

*Reclaiming Davidson's Methodological Rationalism as Galilean Idealization in Psychology*¹

Carole J. Lee

Philosophy of the Social Sciences 2010

40(1) 84-106

In his early experimental work with Suppes, Davidson adopted rationality assumptions, not as necessary constraints on interpretation, but as practical conceits in addressing methodological problems faced by experimenters studying decision-making under uncertainty. Although the content of their theory has since been undermined, their methodological approach – a Galilean form of methodological rationalism – lives on in contemporary psychological research. This paper draws on Max Weber's *verstehen* to articulate an account of Galilean methodological rationalism; explains how anomalies faced by Davidson's early experimental work gave rise to his later, canonical claims about rationality and interpretation; and, reclaims this Galilean framework for use in contemporary psychological research.

Key words: Galilean models, Galilean idealization, methodological rationalism, Davidson, rationality, radical interpretation, interpretation.

Within the social sciences, methodological debate about the epistemic and metaphysical features of Galilean idealization arise primarily in the context of economic theorizing, which adopts distorting idealizing assumptions about the nature of the market (e.g., that it is in perfect equilibrium) and about the nature of economic agents (e.g., that they are perfectly informed, self-interested, and utility-maximizing). However, less work has been done to articulate the epistemological and metaphysical issues that arise with respect to Galilean idealization, as such, in the context of psychological theorizing. My paper aims to open up this avenue of inquiry by articulating a Galilean account of methodological rationalism embraced by Donald Davidson himself.

Galilean idealization involves the deliberate simplification of the properties attributed to a phenomenon in order to gain a better understanding of the features of the real world or, less ambitiously, of the idealized model at hand. Simplification can involve the omission of causal

¹ Problematic copyediting in the published version – in particular, the location of the tables – makes for confusing reading. With the exception of this footnote, this document represents the original, submitted version.

features not relevant to the purposes at hand: for example, Newton's derivation of Kepler's laws in the *Principia* assumed the sun to be at rest, thus ignoring the slight motion of the sun which results from its attraction to the earth.² Simplification can also take place in the conceptualization of the structural features of the phenomenon: for example, the "ideal plane" idealizes (or, some would say, fictionalizes) the factors relevant to studying inertial motion by replacing the "small perturbations produced by friction" with other factors that "are easier to think about, or with which it is easier to calculate."³

Methodological rationalism is generally identified with principles about how one ought to interpret others. The line goes something like this. In explaining human behavior, we should initially seek to represent what people do as rational: we should impute desires, beliefs, and other mental states so their observed behavior is rational in relation to those mental states, and so their mental states are rational in relation to each other. If we cannot arrive at rationalizing explanations that comport with the evidence, methodological rationalism allows explanations construing others as irrational.

We can interpret methodological rationalism's substantive rationality assumptions in at least a couple of different ways. On the one hand, we can interpret these rationality assumptions as being necessary or constitutive of others' beliefs and desires. This is in line with Davidson's canonical work on interpretation. On the other hand, one can interpret these rationality assumptions as useful, simplifying, and falsifiable devices in formulating interpretations and explanations. Under this Galilean approach – in which we adopt idealized rationality assumptions for methodological purposes – we can create increasingly realistic models of others' beliefs and desires by de-idealizing these assumptions in ways suggested by empirical evidence.

In this paper, I will argue that Davidson once adopted methodological rationalism as a form of Galilean Idealization in psychological theorizing. And, I will argue that this approach to methodological rationalism has continuing relevance in contemporary paradigms of psychological research. In the first part of this paper, I will use Max Weber's account of *verstehen* as a template to articulate how methodological rationalism functions within a research paradigm and the Galilean models to which methodological rationalism gives rise. In the second part, I will argue that, in his experimentally-oriented work with Patrick Suppes in the 1950's,

² Ernan McMullin, "Galilean Idealization," *Studies in History and Philosophy of Science* 16, no. 3 (1985): 258.

³ Nancy Cartwright, *Nature's Capacities and Their Measurement* (Oxford, U.K.: Oxford University Press, 1989), 187.

Davidson articulates a methodological rationalist research paradigm that gives rise to explicitly Galilean models for decision-making under uncertainty.

In the third part of this paper, I will discuss Davidson's response to the predictive failure of Davidson and Suppes's Galilean model for decision-making under uncertainty. This failure inspired Davidson to quit his career as an experimental psychologist and, at the same time, led him to making canonical, paradigm-saving claims about the methodologically *necessary* role of rationality assumptions in interpretation. Davidson continued to maintain these stronger claims about the constitutive rationality of beliefs and desires long after his original model's rationality assumptions were undermined by the heuristics and biases research program.

Given the predictive failure of Davidson and Suppes's model, we might very well ask ourselves why we should embrace methodological rationalism and the Galilean models to which they give rise in interpretation and explanation in psychology. To allay these worries, in the fourth part of the paper, I will discuss contemporary research programs in psychology that may be described as forms of Galilean methodological rationalism. In this section, I will also contrast my account to others in the philosophy of social science literature on interpretation, rationality, and psychological explanation.

1. Weber's Galilean Methodological Rationalism

Max Weber's methodological rationalism is justified from the internal logic of his account of interpretation, *verstehen*, which explains social actions in terms of individuals' intentional states.⁴ For Weber, what distinguishes sociology from the natural sciences is a difference in interest and perspective. Unlike biologists, sociologists aim to explain and give meaning to social action by reference to the meaningful, subjective states of mind or intentions leading to an individual's act.⁵ An action is distinguished from mere behavior insofar as we can impute the acting individual as attaching a subjective meaning – or intention – to it. That action is social insofar as the individual, in acting, takes into account the behavior of others. What Weber ultimately seeks in *verstehen* is “subjectively understandable” interpretation, where an

⁴ Max Weber, *The Theory of Social and Economic Organization*, trans. A. M. Henderson and Talcott Parsons (New York, NY: Oxford University Press, 1947), 88.

⁵ *Ibid.*, 101.

action is “subjectively understandable” so long as we can understand “the subjective ‘states of mind’ of actors.”⁶

The basic object of generalization and explanation in Weber’s account of *verstehen* is his notion of the “pure type.” A “pure type” provides a process-model of how – for an average, typical, pure, or ideal kind of case – a type of act follows from a set of imputed intentional states. In constructing a pure type, we seek a vocabulary and conceptual scheme to pick out causally relevant factors and events responsible for an observed type of act. Our background theories suggest relevant factors or events to observe; the same background theories suggest testable generalizations and hypotheses. These factors are described at a level of abstraction appropriate to the conceptual resources and inductive and explanatory practices of the discipline.⁷

Pure types provide process models of how individual agents’ beliefs and desires caused the social action or event in question. In my reading of Weber, these models are Galilean since they have the following features.⁸ They:

- (a) adopt distorting, falsifiable rationality assumptions in modeling an agent’s beliefs or actions, in order to
- (b) address the problem of massive underdetermination of theory by evidence by finding a way to omit some causal features (thereby reducing the domain of admissible explanations),
- (c) make the statement of one’s *ceteris paribus* generalizations or the formalization of one’s theory more perspicuous, and
- (d) simplify one’s conception of the features of a phenomenon with other factors that are easier to think about or with which it is easier to calculate.

And, these assumptions:

- (e) are adopted as practical methodological conceits, rather than as necessary constraints.

In constructing pure types, Weber explicitly suggests that rationalizing assumptions can deal with the problem of massive underdetermination of theory by evidence. Under Weber’s account of *verstehen*, the sociologist’s methodological burden is to solve the question of how to go about picking out pure types that figure in predictions and causal explanations. But, for any single act, we may tell many complex intentionally-laden stories. There are a seemingly endless number of shifting variables that may be combined in different ways: sensory perceptions,

⁶ Ibid., 88.

⁷ Ibid., 103-4.

⁸ Nancy Cartwright, *How the Laws of Physics Lie* (Oxford, UK: Oxford University Press 1983).

desires, presuppositions, beliefs, plans, commitments, idiosyncratic habits of thought or feeling, prejudices, impressions, and memories.

In response to the problem of numerous, “qualitatively heterogeneous” intentional states, Weber suggests we adopt methodological rationalism as an expedient, simplifying conceit.⁹ Attributing rational intentions and actions to agents simplifies process models by omitting causal features that do not serve as rational reasons for an agent’s action. In cases of irrational action, methodological rationalism simplifies the identification and formulation of process models by describing irrational action as a deviation from the rational course, where this deviation can be attributed to a single or small set of factors, such as an agent’s errors or emotions.¹⁰ Thus, Weber’s assumptions about rational versus irrational belief (b) addresses the problem of massive underdetermination of theory by evidence by finding a way to omit some causal features (thereby reducing the domain of admissible explanations).

Weber is also explicit about the methodological preference for methodological rationalism insofar as it (c) makes the statement of *ceteris paribus* generalizations more perspicuous. In cases of rational action, rationalizing pure types, though “abstract and unrealistic,” are thereby able to “perform its methodological functions in formulating the clarification of terminology, and in the formulation of classifications, and of hypotheses.”¹¹ These classificatory schemes are “better” because they have “the merit of clear understandability and lack of ambiguity.”¹²

In cases of irrational action, these rationalizing pure types (d) simplify one’s conception of irrational action so as to make irrational action easier to think about and analyze empirically. In cases of irrational action, “it is convenient to treat all irrational, effectually determined elements of behavior as factors of deviation from a conceptually pure type of rational action” for “the purposes of a typological scientific analysis.”¹³ For example, “panic on the stock exchange can be most conveniently analyzed by attempting to determine first what the course of action would have been if it had not been influenced by irrational affects; it is then possible to introduce the irrational components as accounting for the observed deviations from this hypothetical

⁹ Weber, *The Theory of Social and Economic Organization*, 111.

¹⁰ *Ibid.*, 92.

¹¹ *Ibid.*, 111.

¹² *Ibid.*, 92.

¹³ *Ibid.*

course.”¹⁴ For Weber, it is only by describing irrational action as “deviations from this hypothetical course” that we are able to “assess the causal significance of irrational factors as accounting for the deviations from this type.”¹⁵ These rational and irrational process models serve as *ceteris paribus* generalizations about intentional states and the actions they cause. Unfortunately, empirically testing such process models is “feasible with relative accuracy only in the few very special cases susceptible of psychological experimentation.”¹⁶

Weber’s rationalizing assumptions, as portrayed above, are (e) adopted as methodological conceits, and not as necessary requirements for the possibility of intentional explanation. Weber explicitly states that it is “for these reasons of methodological convenience” that we adopt methodological rationalism. Methodological rationalism is only “a methodological device.”¹⁷ Not only does it “not involve a belief in the actual predominance of rational elements in human life,” it also, by implication, does not require the general rationality of others’ intentional states.¹⁸

Weber explicitly claims that, whatever rationality assumptions one adopts, they are (a) admitted distortions that are empirically falsifiable. He observes that because pure types are abstract and ideal, “it is probably seldom if ever that a real phenomenon can be found which corresponds exactly to one of these ideally constructed pure types.”¹⁹ A pure type might be idealized and distorted in the sense that the process model refers, not to a particular actor or particular case, but to a hypothetical actor with averaged, approximated, or idealized intentions. As such, this pure type need not “refer to an objectively ‘correct’ meaning or one which is ‘true’ in some metaphysical sense.”²⁰ For example, an economic pure type can be formulated in terms of “economic ends alone,” rather than include political, cultural, or other sociological motives.²¹ However, this kind of simplification is a distorting one since “in the majority of cases of action important in history or sociology[,] the motives which determine it are qualitatively heterogenous” (e.g., not involving economic motivations alone).²² Thus, Weber admits that “it is

¹⁴ Ibid.

¹⁵ Ibid.

¹⁶ Ibid., 97. It is important to note that my account of Galilean methodological rationalism is not committed to Weber’s particular claim about how to interpret irrational beliefs and actions. The core part of my account is instead captured by features (a) – (e) articulated in the beginning of this section.

¹⁷ Ibid., 92.

¹⁸ Ibid.

¹⁹ Ibid., 110-1.

²⁰ Ibid., 89.

²¹ Ibid., 97.

²² Ibid., 111.

probably seldom if ever that a real phenomenon can be found which corresponds exactly to one of these ideally constructed pure types.”²³ Tellingly, for Weber, methodological rationalism’s simplifications are legitimate since they are “similar” to “the assumption of an absolute vacuum” in explaining or modeling a physical or chemical reaction.²⁴

2. Davidson’s Galilean Roots: Subjectively Expected Utility Theory

Von Neumann and Morgenstern’s epoch-making book, *Theory of Games and Economic Behavior*, axiomatized expected utility theory in a way that made possible calculations of the expected utility of actions/lotteries in terms of utilities attached to individual states of nature.²⁵ A decade after the first edition of this book, Ward Edwards, often described as the founder of psychological research on judgment and decision-making under uncertainty, attributed von Neumann and Morgenstern’s book as the true beginning of “the modern period in the study of risky decision-making.”²⁶ In 1954, Edwards published an article about von Neumann and Morgenstern’s and earlier economic models of decision-making in a prominent psychology journal, as a way to introduce key theoretical ideas to psychologists for empirical investigation.

At this time, Edwards took the psychologically important question to be whether or not such a model could be used to account for simple experimental examples of risky decisions.²⁷ He recounted economists’ and psychologists’ attempts to formulate a theory of expected utility that would be both normatively and descriptively adequate. Some challenged von Neumann and Morgenstern’s method of utility measurement, which they had simplified as the problem of measuring objective monetary amounts. For example, economists Milton Friedman and Leonard Savage observed that a person who buys insurance (and thus prefers to pay in order not to take risks), will also buy lottery tickets (suggesting a preference to pay in order to take risks). In light of these different sensitivities to risk for monetary losses and gains, they proposed a doubly inflected subjective utility curve for money.²⁸ Edwards himself discovered that participants

²³ Ibid., 110.

²⁴ Ibid.

²⁵ John von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior*, 1 ed. (Princeton, NJ: Princeton University Press, 1944).

²⁶ Ward Edwards, "The Theory of Decision Making," *Psychological Bulletin* 51, no. 4 (1954): 392.

²⁷ Ibid.: 394.

²⁸ Milton Friedman and Leonard J. Savage, "The Utility Analysis of Choices Involving Risk," *The Journal of Political Economy* 56, no. 4 (1948).

preferred bets involving some probability values over others – a finding suggesting that choices between bets are not determined solely by the expected amounts of money involved.²⁹

Davidson and Suppes challenged a different feature of von Neumann and Morgenstern's account. All of the theorists mentioned above took for granted that objective probabilities could be associated with subjective probabilities. However, Davidson and Suppes took this assumption to be "questionable or at least undesirable" since such an assumption required: (1) that there be such a thing as objective probabilities, (2) that agents know what these objective probabilities are, (3) that there be a behaviorist method for testing whether an agent is acting in accordance with objective probabilities, and (4) that there be an explanation invoking objective probabilities for why an agent acts, even when these objective probabilities are not known to the agent.³⁰ For them, a satisfactory empirical interpretation of von Neumann and Morgenstern's account would not only deal with subjective utilities, but also subjective probabilities – a move that aligned them with Frank Ramsey's conceptualization and methodology for measuring subjective utility and subjective probability.³¹

Davidson's early work with Suppes sought to provide a normatively and descriptively adequate account.³² In doing so, they adopted a Galilean interpretation of methodological rationalism. In their models, it was assumed that agents' choices maximize their expected utility, that subjective utilities remain invariant across logically identical outcomes and gambles, and that agents' preferences conform to the transitivity axiom. To get a sense of the methodological role that these rationality assumptions served, we need to take a detailed look at their conception of the goals and problems faced by experimenters studying decision under uncertainty.

According to Davidson and Suppes, experimenters face a problem of epistemic interdependence between subjective probabilities and subjective utilities: "[g]iven the agent's beliefs (his subjective probabilities) it's easy to compute his relative values from his choices;

²⁹ Ward Edwards, "The Reliability of Probability-Preferences," *American Journal of Psychology* 67, no. 1 (1954), ———, "Probability-Preferences in Gambling," *American Journal of Psychology* 66, no. 3 (1953).

³⁰ Donald Davidson, Patrick Suppes, and Sydney Siegel, *Decision Making: An Experimental Approach* (Stanford, CA: Stanford University Press, 1957), 11.

³¹ Frank P. Ramsey, "The Foundations of Mathematics," *Proceedings of the London Mathematical Society* 25 (1925).

³² Davidson, Suppes, and Siegel, *Decision Making: An Experimental Approach*, 3.

given his values, we can infer his beliefs. But given only his choices, how can we work out both his beliefs and his values?”³³

In order to solve the problem of the epistemic interdependence of subjective probabilities and subjective utilities, Davidson suggests that theorists employ a version of Ramsey’s “method for abstracting subjective values and probabilities simultaneously from choice behaviour.”³⁴ We begin by empirically discovering a case in which an agent is indifferent between two gambles, where these gambles have a special type of formal relationship. In this special choice task, the events are the occurrence of some event E and its non-occurrence $\text{not-}E$; there are only 2 possible payoff values; and, these payoff values both show up in each of the two gambles, though the payoffs are reversed for E and $\text{not-}E$. For example, suppose we have the following choice between Gamble One and Gamble Two. In Gamble One, the agent receives \$11.00 if E occurs, and receives nothing if E does not occur. In Gamble Two, the agent receives \$ 0.00 if E occurs, and receives \$11.00 if E does not occur.

| | Gamble One | Gamble Two |
|----------------|------------|------------|
| E | \$11.00 | \$ 0.00 |
| $\text{not-}E$ | \$ 0.00 | \$11.00 |

Here, we do not know what the agent’s subjective utility is for gaining \$11.00 versus gaining \$0.00. However, we can get a toehold into assigning the subject’s subjective utilities and subjective probabilities if we assume that the subjective utility values attached to the objective payoffs \$11.00 and \$0.00 are identical for each gamble. That is, we will assume that the outcomes have the following subjective utilities attached to them:

| | Gamble One | Gamble Two |
|----------------|------------|------------|
| E | $u(x)$ | $u(y)$ |
| $\text{not-}E$ | $u(y)$ | $u(x)$ |

³³ Donald Davidson, "Belief and the Basis of Meaning (1974)," in *Inquiries into Truth and Interpretation* (New York, NY: Oxford University Press, 2001), 145.

³⁴ ———, "Psychology as Philosophy (1974)," in *Essays on Actions and Events* (Oxford, UK: Oxford University Press, 1980), 236.

This turns the individual's expected utility calculations into the following:

$$\text{Expected Utility}_{(\text{Gamble One})}: p(E)u(x) + p(\text{not-}E)u(y)$$

$$\text{Expected Utility}_{(\text{Gamble Two})}: p(E)u(y) + p(\text{not-}E)u(x)$$

With these expected utility calculations in hand, we can infer that the agent's total expected utility of taking Gamble 1 equals her expected utility of taking Gamble 2 in cases where she believes that the probability of E equals the probability of not- E . That is, for cases where $p(E) = p(\text{not-}E)$:

$$\begin{aligned} \text{Expected Utility of Gamble 1} &= \text{Expected Utility of Gamble 2} \\ p(E)u(x) + p(\text{not-}E)u(y) &= p(E)u(y) + p(\text{not-}E)u(x) \end{aligned}$$

This is so because, if the agent thought that E was either more or less likely to occur than not, then she would have preferred the gamble resulting in the higher expected utility.

Once researchers have identified the agent's probabilities for E and not- E , they can estimate the agent's other subjective utility values relative to x and y . Let's say that the agent is faced with the following Gambles, where $u(w)$ and $u(z)$ are attached to two new monetary outcomes:

| | Gamble 3 | Gamble 4 |
|----------|----------|----------|
| E | $u(x)$ | $u(w)$ |
| not- E | $u(y)$ | $u(z)$ |

Again, let's assume that the agent's choices maximize her expected utility. And, let's assume that the agent's values for $u(x)$ and $u(y)$ have not changed since the last choice task. If we discover that the agent is indifferent between Gambles 3 and 4, we can infer that the difference in utility between x and z is equal to the difference in utility between y and w . This follows from the fact that when $p(E) = p(\text{not-}E)$, the expected utility calculation simplifies to: $u(x) + u(y) = u(w) + u(z)$, which is equivalent to $u(x) - u(z) = u(w) - u(y)$. If we continue to study gambles involving the equiprobable events E and not- E , then we can continue this process to "scale" the

agent's other subjective utility values for different monetary outcomes, so long as we assume that her subjective utility values remain invariant across monetary outcomes and logically identical gambles, and are transitive with respect to each other. Once we have established utility values for some finite set of outcomes, we can determine the subjective probabilities of a finite set of events (not just events *E* and not-*E*) by dealing only with choices with outcomes whose subjective utility values are known. However, we can only scale an agent's subjective utilities and map her utility function for various monetary outcomes if and only if those subjective utilities obey the transitivity axiom.

So, Davidson deals with the epistemic interdependence of subjective utilities and subjective probabilities by assuming that choices under uncertainty maximize an agent's expected utility and that her subjective utilities and probabilities remain constant across logically identical outcomes and gambles. These assumptions address the underdetermination of theory by evidence by ruling out interpretations that would assign changing subjective utilities and probabilities to an agent. This approach rules out interpretations that involve violations of the transitivity axiom. And, more generally, this approach rules out interpretations attributing non-utility related reasons for preferring one gamble over another. Thus, Davidson's application of Ramsey's theory (b) addresses the problem of massive underdetermination of theory by evidence by finding a way to omit some causal features. And, Davidson (d) simplifies our conception of the features of these phenomena with factors that are easier to think about and with which it is easier to calculate an agent's subjective utilities and subjective probabilities. In this scheme, expected utility's axioms and rules become "convenient ways of summarizing facts about the structure of basic preferences."³⁵

These simplifications of the phenomena (c) make the formalization of one's theory of interpretation more perspicuous, since it allows for the development of a scale of subjective utilities and probabilities. Rather than deal with absolute utilities or probabilities, Davidson's approach provides numerical assignments for the sake of measuring "comparisons of differences."³⁶ By creating a relative scale of values, Davidson suggests a deeply holist approach to assigning utilities and probabilities. When these utilities are measured in the context of gambles with monetary outcomes, Davidson's assumptions about the fixed nature of an agent's

³⁵ Davidson, "Belief and the Basis of Meaning (1974)," 146.

³⁶ *Ibid.*, 147.

subjective utilities allows us to graph an agent's utility function for different monetary outcomes.³⁷ If Davidson's approach allowed subjective utilities for logically identical outcomes and gambles to change, such a function would be difficult to identify, since we could not infer subjective utility values relative to others that remain fixed between logically identical outcomes and gambles. And, if his approach allowed for intransitive subjective utilities, this function would be impossible to define.

What's most surprising about this work is that Davidson embraced the possibility that his account could (a) adopt distorting, falsifiable rationality assumptions as (e) practical methodological conceits, rather than as necessary constraints. The formal model for decision-making is given "an operational interpretation;" and, the model's hypotheses are simply "logical consequences of the model and interpretation." If the model provides true predictions, then this raises our confidence in the model and interpretation. However, if the model provides false predictions, then "the *conjunction* of formal model and empirical interpretation. . . must be rejected."³⁸ And, the theorists may choose to either "revise the formal model or the empirical interpretation, or both."³⁹ In this passage, Davidson and Suppes seem to admit the possibility of rejecting expected utility theory as a descriptively accurate theory of human choice. And, they acknowledge that their assumptions about the stability of subjective utilities and subjective probabilities, as well as the transitivity of preferences, are distorted: they are "greatly simplified," the "cost" for which is to remove the model "a step from the empirical circumstances of its application."⁴⁰ For them, "whether these conditions are empirically satisfied" by an agent's preferences "is testable by observing his decisions."⁴¹ So, Davidson and Suppes seem open to the possibility that these assumptions (a) are adopted as distorting, falsifiable rationality assumptions.

This is not to say that the rejection of these assumptions would not make interpretation extremely challenging. Under this methodology, "[u]nless a man's preferences remain stable over the time it takes to test the hypothesis that his preferences have a certain structure, it is difficult to know how to test the hypothesis."⁴² This is because, if these assumptions were

³⁷ Davidson, Suppes, and Siegel, *Decision Making: An Experimental Approach*, 63-5, 76-7.

³⁸ *Ibid.*, 4-5.

³⁹ *Ibid.*, 5.

⁴⁰ *Ibid.*, 40.

⁴¹ *Ibid.*

⁴² *Ibid.*

undermined, it would be difficult to get a toehold into interpretation. However, in this context, Davidson and Suppes treat this as a practical problem rather than a deep fact about subjective utilities and subjective probabilities.

3. The Fate of Davidson's Galilean Models

Not surprisingly, Davidson and Suppes's model faced various empirical anomalies. However, as I will explain, it was not the anomalies that undermined their research program. Rather, it was the empirical falsification of their substantive rationality assumptions, coupled with a new research paradigm's success in identifying and explaining new phenomena.

In "Psychology as Philosophy," Davidson remarks that, in his experimental work with Merrill Carlsmith, they discovered that "as time went on, people became steadily more consistent; intransitivities were gradually eliminated; after six sessions, all subjects were close to being perfectly consistent. This was enough to show that a static theory like Ramsey's could not, even under the most carefully controlled conditions, yield accurate predictions: merely making choices (with no reward or feedback) alters future choices." They also discovered that "[i]f the choices of an individual over all trials were combined, on the assumption that his 'real' preference was for the alternative of a pair he chose most often, then there were almost no inconsistencies at all. Apparently, from the start there were underlying and consistent values which were better and better realized in choice. I found it impossible to construct a formal theory that could explain this, and gave up my career as an experimental psychologist."⁴³

From a methodological point of view, this anomaly is not sufficient to undermine the entire research program: participants' increasingly consistent preferences lends credence to the assumption that preferences respect transitivity, at least in the longer term. The localized, predictive inaccuracy of Davidson's model would not be so surprising if he embraced the Galilean features of his approach: he might explain these local, predictive failures by observing that his conceptualization of subjective utilities and subjective probabilities are idealized in ways that make their identification and mathematical scaling easier to calculate, and that the model's assumptions do not apply to every decision under uncertainty.

However, instead of drawing these kinds of conclusions, Davidson moves to a transcendental argument in which rationality assumptions are necessary for the very possibility

⁴³ Davidson, "Psychology as Philosophy (1974)," 236.

of belief and desire attribution. In this framework, his surprising empirical findings serve as affirmation of the general or overall rationality of beliefs and desires.

“The significance of the experiment I described a page or so back is that it demonstrates how easy it is to interpret choice behavior so as to give it a consistent and rational pattern. When we learn that apparent inconsistency fades with repetition but no learning, we are apt to count the inconsistency as merely apparent. When we learn that frequency of choice may be taken as evidence for an underlying consistent disposition, we may decide to write off what seem to be inconsistent choices as failures of perception or execution.”⁴⁴

In this later work, Davidson is not willing to admit the possibility that these assumptions are distorting or falsifiable assumptions, contrary to (a). This is because, for Davidson, the very possibility of attributing subjective utilities and probabilities to an agent requires adopting such assumptions. In later Davidson, these assumptions are constitutive of the science of decision theory itself. And, violations of these assumptions would render impossible the attribution of subjective utilities and subjective probabilities.

From a formal point of view, the situation is analogous to fundamental measurement in physics, say of length, temperature, or mass. The assignment of numbers to measure any of these assumes that a very tight set of conditions holds. And I think that we can treat the cases as parallel in the following respect. Just as the satisfaction of the conditions for measuring length or mass may be viewed as constitutive of the range of application of the sciences that employ these measures, so the satisfaction of conditions of consistency and rational coherence may be viewed as constitutive of the range of applications of such concepts as those of belief, desire, intention and action. It is not easy to describe in convincing detail an experiment that would persuade us that the transitivity of the relation of *heavier than* had failed. Though the case is not as extreme, I do not think we can clearly say what should convince us that a man at a given time (without change of mind) preferred *a to b*, *b to c*, and *c to a*. The reason for our difficulty is that we cannot make good sense of an attribution of preference except against a background of coherent attitudes.⁴⁵

Davidson seems to admit that there is some kind of distortion introduced in the attribution of subjective utilities and subjective probabilities when he says that “the constitutive force in the realm of behaviour” requires that we “*warp the evidence* to fit this frame.”⁴⁶ However, these “warping” assumptions are not, contrary to (e), adopted as practical methodological conceits. Rather, they are considered necessary constraints on the empirical study of choice under uncertainty.

It is this later transcendental move that got taken up into Davidson’s canonical work on radical translation. In his later work, Davidson directly likens the problems and methods in radical interpretation to those in psychological research on choice under uncertainty.⁴⁷ The

⁴⁴ Ibid., 237.

⁴⁵ Ibid., 236-7.

⁴⁶ Italics mine. Ibid., 239.

⁴⁷ Ibid, ———, “Belief and the Basis of Meaning (1974).”

analogies run deep: as in the case of decision under uncertainty, theorists in radical interpretation are faced with a problem of epistemic interdependence – this time between an agent’s beliefs and the meaning of her utterances. As in the case of decision under uncertainty, the task of radical translation is to deliver a theory of both simultaneously. In choice situations, Davidson deals with the epistemic interdependence of subjective utilities and subjective probabilities by assuming that choices under uncertainty maximize an agent’s expected utility and that her subjective utilities and probabilities remain invariant across logically identical choice situations; and, how strong these preferences are get scaled in reference to other preferences with the help of the transitivity assumption. In radical interpretation, the theorist simultaneously delivers belief and meaning by assuming that the speaker’s beliefs and meanings are true; and, which truth conditions constitute the meaning of an utterance gets determined by looking to the agent’s body of beliefs and meanings to create a rational, coherent interpretation. In both cases, these rationality assumptions allow us to infer an agent’s other intentional states by holding various rationality assumptions constant. And, this serves to eliminate interpretations involving radical irrationality.

Most importantly for our purposes here: Davidson’s account of radical interpretation is like his later account of decision in that it claims that norms of truth and rationality inherent in the interpretive stance only allow us to discover that the beliefs, desires, and utterances of others – insofar as they are interpretable at all – are generally or mostly true and rational. “[W]e have no reason to count” an individual as “having beliefs, or as saying anything” if we cannot interpret her as being mostly rational.⁴⁸

Unfortunately for both early and later Davidson, empirical research in the early 1980’s demonstrated that key rationality assumptions of the original decision model – namely, the invariance and transitivity of preferences – were false. These results were discovered by Daniel Kahneman and Amos Tversky, thanks to their decision to move back to associating subjective probabilities with objective ones. They got around the problem of the epistemic interdependence of subjective probabilities and subjective utilities by moving towards a model using objective probabilities: here, experimenters took the single-event probabilities described in the decision problems to be identical to participants’ beliefs about the probabilities in question.

⁴⁸ Donald Davidson, "Radical Interpretation (1973)," in *Inquiries into Truth and Interpretation* (New York, NY: Oxford University Press, 2001), 137.

This methodological move was crucial in the heuristics and biases's eventual success. By embracing objective probabilities, Kahneman and Tversky were freed from having to come up with a solution for dealing with the epistemic interdependence of subjective probabilities and subjective utilities. From the perspective of Davidson and Suppes's model, this meant that Kahneman and Tversky were freed from needing to identify decision situations in which the agent was indifferent between two options. Davidson and Suppes operationalized the notion of "indifference" in the following way: "the relation of indifference between two options" holds for Davidson and Suppes, "if and only if each is weakly preferred to the other."⁴⁹ An agent is said to "weakly prefer" a gamble when she chooses it over the other. So, if an agent's preferences change or reverse between choices, despite being faced with a logically identical set of gambles, Davidson and Suppes would interpret this reversal as the agent's indifference between the gambles. Kahneman and Tversky's model, however, has no need for this operationalized notion of indifference and can describe such cases as reversals of preference. This, of course, allowed them to discover the framing effect, in which agents' preferences reverse across logically identical gambles.⁵⁰ And, by relying on objective probabilities instead of using the transitivity axiom to scale preferences and subjective probabilities, Kahneman and Tversky were able to discover and articulate the endowment effect, in which agents' preferences violate the transitivity axiom.⁵¹

These violations of *a priori* rules of rationality contributed to Kahneman and Tversky's catalogue of irrational biases, many of which were published as a collection in their canonical *Science* paper in 1974.⁵² Their research program shifted the disciplinary focus away from theories that were both normative and descriptive, to theories that could explain irrational judgment. This can be seen in citation patterns in the decade that followed: although articles

⁴⁹ Davidson, "Belief and the Basis of Meaning (1974)," 13.

⁵⁰ Kahneman and Tversky's framing effects literature demonstrates that agents systematically reverse their preferences when a gamble is described in a negative frame versus a positive frame (and vice versa). This reversal violates the invariance principle – the principle that preferences should be invariant under logically equivalent descriptions of the acts, outcomes, and states of nature. Amos Tversky and Daniel Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science* 211, no. 4481 (1981): 453.

⁵¹ The endowment effect, where people demand more to give up an object than they are willing to pay to acquire it, can be manipulated so as to induce subjects to express intransitive preferences. Daniel Kahneman, Jack L. Knetsch, and Richard H. Thaler, "Experimental Tests of the Endowment Effect and the Coase Theorem," *The Journal of Political Economy* 98, no. 6 (1990): 1339.

⁵² Amos Tversky and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science* 185, no. 4157 (1974).

reporting rational and irrational judgment were published at similar rates,⁵³ studies on irrational judgment were more cited at a six to one ratio.⁵⁴ The impact of these results extended into other social scientific disciplines, including research in medical diagnosis, law, economics, management science, and political science.⁵⁵

By this time, Davidson had already made the shift to his stronger claims about the constitutive rationality of beliefs and desires in interpretation and explanation. However, rather than embrace his previous, Galilean approach to rationality assumptions in psychological theorizing, Davidson's subsequent claims have all of the trappings of a paradigm of research confronted by anomaly.

As Kuhn observes, the defenders of a research paradigm faced with anomaly "will devise numerous articulations and *ad hoc* modifications of their theory in order to eliminate any apparent conflict."⁵⁶ Although Davidson does not address or respond to the heuristics and biases literature directly, in the early and mid 1980's, he responds indirectly by developing a somewhat ad hoc account to accommodate some forms of irrationality. To accomplish this, Davidson adopts a Freudian style framework: here, the mind is "partitioned into quasi-independent structures" that can fail to be transparent to each other.⁵⁷ This partitioning allows one part of the mind to have beliefs, desires, or intentions that are inconsistent or irrational with respect to those of another. However, Davidson continues to maintain that "a part of the mind must show a

⁵³ L. Lola Lopes, "The Rhetoric of Irrationality," *Theory and Psychology* 1, no. 1 (1991).

⁵⁴ Jay J. J. Christensen-Szalanski and Lee Roy Beach, "The Citation Bias: Fad and Fashion in the Judgment and Decision Literature," *American Psychologist* 39 (1984).

⁵⁵ M. H. Bazerman, *Judgment in Managerial Decision Making* (New York, NY: Wiley, 1990), W. Casscells, A. Schoenberger, and T. B. Grayboys, "Interpretation by Physicians of Clinical Laboratory Results," *New England Journal of Medicine* 299 (1978), David M. Eddy, "Probabilistic Reasoning in Clinical Medicine," in *Judgment under Uncertainty*, ed. Daniel Kahneman, Paul Slovic, and Amos Tversky (Cambridge, UK: Cambridge University Press, 1982), Arthur S. Elstein, Lee S. Shulman, and Sarah A. Sprafka, "Medical Problem Solving: A Ten-Year Retrospective," *Evaluation and the Health Professions* 13, no. 1 (1990), Daniel Kahneman and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47, no. 2 (1979), R. B. Korobkin and T. S. Ulen, "Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics," *California Law Review* 88 (2000), Michael J. Saks and Robert F. Kidd, "Human Information Processing and Adjudication: Trial by Heuristics," *Law and Society Review* 15, no. 1 (1980), P. M. Sniderman, R. A. Brody, and Philip E. Tetlock, *Reasoning and Choice: Explorations in Political Psychology* (New York, NY: Cambridge University Press, 1991), C. R. Sunstein, "Behavioral Analysis of Law," *University of Chicago Law Review* 64 (1997), William C. Thompson and Edward L. Schumann, "Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy," *Law and Human Behavior* 11 (1987), Amos Tversky and Daniel Kahneman, "Rational Choice and the Framing of Decisions," *The Journal of Business* 59, no. 4, Part 2: The Behavioral Foundations of Economic Theory (1986), Tversky and Kahneman, "The Framing of Decisions and the Psychology of Choice."

⁵⁶ Thomas S. Kuhn, *The Structure of Scientific Revolutions* (Chicago, IL: Chicago University Press, 1962), 78.

⁵⁷ Donald Davidson, "Paradoxes of Irrationality (1982)," in *Problems of Rationality*, ed. Donald Davidson (Oxford, U.K.: Oxford University Press, 2004), 180-1.

larger degree of consistency or rationality than is attributed to the whole” since “large deviations from reality or consistency begin to undermine our ability to describe and explain what is going on in mental terms.”⁵⁸ By partitioning the mind in this way, Davidson allows the mind, as a collectivity of partitions, to have irrational beliefs, desires, and intentions, while still preserving the overall, constitutive rationality of an important partition of the mind. Although different partitions of the mind can have the contradictory beliefs “*p*” and “not-*p*,” he claims it is impossible for the conscious part of the mind to attend to the contradiction “*p* and not-*p*” directly.⁵⁹

The plausibility of Davidson’s revised account rests on its ability to accommodate the empirical results of the heuristics and biases research program. Under Davidson’s modified account, one must construe preference reversals and intransitivities as the handiwork of different partitions of the mind. However, Davidson failed to pursue this line of argumentation in later work. Instead, Davidson seems to drop this approach altogether in favor of simply restating his commitment to the constitutive rationality of thought and action. “[W]e cannot recognize as thought phenomena that are too far out of line” with “logic” and “decision theory.”⁶⁰ Interpretation “is built on the norms of rationality,”⁶¹ which “leaves no room for irrationality, and therefore is powerless to explain it.”⁶² “[C]eteris paribus, the laws of decision theory do describe how people make real choices.”⁶³ Like the measurement of length, “[i]f the law of transitivity fails in a single case, the entire theory. . . is false,” and we are not justified talking of preferences or desires at all.⁶⁴

Unfortunately for Davidson, anomalies have the power to lead to the rejection of a paradigm when “an alternative candidate is available to take its place.”⁶⁵ Davidson and Suppes’s

⁵⁸ Ibid.

⁵⁹ ———, “Incoherence and Irrationality (1985),” in *Problems of Rationality*, ed. Donald Davidson (Oxford, U.K.: Oxford University Press, 2004), 198.

⁶⁰ ———, “Representation and Explanation (1990),” in *Problems of Rationality*, ed. Donald Davidson (Oxford, U.K.: Oxford University Press, 2004), 97-9.

⁶¹ ———, “Could There Be a Science of Rationality? (1995),” in *Problems of Rationality*, ed. Donald Davidson (Oxford, U.K.: Oxford University Press, 2004), 130.

⁶² Ibid., 133.

⁶³ ———, “What Thought Requires (2001),” in *Problems of Rationality*, ed. Donald Davidson (Oxford, U.K.: Oxford University Press, 2004), 146.

⁶⁴ Ibid., 148.

⁶⁵ Kuhn, *The Structure of Scientific Revolutions*, 77.

model, as well as Ward Edward's model of "man as an intuitive statistician,"⁶⁶ were overtaken by the heuristics and biases research program – a research program that undermined the substantive rationality assumptions built into these models and that identified new phenomena their accounts could better explain.

4. Galilean Methodological Rationalism in Psychology

Davidson and Suppes's model fails to provide a *how-actually* style explanation since these static models cannot explain dynamic choice behavior over time. Nor can their models recognize reversals of preference or intransitivities as such. In the face of the predictive failure of Davidson's model, we might very well wonder: why would we *ever* want to use methodological rationalism to create Galilean models of human judgment and decision-making?

Clearly, Davidson's particular rationality assumptions have been empirically undermined. However, I think Davidson's account of Galilean methodological rationalism provides an early prototype for the role that rationality assumptions serve in contemporary, Galilean models of human cognition. These research paradigms adopt substantive rationality assumptions in order to serve methodological roles within particular paradigms of research.

For example, evolutionary psychology exemplifies a form of Galilean methodological rationalism. Evolutionary psychologists engaged in adaptive research programs adopt the distorting, falsifiable notion of ecological rationality.⁶⁷ For evolutionary psychologists, a cognitive process is said to be ecologically rational when it produces mostly true beliefs when employed in adaptive contexts of reasoning. Adaptationist research programs adopt the idealized, distorting assumption that cognitive processes are ecologically rational. This notion of ecological rationality invites psychologists to hypothesize about the problems and information contexts to which our cognitive processes adapted. Information contexts constrain the kinds of representations that our cognitive processes are said to be sensitive to; while, adaptive problems help theorists think about how our minds might have solved those problems given the

⁶⁶ Ward Edwards, "Nonconservative Information Processing Systems," (Ann Arbor, MI: University of Michigan, Institute of Science and Technology, 1966).

⁶⁷ For examples, see Leda Cosmides and John Tooby, "Are Humans Good Intuitive Statisticians after All? Rethinking Some Conclusions from the Literature on Judgment under Uncertainty," *Cognition* 58 (1996), Gerd Gigerenzer and Ulrich Hoffrage, "How to Improve Bayesian Reasoning without Instruction: Frequency Formats," *Psychological Review* 102, no. 4 (1995).

informational (and representational) constraints.⁶⁸ This “fit” between cognitive processes and environmental cues explains how it is that the mind could arrive at accurate judgments on the basis of a few, robust cues, rather than engage forms of information search and inference that involve more resources, including time, attention, and computational ability. “This exploration of boundedly rational heuristics involves (1) designing computational models of candidate simple heuristics, (2) analyzing the environmental structures in which they perform well, (3) testing their performance in real-world environments, and (4) determining whether and when people (and other animals) really use these heuristics. . . . The different stages of this research program rest on multiple methods, including theoretical modeling of heuristics, computer simulation of their performance, mathematical analysis of the fit between heuristics and specific environments, and laboratory experimentation.”⁶⁹

The research program’s assumption about natural selection’s truth-optimizing nature is (a) distorting and false.⁷⁰ However, by framing cognition in these adaptive terms, the notion of ecological rationality makes salient the representation of information and the processes tuned to those representations. In doing so, it leads to (b) the passing over of other causal features (which reduces the domain of admissible explanations). It (d) simplifies one’s conception of cognition. And, it (c) helps to make theory formulation more perspicuous by operationalizing “ecological rationality” as a reliability measurement of individual cognitive processes. Ecological rationality is adopted (e) not as a necessary constraint, but as a methodological conceit that helps to motivate new theories of cognition, discover new effects, and (of course) deal with underdetermination.

In other contemporary research programs, theorists are closer to de-idealizing and naturalizing their rationality assumptions. For example, psychological research on conversational pragmatics adopt naturalized, Gricean conversational norms to explain participant judgments and decisions under uncertainty.⁷¹ This Gricean turn suggests that participants’

⁶⁸ For detailed discussion, see Carole J. Lee, "The Representation of Judgment Heuristics and the Generality Problem," in *Proceedings of the Cognitive Science Society*, ed. Danielle S. McNamara and J. Gregory Trafton (Austin, TX: Cognitive Science Society, 2007).

⁶⁹ Peter M. Todd and Gerd Gigerenzer, "Precis of *Simple Heuristics That Make Us Smart*," *Behavioral and Brain Sciences* 23 (2000).

⁷⁰ Stephen P. Stich, "Dennett on Intentional Systems," in *Mind and Cognition: A Reader*, ed. William Lycan (Oxford, UK: Basil Blackwell Ltd, 1990).

⁷¹ Carole J. Lee, "Gricean Charity: The Gricean Turn in Psychology," *Philosophy of the Social Sciences* 36, no. 2 (2006).

inferences about the intended meaning of the questions and tasks are guided by rational norms of conversation.⁷² Since the conversationally rational interpretation of an experimental task can diverge from the experimenter's intended meaning, the notion of conversational rationality can be used to identify conversationally-related confounds to control for. To avoid the mismatch between the experimenter's intended meaning and participants' interpretations, experimenters have begun to adopt protocols to identify participant interpretations.

This research paradigm has de-idealized its rationality assumptions by moving from Grice's *a priori* norms of conversational rationality towards empirically-supported, naturalized ones. Here, the content of the norms themselves can be revised in light of empirical research on participants' conversational inferences and, thereby, give rise to ever more accurate models of conversational inference. Any particular conversational norm can be (a) "falsified" in the sense that it can be shown not to hold for participants in particular conversational contexts. These naturalized conversational norms make conversationally-relevant features of questionnaires salient at the cost of (b) passing over other causal features (thereby reducing the domain of admissible explanations). They (c) make the statement of one's *ceteris paribus* generalizations more perspicuous. And, they (d) simplify researchers' conception of the phenomena. These conversational norms are (e) not adopted as necessary constraints for the very possibility of interpretation and explanation, but as empirically compelling ways of describing participants' conversational judgments under uncertainty.

The behavioral and social sciences are immature in the sense that a single paradigm cannot, by itself, guide a particular area of research. And, so it should not come as a surprise that multiple research programs in psychology conceive of human judgment as being "rational" in different ways. My account of Galilean methodological rationalism articulates important methodological commonalities (a) – (e) across these research programs, despite their different conceptions of rationality.

My account of Galilean methodological rationalism is also distinctive because it allows for the articulation of unexpected commonalities between Weber's and Davidson's accounts of interpretation and explanation. Others have observed important overlaps between Weber's and Davidson's work. In particular, Stephen Turner has observed that both Weber and Davidson

⁷² Paul Grice, "Logic and Conversation," in *Studies in the Way of Words* (Cambridge, MA: Harvard University Press, 1989). For a review of the Gricean Turn in psychological research, see Denis J. Hilton, "The Social Context of Reasoning: Conversational Inference and Rational Judgment," *Psychological Bulletin* 118, no. 2 (1995).

consider the intelligibility of interpretation to the central; and, for both theorists, one's interpretation is said to be intelligible insofar as it conforms to (or deviates in straightforward ways from) rational norms.⁷³ My account agrees with Turner that Weber's reliance on norms of rationality should be construed as a form of scientific idealization. However, Turner's account invokes later Davidson's claims about intelligibility, rationality, and interpretation. And, the later Davidson does not accord with the Galilean approach embraced by Weber.

Like David Henderson's and Mark Risjord's accounts of interpretation, this account of Galilean methodological rationalism gives rise to naturalized interpretations.⁷⁴ These interpretations can be naturalized insofar as they can be guided by psychological evidence about the ways in which people actually form judgments and decisions.⁷⁵ However, it is important to note that this account of Galilean methodological rationalism has a slightly different focus than these accounts. My suggestion is that our source of psychological evidence – psychological theorizing itself – adopts norms of rationality in constructing Galilean models of human cognition. When research paradigms invoking these rationality assumptions serve to explain judgments or decisions, they give rise to what Henderson calls “modest” rationalizing explanations: viz., explanations whose explanatory force derives from describing the cognitive disposition/process involved, where these happen to be described in rationalizing ways.⁷⁶

5. Conclusions

By reclaiming Davidson's Galilean roots, I have shown that Davidson's early account of interpretation and explanation has not been supplanted by contemporary psychological theory, but has provided an early prototype for methodological rationalism as Galilean idealization in contemporary psychology. Although the content of Davidson and Suppes's rationality assumptions have been shown to be false (and Davidson himself drew the wrong lessons in dealing with unexpected anomalies), their more general experimental *approach* to studying

⁷³ Here, Davidson has a narrower set of “rational norms” than Weber. Unlike Davidson, Weber's account has the freedom to adopt a variety of ideal types that capture different (i.e., not simply decision theoretic) forms of normative action. Stephen P. Turner, “The Continued Relevance of Weber's Philosophy of Social Science,” *Max Weber Studies* 7 (2007).

⁷⁴ Mark Risjord, *Woodcutters and Witchcraft* (Albany, NY: State University of New York Press, 2000), David K. Henderson, *Interpretation and Explanation in the Human Sciences* (Albany, NY: State University of New York Press, 1993).

⁷⁵ Lee, “Gricean Charity: The Gricean Turn in Psychology.”

⁷⁶ Henderson, *Interpretation and Explanation in the Human Sciences*, 135-6.

cognition lives on. This Galilean methodological rationalism is theoretically sympathetic to my naturalized approach to charitable interpretation, in which the norms of rationality used to guide interpretation are empirically falsifiable and not construed as necessary constraints on interpretation.⁷⁷

Bibliography

- Bazerman, M. H. *Judgment in Managerial Decision Making*. New York, NY: Wiley, 1990.
- Cartwright, Nancy. *How the Laws of Physics Lie*. Oxford, UK: Oxford University Press 1983.
- . *Nature's Capacities and Their Measurement*. Oxford, U.K.: Oxford University Press, 1989.
- Casscells, W., A. Schoenberger, and T. B. Grayboys. "Interpretation by Physicians of Clinical Laboratory Results." *New England Journal of Medicine* 299, (1978): 999-1000.
- Christensen-Szalanski, Jay J. J., and Lee Roy Beach. "The Citation Bias: Fad and Fashion in the Judgment and Decision Literature." *American Psychologist* 39, (1984): 75-8.
- Cosmides, Leda, and John Tooby. "Are Humans Good Intuitive Statisticians after All? Rethinking Some Conclusions from the Literature on Judgment under Uncertainty." *Cognition* 58, (1996): 1-73.
- Davidson, Donald. "Belief and the Basis of Meaning (1974)." In *Inquiries into Truth and Interpretation*, 141-54. New York, NY: Oxford University Press, 2001.
- . "Could There Be a Science of Rationality? (1995)." In *Problems of Rationality*, edited by Donald Davidson, 117-34. Oxford, U.K.: Oxford University Press, 2004.
- . "Incoherence and Irrationality (1985)." In *Problems of Rationality*, edited by Donald Davidson, 189-98. Oxford, U.K.: Oxford University Press, 2004.
- . "Paradoxes of Irrationality (1982)." In *Problems of Rationality*, edited by Donald Davidson, 169-88. Oxford, U.K.: Oxford University Press, 2004.
- . "Psychology as Philosophy (1974)." In *Essays on Actions and Events*, 229-38. Oxford, UK: Oxford University Press, 1980.
- . "Radical Interpretation (1973)." In *Inquiries into Truth and Interpretation*, 125-40. New York, NY: Oxford University Press, 2001.
- . "Representation and Explanation (1990)." In *Problems of Rationality*, edited by Donald Davidson, 87-100. Oxford, U.K.: Oxford University Press, 2004.
- . "What Thought Requires (2001)." In *Problems of Rationality*, edited by Donald Davidson, 135-50. Oxford, U.K.: Oxford University Press, 2004.
- Davidson, Donald, Patrick Suppes, and Sydney Siegel. *Decision Making: An Experimental Approach*. Stanford, CA: Stanford University Press, 1957.
- Eddy, David M. "Probabilistic Reasoning in Clinical Medicine." In *Judgment under Uncertainty*, edited by Daniel Kahneman, Paul Slovic and Amos Tversky, 249-67. Cambridge, UK: Cambridge University Press, 1982.
- Edwards, Ward. "Nonconservative Information Processing Systems." Ann Arbor, MI: University of Michigan, Institute of Science and Technology, 1966.
- . "Probability-Preferences in Gambling." *American Journal of Psychology* 66, no. 3 (1953): 349-64.
- . "The Reliability of Probability-Preferences." *American Journal of Psychology* 67, no. 1 (1954): 68-95.
- . "The Theory of Decision Making." *Psychological Bulletin* 51, no. 4 (1954): 380-417.
- Elstein, Arthur S., Lee S. Shulman, and Sarah A. Sprafka. "Medical Problem Solving: A Ten-Year Retrospective." *Evaluation and the Health Professions* 13, no. 1 (1990): 5-36.
- Friedman, Milton, and Leonard J. Savage. "The Utility Analysis of Choices Involving Risk." *The Journal of Political Economy* 56, no. 4 (1948): 279-304.
- Gigerenzer, Gerd, and Ulrich Hoffrage. "How to Improve Bayesian Reasoning without Instruction: Frequency Formats." *Psychological Review* 102, no. 4 (1995): 684-704.

⁷⁷ Lee, "Gricean Charity: The Gricean Turn in Psychology."

- Grice, Paul. "Logic and Conversation." In *Studies in the Way of Words*. Cambridge, MA: Harvard University Press, 1989.
- Henderson, David K. *Interpretation and Explanation in the Human Sciences*. Albany, NY: State University of New York Press, 1993.
- Hilton, Denis J. "The Social Context of Reasoning: Conversational Inference and Rational Judgment." *Psychological Bulletin* 118, no. 2 (1995): 248-71.
- Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. "Experimental Tests of the Endowment Effect and the Coase Theorem." *The Journal of Political Economy* 98, no. 6 (1990): 1325-48.
- Kahneman, Daniel, and Amos Tversky. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47, no. 2 (1979): 263-92.
- Korobkin, R. B., and T. S. Ulen. "Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics." *California Law Review* 88, (2000): 1051-144.
- Kuhn, Thomas S. *The Structure of Scientific Revolutions*. Chicago, IL: Chicago University Press, 1962.
- Lee, Carole J. "Gricean Charity: The Gricean Turn in Psychology." *Philosophy of the Social Sciences* 36, no. 2 (2006): 193-218.
- . "The Representation of Judgment Heuristics and the Generality Problem." In *Proceedings of the Cognitive Science Society*, edited by Danielle S. McNamara and J. Gregory Trafton, 1211-6. Austin, TX: Cognitive Science Society, 2007.
- Lopes, L. Lola. "The Rhetoric of Irrationality." *Theory and Psychology* 1, no. 1 (1991): 65-82.
- McMullin, Ernan. "Galilean Idealization." *Studies in History and Philosophy of Science* 16, no. 3 (1985): 247-73.
- Ramsey, Frank P. "The Foundations of Mathematics." *Proceedings of the London Mathematical Society* 25, (1925): 338-84.
- Risjord, Mark. *Woodcutters and Witchcraft*. Albany, NY: State University of New York Press, 2000.
- Saks, Michael J., and Robert F. Kidd. "Human Information Processing and Adjudication: Trial by Heuristics." *Law and Society Review* 15, no. 1 (1980): 123-60.
- Sniderman, P. M., R. A. Brody, and Philip E. Tetlock. *Reasoning and Choice: Explorations in Political Psychology*. New York, NY: Cambridge University Press, 1991.
- Stich, Stephen P. "Dennett on Intentional Systems." In *Mind and Cognition: A Reader*, edited by William Lycan, 167-84. Oxford, UK: Basil Blackwell Ltd, 1990.
- Sunstein, C. R. "Behavioral Analysis of Law." *University of Chicago Law Review* 64, (1997): 1175-95.
- Thompson, William C., and Edward L. Schumann. "Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy." *Law and Human Behavior* 11, (1987): 167-87.
- Todd, Peter M., and Gerd Gigerenzer. "Precis of *Simple Heuristics That Make Us Smart*." *Behavioral and Brain Sciences* 23, (2000): 727-41.
- Turner, Stephen P. "The Continued Relevance of Weber's Philosophy of Social Science." *Max Weber Studies* 7, (2007): 37-62.
- Tversky, Amos, and Daniel Kahneman. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185, no. 4157 (1974): 1124-31.
- . "Rational Choice and the Framing of Decisions." *The Journal of Business* 59, no. 4, Part 2: The Behavioral Foundations of Economic Theory (1986): S251-S78.
- . "The Framing of Decisions and the Psychology of Choice." *Science* 211, no. 4481 (1981): 453-8.
- von Neumann, John, and Oskar Morgenstern. *Theory of Games and Economic Behavior*. 1 ed. Princeton, NJ: Princeton University Press, 1944.
- Weber, Max. *The Theory of Social and Economic Organization*. Translated by A. M. Henderson and Talcott Parsons. New York, NY: Oxford University Press, 1947.