

BIASES IN THE INTERPRETATION AND USE OF RESEARCH RESULTS

Robert J. MacCoun

Richard and Rhoda Goldman School of Public Policy, University of California,
Berkeley, California 94720-7320; e-mail: maccoun@socrates.berkeley.edu

KEY WORDS: advocacy, ideology, judgment, methodology, politics, values

ABSTRACT

The latter half of this century has seen an erosion in the perceived legitimacy of science as an impartial means of finding truth. Many research topics are the subject of highly politicized dispute; indeed, the objectivity of the entire discipline of psychology has been called into question. This essay examines attempts to use science to study science: specifically, bias in the interpretation and use of empirical research findings. I examine theory and research on a range of cognitive and motivational mechanisms for bias. Interestingly, not all biases are normatively proscribed; biased interpretations are defensible under some conditions, so long as those conditions are made explicit. I consider a variety of potentially corrective mechanisms, evaluate prospects for collective rationality, and compare inquisitorial and adversarial models of science.

CONTENTS

INTRODUCTION	260
<i>Accusations and Controversies</i>	260
<i>Chapter Overview</i>	262
THE SCIENTIFIC STUDY OF BIASED SCIENCE	263
<i>Bias in the Eye of the Beholder?</i>	263
<i>Operationalizing Bias</i>	264
<i>An Experimental Paradigm</i>	265
<i>Overview of Theoretical Perspectives</i>	267
COLD COGNITIVE SOURCES OF BIAS	269
<i>Strategy-Based Errors</i>	269
<i>Mental Contamination</i>	270
MOTIVATED COGNITION	270

<i>The Psychodynamics of Science</i>	270
<i>Cognitive Dissonance Theory</i>	271
<i>Motive-Driven Cognition; Cognition-Constrained Motivation</i>	271
BAYESIAN PRIORS AND ASYMMETRIC STANDARDS	273
CORRECTIVE PRACTICES	275
<i>Debiasing</i>	275
<i>Falsification, Strong Inference, and Condition Seeking</i>	276
<i>Peer Reviewing, Replication, Meta-Analysis, Expert Panels</i>	277
<i>Will “Truth Win” Via Collective Rationality?</i>	278
<i>Adversarial Science</i>	279
CONCLUSION	281

INTRODUCTION

The claim that a social scientist is “biased” is rarely a neutral observation. In our culture, it can be a scathing criticism, a devastating attack on the target’s credibility, integrity, and honor. Rather than coolly observing that “Professor Doe’s work is biased,” we are apt to spit out a phrase like “. . . is *completely* biased” or “. . . is biased *as hell*.” Such expressions of righteous indignation are generally a sure sign that some kind of norm has been violated. The sociologist Robert Merton (1973) articulated four norms of science that are widely shared in our culture by both scientists and nonscientists alike. *Universalism* stipulates that scientific accomplishments must be judged by impersonal criteria; the personal attributes of the investigator are irrelevant. *Communism* (as in “communalism”) requires scientific information to be publicly shared. *Disinterestedness* admonishes investigators to proceed objectively, putting aside personal biases and prejudices. Finally, *organized skepticism* requires the scientific community to hold new findings to strict levels of scrutiny, through peer review, replication, and the testing of rival hypotheses.¹ This chapter examines theory and research on the violation—or the perceived violation—of Merton’s norms, by social scientists and by those who use our research.

Accusations and Controversies

In recent years, psychological researchers have been criticized for interpreting our data in ways that promote liberal political views, disparage conservative views (Ray 1989, Suedfeld & Tetlock 1991, Tetlock 1994, Tetlock & Mitchell 1993), and ignore radical views (e.g. Fox 1993). The psychological research literature has been criticized for being sexist (see Eagly 1995, Gannon et al 1992, Tavis 1992), racist (see Yee et al 1993), anti-Semitic (Greenwald & Schuh 1994), homophobic (Herek et al 1991), ageist (Schaie 1988), antireligious (e.g. Richards & Davison 1992), and biased toward a Western individual-

¹Koehler (1993) presents evidence that scientists endorse such norms.

ist world view (e.g. Sampson 1989). Within the American Psychological Association, there have been spirited debates about the propriety of legislative and judicial advocacy by the organization and its members (Barrett & Morris 1993; Fiske et al 1991; Jarrett & Fairbank 1987; Saks 1990, 1993).

Sadly, there is no shortage of politicized research topics, where the motives of researchers and the interpretation of their findings are fiercely disputed (Alonso & Starr 1987, Maier 1991, Porter 1995, Suedfeld & Tetlock 1991). Some topics are matters of perpetual dispute; examples include research on the effects of gun control (Nisbet 1990), the death penalty (Costanzo & White 1994), pornography (Linz & Malamuth 1993), and drug prohibition (MacCoun 1993a, MacCoun & Reuter 1997). And recent years have seen the emergence of new battlegrounds involving research on global warming (Gelbspan 1997), HIV/AIDS (Epstein 1996), the addictiveness of tobacco (Cummings et al 1991, Glantz 1996), the biological basis of sexual orientation (LeVay 1996), the effects of gay and lesbian service personnel on military cohesion (Herek et al 1996, MacCoun 1993b), and the validity of therapeutically elicited repressed memories (Pezdek & Banks 1996) and racial stereotypes (Gilbert 1995, Lee et al 1995). But surely the most explosive example involving our own discipline is the long-standing dispute about racial differences and their heritability (Gould 1981), most recently resurrected by Herrnstein & Murray's (1994) *The Bell Curve* and the huge critical literature that has emerged in response (e.g. Fraser 1995, Fischer et al 1996, Neisser et al 1996).

The very decision to study certain topics is sufficient to prompt some observers to infer that the investigator is biased. Not infrequently, government officials have denounced or attempted to ban entire topics of research. A notorious example involves federal efforts to discredit early studies documenting that some diagnosed alcoholics are able to engage in sustained "controlled drinking"—drinking at reduced and less problematic levels (see Chiauuzzi & Liljegren 1993, Marlatt et al 1993). Other examples include the cancellation of an NIH-funded conference on genetic influences on violence (see Johnson 1993), congressional efforts to end epidemiological research on gun violence by the Centers for Disease Control (see Herbert 1996), various congressional attempts to block survey research on adolescent and adult sexual behavior (see Gardner & Wilcox 1993, Miller 1995), and Representative Dick Solomon's ongoing efforts to pass the Anti-Drug Legalization Act, which states that "no department or agency of the United States Government shall conduct or finance, in whole or in part, any study or research involving the legalization of drugs."² The private sector is also guilty of research censorship, as illustrated

²My own research on this topic (MacCoun 1993a, MacCoun & Reuter 1997) is funded by a private source, the Alfred P. Sloan Foundation.

by the recent disclosure that a pharmaceutical company blocked publication of a study equating the effectiveness of its drug and less expensive generic alternatives (Dong et al 1997, Rennie 1997).

Chapter Overview

The focus of this essay is on actual and perceived violations of these norms; specifically, judgmental biases in the selection and interpretation of research evidence. I focus on psychological theory and research on biases in the interpretation and use of scientific evidence, by both scientists and nonscientists. For convenience, I use the term “judge” to refer generically to both scientists who collect and present data and consumers who interpret it; context will make clear that some discussions are more relevant for the former role, others the latter.

I refer the interested reader elsewhere for discussion of bias in the *conduct* of research, including research design (Campbell & Stanley 1963), choice of study populations (e.g. Graham 1992, Hambrecht et al 1993), statistical analysis (e.g. Abelson 1995, Rosenthal 1994), data presentation (e.g. Huff 1954, Monmonier 1996), experimenter gender (e.g. Eagly & Carli 1981), and experimenter expectancies (Campbell 1993, Harris 1991, Rosenthal 1994).

Given the need to bound this topic, I must give short shrift to important contributions that other academic disciplines have made to our understanding of biases in the interpretation and use of research findings. For example, I will only briefly touch on findings from the extensive sociology literature on the effects of institutional factors, professional incentives, social networks, and demographic stratification on the scientific research process (see Cole 1992, Merton 1973, Zuckerman 1988). Similarly, I assume most readers of the *Annual Review* have at least a passing familiarity with the major developments in twentieth century philosophy of science (see Gholsen & Barker 1985, Laudan 1990, Shadish 1995, Thagard 1992), so I limit my discussion to a few points where developments in philosophy inform a psychological debate, or vice versa.

I also sidestep the burgeoning postmodernist literatures on social constructivism, deconstructionism, hermeneutics, and the like (Best & Kellner 1991; cf Gross & Levitt 1994). Much of the material discussed here resonates with those perspectives, but there are also fundamental philosophical differences. While postmodernist thought has profoundly stirred many academic disciplines, it has created only modest ripples in mainstream psychology (see Ger-gen 1994, Smith 1994, Wallach & Wallach 1994). Though some attribute this to our discipline’s collective naïveté about these heady intellectual currents, there are other plausible explanations. The mind’s role in constructing our world is already a focal concern of all but the most radically behaviorist psy-

chologists. Yet our discipline attracts few who endorse a radically idealist ontology; if there's no *there* out there to study, the practice of scientific psychology would seem fairly pointless beyond the fun and profit, or at least the fun. Nevertheless, few mainstream psychologists resemble the hypothetico-deductive “straw positivist” depicted in postmodern critiques (see Shadish 1995). Cook (1985) has provided a stylish label—“post-positivist critical multipism”—for the dominant view in empirical psychology at least since Campbell & Stanley (1963): The choice of research methods poses inevitable trade-offs, and we can at best hope to approximate truth through a strategy of triangulation across multiple studies, investigators, and fallible methodologies. The psychological study of biased use of evidence is an essential part of that program, and systematic empirical methods have played a crucial role in identifying those biases.

In the following section, I briefly highlight the hazards of attributing bias and summarize the dominant strategies for scientifically studying the biased use of science. In the next three sections, I briefly review theory and research on three different sources of biased evidence processing: cold cognitive mechanisms, motivated cognition, and asymmetric standards of proof. Explicating these mechanisms makes clear that some forms of bias are more forgivable than others; indeed, some seem normatively defensible. In the final section, I discuss corrective practices for mitigating bias, including debiasing techniques, falsification and other hypothesis testing strategies, and institutional practices like peer reviewing, replication, meta-analysis, and expert panels. I consider the conditions under which collective judgment attenuates or exacerbates bias, and I compare adversarial and inquisitorial models of science.

THE SCIENTIFIC STUDY OF BIASED SCIENCE

Bias in the Eye of the Beholder?

The notion that observers' personal prejudices and interests might influence their interpretation of scientific evidence dates back at least to Francis Bacon (Lord et al 1979). But talk is cheap—it is easier to accuse someone of bias than to actually establish that a judgment is in fact biased. Moreover, it is always possible that the bias lies in the accuser rather than (or in addition to) the accused. There are ample psychological grounds for taking such attributions with a grain of salt.

For example, research using the attitude attribution paradigm (see Nisbett & Ross 1980) suggests that we might be quick to “shoot the messenger,” viewing unpalatable research findings as products of the investigator's personal dispositions rather than properties of the world under study. Research on the “hostile media phenomenon” (Vallone et al 1985, Giner-Sorolla & Chaiken

1994) shows that partisans on both sides of a dispute tend to see the exact same media coverage as favoring their opponents' position. Keltner & Robinson (1996) argue that partisans are predisposed to a process of *naïve realism*; by assuming that their own views of the world are objective, they infer that subjectivity (e.g. due to personal ideology) is the most likely explanation for their opponents' conflicting perceptions. Because this process tends to affect both sides of a dispute, Robinson, Keltner, and their colleagues have demonstrated that the gap between partisans' perceptions in a variety of settings are objectively much smaller than each side believes.

Thus, we should be wary about quickly jumping to conclusions about others' biases. For example, "everyone knows" that scientists sponsored by tobacco companies are biased, having sold out their objectivity for a lucrative salary. Thus, it may come as a surprise—it did to me—to learn from surveys that a majority of these scientists acknowledge that cigarette smoking is addictive and a cause of lung cancer (Cummings et al 1991)—though this finding does not exonerate them from responsibility for their professional conduct. In a related vein, lawyer Peter Huber (e.g. Foster et al 1993) has received considerable attention for his argument that scientists who serve as expert witnesses are guilty of "junk science," spewing out whatever pseudoscientific conclusions are needed to support their partisan sponsors. While Huber's general conclusion might be correct, the cases he makes against specific experts are vulnerable to a host of inferential biases, including many of the same methodological shortcomings he identifies in their research (MacCoun 1995).

Operationalizing Bias

Rather than attempt a theory-free definition of bias, I'll make use of Hastie's & Rasinski's (1988, Kerr et al 1996) taxonomy of "logics" for demonstrating that judgments are biased.³

One such logic is the one most observers use in attributing bias to others—a direct comparison of judgments across judges. (Or across groups of judges that differ in some attribute, e.g. men vs women, liberals vs conservatives, etc.) If the judgments are discrepant, then even in the absence of external criteria one can arguably infer bias. (As we see, Bayesians may disagree.) The problem is, bias on who's part? A weakness of this logic is that the observed discrepancy tells us nothing about whether either judge (or group of judges) is actually correct; both could be wrong. I'll return to this logic at the end of this essay when I examine the efficacy of collective strategies for bias reduction.

³As distinguished from taxonomies of bias *mechanisms*, discussed below.

In a second logic, bias or error is established directly by measuring the discrepancy between the judgment and the true state being judged. This logic has been quite fruitful in psychophysics, perhaps less so in social psychology, where we often lack objective measures of the “true” state of the sociopolitical environment. The third and fourth logics have been most productive in cognitive and social psychology and form the basis for much of the research discussed here. In these logics, the presence and content of various informational cues is manipulated in a between- or within-subjects experiment. In the third logic, a bias is established by showing that a judge is “using a bad cue”—i.e. overutilizing a cue relative to normative standards (e.g. legal rules of evidence, a rational choice model, or the cue’s objective predictive validity). In the fourth, a bias is established by demonstrating that the judge is “missing a good cue”—underutilizing a cue relative to normative standards. These are “sins of commission” and “sins of omission,” respectively (Kerr et al 1996).

An Experimental Paradigm

Mahoney (1977) conducted the earliest rigorous demonstration of biased evidence processing using the experimental approach. Behavioral modification experts evaluated one of five randomly assigned versions of a research manuscript on the “effects of extrinsic reward on intrinsic motivation,” a hypothesis in potential conflict with the experts’ own paradigm. The five versions described an identical methodology but varied with respect to the study’s results and discussion section. Mahoney found that the methodology and findings were evaluated more favorably, and were more likely to be accepted for publication, when they supported the experts’ views. Perhaps the most intriguing finding of Mahoney’s study was unintentional; reviewers who received a version of the manuscript with undesirable results were significantly more likely to detect a truly accidental, but technically relevant, typographical mistake.

Lord et al (1979) conceptually replicated Mahoney’s results and extended them in several important ways. Because their study has inspired considerable research on these phenomena, it is worth describing their paradigm in some detail. Based on pretesting results, 24 students favoring capital punishment and 24 opposing it were recruited; each group believed the existing evidence favored their views. They were then given descriptions of two fictitious studies, one supporting the deterrence hypothesis, the other failing to support it. For half the respondents, the prodeterrence paper used a cross-sectional methodology (cross-state homicide rates) and the antideterrent paper used a longitudinal methodology (within-state rates before and after capital punishment was adopted); for the remaining respondents, the methodologies were reversed. Each description contained a defense of the particular methodology and a critique of the opposing approach. Students received and provided initial reac-

tions to each study's results before being given methodological details to evaluate.

Analyses of student ratings of the quality and persuasiveness of these studies revealed a *biased assimilation* effect—students more favorably evaluated whichever study supported their initial views on the deterrent effect, irrespective of research methodology. Students' open-ended comments reveal how either methodology—cross-sectional or longitudinal—could be seen as superior or inferior, depending on how well its results accorded with one's initial views. For example, when the cross-sectional design yielded prodeterrence results, a death-penalty proponent praised the way “the researchers studied a carefully selected group of states...,” but when the same design yielded antideterrence results, another death-penalty advocate argued that “there were too many flaws in the picking of the states...” Having been exposed to two studies with imperfect designs yielding contradictory results, one might expect that Lord et al's participants would have become more moderate in their views; if not coming to an agreement, at least shifting toward the grey middle zone of the topic. But Lord et al argue that such situations actually produce *attitude polarization*. Thus, in their study, respondents in each group actually became *more* extreme in the direction of their initial views. Lord and colleagues argued that “our subjects' main inferential shortcoming...did not lie in their inclination to process evidence in a biased manner...Rather, their sin lay in their readiness to use evidence to bolster the very theory or belief that initially 'justified' the processing bias.”

There have been numerous conceptual replications and extensions of the Lord et al findings (Ditto & Lopez 1992, Edwards & Smith 1996, Koehler 1993, Kuhn & Lao 1996, Lord et al 1985, Miller et al 1993, Munro & Ditto 1997, Plous 1991, Sherman & Kunda, cited in Kunda 1990). For example, Plous (1991) noted that biased assimilation and attitude polarization imply that “people will feel *less* safe after a noncatastrophic technological breakdown if they already oppose the particular technology, but will feel *more* safe after such a breakdown if they support the technology.” He supported this prediction in several studies of the reactions of psychology students, ROTC cadets, and professional antinuclear activists to information about a noncatastrophic nuclear breakdown. In a variation on the Lord et al paradigm, Koehler (1993) instilled weak or strong beliefs regarding two fictitious issues, then exposed respondents to studies with either low- or high-quality evidence. Studies that were consistent with instilled beliefs were rated more favorably, and the effect was stronger for those with strong beliefs (see also Miller et al 1993). In a second study, Koehler (1993) replicated the biased assimilation effect with professional experts on opposite sides of ESP debate; intriguingly, the effect was stronger among “hard-nosed” skeptics than among the parapsychologists. McHoskey

(1995) found that identical evidence regarding the JFK assassination was judged to be supportive by both conspiracy theorists and their detractors.

The biased assimilation phenomenon fit comfortably into an already burgeoning literature on biased information processing. The attitude polarization finding—the notion that exposure to mixed evidence moves opposing groups farther apart rather than closer together—was more novel. Yet Miller et al (1993) and Kuhn & Lao (1996) each note with surprise that this finding was so widely cited and accepted without critical challenge. As Kuhn & Lao note, “the findings contradict an assumption basic to much educational thought and prevalent in our culture more broadly—the assumption that engaging people in thinking about an issue will lead them to think better about the issue” (p. 115).

Happily, subsequent studies suggest possible boundary conditions on these phenomena. The biased assimilation effect is robust among judges with extreme attitudes but difficult to replicate among those with moderate views (Edwards & Smith 1996, McHoskey 1995, Miller et al 1993). Several studies have found that attitude polarization is limited to self-reported change ratings (Kuhn & Lao 1996, Miller et al 1993, Munro & Ditto 1997), though McHoskey (1995) found polarization in direct measures of attitude change. Miller et al (1993) found that neutral raters did not perceive any significant attitude polarization in essays written by the judges. Kuhn & Lao (1996) also found that polarization was just as common among respondents who wrote essays and/or discussed the topic in lieu of examining mixed research evidence.

In a paper on biased evidence evaluation, one offers conclusions about mixed evidence with some trepidation! But I think it is safe to say that the studies just cited, and additional evidence reviewed elsewhere in this chapter, provide strong support for the existence of biased assimilation effects and weak support for attitude polarization effects. Attitude polarization in response to mixed evidence, if it does exist, is a remarkable (and remarkably perverse) fact about human nature, but the mere fact that participants *believe* it is occurring is itself noteworthy. And even in the absence of attitude polarization, biased assimilation is an established phenomenon with troubling implications for efforts to ground contemporary policy debates in empirical analysis.

Overview of Theoretical Perspectives

We are blessed with a wealth of theoretical perspectives for explaining biased evidence processing. As we shall see, these accounts are not mutually exclusive (and are probably not mutually exhaustive). Integrating them into a grand theory seems premature. Instead, I first sketch five prototypes of biased evidence processing. The prototypes vary with respect to intentionality, motivation, and normative justifiability. By *intentionality*, I refer to the combination of consciousness and controllability; a bias is intentional when the judge is

aware of a bias yet chooses to express it when she could do otherwise (see Fiske 1989). *Motivation* is shorthand for the degree to which the bias has its origins in the judge's preferences, goals, or values; intentional bias is motivated, but not all motivated biases are intentional. Finally, *normative justification* distinguishes appropriate or defensible biases from inappropriate or indefensible biases; justification is always relative to some normative system, and I'll refer to several, including Merton's norms of science, Bayesian and decision theoretic norms of inference, ethical norms, and legal norms.

The first prototype is *fraud*: intentional, conscious efforts to fabricate, conceal, or distort evidence, for whatever reason—material gain, enhancing one's professional reputation, protecting one's theories, or influencing a political debate. There is a growing literature on such cases (see Fuchs & Westervelt 1996, Woodward & Goodstein 1996), though we still lack estimates on their prevalence. At a macro level, they are often explicable from sociological, economic, or historical perspectives (Cole 1992, Zuckerman 1988). At a micro level, they are sometimes explicable in terms of individual psychopathology. These cases are extremely serious, but I give them short shrift here, focusing instead on generic psychological processes that leave us all vulnerable to bias. I should note, however, that scarce funding and other institutional pressures can blur the lines between fraud and less blatant sources of bias; see recent examinations of tobacco industry research (Cummings et al 1991, Glantz 1996), drug prevention evaluations (Moskowitz 1993), risk prevention research (Fischhoff 1990), global warming testimony (Gelbspan 1997), and the Challenger disaster (Vaughan 1996).

A second prototype is *advocacy*: the selective use and emphasis of evidence to promote a hypothesis, without outright concealment or fabrication. As I discuss below, advocacy is normatively defensible provided that it occurs within an explicitly advocacy-based organization, or an explicitly adversarial system of disputing. Trouble arises when there is no shared agreement that such adversarial normative system is in effect.

I suspect the general public tends to jump to fraud or advocacy as explanations for findings they find "fishy," but contemporary psychologists recognize that most biased evidence processing can occur quite unintentionally through some combination of "hot" (i.e. motivated or affectively charged) and "cold" cognitive mechanisms. The prototypical *cold bias* is unintentional, and unconscious, and it occurs even when the judge is earnestly striving for accuracy. The prototypical *hot bias* is unintentional and perhaps unconscious, but it is directionally motivated—the judge wants a certain outcome to prevail. Though the distinction is useful, Tetlock & Levi (1982) made a persuasive case for the difficulty of definitively establishing whether an observed bias is due to hot vs cold cognition; the recent trend has been toward integrative "warm" theories.

Research on biased processing of scientific evidence has given somewhat less attention to the final prototype, which might be called *skeptical* processing. In skeptical processing, the judge interprets the evidence in an unbiased manner, but her conclusions may differ from those of other judges because of her prior probability estimate, her asymmetric standard of proof, or both. This is arguably normative on decision theoretic grounds, but those grounds are controversial.

COLD COGNITIVE SOURCES OF BIAS

Strategy-Based Errors

Numerous mechanisms have been identified in basic cognitive psychological research on memory storage and retrieval, inductive inference, and deductive inference that can produce biased evidence processing even when the judge is motivated to be accurate and is indifferent to the outcome. Arkes (1991) and Wilson & Brekke (1994) offer taxonomies for organizing these different sources of judgmental bias or error and detailed reviews of the relevant research.

For Arkes (1991), *strategy-based errors* occur when the judge, due to ignorance or mental economy, uses “suboptimal” cognitive algorithms. (Wilson & Brekke 1994 offer a similar category of “failures to know or apply normative rules of inference.”) Examples that might influence the interpretation of research findings include: (a) using fallacious deductive syllogisms (e.g. affirming the consequent, denying the antecedent), (b) failing to adjust for nonindependence among evidentiary items, (c) confusing correlation with causation, and (d) relying on heuristic persuasive cues (e.g. appeals to an investigator’s prestige or credentials).

One pervasive mental heuristic with special relevance to scientific evidence processing is *positive test strategy* (Klayman & Ha 1987), whereby hypotheses are tested by exclusively (or primarily) searching for events that occur when the hypothesis says they should occur. For example, to test the hypothesis that environmental regulations reduce employment rates, one simply cites jurisdictions with strict regulations and high unemployment (and, perhaps, jurisdictions with lax regulations and low unemployment). The evidence suggests that this kind of strategy is pervasive even in the absence of any particular outcome motivations (Fischhoff & Beyth-Marom 1983, Nisbett & Ross 1980, Snyder 1981). Positive test strategy clearly falls short of normative standards of inference, which would require data analysis strategies that take equal account of jurisdictions with strict regulations and low unemployment, or lax regulations and high unemployment. This kind of hypothesis testing is often called *confirmatory bias* (or confirmation bias), because the hypothesis is more likely to be confirmed than disconfirmed irrespective of its truth value. But in an insightful set-theoretic analysis, Klayman & Ha (1987; also see Friedrich 1993) demonstrate

that in some classes of situations, the positive strategy can be an efficient means of reaching correct conclusions. Of course, we seldom know when we are in such situations.

Mental Contamination

Wilson & Brekke (1994) call their category of nonstrategic error *mental contamination*, which they define as “the process whereby a person has an unwanted judgment, emotion, or behavior because of mental processing that is unconscious or uncontrollable” (p. 117). One type of mental contamination involves the unwanted consequences of automatic cognitive processing. For example, schematic principles of memory suggest that once a particular theory about the world becomes well learned, it filters our attention to and interpretation of incoming data (e.g. Nisbett & Ross 1980). A second subcategory is *source confusion*, whereby disassociation or misattribution breaks the link between information and its source. From a scientific perspective, this separation is arguably a good thing if the source information in question involves things like a study author’s race, gender, or nationality; the separation is much more serious when the source information includes key caveats about the study’s methodology.

Arkes (1991) also identifies two categories of nonstrategic error, both of which might be classified as sources of mental contamination. *Psychophysical-based errors* stem from nonlinear relationships between objective stimuli and their subjective representations. Examples include framing effects, anchoring effects, and context effects. *Association-based errors* are perverse side effects of otherwise adaptive principles of spreading activation in semantic memory. One example might be *hindsight bias*; e.g. the exaggerated tendency for research results to seem “obvious” ex post, relative to ex ante predictions (Slovic & Fischhoff 1977). Other examples might include priming effects, and perhaps the availability and representativeness heuristics (Kahneman et al 1982, Nisbett & Ross 1980).

MOTIVATED COGNITION

The Psychodynamics of Science

A recent paper by Elms (1988) helps to illustrate the limitations of the psychodynamic literature on scientific practice. Elms argues that Freud’s psychobiographical analysis of the scientific career of Leonardo da Vinci was distorted by Freud’s own “projected identification with Leonardo, incorporating aspects of his own sexual history and his anxieties about the future of the psychoanalytic movement.” But it is difficult to see how one might falsify such hypotheses, even in principle (Popper 1959). Moreover, Elms’ psychodynamic analy-

sis of Freud's psychodynamic analysis of da Vinci opens up an infinite regress, challenging us to analyze Elm's own motivations (an opportunity I'll forgo.)

Cognitive Dissonance Theory

A more tractable motivational account is Festinger's (1957) theory of cognitive dissonance. An early prediction was that dissonance aversion should encourage judges to seek out supportive information and shun potentially unsupportive information—the "selective exposure" hypothesis. In essence, this is a motivationally driven form of confirmatory bias. Despite a skeptical early review (Freedman & Sears 1965), subsequent research has shown that these effects do occur when judges have freely chosen to commit to a decision and the decision is irreversible—two conditions that should promote maximal dissonance and discourage belief change as its mode of resolution (Frey 1986). While this research shows that dissonance reduction is *sufficient* to produce confirmatory biases, research cited earlier shows that it isn't necessary.

Berkowitz & Devine (1989; Munro & Ditto 1997) argue that dissonance theory provides a parsimonious account of biased evidence assimilation. In brief, the notion is that discovering that research findings contradict one's hypothesis may well create dissonance, which might be resolved by discrediting the research that produced the findings. But dissonance could also be resolved by changing one's belief in the hypothesis. A weakness of the theory is its inability to clearly predict the choice of resolution mode (see Kunda 1990, Lord 1989, Schlenker 1992). Lord (1989) contends that biased assimilation is cognitive rather than motivational in nature. He notes that Lord et al (1985) were able to eliminate the effect using cognitive instructions (consider how you'd evaluate the study given opposite results) but not motivational instructions (try to be unbiased)—though his implicit argument that cognitive instructions can only eliminate cognitive biases seems questionable. At any rate, this kind of motivational vs cognitive debate rarely produces clear winners (Tetlock & Levi 1982). It may be the case that purely motivational biases play their strongest role not in the initial evaluation of evidence but rather in researchers' resistance to reconsidering positions they've publically endorsed in the past (see, for example, Staw & Ross 1989, on organizational research on sunk costs and escalating commitments).

Motive-Driven Cognition; Cognition-Constrained Motivation

Recent theories of motivated cognition are notable for integrating motivational and cognitive processes. For example, dual process theories of persuasion (e.g. Chaiken et al 1989, Petty & Cacioppo 1986) propose that a judge will only evaluate information rigorously and systematically if she is both motivated and able to do so; if both conditions aren't met, judgments will be formed heuristically using superficial cues or cognitive shortcuts. Although the motiva-

tion posited in these models was a desire for accuracy, Chaiken and her colleagues (e.g. Chaiken et al 1989, Giner-Sorolla & Chaiken 1997, Liberman & Chaiken 1992) have extended this work by examining the effects of defensive and impression management motives. For instance, under defensive motivation, judges will use heuristic processing if it leads to congenial conclusions, only resorting to systematic processing if it does not.

Kruglanski (1989, Kruglanski & Webster 1996) has offered a taxonomy of motives organized around their epistemic objectives, rather than their psychological origins. At one extreme of a continuum, one has a *need for cognitive closure*; at the other, a *need to avoid closure*. The closure that is desired or avoided is *specific* when one seeks or shuns a particular answer, or *nonspecific* if one seeks or avoids closure irrespective of its content. The need for closure creates tendencies to reach a conclusion as quickly as possible (“seizing”) and stick to it as long as possible (“freezing”). In an imaginative research program, Kruglanski and his colleagues have demonstrated a variety of ways in which these motives influence information search, hypothesis formation, causal attributions, and inductive and deductive inference. Kruglanski & Webster (1996) discuss advantages of this framework over earlier concepts such as intolerance of ambiguity, authoritarianism, and dogmatism.

Pyszczynski & Greenberg (1987) and Kunda (1990) review much of the recent work on the effects of directional motives—where the judge prefers a particular outcome—on the generation and evaluation of hypotheses about the world. Pyszczynski & Greenberg (1987) argue that while motivation influences hypothesis testing, most of us feel constrained by the desire to maintain an “illusion of objectivity.” Similarly, Kunda (1990, p. 482) argues that directional biases “are not unconstrained: People do not seem to be at liberty to conclude whatever they want to conclude merely because they want to. Rather...people motivated to arrive at a particular conclusion attempt to be rational and to construct a justification of their desired conclusion that would persuade a dispassionate observer.” For example, Sherman & Kunda (cited in Kunda 1990) found that caffeine drinkers’ prior understanding of research methodology constrained their willingness to reject findings about the hazards of caffeine. Along similar lines, McGuire & McGuire (1991) have found only weak support for a “wishful thinking” effect, in which the desirability of a proposition enhances perceptions of its likelihood; they argue that this “autistic” effect is largely offset by other, more rational cognitive principles.

Kalven & Zeisel’s (1966) “liberation hypothesis” is essentially a corollary of the principle that the expression of bias is constrained by objective evidence. They argued that jurors are most likely to allow personal sentiments to influence their verdicts when the trial evidence is ambiguous. In support, MacCoun (1990,

Kerr et al 1996) cites several lines of individual- and group-level research demonstrating enhanced extra-evidentiary bias when evidence is equivocal.

Two recent studies indicate that the kind of biased assimilation effect documented by Lord et al is largely mediated by more stringent processing of evidence supporting views contrary to one's own. Ditto & Lopez (1992) found that students were significantly more likely to scrutinize a medical test when they tested positive for a potentially dangerous (fictitious) enzyme; they were also more than twice as likely to retest themselves. These reactions might appear to be normatively reasonable, but Ditto & Lopez also found that relative to students testing negative, students testing positive perceived the disease as less serious and more common; findings that argue in favor of a defensive motivational account and against a rational interpretation. Similarly, Edwards & Smith (1996) found support for a "disconfirmation bias," in which evidence inconsistent with the judge's prior beliefs was scrutinized more extensively. Moreover, this effect was heightened among participants with strongest emotional convictions about the issue. Munro & Ditto (1997) present evidence that affective responses play a significant role in mediating biased evidence assimilation.

BAYESIAN PRIORS AND ASYMMETRIC STANDARDS

Some of the most sophisticated thinking about evidence evaluation has come from the decision theory tradition, especially in the domains of medical and legal decision making, signal detection theory, and statistical inference. Psychologists are especially well acquainted with the latter domain. Our slavish adherence to the conventional 0.05 alpha level has been blamed for many sins, and here I'll add one more. By fixing alpha, we've basically opted out of the most interesting part of the decision theoretic process: deciding how we should best trade off errors in a particular judgment context. This may explain why psychological explanations of biased evidence processing have largely overlooked the decision theoretic distinction between inductive judgments and standards of proof.

In a highly simplified decision theoretic analysis of scientific evidence evaluation, the judge assesses $p(H|D)$, the conditional probability of the hypothesis (H) given the data (D). Of course, in a simplified Bayesian model, $p(H|D)$ equals the product of a likelihood ratio denoting the diagnosticity of the evidence, $p(D|H)/p(D)$, and the judge's *prior probability* (or "prior"), $p(H)$. (More sophisticated models appear in Howson & Urbach 1993, Schum & Martin 1982). For a Bayesian, the prior probability component is an open door to personal bias; so long as diagnosticity is estimated in a sound manner and integrated coherently with one's prior, the updated judgment is normatively defensible (see Koehler 1993). Of course, the normative status of this framework is a

source of continuing controversy among philosophers and statisticians (see Cohen 1989, Mayo 1996), especially the notion of subjective priors. Moreover, challenges to the theory's descriptive status (Arkes 1991, Kahneman et al 1982, Nisbett & Ross 1980, Pennington & Hastie 1993) leave its normative applicability in doubt. And much of the evidence reviewed here implies that the diagnosticity component is itself a major locus of bias, irrespective of the judge's prior.

But decision theory also identifies a second, less-controversial locus of potentially defensible "bias." Our probabilistic assessment of the hypothesis yields a continuous judgment on a 0-1 metric, yet circumstances often demand that we reach a *categorical* verdict: Will we accept or reject the hypothesis? This conversion process requires a *standard of proof*. Statistical decision theory, signal detection theory, and formal theories of jurisprudence share a notion that this standard should reflect a tradeoff among potential decision errors. A simple decision-theoretic threshold for minimizing one's regret is $p^* = u(\text{FN})/[u(\text{FN}) + u(\text{FP})]$, where $u(\text{FP})$ equals one's aversion to false positive errors, and $u(\text{FN})$ denotes one's aversion to false negative errors (see DeKay 1996, MacCoun 1984). The standard of proof, p^* , cleaves the assessment continuum into rejection and acceptance regions. Thus the standard of proof reflects one's evaluation of potential errors. This evaluation is a policy decision, not just a scientific decision, arguably even in the case of the conventional 0.05 alpha level.

When one error is deemed more serious than the other, the standard of proof becomes asymmetrical and can easily produce greater scrutiny of arguments favoring one position over another. Thus, even for most non-Bayesians, there is a plausible normative basis for "bias" in assessments of scientific research (see Hammond et al 1992). Note, however, that this form of bias is limited to qualitative, categorical decisions ("it's true," "he's wrong"); it cannot justify discrepancies across judges (or across experimental manipulations of normatively irrelevant factors) in their quantitative interpretations of the diagnosticity of evidence, $p(\text{D}/\text{H})/p(\text{D})$.

Mock jury research has established that various prejudicial factors influence jurors' standards of proof (Kerr 1993). A variety of methods have been developed for estimating mock jurors' p^* values. Interestingly, p^* as estimated indirectly from self-reported aversion to decision errors allows more accurate prediction of verdicts than direct self-reports of p^* , suggesting that jurors may be unwilling or unable to articulate their standards (see Hastie 1993, MacCoun 1984). Yet even the best estimates of p^* have fairly poor predictive power. Pennington & Hastie's (1993a) story model departs from this decision theoretic framework, replacing the $p(\text{H})$ vs p^* comparison with a more complex cognitive process of mapping the evidence onto alternative narrative structures and selecting the one with the best "goodness of fit." Thagard's (1992) ex-

planatory coherence model (ECHO) offers a similar interpretation using a connectionist constraint satisfaction network. Interestingly, Thagard describes his model as being purely cognitive; he considers a “Motiv-ECHO” model incorporating motivational postulates but ultimately rejects it as being superfluous. Still, it should be noted that several similar constraint satisfaction models have incorporated strong motivational components (see Read et al 1997).

This error tradeoff might explain Wilson et al’s (1993) demonstration of a “leniency bias,” such that professional scientists were more willing to publish studies with important findings, and an “oversight bias,” in which the scientists actually rated the identical methodology more favorably when the topic was important. The oversight bias is difficult to justify, but the leniency bias is arguably normative. In general, scientists seem to believe the decision to publish findings should be influenced by their perceived importance, but only up to a point. Studies reporting truly revolutionary findings are held to perhaps the highest standards of all, leaving a field open to claims that it is biased against novel or radical ideas. In a remarkable journal editorial, Russett (1988) described the angst involved in his decision to publish a paper asserting that group transcendental meditation reduced regional violence in the Middle East. Bem & Honorton (1994) describe a similar dilemma. Bem, who considered himself a skeptic regarding telepathy, joined forces with nonskeptic Charles Honorton to conduct a rigorous meta-analytic review of studies using the ganzfeld procedure. Honorton then passed away before the conclusion of the research, leaving Bem in the personally awkward position of deciding whether to try to publish results that seemingly document the existence of telepathy. [He did (see Bem & Honorton 1994).]

CORRECTIVE PRACTICES

Debiasing

Behavioral decision researchers have produced a burgeoning literature on *de-biasing* techniques (Arkes 1991, Koehler 1991, Nisbett 1993, Lerner & Tetlock 1994, Lord et al 1985, Schum & Martin 1982, Wilson & Brekke 1994). Examples include increasing incentives for accuracy, holding judges accountable for their judgments, enhancing outcome feedback, providing inferential training, task decomposition, and encouraging the consideration of alternative hypotheses. It should be noted that none of these techniques provides “silver bullet” solutions to the bias problem. Researchers are still trying to understand why some techniques work for some biases but not others (Arkes 1991, Wilson & Brekke 1994). Limited forms of these debiasing techniques are already built into traditional scientific practice through methodological training and professional socialization, replication, peer review, and theory competition.

Falsification, Strong Inference, and Condition Seeking

Scientific training and socialization emphasize self scrutiny, rooted in part in Popper's (1959) principle of falsificationism. Acknowledging Hume's argument that induction can never confirm a hypothesis, Popper contended that induction might permit one to *falsify* a hypothesis, via the *modus tollens* syllogism: "If p then q; not q; therefore, not p." For Popper, falsification permits a particular sort of scientific progress; at best we can weed out bad ideas while seeing how our leading hypotheses hold up under attack. Popper's claim that falsificationism distinguishes science from pseudoscience has comforted psychologists seeking to distinguish our efforts from those of self-help gurus, astrologers, and the like. But many have noted that in practice, it is exceedingly difficult to achieve agreement that one has falsified a hypothesis (see Greenwald et al 1986, Julnes & Mohr 1989, Laudan 1990, McGuire 1983, cf Klayman & Ha 1987). A resourceful theorist can generally invoke ancillary theoretical principles to explain away a disconfirming finding, often with justification. McGuire (1983) goes so far as to conjecture that all psychological hypotheses are correct under *some* conditions, "provided that the researcher has sufficient stubbornness, stage management skills, resources, and stamina" (p. 15) to find those conditions.

In a classic paper, Platt (1964), a practicing biologist, argued that our personal attachment to our hypotheses clouds our judgment and sets science up as a conflict among scientists, rather than among ideas. He suggested that rapidly advancing research programs share a common strategy that mitigates these confirmationist tendencies. Under this *strong inference* strategy, the researcher designs studies to test not a single hypothesis, but an array of plausible competitors. Greenwald et al (1986, Greenwald & Pratkanis 1988) applaud Platt's intent but suggest that his strategy is rooted in a naïve faith in falsificationism. Instead, they recommend a strategy they called *condition seeking*, in which a researcher deliberately attempts to "discover which, of the many conditions that were confounded together in procedures that have obtained a finding, are indeed necessary or sufficient" (McGuire 1983, p. 223). Condition seeking is data driven rather than theory driven. Critics have countered that condition seeking will lead to a proliferation of special-case findings, undermining the development of more general theories (Greenberg et al 1986, MacKay 1988, Moser et al 1988). Greenwald & Pratkanis (1988) reply that results-centered research strategies will yield findings with greater shelf life than theory-centered research findings and ultimately provide the grist for better theory formulation. Related strategies that deserve wider recognition include devil's advocacy (Schwenk 1990), the "consider the opposite" heuristic (Koehler 1991, Lord et al 1985), and Anderson's & Anderson's (1996) "destructive testing" approach.

Peer Reviewing, Replication, Meta-Analysis, Expert Panels

When self-scrutiny fails, we rely on institutional safeguards such as peer reviewing, research replication, meta-analysis, expert panels, and so on. A detailed review of these topics is beyond the scope of this essay, but it should be noted that many of these practices have themselves been scrutinized using empirical research methods. For example, Peters & Ceci (1982) provided a dramatic demonstration of the unreliability of the peer review process. A dozen scientific articles were retyped and resubmitted (with fictitious names and institutions) to the prestigious journals that had published them 18–32 months earlier. Three were recognized by the editors; eight of the remaining nine not only went unrecognized but got rejected the second time around. (One suspects that many articles would get rejected the second time around even when recognized.) Cicchetti (1991) and Cole (1992) provide equally sobering but more rigorously derived evidence on the noisiness of the peer review process, citing dismally low interreferee reliabilities in psychology journals (in the 0.19 to 0.54 range), medical journals (0.31 to 0.37), and the NSF grant reviewing process (0.25 in economics, 0.32 in physics). To make matters worse, at least some of this small proportion of stable variance in ratings is probably attributable to systematic bias, though the limited research base precludes any strong conclusions (see Blank 1991, Gardner & Wilcox 1993, Gilbert et al 1994, Laband & Piette 1994, Rennie 1997).

Traditionally, replications have been viewed as the most essential safeguard against researcher bias. Of course, this can only work if replications are attempted, and in fact, exact replications are fairly rare (Bornstein 1990), in part because editors and reviewers are biased against publishing replications (Neuliep & Crandall 1990, 1993). Moreover, replications can't eliminate any bias that's built into a study's methodology. In keeping with the critical multiplist perspective noted earlier (Cook 1985), the fact that most replications in the social sciences are "conceptual" rather than exact is probably a healthy thing (Berkowitz 1992), providing the opportunity for triangulation across diverse methodologies.

Despite some initial resistance, social scientists have come to recognize the tremendous corrective benefits of metaanalysis, the statistical aggregation of results across studies (e.g. Cooper & Hedges 1994, Schmidt 1992). Conducting a meta-analysis frequently uncovers errors or questionable practices missed by journal referees. And meta-analyses are sufficiently explicit that dubious readers who dispute a meta-analyst's conclusions can readily conduct their own reanalysis, adding or subtracting studies or coding new moderator variables. Most importantly, early conclusions about the effects of publication bias on meta-analytic results have led to new standards for literature reviewing

that seem likely to attenuate the citation biases that plague traditional reviews (e.g. Greenwald & Schuh 1994).

Will “Truth Win” Via Collective Rationality?

Our reliance on replication, peer review, and expert panels reveals that we are unwilling to place all our faith in training and socialization as means for guaranteeing unbiased judgments by individual researchers. Institutional practices like peer review, expert panels (e.g. Neisser et al 1996), and expert surveys (e.g. Kassin et al 1989) are premised on a belief that collective judgment can overcome individual error, a principle familiar to small-group psychologists as the *Lorge-Solomon Model A* (Lorge & Solomon 1955) (Model B having long since been forgotten). In this model, if p is the probability that any given individual will find the “correct” answer, then the predicted probability that a collectivity of size r will find the answer is $P = 1 - (1 - p)^r$. Implicit in this equation is the assumption that if at least one member finds the answer, it will be accepted as the collectivity’s solution—the so-called *Truth Wins* assumption (e.g. Laughlin 1996). This can only occur to the extent that group members share a normative framework that establishes the “correctness” of the solution. That framework might be acknowledged by most academicians (the predicate calculus, Bayes Theorem, organic chemistry), or it might not (e.g. astrology, numerology, the I Ching).

For almost half a century, social psychologists have tested the “truth wins” assumption for a variety of decision tasks (see Kerr et al 1996, Laughlin 1996). Though much of this work involves the domain of small, face-to-face group discussion, the basic social aggregation framework and many of the findings can and have been generalized to wider and more diffuse social networks (e.g. Latané 1996). First and foremost, “truth” rarely wins, at least not in the strict version where a solution is adopted if at least a single member identifies or proposes it. At best, “truth supported wins”—at least some social support is needed for a solution to gain momentum, indicating that truth seeking is a social as well as intellectual process (see Laughlin 1996, Nemeth 1986). Second, when members lack a shared conceptual scheme for identifying and verifying solutions—what Laughlin calls “judgmental” as opposed to “intellective” tasks—the typical influence pattern is *majority amplification*, in which a majority faction’s influence is disproportionate to their size, irrespective of the truth value of their position (see Kerr et al 1996).

Collective decision making (or statistical aggregation of individual judgments) is well suited for reducing *random error* in individual judgments. Indeed, this is a major rationale underlying the practices of replication and meta-analysis (Schmidt 1992). What about bias? A common assertion is that group decision making will correct individual biases, but whether in fact this actually

occurs depends on many factors, including the strength of the individual bias, its prevalence across group members, heterogeneity due to countervailing biases, and the degree to which a normative framework for recognizing and correcting the bias is shared among group members (see Kerr et al 1996, Tindale et al 1996). Elsewhere, my colleagues and I (Kerr et al 1996) demonstrate that under a wide variety of circumstances, collective decision making will significantly amplify individual bias, rather than attenuate it.

Adversarial Science

Collective decision making is most likely to amplify bias when it is *homogeneous* across participants. Heterogeneous biases create the potential for bias correction through constructive conflict. In the Anglo-American adversarial legal system, this notion is captured by the phrase “truth will out.” Yet the Western scientific tradition is quite self-consciously inquisitorial, rather than adversarial (Burk 1993, Lind & Tyler 1988, Thibaut & Walker 1978). In the inquisitorial model, the investigator strives to be neutral and objective, actively seeking the most unbiased methods for arriving at the truth. This dispassionate approach extends to the presentation of results; ideally, the investigator simply “tells it like it is” irrespective of who’s ox gets gored. She “calls it like she sees it” yet in theory anyone else should see it and call it the same way if they examine the evidence she’s gathered. In contrast, in an adversarial system, the investigator is an explicit advocate, actively seeking and selectively reporting the most favorable evidence for her position. Sociologists of science (see Cole 1992, Zuckerman 1988) paint a picture of scientific practice that is a dissonant blend of these seemingly unblendable models.

Thibaut & Walker (1978) proposed a normative “theory of procedure” for choosing between inquisitorial and adversarial processes. They argued that the inquisitorial method is to be preferred for “truth conflicts,” purely cognitive disagreements in which the parties are disinterested (or have shared interests) and simply want to discover the correct answer. The adversarial approach is to be preferred for “conflicts of interest” in which the parties face a zero-sum (or constant sum) distribution of outcomes. According to Thibaut & Walker, in the latter context, the goal is not to find truth but to provide justice—a fair procedure for resolving the conflict.

The problem is that social science research problems rarely fit into this tidy dichotomy. Many of the issues we study involve a messy blend of truth conflicts and conflicts of interest, making it difficult to separate factual disputes from value disputes (see Hammond 1996, Hammond et al 1992). Many researchers (e.g. Sears 1994) and research organizations (e.g. the Society for Psychological Study of Social Issues; see Levinger 1986) have explicitly embraced an adversarial or advocacy-oriented view of social research, and many

of us were attracted to the social sciences by social activist motives. But merging the adversarial and inquisitorial modes is problematic (see Burk 1993, Foster et al 1993). The adversarial legal system has many key features that are lacking in scientific practice. Here, I'll note four. First, the adversarial roles of the participants are quite explicit; no one mistakes an American trial lawyer for a dispassionate inquisitor. Second, at least two opposing sides are represented in the forum—though their resources may differ profoundly. Third, there is explicit agreement about the standard of proof, burden of proof (who wins in a tie?), and ultimate decision maker (i.e. the judge or jury). And fourth, in many (though not all) legal disputes, the opposing positions “bound” the truth, either because one of the positions is in fact true, or because the truth lies somewhere between the two positions.

Scientific practice is clearly very different. As expressed by Merton's (1973) norms, citizens in our culture have very clear role expectations for scientists; if one claims the authority of that role, one is bound to abide by its norms or risk misleading the public. This surely doesn't preclude advocacy activities on the part of scientists, but it does mean that we must be quite explicit about which hat we are wearing when we speak out, and whether we are asserting our facts (e.g. the death penalty has no marginal deterrent effect) or asserting our values (e.g. the death penalty degrades human life). Graduate training in schools of public policy analysis is much more explicit about managing these conflicting roles. For example, Weimer & Vining's (1992) textbook provides a neutral discussion of three different professional models: the *objective technician* who maintains a distance from clients but lets the data “speak for itself,” avoiding recommendations; the *client's advocate* who exploits ambiguity in the data to strike a balance between loyalty to the facts and loyalty to a client's interests; and the *issue advocate* who explicitly draws on research opportunistically in order to promote broader values or policy objectives.

Moreover, as noted at the outset, many have argued that social science as represented in our major journals is too homogeneous—too liberal, too Anglo-centric, too male, and so on. It should be noted that the viewpoints reflected in published research are surely endogenous; if our leanings influence our findings, our findings surely influence our leanings as well. But if scientists' prejudices influence their research, there is little hope that “truth will out” in the absence of a sizable or at least vigorous representation of alternative viewpoints (see Brenner et al 1996, Nemeth 1986). But as Latane' (1996) demonstrated, minority viewpoints often survive via processes of clustering and isolation; in the social sciences, this seems to manifest itself in separate journals, separate conferences, separate networks, and even separate academic departments.

Third, disputes over scientific findings typically lack an explicit burden and standard of proof, and an explicit final decision maker. This contributes to the

seeming intractability of many debates; when each observer is free to establish her own p^* , there is no grounds for consensus on who “won.” Expert panels assembled by the National Academy of Sciences and other organizations attempt to circumvent this problem, with mixed success. This is surely a blessing as well as a curse. In a democratic society, we should be wary of philosopher kings. Research findings are rarely a direct determinant of policy decisions, a fact that is only partially attributable to policymakers’ self-serving selectivity (Weiss & Bucuvalas 1980). Social scientists are sometimes strikingly naïve about the gaps between our research findings and the inputs needed for sound policy formation (see MacCoun & Reuter 1997, Weimer & Vining 1992).

But more importantly, the history of science (e.g. Gholson & Barker 1985, Thagard 1992) reveals little basis for assuming that the truth is represented among those factual positions under dispute at any given moment (also see Klayman & Ha 1987). This underscores the inherent ambiguity of using discrepancies among judges to locate and measure bias (Kerr et al 1996)—all of us might be completely off target.

CONCLUSION

I have cited a wealth of evidence that biased research interpretation is a common phenomenon, and an overdetermined one, with a variety of intentional, motivational, and purely cognitive determinants. But there is danger of excessive cynicism here. The evidence suggests that the biases are often subtle and small in magnitude; few research consumers see whatever they want in the data. The available evidence constrains our interpretations—even when intentions are fraudulent—and the stronger and more comprehensive the evidence, the less wiggle room available for bias. In addition, far from condemning the research enterprise, the evidence cited here provides grounds for celebrating it; systematic empirical research methods have played a powerful role in identifying biased research interpretation and uncovering its courses.

Nor are all biases indefensible. There are some normative grounds for accepting differing opinions about imperfect and limited research on complex, multifaceted issues. There is nothing inherently wrong with differing standards of proof and nothing shameful about taking an advocacy role—provided we are self-conscious about our standards and our stance and make them explicit. Fostering hypothesis competition and a heterogeneity of views and methods can simultaneously serve the search for the truth and the search for the good. But there is a pressing need to better articulate the boundary between adversarialism and what might be called heterogeneous inquisitorialism—a partnership of rigorous methodological standards, a willingness to tolerate un-

certainty, a relentless honesty, and the encouragement of a diversity of hypotheses and perspectives.

Visit the *Annual Reviews* home page at
<http://www.AnnualReviews.org>.

Literature Cited

- Abelson RP. 1995. *Statistics as Principled Argument*. Hillsdale, NJ: Erlbaum
- Alonso W, Starr P. 1987. *The Politics of Numbers*. New York: Russell Sage
- Anderson CA, Anderson KB. 1996. Violent crime rate studies in philosophical context: a destructive testing approach to heat and southern culture of violence effects. *J. Pers. Soc. Psychol.* 70:740–56
- Arkes HR. 1991. Costs and benefits of judgment errors: Implications for debiasing. *Psychol. Rev.* 110:486–98
- Barrett GV, Morris SB. 1993. The American Psychological Association's amicus curiae brief in *Price Waterhouse v. Hopkins*: the values of science versus the values of the law. *Law Hum. Behav.* 17:201–15
- Bem DJ, Honorton C. 1994. Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychol. Bull.* 115:4–18
- Berkowitz L. 1992. Some thoughts about conservative evaluations of replications. *Pers. Soc. Psychol. Bull.* 18:319–24
- Berkowitz L, Devine PG. 1989. Research traditions, analysis, and synthesis in social psychological theories: the case of dissonance theory. *Pers. Soc. Psychol. Bull.* 15:493–507
- Best S, Kellner D. 1991. *Postmodern Theory*. New York: Guilford
- Blank RM. 1991. The effects of double-blind versus single-blind reviewing: experimental evidence from *The American Economic Review*. *Am. Econ. Rev.* 81:1041–67
- Bornstein RF. 1990. Publication politics, experimenter bias and the replication process in social science research. *J. Soc. Behav. Pers.* 5:71–81
- Brenner LA, Koehler DJ, Tversky A. 1996. On the evaluation of one-sided evidence. *J. Behav. Decis. Mak.* 9:59–70
- Burk DL. 1993. When scientists act like lawyers: the problem of adversary science. *Jurimetr. J.* 33:363–76
- Campbell DT. 1993. Systematic errors to be expected of the social scientist on the basis of a general psychology of cognitive bias. In *Interpersonal Expectations: Theory, Research, and Applications*, ed. PD Blanck, pp. 25–41. Cambridge, UK: Cambridge Univ. Press
- Campbell DT, Stanley JC. 1963. *Experimental and Quasi-Experimental Designs for Research*. New York: Houghton Mifflin
- Chaiken S, Liberman A, Eagly AH. 1989. Heuristic and systematic information processing within and beyond the persuasion context. See Uleman & Bargh 1989, pp. 212–52
- Chiauzzi EJ, Liljegren S. 1993. Taboo topics in addiction treatment: an empirical review of clinical folklore. *J. Subst. Abuse Treat.* 10:303–16
- Cicchetti DV. 1991. The reliability of peer review for manuscript and grant submissions: a cross-disciplinary investigation. *Behav. Brain Sci.* 14:119–86
- Cohen LJ. 1989. *An Introduction to the Philosophy of Induction and Probability*. Oxford, UK: Oxford Univ. Press
- Cole S. 1992. *Making Science: Between Nature and Society*. Cambridge, MA: Harvard
- Cook TD. 1985. Post-positive critical multiplicity. In *Social Science and Social Policy*, ed. L Shotland, MM Mark, pp. 21–62. Beverly Hills, CA: Sage
- Cooper H, Hedges LV. 1994. *The Handbook of Research Synthesis*. New York: Russell Sage Found.
- Costanzo M, White LT, eds. 1994. The death penalty in the United States (special issue). *J. Soc. Issues* 50:1–197
- Cummings KM, Russell S, Gingrass A, Davis R. 1991. What scientists funded by the to-

- bacco industry believe about the hazards of cigarette smoking. *Am. J. Public Health* 81:894-96
- DeKay ML. 1996. The difference between Blackstone-like error ratios and probabilistic standards of proof. *Law Soc. Inq.* 21: 95-132
- Ditto PH, Lopez DF. 1992. Motivated skepticism: use of differential decision criteria for preferred and nonpreferred conclusions. *J. Pers. Soc. Psychol.* 63: 568-84
- Dong BJ, Hauck WW, Gambertoglio JG, Gee L, White JR. 1997. Bioequivalence of generic and brand-name levothyroxine products in the treatment of hypothyroidism. *JAMA* 277:1205-13
- Eagly AH. 1995. The science and politics of comparing women and men. *Am. Psychol.* 50:145-58
- Eagly AH, Carli LL. 1981. Sex of researchers and sex-typed communications as determinants of sex differences in influenceability: a meta-analysis of social influence studies. *Psychol. Bull.* 90:1-20
- Edwards K, Smith EE. 1996. A disconfirmation bias in the evaluation of arguments. *J. Pers. Soc. Psychol.* 71:5-24
- Elms AC. 1988. Freud as Leonardo: Why the first psychobiography went wrong. *J. Pers.* 56:19-40
- Epstein S. 1996. *Impure Science: AIDS, Activism, and the Politics of Knowledge*. Berkeley: Univ. Calif. Press
- Festinger L. 1957. *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford Univ. Press
- Fischer CS, Hout M, Jankowski MS, Lucas SR, Swidler A, Voss K. 1996. *Inequality by Design: Cracking the Bell Curve Myth*. Princeton, NJ: Princeton
- Fischhoff B. 1990. Psychology and public policy: tool or toolmaker? *Am. Psychol.* 45: 647-53
- Fischhoff B, Beyth-Marom R. 1983. Hypothesis evaluation from a Bayesian perspective. *Psychol. Rev.* 90:239-60
- Fiske ST. 1989. Examining the role of intent: toward understanding its role in stereotyping and prejudice. See Uleman & Bargh 1989, pp. 253-83
- Fiske ST, Bersoff DN, Borgida E, Deaux K, Heilman ME. 1991. Social science research on trial: use of sex stereotyping research in Price Waterhouse v. Hopkins. *Am. Psychol.* 46:1049-60
- Foster KR, Bernstein DE, Huber PW. 1993. *Phantom Risk: Scientific Inference and the Law*. Cambridge, MA: MIT Press
- Fox DR. 1993. Psychological jurisprudence and radical social change. *Am. Psychol.* 48:234-41
- Fraser S, ed. 1995. *The Bell Curve Wars: Race, Intelligence, and the Future of America*. New York: Basic Books
- Freedman JL, Sears DO. 1965. Selective exposure. *Adv. Exp. Soc. Psychol.* 2:57-97
- Frey D. 1986. Recent research on selective exposure to information. *Adv. Exp. Soc. Psychol.* 19:41-80
- Friedrich J. 1993. Primary error detection and minimization (PEDMIN) strategies in social cognition: a reinterpretation of confirmation bias phenomena. *Psychol. Rev.* 100:298-319
- Fuchs S, Westervelt SD. 1996. Fraud and trust in science. *Perspect. Biol. Med.* 39:248-70
- Gannon L, Luchetta T, Rhodes K, Pardie L, Segrist D. 1992. Sex bias in psychological research: progress or complacency? *Am. Psychol.* 47:389-96
- Gardner W, Wilcox BL. 1993. Political intervention in scientific peer review: research on adolescent sexual behavior. *Am. Psychol.* 48:972-83
- Gelbspan R. 1997. *The Heat is On: The High Stakes Battle Over Earth's Threatened Climate*. New York: Addison Wesley
- Gergen KJ. 1994. Exploring the postmodern: perils or potentials? *Am. Psychol.* 49:412-16
- Gholson B, Barker P. 1985. Kuhn, Lakatos, and Laudan: applications in the history of physics and psychology. *Am. Psychol.* 40: 755-69
- Gilbert JR, Williams ES, Lundberg GD. 1994. Is there gender bias in *JAMA's* peer review process? *JAMA* 272:139-42
- Gilbert N. 1995. *Was It Rape? An Examination of Sexual Assault Statistics*. Menlo Park, CA: Henry J Kaiser Family Found.
- Giner-Sorolla R, Chaiken S. 1994. The causes of hostile media judgments. *J. Exp. Soc. Psychol.* 30:165-80
- Giner-Sorolla R, Chaiken S. 1997. Selective use of heuristic and systematic processing under defense motivation. *Pers. Soc. Psychol. Bull.* 23:84-97
- Glantz SA. 1996. *The Cigarette Papers*. Berkeley: Univ. Calif. Press
- Gould SJ. 1981. *The Mismeasure of Man*. New York: Norton
- Graham S. 1992. "Most of the subjects were white and middle class": trends in published research on African Americans in selected APA journals 1970-1989. *Am. Psychol.* 47:629-39
- Greenberg J, Solomon S, Pyszczynski T, Steinerberg L. 1988. A reaction to Greenwald, Pratkanis, Leippe, and Baumgardner (1986): Under what conditions does re-

- search obstruct theory progress? *Psychol. Rev.* 95:566–71
- Greenwald AG, Pratkanis AR. 1988. On the use of “theory” and the usefulness of theory. *Psychol. Rev.* 95:575–79
- Greenwald AG, Pratkanis AR, Leippe MR, Baumgardner MH. 1986. Under what conditions does theory obstruct research progress? *Psychol. Rev.* 93:216–29
- Greenwald AG, Schuh ES. 1994. An ethnic bias in scientific citations. *Eur. J. Soc. Psychol.* 24:623–39
- Gross PR, Levitt N. 1994. *Higher Superstition: The Academic Left and Its Quarrels with Science*. Baltimore: Johns Hopkins Univ. Press
- Hambrecht M, Maurer K, Hafner H. 1993. Evidence for a gender bias in epidemiological studies of schizophrenia. *Schizophr. Res.* 8:223–31
- Hammond KR. 1996. *Human Judgment and Social Policy: Irreducible Uncertainty, Inevitable Error, Unavoidable Injustice*. New York: Oxford Univ. Press
- Hammond KR, Harvey LO, Hastie R. 1992. Making better use of scientific knowledge: separating truth from justice. *Psychol. Sci.* 3:80–87
- Harris MJ. 1991. Controversy and cumulation: meta-analysis and research on interpersonal expectancy effects. *Pers. Soc. Psychol. Bull.* 17:316–22
- Hastie R. 1993a. Algebraic models of juror decision processes. See Hastie 1993b, pp. 84–115
- Hastie R, ed. 1993b. *Inside the Juror: the Psychology of Juror Decision Making*. New York: Cambridge Univ. Press
- Hastie R, Rasinski KA. 1988. The concept of accuracy in social judgment. In *The Social Psychology of Knowledge*, ed. D Bar-Tal, AW Kruglanski, pp. 193–208. New York: Cambridge Univ. Press
- Herbert B. 1996. More N. R. A. mischief. *New York Times*, July 5, p. A15
- Herek GM, Jobe JB, Carney R, eds. 1996. *Out in Force: Sexual Orientation and the Military*. Chicago: Univ. Chicago Press
- Herek GM, Kimmel DC, Amaro H, Melton GB. 1991. Avoiding heterosexist bias in psychological research. *Am. Psychol.* 46: 957–63
- Herrnstein RJ, Murray C. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free
- Howson C, Urbach P. 1993. *Scientific Reasoning: The Bayesian Approach*. Chicago: Open Court. 2nd ed.
- Huff D. 1954. *How to Lie with Statistics*. New York: Norton. 142 pp.
- Jarrett RB, Fairbank JA. 1987. Psychologists’ views: APA’s advocacy of and resource expenditure on social and professional issues. *Prof. Psychol. Res. Pract.* 18:643–46
- Johnson D. 1993. The politics of violence research. *Psychol. Sci.* 4:131–33
- Julnes G, Mohr LB. 1989. Analysis of no-difference findings in evaluation research. *Eval. Rev.* 13:628–55
- Kahneman D, Slovic P, Tversky A, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge Univ. Press
- Kalven H, Zeisel H. 1966. *The American Jury*. Boston: Little, Brown
- Kassin SM, Ellsworth PC, Smith VL. 1989. The “general acceptance” of psychological research on eyewitness testimony: a survey of the experts. *Am. Psychol.* 44: 1089–98
- Keltner D, Robinson RJ. 1996. Extremism, power, and the imagined basis of social conflict. *Curr. Direct. Psychol. Sci.* 5:101–5
- Kerr NL. 1993. Stochastic models of juror decision making. See Hastie 1993b, pp. 116–35
- Kerr NL, MacCoun RJ, Kramer G. 1996. Bias in judgment: comparing individuals and groups. *Psychol. Rev.* 103:687–719
- Klayman J, Ha YW. 1987. Confirmation, disconfirmation, and information in hypothesis testing. *Psychol. Rev.* 94:211–28
- Koehler JJ. 1991. Explanation, imagination, and confidence in judgment. *Psychol. Bull.* 110:499–519
- Koehler JJ. 1993. The influence of prior beliefs on scientific judgments of evidence quality. *Organ. Behav. Hum. Decis. Proc.* 56:28–55
- Kruglanski AW. 1989. *Lay Epistemics and Human Knowledge: Cognitive and Motivational Bases*. New York: Plenum
- Kruglanski AW, Webster DM. 1996. Motivated closing of the mind: “seizing” and “freezing”. *Psychol. Rev.* 103:263–83
- Kuhn D, Lao J. 1996. Effects of evidence on attitudes: Is polarization the norm? *Psychol. Sci.* 7:115–20
- Kunda Z. 1990. The case for motivated reasoning. *Psychol. Bull.* 108:480–98
- Laband DN, Piette MJ. 1994. A citation analysis of the impact of blinded peer review. *JAMA* 272:147–49
- Latané B. 1996. Strength from weakness: the fate of opinion minorities in spatially distributed groups. See Witte & Davis 1996, pp. 193–220
- Laudan L. 1990. *Science and Relativism: Controversies in the Philosophy of Science*. Chicago: Univ. Chicago Press

- Laughlin PR. 1996. Group decision making and collective induction. See Witte & Davis 1996, pp. 61–80
- Lee YT, Jussim LJ, McCauley CR, eds. 1995. *Stereotyped Accuracy: Toward Appreciating Group Differences*. Washington, DC: Am. Psychol. Assoc.
- Lerner JS, Tetlock PE. 1994. Accountability and social cognition. *Enycl. Hum. Behav.* 1:1–10
- LeVay S. 1996. *Queer Science: The Use and Abuse of Research into Homosexuality*. Cambridge, MA: MIT Press
- Levinger G, ed. 1986. SPSSI at 50: historical accounts and selected appraisals (special issue). *J. Soc. Issues* 42:1–147
- Liberman A, Chaiken S. 1992. Defensive processing of personally relevant health messages. *Pers. Soc. Psychol. Bull.* 18: 669–79
- Lind EA, Tyler TR. 1988. *The Social Psychology of Procedural Justice*. New York: Plenum
- Linz D, Malamuth N. 1993. *Pornography*. Newbury Park, CA: Sage
- Lord CG. 1989. The “disappearance” of dissonance in an age of relativism. *Pers. Soc. Psychol. Bull.* 15:513–18
- Lord CG, Lepper MR, Preston E. 1985. Considering the opposite: a corrective strategy for social judgment. *J. Pers. Soc. Psychol.* 47:1231–43
- Lord CG, Ross L, Lepper MR. 1979. Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *J. Pers. Soc. Psychol.* 37:2098–109
- Lorge I, Solomon H. 1955. Two models of group behavior in the solution of Eureka-type problems. *Psychometrika* 20:139–48
- MacCoun RJ. 1984. Modeling the impact of extralegal bias and defined standards of proof on the decisions of mock jurors and juries. *Dissert. Abstr. Int.* 46:700B
- MacCoun RJ. 1990. The emergence of extralegal bias during jury deliberation. *Crim. Just. Behav.* 17:303–14
- MacCoun RJ. 1993a. Drugs and the law: a psychological analysis of drug prohibition. *Psychol. Bull.* 113:497–512
- MacCoun RJ. 1993b. Unit cohesion and military performance. In *Sexual Orientation and U. S. Military Personnel Policy: Policy Options and Assessment*, pp. 283–331. Santa Monica, CA: RAND
- MacCoun RJ. 1995. Review of K. R. Foster, D. E. Bernstein, & P. W. Huber (1993). *J. Policy Anal. Manage.* 14:168–71
- MacCoun RJ, Reuter P. 1997. Interpreting Dutch cannabis policy: reasoning by analogy in the legalization debate. *Science* 278:47–52
- MacKay DG. 1988. Under what conditions can theoretical psychology survive and prosper? Integrating the rational and empirical epistemologies. *Psychol. Rev.* 95: 559–65
- Mahoney MJ. 1977. Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cogn. Ther. Res.* 1:161–75
- Maier MH. 1991. *The Data Game: Controversies in Social Science Statistics*. New York: Sharpe
- Marlatt GA, Larimer ME, Baer JS, Quigley LA. 1993. Harm reduction for alcohol problems: moving beyond the controlled drinking controversy. *Behav. Ther.* 24: 461–504
- Mayo DG. 1996. *Error and the Growth of Experimental Knowledge*. Chicago: Univ. Chicago Press
- McGuire WJ. 1983. A contextualist theory of knowledge: its implications for innovation and reform in psychological research. *Adv. Exp. Soc. Psychol.* 16:1–47
- McGuire WJ, McGuire CV. 1991. The content, structure, and operation of thought systems. In *The Content, Structure, and Operation of Thought Systems*, ed. RS Wyer, TK Srull, pp. 1–78. Hillsdale, NJ: Erlbaum
- McHoskey JW. 1995. Case closed? On the John F. Kennedy assassination: biased assimilation of evidence and attitude polarization. *Basic Appl. Soc. Psychol.* 17: 395–409
- Merton RK. 1973. *The Sociology of Science*. Chicago: Univ. Chicago Press
- Miller AG, McHoskey JW, Bane CM, Dowd TG. 1993. The attitude polarization phenomenon: role of response measure, attitude extremity, and behavioral consequences of reported attitude change. *J. Pers. Soc. Psychol.* 64:561–74
- Miller PV. 1995. They said it couldn't be done: the National Health and Social Life Survey. *Public Opin. Q.* 59:404–19
- Monmonier M. 1996. *How to Lie with Maps*. Chicago: Univ. Chicago Press. 2nd ed.
- Moser K, Gadenne V, Schroder J. 1988. Under what conditions does confirmation seeking obstruct scientific progress? *Psychol. Rev.* 95:572–74
- Moskowitz JM. 1993. Why reports of outcome evaluations are often biased or uninterpretable: examples from evaluations of drug abuse prevention programs. *Eval. Plan.* 16:1–9

- Munro GD, Ditto PH. 1997. Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information. *Pers. Soc. Psychol. Bull.* 23: 636–53
- Neisser U, Boodoo G, Bouchard TJ, Boykin AW, Brody N, et al. 1996. Intelligence: knowns and unknowns. *Am. Psychol.* 51: 77–101
- Nemeth CJ. 1986. Differential contributions of majority and minority influence. *Psychol. Rev.* 93:23–32
- Neuliep JW, Crandall R. 1990. Editorial bias against replication research. *J. Soc. Behav. Pers.* 5:85–90
- Neuliep JW, Crandall R. 1993. Reviewer bias against replication research. *J. Soc. Behav. Pers.* 8:21–29
- Nisbet L, ed. 1990. *The Gun Control Debate*. Buffalo, NY: Prometheus Books
- Nisbett RE. 1993. *Rules for Reasoning*. Hillsdale, NJ: Erlbaum
- Nisbett RE, Ross L. 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice Hall
- Pennington N, Hastie R. 1993. The story model of juror decision making. See Hastie 1993b, pp. 192–221
- Peters DP, Ceci SJ. 1982. Peer-reviewed practices of psychological journals: the fate of accepted published articles, submitted again. *Behav. Brain Sci.* 5:187–95
- Petty RE, Cacioppo JT. 1986. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag
- Pezdek K, Banks WP, eds. 1996. *The Recovered Memory/False Memory Debate*. San Diego: Academic
- Platt JR. 1964. Strong inference. *Science* 146: 347–53
- Plous S. 1991. Biases in the assimilation of technological breakdowns: Do accidents make us safer? *J. Appl. Soc. Psychol.* 21: 1058–82
- Popper KR. 1959. *The Logic of Scientific Discovery*. New York: Basic Books
- Porter TM. 1995. *Trust in Numbers: the Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton Univ. Press
- Pyszczynski T, Greenberg J. 1987. Toward an integration of cognitive and motivational perspectives on social inference: a biased hypothesis-testing model. *Adv. Exp. Soc. Psychol.* 20:297–340
- Ray JJ. 1989. The scientific study of ideology is too often more ideological than scientific. *Pers. Individ. Differ.* 10:331–36
- Read SJ, Vanman EJ, Miller LC. 1997. Connectionism, parallel constraint satisfaction processes, and gestalt principles: (Re)Introducing cognitive dynamics to social psychology. *Pers. Soc. Psychol. Rev.* 1:26–53
- Rennie D. 1997. Thyroid storm. *JAMA* 277: 1238–42
- Richards PS, Davison ML. 1992. Religious bias in moral development research: a psychometric investigation. *J. Sci. Study Relig.* 31:467–85
- Rosenthal R. 1994. Science and ethics in conducting, analyzing, and reporting psychological research. *Psychol. Sci.* 5:127–34
- Russett B. 1988. Editor's comment. *J. Confl. Resolut.* 32:773–75
- Saks MJ. 1990. Expert witnesses, nonexpert witnesses, and nonwitness experts. *Law Hum. Behav.* 14:291–313
- Saks MJ. 1993. Improving APA science translation amicus briefs. *Law Hum. Behav.* 17: 235–47
- Sampson EE. 1989. The challenge of social change for psychology: globalization and psychology's theory of the person. *Am. Psychol.* 44:914–21
- Schaie KW. 1988. Ageism in psychological research. *Am. Psychol.* 43:179–83
- Schlenker BR. 1992. Of shape shifters and theories. *Psychol. Inq.* 3:342–45
- Schmidt FL. 1992. What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *Am. Psychol.* 47:1173–81
- Schum DA, Martin AW. 1982. Formal and empirical research on cascaded inference. *Law Soc. Rev.* 17:105–51
- Schwenk CR. 1990. Effects of devil's advocacy and dialectical inquiry on decision making: a meta-analysis. *Organ. Behav. Hum. Decis. Process.* 47:161–76
- Sears DO. 1994. Ideological bias in political psychology: the view from scientific hell. *Polit. Psychol.* 15:547–56
- Shadish WR. 1995. Philosophy of science and the quantitative-qualitative debates: thirteen common errors. *Eval. Program Plan.* 18:63–75
- Slovic P, Fischhoff B. 1977. On the psychology of experimental surprises. *J. Exp. Psychol.: Hum. Percept. Perform.* 3:544–51
- Smith MB. 1994. Selfhood at risk: postmodern perils and the perils of postmodernism. *Am. Psychol.* 49:405–11
- Snyder M. 1981. Seek and ye shall find: testing hypotheses about other people. In *Social Cognition: The Ontario Symposium on Personality and Social Psychology*, ed. ET Higgins, CP Heiman, MP Zanna, pp. 277–303. Hillsdale, NJ: Erlbaum
- Staw BM, Ross J. 1989. Understanding behavior

- ior in escalation situations. *Science* 246: 216–46
- Suedfeld P, Tetlock PE, eds. 1991. *Psychology and Social Policy*. New York: Hemisphere
- Tavris C. 1992. *The Mismeasure of Woman*. New York: Simon & Schuster
- Tetlock PE. 1994. Political psychology or politicized psychology: Is the road to scientific hell paved with good moral intentions? *Polit. Psychol.* 15:509–29
- Tetlock PE, Levi A. 1982. Attribution bias: on the inconclusiveness of the cognition-motivation debate. *J. Exp. Soc. Psychol.* 18:68–88
- Tetlock PE, Mitchell G. 1993. Liberal and conservative approaches to justice: conflicting psychological portraits. In *Psychological Perspectives on Justice: Theory and Applications*, ed. BA Mellers, J Baron, pp. 234–55. New York: Cambridge Univ. Press
- Thagard P. 1992. *Conceptual Revolutions*. Princeton, NJ: Princeton Univ. Press
- Thibaut J, Walker L. 1978. A theory of procedure. *Calif. Law Rev.* 26:1271–89
- Tindale RS, Smith CM, Thomas LS, Filkins J, Sheffey S. 1996. Shared representations and asymmetric social influence processes in small groups. See Witte & Davis 1996, pp. 81–104
- Uleman JS, Bargh JA, eds. 1989. *Unintended Thought*. New York: Guilford
- Vallone RP, Ross L, Lepper MR. 1985. The hostile media phenomenon: biased perception and perceptions of media bias in coverage of the Beirut massacre. *J. Pers. Soc. Psychol.* 49:577–85
- Vaughan D. 1996. *The Challenger Launch Decision*. Chicago: Univ. Chicago Press
- Wallach L, Wallach MA. 1994. Gergen versus the mainstream: Are hypotheses in social psychology subject to empirical test? *J. Pers. Soc. Psychol.* 67:233–42
- Weimer DL, Vining AR. 1992. *Policy Analysis: Concepts and Practice*. Englewood Cliffs, NJ: Prentice Hall. 2nd ed.
- Weiss CH, Bucuvalas MJ. 1980. Truth tests and utility tests: decision-makers' frames of reference for social science research. *Am. Sociol. Rev.* 45:302–13
- Wilson TD, Brekke N. 1994. Mental contamination and mental correction: unwanted influences on judgments and evaluation. *Psychol. Bull.* 116:117–42
- Wilson TD, DePaulo BM, Mook DG, Klaaren KJ. 1993. Scientists' evaluations of research: the biasing effects of the importance of the topic. *Psychol. Sci.* 4:322–25
- Witte E, Davis JH, eds. 1996. *Understanding Group Behavior*, Vol. 1: *Consensual Action by Small Group*. Matwah, NJ: Erlbaum
- Woodward J, Goodstein D. 1996. Conduct, misconduct and the structure of science. *Am. Sci.* 84:479–90
- Yee AH, Fairchild HH, Weizmann F, Wyatt GE. 1993. Addressing psychology's problem with race. *Am. Psychol.* 48:1132–40
- Zuckerman H. 1988. The sociology of science. In *Handbook of Sociology*, ed. NJ Smelser, pp. 511–74. Beverly Hills, CA: Sage