative mechanism would remain unknown. A finding that judges show greater deference to a precedent might or might be indicative of superior second order reasoning. It could also be due to the fact that lay people lack a sense of the judicial conventions and practices pertaining to the assessment of case similarity, the distinguishing of precedents, and more. By the same token, a finding that judges display a heightened respect for jurisdictional limitations need not stand for the proposition that judges engage in different cognitive processes. It could be readily interpreted as standing for the proposition that lay people lack familiarity with the constitutional principles that underlie the rules of federalism,[7] lack the nuanced knowledge of how to weight them against competing considerations, and the like.[8] Lay people are likely to be unfamiliar with the profession's response for disregarding the hierarchy of reasons, such as reputational damage and being overturned. They are also not likely to be familiar with the dilemmas surrounding the personal commitment to the judicial role.[9]

One alternative, and plausible, explanation for hypothesized findings of judicial superiority is that judges are experts at judging, whereas lay people are novices. Psychological research on expertise shows that experts perform differently—in certain ways, better—than novices. (Chi, 2006; Ericsson & Ward, 2007). The possibility that judges have a better grasp of the conventions and practices of the judicial role, and are better in weighting and trading off the competing rules is consistent with the advantages of expertise.[10] Moreover, the determinants of judicial uniqueness noted by Schauer—namely, selection, training, and experience—closely resemble the

[7] Schauer actually discusses the potential problem with lay understandings of the concepts of federalism and jurisdictions (fn. 29), but does not treat it as a potential methodological problem with the study.

[8] For an insightful demonstration of evaluating and weighting judicial goals see Robbennolt, MacCoun & Darley (this volume).

[9] A tempting solution to these difficulties would be to provide lay subjects with special training about these matters in preparation for the experiment. To the extent that the training would be feasible and effective, it could jeopardize the study's internal validity. For example, lay subjects might interpret the instruction as a cue for a desired decision.

[10] It must be noted, however, that expertise is also characterized by sub-par functioning. Experts tend to display over-confidence, fail to notice details, and are less agile in adapting to change. Expertise is also no guarantee against the effects of bias (Chi, 2006; Koehler, Brenner, & Griffin, 2002). It is also important to note that expertise are typically narrow in scope. While experts perform differently on one type of task, the do not necessarily perform differently on adjacent tasks, even within the domain of their expertise.

factors that have been found to develop expertise (Feltovich, Prietula & Ericsson, 2006).[11] While Schauer explicitly rejects the explanation based on expertise,[12] the proposed studies do not seem to provide a way to discriminate between the two explanations.[13]

Construct validity requires also correct identification of the operative construct. Underlying the proposed hypothesis is the belief that judges are considerably superior to lay people in following second-order rules. Schauer is skeptical of lay people's ability to abide by second-order rules, and goes so far as to suggest that they might not know how to engage in this form of reasoning at all.[14] The explanation offered for lay people's low capabilities is that they have little experience making decisions of this kind.

Yet, there is reason to doubt that second-order reasoning is rare or undeveloped in non-judicial settings. People engage in some form of second-order reasoning every time they walk by a store window displaying a coveted item that exceeds their budget. Dentists oblige their patients' requests even when they recommend a different course of action. Plumbers heed the building code, the architect's plans, or the client's preferences, even when a different solution seems to them to be more effective, sensible, or aesthetic. Second order reasoning plays a central role in the world of other professionals. For example, human resource personnel are habitually confronted with considerations such as employment laws, company policies, maintaining consistency with prior cases, and setting an example for future ones. Other examples abound.[15]

While it is clear that judges do indeed engage in a considerable amount of second-order reasoning, there is reason to suspect that their fidelity is less than perfect. A substantial and growing body of quantitative analyses of judicial behavior shows that judicial decisions are systematically biased by judges' personal attitudes (Klein, 2002; Segal & Spaeth, 1993, 2002),

---

[11] The expertise explanation would also explain the examples Schauer uses to illustrate the uniqueness of judging: the difference in the mathematical skills of a Harvard professor of mathematics and lay people, and differences between a psychiatrist and a mechanic when it comes to providing psychiatric care or to fixing cars.

[12] See section IV. The Question of expertise.

[13] The proposed comparison of judges to non-judges parallels what researchers in the field of expertise call the *relative* line of inquiry. A different type of research takes an *objective* approach, namely, focusing on how well the experts' performance stacks up against predetermined measures of excellence (Chi, 2006). By adopting the former approach, the proposed studies will, at best, indicate that judges are somewhat better at some aspects of judicial decisions than lay people. But that finding falls short of answering what is arguably the crucial questions: whether judges are good enough relative to some objective expectation, and whether they fulfill their constitutional role satisfactorily.

[14] Referring to second order reasoning, Schauer characterizes judges as "people who know how to x" and contrasts them with lay people "who do not know how to x at all" (p. 20)

[15] It should be noted that the article by Sunstein and Ullman-Margalit (1999) cited by Schauer pertains to second-order decisions made in non judicial contexts.

the composition of panels (Cross & Tiller, 1998), personal prestige (Klein & Morrisroe, 1999), and more.[16] These observations suggest that judges either do not fully adhere to the second-order reasons, or that they interpret and apply those reasons in a biased manner that effectively undermines the rules' normative import.[17] It should be noted that one experiment that tested judicial adherence to second-order rules revealed a rather lackluster level of compliance. In this study, the judges' decisions tended to be influenced by information which they themselves ruled to be inadmissible (Wistrich et al., 2005).[18]

Moreover, it is worth noting that judicial experience might actually make judges feel less bound by second-order rules than one might otherwise believe. To a large extent, judging requires deciding not only which rules and precedents ought to be followed, but also which ones ought *not* to be followed. In most complex cases, judges are faced with multitudes of powerful reasons (Schauer, 1988), some of which are likely to pose a conflict between two or more second-order rules. When such a conflict occurs, the judicial dilemma requires a determination as to which second-order principle ought to be followed and which one ought to be rejected. In such situations, the fidelity to second-order rules necessarily entails a rejection of (other) second-order rules, as indicated in Karl Llewellyn's conception of "dueling canons."[19] It appears, then, that finding acceptable ways to ignore, dismiss, or interpret away second-order rules is yet another facet of judicial expertise.

In Praise of Eclecticism

In the overview of the extant psychology of judging, Schauer surveys two existing types of research. The one body of research tests real judges in a variety of experimental tasks that have previously been tested with lay subjects. These studies tend to demonstrate that judges are indeed prone to most of the same biases and errors as people in general (Guthrie, Rachlinski, & Wistrich, 2001; Wistrich, Guthrie, & Rachlinski, 2005). While Schauer is appreciative of this

[16] This body of research is a good example of the productive use of cross disciplinary research. While the constructs underlying these findings are psychological, much of this research has been performed by political scientists, in non-experimental settings.

[17] A recent experiment conducted with law students demonstrated how second order rules are distorted by the ideology of the participants. Furgeson, J. R., Babcock,. L., & Shane, P. M. (in press). Behind the Mask of Method: Political Orientation and Constitutional Interpretive Preferences. *Law & Human Behavior*.

[18] True, the tasks involved in these studies were more akin to fact-finding and rendering of verdicts, but there is no obvious reason to believe that the performance would be better in appellate-like decision making.

[19] Llewellyn, K. N. (1950). Remarks on the Theory of Appellate Decision and the Rules or Canons About How Statutes Are to Be Construed. *Vanderbilt Law Review, 3*, 395.

research, he considers it to be of lesser interest to his project because they focus on the fact-finding and verdict-rendering dimensions of the judicial role, which are performed also by lay jurors. Because the studied tasks are not performed exclusively by judges, this research fails to meet the proposed standard of uniqueness.

Schauer is far more critical of the literature that applies basic psychology to judging.[20] This research fails on both dimensions of the proposed uniqueness in that it uses ordinary people as subjects in the performance non-judicial decisions. In other words, the external validity gap is seen to render this work invalid. Schauer briskly dismisses the application of this work for being axiomatic, unargued, and unresearched.[21]

This brings us back to the critiques of the external validity of basic psychological research. With so many degrees of freedom separating the laboratory environment from real world contexts of human action, one might wonder how experimental research can ever be deemed to bear any practical relevance to real life. Yet, an array of experimental findings are notably present outside the confines of the laboratory. As it turns out, the gathering of intelligence by the CIA in preparation for the Iraq War[22] bears an eerie similarity to various forms of biased reasoning generated in the laboratory (*e.g.*, Frey, 1986; Klayman, 1995; Kunda, 1990; Nickerson, 1998). The behavior of nations and ethnic groups entangled in real conflicts corresponds closely to the behavior of arbitrarily formed groups in the laboratory (*e.g.*, Brewer, 1979). Marketers and political consultants routinely exploit human judgment processes gleaned from the laboratory (*e.g.*, Rozin, & Royzman, 2001; Nisbett & Wilson, 1977). Prejudicial behavior by employers (Bertrand & Mullainathan, 2004) mirrors behavior observed in experimental settings (*e.g.*, Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997). Likewise, jury decisions to send convicted inmates to their death (Eisenberg, Garvey, & Wells, 2001) appear to be influenced by the same attitudes that affect mock jurors' decisions in hypothetical cases (Thompson, Cowan, Ellsworth, & Harrington, 1984). A meta-analysis of experimentation is social psychology has

---

[20] *Basic psychology* focuses on phenomena that are relatively generalizable across people, situations, tasks, and contexts. Basic psychologists research phenomena like memory, reasoning, and persuasion. *Applied psychology* research focuses on testing psychological phenomena as they are performed in particular contexts. For example, applied psychology tests memory performance in the context witness testimony, reasoning in police investigations, and persuasion in political campaigning.

[21] Specifically, the essay states that the premise underlying this literature is an "undocumented and unargued premise" that lurks in the background (p. 2). It is described as something that researchers merely assume (p.2); an "unexpressed" and "typically unresearched" outlook (p.14); and a viewpoint that is taken as "axiomatic" but "hardly based on systematic research directed precisely at that question" (p. 20).

[22] See Senate Report 108-301. Report of the Select Committee on Intelligence on the U.S. Intelligence Community's Prewar Intelligence Assessments on Iraq. July 7, 2004 (http://intelligence.senate.gov/pub108thcongress.html).

shown a rather strong correspondence between findings obtained in the laboratory and in the field.[23]

Thus, it seems that applying psychological research to capture real world phenomena is a complicated fete, which can be neither ascertained nor controverted offhandedly. As discussed below, subject to a careful and methodical examination, psychological research can be safely applied to some real life situations, but not to others. The domain of judicial decision making presents a most difficult application, and thus requires a heightened level of scrutiny.

For the research to be deemed useful outside the confines of the controlled laboratory setting, it must first be shown that the observed phenomenon was not an artifact of the specific experiment. One important way to allay this concern is by replicating the finding under similar and different experimental settings. Validity increases when the same finding is observed using different populations of subjects, stimulus materials, instructions, and tasks. It increases also if replications are conducted in different laboratories. Validity is further increased by the robustness of the finding, that is, its recurrence under various manipulations, across wide ranges of values, and in the presence of counter forces.[24]

Still, external validity does not guarantee that the findings apply equally to every domain of human behavior. To be deemed applicable to a particular real world practice, the finding must not be trumped, weakened, or distorted by particular features of the domain, as these were not present in the experimental setting and their potential influence on the finding is unknown to the experimenter. This last link in the applicability chain has been coined *contextual attentiveness* (Arlen & Talley, in press). Contextual attentiveness stands for the degree to which the experimental findings map onto the rich context of the real world, and thus can be deemed as an offshoot of external validity. A threshold criterion for establishing contextual attentiveness is the facial similarity between the behavior captured by the laboratory finding and the behavior

---

[23] The correlate coefficient of the findings was found to be about 0.73. Anderson, A. A., Lindsay, J. L., & Bushman, B. J. (1999). Research in the psychological laboratory: Truth or triviality? *Current Directions in Psychological Science, 8*, 3-9. The similarity between laboratory findings and field findings does not ensure that the findings are applicable to real world applications, but it does allay some of the concerns about the artificiality of the laboratory setting.

Another aspect of validity is ecological validity, which captures the similarity between the experimental setting and the real life domain. It is undeniable that the bulk of basic psychological research has little ecological validity with respect to the practice of judging, or to any other real world practice for that matter. Still, external validity ought not be confused with ecological validity. High ecological validity naturally increases external validity, but the latter is not dependant on the former.

[24] To follow the abovementioned example, the external validity of the study of the effect of gruesome photographs will be increased by showing that similar , one would need to show that the same effect is observed in other experimental variations, such as using different photographs, media, populations, factual patterns, judicial instructions, and the like.

observed in real life. It is enhanced by the extent to which the psychological theory that underlies the phenomenon is deemed germane to the domain, particularly in the absence of competing theories. Strong support can be derived from corroborative evidence derived from sources other than the experiment itself, such as when archival or field data reveal behaviors that are consistent with the experimentally observed phenomenon. Finally, one might also look, with caution, to self reports by people working in the domain, particularly to those who are considered to be introspective.

It is important to note that there is no established gold standard for determining external validity. None of the abovementioned features can carry the day by itself, nor can any body of research be expected to fit them all. The guiding principle is convergent validity: the more of the noted features that converge towards validity, the more reliable the conclusion.

With these methodological guidelines in mind, we can return to assess Schauer's objection to the application of basic psychological findings to judging. It is beyond the scope of this Chapter to examine the merits of this charge with respect to the various bodies of work Schauer mentions (in which he candidly includes his own previous work. See Arlen, 1998; Hanson & Yosifon, 2004; and Schauer, 2006a, 2006b). It is, however, feasible to assess the objection as it pertains to the applicability of one of the bodies of work, *coherence based reasoning*.[25] To do so, it would be helpful first to review this line of research.

Coherence based reasoning examines a particular characteristic of opinions rendered by American appellate courts. Specifically, it seeks to explain the enigmatic, yet prevalent mode of reasoning in judicial opinions. Even the casual reader of judicial opinions is likely familiar with the experience of being strongly persuaded by an opinion, with all of its components converging to provide overwhelming support for the outcome. The facts of the case, authoritative texts, governing precedents, legal principles, public policies, as well as sheer logic and common sense, all come together in a coherent whole to make for the inevitable and undeniably correct result. The sense of correctness is bolstered by the dearth or absence of arguments to the contrary. By the culmination of the opinion, one might wonder how the decision could be considered to have been anything but obvious in the first place. This sense of obviousness, however, quickly dissolves upon turning to the opinion of the dissenting judges. Dissenting opinions too tend to be supported by a slew of authoritative texts, and they too are strongly coherent, inevitable, and persuasive in defending the opposite conclusion. Thus, while the opinions are exceedingly

[25] Schauer criticizes the application of coherence based reasoning to judging (Simon, 1998, 2002, 2004), and also singles out some of the underlying empirical work, namely Holyoak & Simon (1999) and Simon, Krawczyk, & Holyoak (2004).

coherent internally, they are radically inconsistent with a slew of seemingly plausible arguments contained in the opposing opinion. The divergence between opposing judges' views of a case can lead them to stake remarkable positions. For example, when interpreting statutes, it is not unusual for judges to deny outright that the there is any ambiguity in the statutory text, notwithstanding the fact that similarly positioned judges read the text to mean the very opposite.[26] This mode of reasoning is apparent in almost every appellate case, and is most pronounced in cases that contain dissents.

As a matter of legal policy, this phenomenon has mixed effects that exceed the scope of this Chapter.[27] As a theoretical matter, it offers an opportunity to peer into the judicial decision making process and better understand the extent to which it is constrained by the law, as judges routinely claim it to be. If indeed the coalitions of reasons mustered by judges accurately represent the state of the law, one ought to infer that judicial decision making is indeed tightly constrained by the law, and more importantly, that the law is determinative of single right answers. If, however, we find an alternative explanation for this mode of reasoning, one might call into question the professed constraint and, by implication, the binding nature of the legal materials.

There are good reasons to suspect that the legal materials are considerably less determinative than they are portrayed in judicial opinions. For one, most cases decided by appellate courts are truly complicated and difficult, as they contain sound arguments supporting each side of the issue (Schauer, 1988). The suspicion intensifies once opinions are dissected and stripped down to their constitutive arguments. Relatively short US Supreme Court opinions typically contain a handful of core issues, each of which is supported by an array of arguments, which can easily total fifty arguments or more. The most remarkable feature of the opinions is that virtually every single argument supports the corresponding conclusion and contradicts the opposite opinion.[28] A notable feature of the opinions is that virtually every one of the dozens of arguments supports the corresponding conclusion and contradicts the opposite one. This perfect alignment of reasons is plainly implausible. Assuming that the soundness of the opposing arguments are roughly similar, the mathematical probability that each of the fifty or more arguments line up perfectly is

[26] See Simon (1998), p. 71.

[27] While it can be said to increase the acceptability of the opinions, this mode of argument can be deemed detrimental to adjudication in that it obscures the complexity of the issues involved, blunts the thoroughness of judicial analysis, and unduly devoids the validity of the losing side. For a discussion, see Simon (1998), pp. 129-134.

[28] There are of course more than one way to break down a case and to enumerate its components. Alternative ways will always be possible, though the differences should not bear on the substantive conclusions of the analysis.

astronomically minute.[29] This observation suggests that the overall conclusion of the case plays a role in determining which arguments are endorsed and which are rejected. While judicial decisions are most likely affected by their underlying reasons, there appears also to be an effect in the opposite direction, by which decisions affect the reasons that are claimed to support them. This calls into question the avowed uni-directional relationship between reasons and conclusions, namely that the former should affect the latter, but not the other way round. Judicial reasoning, it would seem, operates bi-directionally, from reasons to decisions, and back in reverse.[30]

This mode of judicial reasoning cannot be explained by the prevailing theories of decision making (e.g., von Neumann, & Morgenstern, 1944; Edwards & Newman, 1982), which seem ill suited to handle complex decisions of the kind that judges face—where the variables are numerous, conflicting, ambiguous, and incommensurable. The phenomenon could, however, be consistent with a body of psychological research that shows that certain cognitive tasks are driven by coherence-maximizing processes.[31] This line of research follows the tradition of cognitive consistency theories—notably balance theory (Abelson & Rosenberg, 1958; Heider, 1946, 1958) and cognitive dissonance theory (Festinger, 1957)—which are based, in turn, on Gestalt psychology (Wertheimer, 1923/1938). Cognitive consistency theories were animated by principles of structural dynamics, that posits that relevant cognitive processes are determined holistically, rather than elementally. The holistic structural properties are deemed to be dynamic,

---

[29] This form of inquiry was first demonstrated in the analysis of *Ratzlaf v. United States*, 510 U.S. 135 (1994), a relatively mundane case chosen almost at random. See Simon (1998, pp. 61-72, and 73-102 *passim*). For another example, analyzing *Rogers v. Tennessee*, 532 U.S. 451 (2001), see Simon (2002). The *Bush v. Gore* case provides a more familiar manifestation of the phenomenon (draft on file with author).

[30] The bi-directional nature of the process offers an alternative explanation to the Realist charge that judges merely stack their opinions with whichever arguments support their preferred choices. See Simon (1998), pp. 134-37.

An alternative to the bi-directional explanation is that the strong alignment of arguments is a product of "padding" of opinions (Posner, 1995). The concept of padding implies that not all reasons play the same role in the decision, as some are deemed to actually exert power on the decision maker, whereas others merely serve as *ex post* justifications. This seems true, though it is doubtful that padding could fully account for the observed coherence (see Simon, 1998, pp. 35-36). . To illustrate, even if all but the handful of core issues served as mere justifications, one would still need to explain how the handful of core arguments lined up to cohere with the decision. For example, in the Ratzlaf case there were six core issues. Assuming that they were about equally plausible, the probability of all six lining up to support the respective conclusion is one in sixty-four. A precise assessment of likelihoods would depend on the degree to which the issues and arguments are independent of one another.

[31] The processes include vision (McClelland & Rumelhart, 1981), social reasoning (Read & Miller, 1998; Read, Vanman, & Miller, 1997), analogical reasoning (Holyoak & Thagard, 1989; Spellman & Holyoak, 1992), relational inference (Hummel & Holyoak, 2003), text comprehension (Kintsch, 1988), and more. These strands of research are based on a connectionist architecture, and are resolved via Parallel Constraint Satisfaction Mechanisms (Holyoak & Thagard, 1989; Read, Vanman, & Miller, 1997).

so that interrelatedness of the elements generates forces that determine the configuration of the structure. Some things "go together," that is, they are related by cohesive forces, whereas other things tend to disperse. These forces determine the stability of the structure and drive it towards a state of equilibrium, or Gestalt. Perhaps most importantly, the dynamic forces that occur at the structural level involve changes, or "reconstructions" of the cognitive elements (Rosenberg and Abelson, 1960), that is, by "distorting the state of affairs" (Asch, 1940, 454)

A series of experiments conducted by Keith Holyoak, Stephen Read and myself was designed to explore the possibility that a theory of coherence-maximization would apply also to the domain of decision making, in particular, to complex decisions like the ones judges make. To understand the concept of coherence based reasoning, it would be helpful to briefly describe the common design of the studies. In the first study, participants were first presented with a pretest that contained a number of apparently unrelated vignettes, that were followed by a statement or two that could be inferred from them. Participants were asked to rate their agreement with the total of twelve such inferences. Some vignettes involved factual judgments, and some involved more abstract issues such as analogies and issues of social policy. In a separate phase of the experiment, participants were asked to play the role of a young judge, assigned to decide a civil case in which Quest, a software company, filed a libel lawsuit against one of its shareholders, Jack Smith. The company alleged that Smith posted a libelous statement about the company that caused it to go bankrupt. The case revolved around six core points of dispute, with each party making an argument on each of the six issues. The key feature in the design was that the case was constructed from all of the vignettes that were used in the first phase of the experiment, and the litigants' arguments were virtually identical to the inferences that followed those vignettes. Participants were asked to render a verdict and to report their confidence in the verdict. They were also asked to rate their agreement with the twelve arguments made by the parties.

The central finding in these studies is the derived from the comparison between the ratings on the virtually identical items obtained at the two different phases of the study. Consistent with the prediction from coherence-based reasoning, participants were found to have made confident decisions despite the complexity and difficulty of the case. In comparison to the moderate and noisy ratings given in the first phase of the study, at the point of decision, the ratings manifested polarized states of coherence: participants who voted for Smith reported strong agreement with the arguments that supported his case and disagreement with the arguments that supported Quest's position, while opposite ratings were reported by participants who decided for Quest. In other words, participants changed their view of the issues over the course the decision from a state of cacophony to one of steadfast coherence (Holyoak & Simon, 1999).

The findings from this and other studies support the conclusion that complex decisions are driven by coherence-maximizing processes, by which people's perceptions of the tasks shift during the decision making process from an initial state of conflict to a final state of steadfast coherence. By the culmination of the process, the arguments involved in the task have shifted and spread apart into two or more coherent subsets, one providing overwhelming support for the emerging decision, and the other providing depressed support for the rejected decision choice. This spreading apart enables a comfortable and confident decision. It follows, then, that the state of coherence is not a property of the arguments themselves, but rather an artificial cognitive state imposed by the decision making process. This process is understood to be adaptive in that it enables people to make decisions and conduct their affairs even in the face of stifling complexity. It must, however, be acknowledged that it does so by means of distorting the factors involved in the decision.

Though inspired by judicial decision making, these experimental findings were borne by a basic-psychological research program, and thus cannot automatically be said to pertain to the domain of judging. To support the claim of applicability, the research must be shown to withstand the test of external validity.[32] Coherence effects have been observed repeatedly, without fail, in a wide range of studies (Simon, Krawczyk, & Holyoak, 2004; Simon, Krawczyk, Bleicher, & Holyoak, 2008; Simon, Pham, Le, & Holyoak, 2001; Simon, Snow, & Read, 2004. For reviews, see Simon, 2002, 2004). The studies have been replicated by other researchers in the United States and in Europe. One study replicated the findings using the same materials as used in the original research (Glöckner, 2007), while others tested a variety of different tasks including financial auditing decisions (Lundberg, 2004, 2007; Phillips, 2002), judgment and decision making (Glöckner, Betsch, & Schindler, under review), legal-economic behavior (Landeo, under review), and evidence evaluation (Lundberg, 2004). The robustness of coherence effects is manifested by the fact that it has been manipulated in numerous ways, always yielding the hypothesized results.[33] Across the various studies, the subjects have role-played young judges, jurors, arbitrators, auditors, and job applicants, while other studies involved no role-playing at all. In all, the studies have been tested with some 3,000 subjects, including undergraduate students, graduate business students, a general sample of Internet users, and

[32] The studies' internal validity and construct validity are beyond the scope of this Chapter. The reader is invited to accept these issues at face value, given the extensive peer review that the studies have undergone.

[33] Manipulations can be found in Holyoak & Simon (1999), studies 1, 3; Simon et al. (2001), studies 1, 2, 3; Simon et al. (2004a), Study 1; Simon et al. (2004b,) studies 1, 2, 3, 4.

experienced professional auditors.[34] The studies have tested a wide range of reasoning tasks, including high level inferences, analogies, rule application, policy decisions, factual judgments, social judgments, probability assessments, and personal preferences. The studies have also been applied to other domains by scholars in the respective fields, including medical decision making (Levy & Hershey, 2008), narrative studies (Foroni & Mayr (2005), and philosophy (Harman & Kulkarni, 2006),

Recall that to apply a body of research to the real world, it must be shown also that the research is contextually attentive to the particular domain. Support for the applicability of coherence based reasoning to judging is derived from the close resemblance between the coherence that is present in judicial opinions and the coherence that is found in the various reasoning tasks performed in the laboratory: very complex decision tasks are resolved successfully, resulting in lopsided and coherent sets of arguments, accompanied by high levels of confidence.[35] The theory underlying the laboratory results—namely, that the cognitive system imposes coherence to facilitate choice—is consistent with the judicial function of producing compelling decisions even for the most close and contested of cases. It should also be noted that coherence maximizing processing operates mostly at a low level of awareness.[36] This lack of awareness further enables the endurance of this type of reasoning in the judicial practice (Wilson & Brekke 1994). It is noteworthy also that key components of coherence based reasoning appear in the theorizing of the some of the notable commentators on the judicial practice, including Holmes (1881, 1897), Cardozo (1921) and Llewellyn (1960).[37] Furthermore, the effects of coherence based reasoning on appellate judging has been supported by a study that examined data from actual court decisions (Beebe, 2006).

Another possible overlap with judging emanates from a recent study that shows that coherence can dissipate soon after the decision has been completed (Simon et al., 2008). Coherence seems to be an ad hoc construct, that appears around the time of making the choice, but does not linger on to limit the decision maker in future cases where the constellation of variables will not necessarily align in the same manner. Thus, while coherence tends to be very strong within each decision, it can be rather weak across cases. This finding suggests that people are capable of applying a particular rule or principle in one case, but not follow it in the next. This observation is consistent with the view that judges alternate between different—even

[34] The professional auditors were tested performing an actuarial task (Lundberg, 2007).

[35] See Simon (1998), pp. 61-102.

[36] Holyoak & Simon (1999), Study 2.

[37] See Simon (1998), pp. 102-121.

opposing—rules, policies, and interpretive principles from case to case (Llewellyn, 1950, 1960; *cf*. Schauer, 2007).

To the extent that one is persuaded by the applicability of coherence based reasoning, its insights can be used to illuminate the normally obscured question of freedom and constraint in judicial decision making. The exceptional and otherwise unexplained coherence reported in judicial opinions is best understood not as a true representation of the constraining nature of the legal materials. The professed constraint is mostly an artifact of the cognitive process that people employ in the making of complex decisions, judicial and otherwise.

Schauer is of course correct in objecting to facile applications of basic research to judicial decision making. Yet, there seems good reason to conclude that coherence based reasoning meets the requisite standards of applicability. Whether one is persuaded by it or not, this application ought not to be regarded as unargued, axiomatic, or unresearched.

Valid and pertinent basic psychological research should not be discounted too readily. Given the difficulties involved in understanding the judicial process, one ought to adopt an eclectic stance, accepting any valid and informative research. An illustrative example can be borrowed from the application of psychology to medicine, specifically, the growing field of medical decision making. This field has been based on a variety of methodological approaches. Some of the studies test the actual performance of doctors in tasks that have been tested on non-physicians (e.g., Leblanc, Brooks, & Norman, 2002; Wallsten, 1981). Other research projects simply apply basic-psychological to medical decisions (Croskerry, 2002; Graber, Franklin, & Gordon, 2005; Pines, 2005; Redelmeier, 2005). The reliance on generic psychological research in this highly specialized field manifests a belief that any type of valid psychological research ought to be brought to bear to better understand how medical decisions are made and how they can be made better (Groopman, 2007). This prioritization of the usefulness of the research over its uniqueness could likewise benefit the study of judicial decision making.

Conclusion

The psychology of judging is poised to benefit much from a call for more experimentation, especially when it comes from a scholar of Schauer's stature. Schauer's innovative and provocative essay provides a good opportunity to force people working in the field to think hard about the strengths and limitations of their methodological choices. Hopefully, it will also attract others to engage these issues and contribute to the development of the field.

Still, the specifics of Schauer's critiques and the proposed research are not free of objections. The laudable attempt to increase the external validity of the experimentation does not come

without costs. First, the insistence on uniqueness ends up compromising the proposed studies' construct validity, and thus muddies up the conclusions that could be drawn from them. Methodological tradeoffs of this kind hound experimental psychologists on a regular basis.[38] It is possible also that like many other important aspects of human behavior, the hypotheses posed by Schauer simply do not lend themselves to experimental testing. Second, the insistence on uniqueness might lead to undue discounting of the potential contributions from non-unique research, especially basic psychology. Instead of uniqueness, the field has most to gain from a pedantic attention to the experimental design coupled with open-mindedness to the types of methodologies used.

The answer to the question posed in the title of Schauer's essay ought to be: yes, there is a psychology of judging. Admittedly, it is underdeveloped. To foster its growth, the field should be guided by the understanding that the practice of adjudication comprises of a wide and diverse range of decision making processes. In some facets of their work, judges exert judicial expertise, in others they behave just like ordinary people, and it is possible that in some facets they engage in processes that are unique to them. Researchers should opportunistically employ whichever methods are best suited for the subject of inquiry. Looking forward, the field stands to be enriched by carefully validated findings from all strands of psychology: basic psychology, applied psychology, the psychology of expertise, and perhaps also a unique psychology of judging.

[38] A key to the success of research projects is the critical examination that takes place in lab meetings at the early stages of experimental design. Of the many seemingly good ideas proposed at these meetings, only a small fraction make it to the phase of experimentation, not to mention onto the pages of scientific journals.

References

Abelson, R. P., & Rosenberg, M. (1958). Symbolic psycho-logic: A model of attitudinal cognition. *Behavioral Science, 3*, 1-8.

Altman, S. (1990). Beyond candor. *Michigan Law Review, 89*, 296-351.

Arlen, J. (1998). The future of behavioral economic analysis of law. *Vanderbilt Law Review, 51*, 1765-1788.

Arlen, J., & Talley, E. (in press). *Experimental law and economics*.

Aronson, E., Wilson, T. D., & Brewer, M. B. (1998). Experimentation in social psychology. In Gilbert, D. T., Fiske, S. T., & Lindzey, G. (Eds), *The handbook of social psychology,* Vol. 1 (4th ed.), pp. 99-142. New York: McGraw-Hill..

Asch, S. E. (1940). Studies in the principles of judgments and attitudes: II. Determination of judgments by group and by ego standards. *The Journal of Social Psychology*, *12*, 433-465.

Beebe, B. (2006). An empirical study of the multifactor tests for trademark infringement *California Law Review, 94*, 1581-1654.

Bertrand, M. & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *The American Economic Review*, 94, 991-1013

Brewer, M. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86,* 307-324.

Bushman, B. J. (1995). Moderating role of trait aggressiveness in the effects of violent media on aggression. *Journal of Personality and Social Psychology, 69*, 950-960.

Cardozo, B. N. (1921). *The nature of the judicial process*. New Haven: Yale University Press.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance*, pp. 21-30. Cambridge University Press.

Croskerry, P. (2002). Achieving quality in clinical decision making: Cognitive strategies and detection of bias. Academic Emergency medicine, 9, 1184–1204.

Cross, F. & Tiller, E. (1998). Judicial partisanship and obedience to legal doctrine: Whistleblowing on the Federal Courts of Appeals, *Yale Law Journal, 107*, 2155-2176.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33,* 510-540.

Edwards, H. (1998). Collegiality and Decision Making on the D.C. Circuit. *Virginia Law Review, 83*, 1335-70.

Edwards, W., & Newman, J. R. (1982). *Multiattribute evaluation*. Beverly Hills, CA: Sage.

Eisenberg, T., Garvey, S. P., & Wells, M. T. (2001). Forecasting life and death: Juror race, religion, and attitude toward the death penalty. *The Journal of Legal Studies, 30,* 277-311.

Ericsson, K. A., & Ward, P. (2007). Capturing the naturally occurring superior performance of experts in the laboratory: Toward a science of expert and exceptional performance. *Current Directions in Psychological Science, 16,* 346-350.

Frank, J. (1930). Law and the modern mind. NY: Brentano's.

Feltovich, P. J., Prietula, M. J., & Ericsson, K. A. (2006). Studies of expertise from psychological perspectives. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance,* pp. 41-67. Cambridge University Press.

Festinger, L (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Foroni, F. & Mayr, U (2005). The power of a story: New, automatic associations from a single reading of a short scenario. Psychonomic Bulletin & Review, 12, 139-144.

Frey, D. (1986). Recent research on selective exposure to information. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 19, pp. 41–80). New York: Academic Press.

Glöckner, A. (2007). *Solving complex legal cases using the intuitive system.* Paper presented at the Subjective Probability, Utility and Decision Making, Warsaw, Poland, August 2007.

Glöckner, A., Betsch, T., & Schindler, N. (under review). Construction of probabilistic inferences by constraint satisfaction.

Graber, M. L., Franklin, N., & Gordon, R. (2005). Diagnostic error in internal medicine. *Archives of Internal Medicine, 165,* 1493-1499.

Groopman, J. (2007). *How doctors think*. New York: Houghton Mifflin.

Guthrie, C., Rachlinski, J.J., & Wistrich, A.J. (2001). Inside the judicial mind. *Cornell Law Review, 86,* 777-830.

Harman, G., & Kulkarni, S. R. (2006). The problem of induction. *Philosophy and Phenomenological Research, 72,* 559-575.

Hanson, J., & Yosifon, D. (2004). The situational character: A critical realist perspective on the human animal. *Georgetown Law Journal, 93,* 1-177.

Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology, 21,* 107-111.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Holmes, O. W. (1881). *The common law*. Boston: Little Brown.

Holmes, O. W. (1897). The path of the law. *Harvard Law Review, 10*, 457-478.

Holyoak, K. J. & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General, 128*, 3-31.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science: A Multidisciplinary Journal, 13*, 295-355.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110*, 220-264.

Kalven, H,. & Zeisel, H. (1966). *The American jury*. Boston: Little, Brown.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163-182.

Klayman, J. (1995). Varieties of confirmation bias. *The Psychology of Learning and Motivation, 32*, 385-418.

Klein, D. E. (2002). *Making law in the United States Courts of Appeals*. New York: Cambridge University Press.

Klein, D. E. & Morrisroe, Darby (1999). The prestige and influence of individual judges on the U. S. Courts of Appeals. *The Journal of Legal Studies, 28*, 371-391

Koehler, D. J., Brenner, L., & Griffin, D. (2002). The calibration of expert judgment: Heuristics and biases beyond the laboratory. In T., Gilovich, D. Griffin, & D. Kahneman (Eds), (pp. 686-715). *Heuristics and biases: The psychology of intuitive judgment.* New York: Cambridge University Press.

Konecni, V. J., & Ebbesen, E. B. (1986). Courtroom testimony by psychologists on eyewitness identification issues: Critical notes and reflections. *Law and Human Behavior, 10*, 117-126.

Kunda, Z. (1990), The case for motivated reasoning. *Psychological Bulletin, 108*, 480-498.

Landeo, C. M. (under review). Tort reform and disputes under endogenous beliefs.

Leblanc, V. R., Brooks, L. R., & Norman, G. R. (2002). Believing is seeing: The influence of a diagnostic hypothesis on the interpretation of clinical features. *Academic Medicine, 77*, No. 10/October Supplement.

Lerner, J. S., Goldberg, J. H., & Tetlock, P. E. (1998). Sober second thought: The effects of accountability, anger, and authoritarianism on attributions of responsibility. *Personality and Social Psychology Bulletin, 24*, 563-574

Levy, A. G., & Hershey, J. C. (2008). Value-induced bias in medical decision making. Medical Decision Making, 28, 269-276.

Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw Hill.

Llewellyn, K. N. (1950). Remarks on the theory of appellate decision and the rules or canons about how statutes are to be construed. *Vanderbilt Law Review, 3*, 395-406.

Llewellyn, K. N. (1960). *The common law tradition: Deciding appeals*. Boston: Little, Brown.

Lundberg, C. G. (2004). Modeling and predicting emerging inference-based decisions in complex and ambiguous legal settings. *European Journal of Operational Research, 153*, 417–432.

Lundberg, C. G. (2007). Models of emerging contexts in risky and complex decision settings. *European Journal of Operational Research, 177*, 1363–1374.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*, 375-407.

Mitchell, G. (2002). Taking behavioralism too seriously? The unwarranted pessimism of the new behavioral analysis of law. *William and Mary Law Review, 43*, 1907-2021.

Mitchell, G. (2003). Tendencies versus boundaries: Levels of generality in behavioral law and economics. *Vanderbilt Law Review, 56*, 1781-1812.

Mitchell, M. L. & Jolley, J. M. (2007). Research design explained, 6th Ed.. Belmont, CA: Wadsworth/Thompson Learning.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175-220.

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*, 250-256.

Phillips, F. (2002). The distortion of criteria after decision-making. *Organizational Behavior and Human Decision Processes, 88*, 769-784.

Pines, J. M. (2005). Profiles in Patient Safety: Confirmation bias in emergency medicine. *Academic Emergency Medicine, 13*, 90-94.

Posner, R. A. (1995). Judges' writing styles (and do they matter?). *University of Chicago Law Review, 62*, 1421-1449.

Posner, R. A. (2008). *How judges think*. Cambridge, MA: Harvard University Press.

Read, S. J., & Miller, L. C. (1998). On the dynamic construction of meaning: An interactive activation and competition model of social perception. In S. J. Read & L. C. Miller (Eds). *Connectionist models of social reasoning and social behavior* (pp. 27-68). Mahwah, NJ: Lawrence Erlbaum.

Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and Gestalt principles: (Re)introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review, 1*, 26-53.

Redelmeier, D. A. (2005). The cognitive psychology of missed diagnoses. *Annals of Internal Medicine, 142*, 115-120.

Rosenberg, M. J., & Abelson, R. P. (1960). An analysis of cognitive balancing. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, Abelson, R. P.., & J. W. Brehm (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 112-163. Oxford, England: Yale University Press.

Ross, L., & Nisbett, R. E. (1991). *The Person and the situation: Perspectives of social psychology*. New York: McGraw Hill.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance and contagion. *Personality and Social Psychology Review, 5*, 296-320.

Schauer, F. (1988). Judging in the corner of the law. *Southern California Law Review, 61*, 1717-1733.

Schauer, F. (2006a). Do cases make bad law? *University of Chicago Law Review, 73*, 883-918.

Schauer, F. (2006b). On the supposed jury-dependence of evidence law. University of Pennsylvania Law Review, 155, 165-202.

Schauer, F. (2007). Why precedent in law (and elsewhere) is not totally (or even substantially) about analogy. SSRN.

Schroeder, T. (1918). Psychologic Study of Judicial Opinions. *California Law Review, 6*, 89-113.

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology, 51*, 515-530.

Segal, J. A., & Spaeth, H. J. (1993). The supreme court and the attitudinal model. New York: Cambridge University Press.

Segal, J. A., & Spaeth, H. J. (2002). The supreme court and the attitudinal model revisited. New York: Cambridge University Press.

Simon, D. (1998). A psychological model of judicial decision making. *Rutgers Law Journal, 30*, 1-142.

Simon, D. (2002). Freedom and constraint in adjudication: A look through the lends of cognitive psychology. *Brooklyn Law Review, 67*, 1097-1139.

Simon, D. (2004). A third view of the black box: Cognitive coherence in legal decision making. *The University of Chicago Law Review, 71*, 511-586.

Simon, D., Pham, L. B., Le, Q. A., & Holyoak, K. J. (2001). The emergence of coherence over the course of decision making. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 1250-1260.

Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004a). Construction of preferences by constraint satisfaction. *Psychological Science, 15*, 331-336.

Simon, D., Daniel C. Krawczyk, Bleicher, A., & Keith J. Holyoak (2008). The transience of constructed preferences. *Journal of Behavioral Decision Making, 21*, 1-14.

Simon, D., Snow, C., J., & Read, S. J. (2004b). The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology, 86*, 814-837.

Spellman, B. A., & Holyoak, K. J. (1992). If Saddam is Hitler then who is George Bush? Analogical mapping between systems of social roles. *Journal of Personality and Social Psychology, 62*, 913-933.

Sunstein, C., & Ullman-Margalit, E. (1999). Second-order decisions. *Ethics, 110*, 5-26.

Thompson, W. C., Cowan, C. L., Ellsworth, P. C., & Harrington, J. C. (1984). Death penalty attitudes and conviction proneness: The translation of attitudes into verdicts. *Law and Human Behavior, 8*, 95-113.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wallsten, T. (1981). Physician and medical student bias in evaluating diagnostic information. *Medical Decision Making, 1*, 145-164.

Wertheimer, W (1923/1938). Laws in organization of perceptual forms. In W. D. Ellis, (Ed.), *A source book of Gestalt psychology*, pp. 71-88. London: Routledge & Kegan Paul.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116*, 117-142.

Wistrich, A.J., Guthrie, C., & Rachlinski, J.J. (2005). Can judges ignore inadmissible information?: The difficulty of deliberately disregarding. *University of Pennsylvania Law Review, 153*, 1251-1345.

Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology, 71*, 291-301.