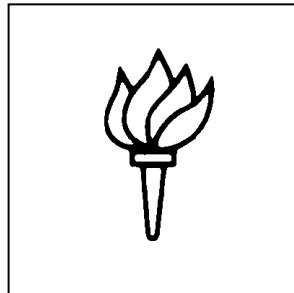


NEW YORK UNIVERSITY SCHOOL OF LAW

NYU Center for Law and Economics



EXPERIMENTAL LAW AND ECONOMICS

Jennifer Arlen and Eric Talley

June 2008

LAW & ECONOMICS RESEARCH PAPER SERIES

WORKING PAPER NO. 08-30

<http://ssrn.com/abstract=1144086>

EXPERIMENTAL LAW AND ECONOMICS

Jennifer Arlen and Eric Talley*

This chapter provides a framework for assessing the contributions of experiments in Law and Economics. We identify criteria for determining the validity of an experiment and find that these criteria depend upon both the purpose of the experiment and the theory of behavior implicated by the experiment. While all experiments must satisfy the standard experimental desiderata of control, falsifiability of theory, internal consistency, external consistency and replicability, the question of whether an experiment also must be “contextually attentive” – in the sense of matching the real world choice being studied -- depends on the underlying theory of decision-making being tested or implicated by the experiment.

We find that the importance of contextual attentiveness depends on whether the experiment tests or implicates a “unitary-process” theory of decision-making or a “multiple-process” theory. Unitary-process theories posit that people employ a single operational approach to make decisions across a broad (or universal) domain of activity. Rational Choice Theory is a unitary-process theory. Because unitary-process theories posit that people employ the same decision-making program in all contexts, experimenters can falsify a unitary-process theory using an experimental choice which bears little resemblance to any real-world choice. Faith in a unitary process account also permits legal policymakers to draw broad normative implications from experiments involving quite artificial choices. By contrast, multiple-process theories hold that people employ multiple decision-making programs when they make choices. Moreover, the relative impact of these programs can depend on the context of the decision. This posited interaction between context and decision-making implies that experimentalists seeking to examine legal decision-making must be sensitive to contextual factors likely to affect deliberative and non-conscious programs in the real world. In addition, policymakers must proceed cautiously before using experimental evidence to draw normative policy conclusions because experimental results may not be robust across contexts.

A central function of law is to influence and shape human behavior. Liability rules, negligence standards, punitive damages, evidentiary presumptions, criminal penalties, default rules of contracting, rules of jurisdictional reach; each plays an important role within a larger normative enterprise of law: discouraging socially undesirable activities while encouraging socially desirable ones. Any legal scholar, policy maker, or judge who seeks to use law for instrumental purposes must inevitably contend with understanding and forecasting how law and human behavior interact. Such forecasts are not easy. Among other things, they minimally require a coherent theory of how people behave in legal environments. Such a theory, moreover, must be capable of predicting how legal structures shape incentives, coordinate beliefs, and communicate expectations, and also how such factors are likely to interact to affect actors’ decisions.

* Arlen is the Norma Z. Paige Professor of Law, New York University School of Law. Talley is Professor of Law, University of California at Berkeley Law School and Senior Economist, RAND Corporation. We would like to thank the following for helpful comments and discussions: Oren Bar-Gill, Colin Camerer, Robert Lee Hotz, George Loewenstein, Elizabeth Phelps, Kathryn Zeiler and participants at a Faculty Workshop at the UC Davis Law School. Thanks as well to Josh Weigensberg for research assistance. All errors are ours.

For nearly a half century, the field of Law and Economics has embraced a specific theoretical account of how legal environments affect behavior. A central analytic concept of that account is *Expected Utility Theory* (and its slightly more generalized cousin, Rational Choice Theory).¹ Informally, Expected Utility Theory maintains that (at least under certain assumptions) individuals' preferences can be represented through a well-defined, consistent mathematical function that translates specific features of a person's current status (wealth, income, health, consumption, etc.) into a scalar value reflecting her level of satisfaction – or “utility.” The theory presumes that the function that translates attributes into utility is stable; thus, a given set of attributes always yields the same utility. It also presumes that in risky environments, an individual uses all available information to form unbiased, consistent estimates about possible outcomes, ultimately making the choice that maximizes the probability-weighted sum of her utility across contingencies. In addition, many applications of Expected Utility Theory assume that individuals are self-interested and indifferent to the welfare of others, although this is not a requirement of the framework.

Since the early 1960s, Law and Economics scholars have employed the Expected Utility framework to great effect, producing insights about the ways that law can alter behavior. In addition, at least some Law and Economics scholars have further maintained that such insights can and should serve as the normative basis for legal policy reforms, through the lens of efficiency. Although the approach is a well-accepted form of academic discourse, the practical import of any Law and Economics argument still must turn on whether (or when) Expected Utility Theory in fact accurately predicts how individuals respond to their legal environment. In other words, to have value as a predictive or normative enterprise, the underlying tenets of Expected Utility Theory must be generally sound (at least on average).

Of course, the soundness of any predictive framework in social science cannot be verified with theory alone; rather, those predictions must be tested externally. There are two generally accepted (and somewhat related) methodologies for testing theories of human behavior: (1) Empiricism, which analyzes data drawn from the “real world”; and (2) Experimentalism, which analyzes data drawn from the laboratory.² This book is dedicated to the latter.

Experimentalists study human decision-making by placing subjects in tightly orchestrated environments and providing them with a set of experiences and/or choices whose presentation, features, and variations are all within the researchers' control. For example, researchers conducting experiments frequently seek to manipulate subjects' initial reference points, the riskiness of their choices, the content of available information, the manner in which such information is presented, and the economic stakes that subjects realize from their actions. By varying critical parameters of choice,

¹ Expected Utility Theory has been one of the foundational bedrocks of economics since the mid-twentieth century (and arguably even before). For an overview of the technical requirements in Expected Utility Theory, as well as a framework generalizing it to a world of cognitive biases, see Camerer & Talley (2007). While the concept of expected utility conventionally is viewed as synonymous with Rational Choice Theory, the latter concept is at least formally a broader one. For the purposes of this book, however, the concepts may be treated as largely interchangeable.

² Notably, both empiricism and experimentalism use accepted statistical techniques both to summarize data and to test hypotheses. Their core distinction comes from the distinct underlying sources of that data.

experimentalists test whether individual responses to such stimuli coincide with the predictions of benchmark theories of behavior. When experimental results do not support existing theories, the results can be used either to modify existing theories or to reject them. In the latter case, the experiments then can serve as a potential basis for developing new theories of behavior. The capacity for experimental work to test the validity of behavioral theories is vital to legal policy, since behavioral theories generate predictions about both likely responses to legal regime change and (in some instances) the desirability of those responses.

Experimentalism is an important complement to empiricism in a number of ways. Most notably, outside of the rare natural experiment, it is frequently difficult – perhaps prohibitively so – to make *causal* inferences from real-world data. Although a researcher may observe that a legal change appears correlated to some subsequent behavioral change, it is difficult to determine, with appreciable confidence, whether that change in behavior was *caused* by the reform or some other contemporaneous event, whether the reform was in fact caused by the anticipated change in behavior, or whether some unobservable phenomenon caused both. Moreover, outside of the laboratory, there are significant sources of noise that can convolute the process of making inferences. Researchers often do not know, for example, precisely what information people had about the choice, how much time they had to focus on it, or whether there were external considerations affecting their choice (e.g., reputational concerns).³

Experimental approaches are often an apt vehicle for addressing the shortcomings of real-world empiricism. Indeed, well-designed experiments reduce (and frequently eliminate) observational noise by placing subjects in a tightly controlled environment. Control, in turn, allows researchers to manipulate exactly what factors could be relevant to subjects' choices, facilitating significantly stronger causal inferences about how specific environmental stimuli affect behavior. Moreover, experimental results can serve as the most compelling basis for testing existing theories of decision-making, generalizing such theories, or even generating new ones.

The Theme(s) of this Book

This book collects some of the leading articles on Experimental Law and Economics, which is a special subset of experimental analysis proper. Experimental Law and Economics encompasses a set of experiments that focus specifically on how individual actors respond to their legal environment (as reflected in rules and institutions). Nevertheless, experiments rooted within the Law and Economics tradition typically transcend their nominal moniker of “legal scholarship,” making substantive contributions more generally to behavioral sciences. This is for good reason: legal institutions are a prime venue within which people interact with one another and the state, and are therefore a focal vehicle for effecting (and affecting) policy commitments.

As with much policy-relevant academic work, experimentalists in Law and Economics frequently strive for immediate relevance and effect. Many studies (both

³ Thus, when confronted with evidence of actual decisions that appear to be inconsistent with Expected Utility Theory, Law and Economics scholars sometimes argue that the evidence does not show irrationality *per se*. Instead, some argue that people's choices were distorted by some imperfection external to the decision-making program (e.g., they may be misinformed).

within and outside of this book) offer results that – at least according to some – provide the basis for either immediate or prospective legal reforms. As noted above, however, the soundness of such conclusions ultimately turns on a collection of factors, the identification and exploration of which is a central purpose of this introduction.

Specifically, our goal is to provide the reader with a toolkit to digest and evaluate not only the studies included in this book, but also the literature more generally, and the strength of any prescriptive inferences she might extract from it. While we shall explore specific implements of this toolkit in the following sections, it is perhaps prudent to highlight three of the most prominent ones at the onset. The first is the necessity of determining the underlying *purpose of an experiment*. Some experiments are designed to test existing theories of decision-making. Others are designed to construct theories, either by calibrating underdetermined accounts, or by providing basic insights that could serve as the kernel of some new future theory. Most of the Experimental Law and Economics literature has been preoccupied (directly or indirectly) with the former goal: testing existing theories of decision-making. Far fewer have focused (at least consciously) on calibration or development of alternative theories of behavior. In particular, experimental tests within this field have generally focused on testing predictions of Expected Utility Theory, the analytical cornerstone of the classical Law and Economics approach.

In many ways, this focus is eminently sensible. Experiments that test existing theories are often perceived to have the broadest potential implications for policy questions (at least within the domain of the theory being tested). An experiment whose results provide support for a posited theory also (indirectly) lends support for subsequently using that theory to predict reactions to future policy reforms – even if such manipulations lie outside the specific experimental design. For example, should an experiment (or series thereof) conclusively demonstrate that people always make decisions consistent with Expected Utility Theory, this experiment might have broad implications for legal policy, for it would imply that one could more comfortably use Expected Utility Theory as a general purpose vehicle for evaluating prospective reforms. Similarly, experimental results that principally falsify Expected Utility Theory *could* cast powerful skepticism on the theory's general predictive powers, even outside the specific experimental domain in question.

Of course, the policy implications of a theory's evident refutation are elusive in the absence of a plausible, falsifiable alternative theoretical account that is more consistent with observed data.⁴ Within Experimental Law and Economics, most well-known studies purport to test Expected Utility Theory against its chief conceptual rival, *Prospect Theory*. Briefly, Prospect Theory posits that decision-makers systematically and predictably deviate from Expected Utility Theory predictions in at least three ways (Kahnemann & Tversky, 1979). First, Prospect Theory maintains that people tend to make decisions relative to a status quo ante reference point, evaluating relative gains

⁴ To be sure, deconstructive results can be helpful in motivating the formulation of alternative theories, capable of themselves being tested. But these results cannot be used as the basis for legal policy until we have developed a theory that helps us predict what factors influence decision-making, as without such a theory we cannot be certain whether real world decisions present subjects with the same choice as the experiments.

and losses differentially.⁵ Changing their perceived reference point may alter their revealed preferences. Second, Prospect Theory maintains that people do not use actuarially accurate, consistent probability assessments when they assess the likelihood of gains and losses, but instead employ subjective beliefs that overvalue small probabilities and undervalue large ones; they also update probabilities in a distinctly non-Bayesian manner. Third, Prospect Theory maintains that individuals exhibit dynamic preference inconsistencies, excessively discounting future payoffs against short term gains or losses (Loewenstein and Prelec, 1992).⁶

The emergence of Prospect Theory as a legitimate rival to Expected Utility Theory raises a second ingredient of the toolkit for readers of experimental work: understanding the *domain of the theory being developed / calibrated*. Both Expected Utility and Prospect Theory share the premise that individuals employ a single program to make all decisions. Yet not all theories of decision-making accept this premise. Indeed, the theories coming to dominate neuroscience and neuroeconomics are based on a quite different view of decision-making. At the risk of oversimplifying, two canonical families of decision-making theory now appear to characterize the literature. The first might be called “unitary-process” theories. Such accounts posit that people employ a common operational approach to make decisions across a broad (or universal) domain of activity. For example, a unitary-process thinker would use a similar cognitive calculus across a variety of contexts, e.g., deciding what to eat for breakfast, whether to marry/divorce, whether to quit smoking, what kind of car to buy, whether to commit a felony, and so forth. Traditionally, both Expected Utility Theory and Prospect Theory are broadly perceived to be unitary-process theories. By contrast, “multiple-process” theories posit that people employ multiple decision-making programs when they make choices. One leading multiple-process account of decision-making holds that decisions result from the concurrent operation of distinct cognitive “programs,” some of which affect our conscious deliberation while others are intuitive, non-conscious processes that intervene prior to conscious decision-making. (Much of the emerging field of neuroeconomics, for example, illustrates how different parts of the brain become activated depending on the type of decision at issue, a finding that is consistent with multiple-process cognition). As a result, individual cognitive processes may become dominant in different context-specific situations, and cross-context comparisons can potentially lead to apparent inconsistencies in behavior.

The distinction between unitary-process and multiple-process theories has a number of important implications for experimental work in Law and Economics. Primarily, recognition that people may employ multiple processes seriously complicates efforts to derive broad normative policy prescriptions from isolated experimental results. The enterprise of testing theories is significantly more apparent within the domain of unitary-process theories – because behavior in one context by definition provides information about the decision-making program people use in other circumstances. Thus, one need only find a *single replicable experimental context* (even

⁵ Specifically, whereas Expected Utility Theory assumes that decision-makers exhibit utility functions that are universally concave, Prospect Theory assumes that decision-makers possess a utility function that has a kink at whatever reference point that person identifies with the status quo, and is convex below that reference point.

⁶ Loewenstein and Prelec (1992).

if somewhat artificial) that is inconsistent with a unitary-process theory's predictions in order to falsify it *writ large*. By contrast, experiments cannot be assumed to have such broad implications if people employ multiple processes, whose relative impact varies in different circumstances. If these theories are correct, then an experiment demonstrating behavioral inconsistencies with a particular decision-making program in one circumstance may demonstrate only that the underlying program is not a unitary process, but it does not imply that people *never* use that program. Similarly, evidence *consistent with* a particular unitary-process theory cannot necessarily be assumed to carry over to other decision-making contexts.

Recognition of the existence of multiple-process theories is particularly important to the now well-known debate over whether Expected Utility Theory or Prospect Theory best explains human behavior. Because these two theories generally are viewed as unitary-process theories, they tend to be treated as mutually exclusive foils. This tradition has arguably led researchers – perhaps too anxiously – to conclude that experimental results inconsistent with one theory both refute that theory and validate its foil. Such conclusions become suspect once we expand the range of plausible theories to include multiple-process theories.

The third implication that emerges from this discussion relates to experimental design. Experiments should pay particular care to control context because many experiments intended to test existing theories may, in the end, have an important role to play in developing new ones, and must be designed (or read) with that goal in mind. For example, suppose one were to discover evidence that within some particular arena (e.g., learning about the risks of automobile accidents), actors do not formulate and update beliefs consistent with Bayes' rule. Under a unitary-process framework, one might portray such a finding as a test (and indeed a refutation) of Expected Utility Theory. Alternatively, when viewed through the lens of multi-process accounts, the finding would constitute a contribution towards locating the boundary between deliberative (“Bayesian”) and unconscious (“heuristic”) cognitive processes. Viewed in this light, the experiment's purpose would be not as much to test a theory as it would be to develop and/or calibrate a new one.

In light of the above distinctions, the collection of experiments in this volume carry a critical implication for policy-oriented researchers interested in Experimental Law and Economics. Our assessment is that unitary-process accounts of *either* Expected Utility Theory *or* Prospect Theory are likely implausible and invalid on those terms. There are numerous, replicable, examples in the literature (some of which are included in the chapters that follow) that reject either Expected Utility Theory or Prospect Theory (or both) within certain experimental settings.

This collection of mutually conflicting results can only be consistent with one of three hypotheses. First, it may be that there exists a yet-to-be-conceived unitary-process theory that unifies all (or most) apparent anomalies (the behavioral analog of Einstein's elusive Unified Field Theory). This explanation, while entirely possible, must await additional contributions by theorists, and has little immediate value to policy makers. Second, it may be that experiments purporting to falsify one side of the unitary-process debate (or the other) are systematically ill-conceived or invalid as a whole. While such

claims warrant consideration,⁷ our view (and that of many others) is that even if some experiments have flawed experimental designs, the sheer depth of evidence against both existing unitary-process accounts makes it unlikely that experimental design error systematically pervades the literature.

The final – and in our view most plausible – hypothesis is that human decision-making is simply a more nuanced phenomenon than unitary-process theories permit. Rather, actors may employ multiple decision-making programs concurrently, and the actuation of each program may depend on the underlying context in systematic ways. In some contexts, conscious decision-making may share many features of Rational Choice Theory. In other moments, unconscious or intuitive processes may intervene, affecting the information that reaches our deliberative processes, the weight we give to various pieces of information, the time and attention devoted to choosing through deliberation, and our willingness to choose based on the outcome of deliberation instead of an “intuition” about what is right.⁸

If the multiple-process account of behavior is a sound one, it would clearly have important implications for both scholars and policy makers attempting to develop and interpret experimental work. For scholars, it would mean that one cannot simply dismiss either Expected Utility Theory or Prospect Theory as a framework to explain decision-making in all contexts based on evidence that falsifies one (or the other) in a specific experimental context. Rather, both theories may predict choices well in some circumstances and poorly in others. If this is true, at least two key tasks for researchers are to (1) isolate / characterize the contextual triggers that induce individuals to switch between distinct cognitive programs; and (2) understand how policy makers can manipulate those triggers, either by dampening or encouraging non-rational cognition for instrumental design purposes. Much of this work has yet to be done.⁹

For policy makers (or those seeking to influence them), the potential validity of multiple-process accounts has a more important implication: it suggests that one must exercise tremendous caution in translating experimental results into policy prescriptions. Greater care, for example, must be taken when extrapolating from the laboratory to real-world settings, since such settings alter context, a difference that itself may actuate wholly distinct cognitive processes, rendering experimental predictions based on more artificial environments untrustworthy. This possibility implies that one must pay significant attention to whether an experiment captures the essential features of a real-world situation of interest. (Such “reality check” desiderata are frequently unpopular (and even unwise) when testing unitary-process theories). In short, careless

⁷ This is a possibility that warrants consideration since many of the experiments are indeed invalid. Moreover, recent work by Professors Plott and Zeiler has shown that certain long-accepted experimental results supporting alternatives to Expected Utility Theory appear to be the product of subject misconceptions (Plott & Zeiler, 2005).

⁸ One suggestive hypothesis, for example, is that deliberative decision-making may conform to traditional Rational Choice Theory (although it may exhibit some elements of loss aversion), while intuitive processes may be more akin to traditional Prospect Theory.

⁹ A significant potential limitation of the multi-process theory of decision-making is the challenge associated with testing such a theory. Indeed, because multi-process theory conceives of decision makers as “switching” cognitive frames from one context to another, it injects degrees of freedom that potentially render it less falsifiable, violating a core Popperian desideratum that remains (albeit with some criticism) at the core of most hard social science research. We revisit this issue – along with measures that researchers might take in addressing it – in Part II of this introduction.

extrapolation from specific experimental results can lead to analogously careless – or even dangerous – policy commitments.

The remainder of this introduction is organized as follows. Section I discusses in greater detail the potential implications that can be drawn from an experiment depending on whether it (1) is designed to test an existing theory or develop a new one and (2) implicates a unitary-process theory or a multiple-process theory. The section illustrates how experiments have broad negative implications for legal policy to the extent that they test and invalidate unitary-process theories. The positive implications of experiments are less clear, however, depending once again on the articulated domain of the theory being tested / developed. Section II presents our claim that the existing experimental literature appears largely to falsify both leading unitary-process accounts (Expected Utility and Prospect Theory) when viewed as unitary-process theories. Section II then describes recent evidence from neuroscience that – in our view – is strongly suggestive of multiple-processes accounts. Section III presents the implications of this analysis for experimental design, presenting context-specific criteria relevant to determining the validity of an experiment. Finally, Section IV overviews the contributions of the Experimental Law and Economics literature contained in this book, highlighting considerations that we believe to be relevant for legal policy. The articles included in this book fall into four areas: (1) Contracting, Legal Entitlement and the Coase Theorem (including the endowment effect); (2) Self-Serving Biases; (3) Other-Regarding Preferences; and (4) Decision-making Under Risk and Uncertainty (focusing particularly on Judges and Juries).

I. IMPLICATIONS OF EXPERIMENTS FOR LEGAL POLICY

Virtually all social scientific efforts to discuss the effects of legal rules on human behavior depend, explicitly or implicitly, on a model (formal or informal) of how people respond to their legal environment. The validity of these analyses therefore depends on whether the underlying model accurately predicts (perhaps with some noise) individual actions. If it fails to predict people's behavior accurately, then any reform prescriptions emanating from that model become untrustworthy.

The task of testing existing theories of human behavior and developing new ones requires the contribution of empiricists and experimentalists. Empirical analysis of behavior in the real world provides insights into *how* people behave in the real world, but it often cannot tell us *why* people behave a particular way. In particular, it often is difficult to use real-world choices to test any particular theory of decision-making because a variety of factors – such as information problems, institutions or third parties – could cause people's choices in the real world to deviate from the predictions of the theory. We cannot easily access whether behavior conforms to theory unless we can control the factors affecting decision-making.

By contrast, experimental analysis allows researchers to examine behavior in a controlled environment, often (though not exclusively) in a lab. Experiments are an important complement to empirical analysis using real-world data because researchers can design experiments to isolate and examine particular causal drivers of human decision-making – for example, responses to sanctions – and examine them separately from other factors that might affect behavior in the real world. Moreover, the ability to

control the decision-making context allows researchers to test the predictions of existing theories with a minimal amount of measurement error due to other forms of observational noise. And, when existing theories are found wanting, experimental results can also provide a portfolio of stylized facts for developing new theories of behavior.

Experimental Law and Economics encompasses a set of experiments that focus on how people respond to legal rules and institutions. As noted in the introduction, however, many such experiments seek to transcend this subject-matter constraint and contribute to the broader quest of formulating robust and predictive general theories of human decision-making. Experiments contribute to this project both by testing existing theories of decision-making (such as Expected Utility Theory) and by providing evidence to develop new theories. In addition, many experiments in Law and Economics endeavor to provide results that can be used to formulate existing legal policy, even in advance of the development of a generally agreed upon theory of decision-making that consistently predicts behavior.

Whether an experiment provides insights that should be used as the basis for legal policy depends on the purposes of the experiment, the domain of the theory it addresses, the results of the experiment, its relationship to other experimental results, and the validity of the experimental design. Perhaps most centrally, the implications of an experiment for legal policy depends on whether the experiment is testing an established theory or is simply providing results that can be used to develop and/or calibrate an emerging theory. In considering experiments that test theories, it is important to consider whether the experiment only is relevant to testing/developing a unitary-process theory (which assumes people employ one decision-making program to make all decisions) or also has implications for multiple-process theories (which assumes that people employ multiple, interacting decision-making programs). This latter consideration can affect both the design of the experiment itself and its interpretation against the existing experimental literature.

A. WHY DO EXPERIMENTS?

Experiments can play important roles in both testing and developing theories. Their role in testing theories is self-evident. Experimental Law and Economics has been used to test numerous theoretical predictions. For example, Hoffman & Spitzer (1982; 1985) devised a number of experimental protocols to test the much heralded prediction of the Coase Theorem (Coase 1960), that in the absence of transaction costs and wealth effects, the initial allocation of a legal rule is irrelevant for purposes of allocational efficiency because people, through bargaining, would allocate goods to those who valued them most. Those early experiments of the Coase theorem largely confirmed its predictions, while also pointing out deviations from the theory (an observation pursued at length by others, as we shall elaborate below). Other experiments have more consistently identified situations in which people deviate from accepted theories, as discussed below.

Experimental results also can be used to develop or calibrate theories of human decision-making. Indeed, it is relatively common for surprising experimental findings – often initially cast as “anomalies” or “paradoxes” – to have powerful feedback effects

on theory, particularly when such apparent anomalies are widespread and routine in their occurrence. In such instances, theorists often feel compelled to search for a new theory of behavior that explains the experimental results. Examples abound in economics where subjects' behavior in experiments induced monumental theoretical reformulations. The well-known St. Petersburg paradox from the mid-eighteenth century,¹⁰ for example, motivated much of the original formulation of risk aversion and the so-called "expected utility" hypothesis about behavior (described in greater detail below). The Allais (1953) paradox demonstrating troubling apparent preference reversals in expected utility contexts provided much of the motivation behind Kahnemann & Tversky's (1979) seminal work developing Prospect Theory.

The distinction between these two goals of testing an existing theory and developing/calibrating a newer one is important because the underlying goal of an experiment can substantively affect not only its design, but also the implications that can be drawn from experimental results.¹¹ Experiments designed to test existing theories have the broadest potential implications for legal policy, depending on the results of the experiment and how they interact with the findings of other valid experiments. An experiment that invalidates a theory of decision-making has broad implications because it suggests that policy makers cannot rely on that theory to design legal policy. Moreover, depending on the domain of the underlying theory (see below), one experiment can undermine the use of a theory in situations reaching far beyond the specific choice considered in the experiment. The reverberations of a negative result can cascade out to all forms of decision-making, not just those covered by the experiment.

On the other hand, experiments may not be able to offer immediate implications for legal policy if their principal aim (or legitimate implication) is to provide results that may be used to develop or calibrate new theories. Care must be taken in extrapolating from the results of experiments endeavoring to develop new theories of behavior, because it is difficult to know the predictive power of the experimental results until the theory of behavior is fully developed and tested. Thus, experiments devised to develop or calibrate theories, while still useful for prospective policy making, may only begin to pay normative dividends after months or years of calibration.

¹⁰ Bernoulli, Daniel (1738), "Exposition of a New Theory on the Measurement of Risk", English Translation: *Econometrica* 22 (1954), 23-36.

¹¹ It is important to recognize at the outset that this distinction is often one of degree, and is not absolute. Many experiments are designed to test a theory in the expectation that it will be falsified, and also provide data that can be used to build a new theory. An obvious example of this is the seminal paper by Kahneman, Knetsch and Thaler (1990) on the endowment effect that both provided data that is inconsistent with Rational Choice Theory but also contributed to the development of Endowment Effect Theory. Other examples of such studies see Plott & Zeiler (2005) and Fehr et al (2007) found within this volume. These sorts of "dual" studies are among the more difficult to evaluate, since – as described below – the desiderata for theory-generating work and theory-testing work are likely not the same. Consequently, it is important to read such studies in a sort of contingent fashion, with the context of the underlying goals that one might ascribe to them firmly in mind. Some dual studies may, for example, attempt exhaustively to falsify an existing theory (or at least narrow its predictive domain), but then only posit informally possible alternative explanations. Others may show only specific examples where a theoretical prediction is not borne out, but then formulate elaborate theoretical frameworks that are consistent with those examples.

B. THE DOMAIN OF A THEORY (AND THE THEORY OF DOMAIN)

In addition, the inferences that can be drawn from an experiment also depend on the precise domain of the theory tested, developed or calibrated by the experiment. These inferences also depend on the relevance for experimental interpretation of other theories of behavior, which the experimenter might not have tested, but which might be plausible nonetheless. In particular, it is important to evaluate experimental results within Law and Economics with respect to two theoretical domains: “unitary-process” and “multiple-process” models of decision-making.¹² *Unitary-process* theories of decision-making, as defined here, hold that people tend to adhere to a particular decision-making program across the entire set of contextual domains in which choices are made. By contrast, *multiple-process* theories of decision-making posit that people employ many decision-making programs, either simultaneously or in a context-contingent fashion. To be fully specified, a multiple-process theory also must identify all possible decision-making contexts with a decision-making program (or combination programs) that are activated within that context.

Much of the seminal experimental work in Law and Economics has been focused on testing the leading theories of behavior currently employed to predict the consequences of legal rules. These theories can fairly be characterized as unitary-process theories. The first of these is Expected Utility Theory, which provides the foundation for traditional economic analysis and most Law and Economics. Informally, Expected Utility Theory holds that people rationally act in a manner consistent with maximizing their expected utility.¹³ This theory conventionally presumes that individuals possess stable, consistent, well-behaved preferences and the capacity to optimally accumulate and assess information. It further assumes that when people evaluate choices in a risky environment, they do so by weighing both the probability and magnitude of the potential outcomes associated with each choice.¹⁴ Finally, although not a requirement of Expected Utility Theory, many applications also assume individuals to be self-interested, indifferent to the welfare of others. Most who write in the tradition of Expected Utility Theory tend to view it as a unitary-process theory, implicitly positing that decision-making in all contexts should satisfy the tenets of Expected Utility Theory, at least on average.

There are a number of leading challenges to conventional Expected Utility Theory, the most notable of which is Prospect Theory, which was initially posited by Kahnemann & Tversky (1979). According to Kahnemann & Tversky, people do not

¹² To be sure, the dichotomy we describe between unitary and multiple-process theories is an exaggerated one, as some work likely falls somewhere in between. We have used this categorical distinction solely for expositional purposes, realizing that these points probably demarcate ends of a fluid spectrum.

¹³ Expected Utility Theory has been one of the foundational bedrocks of economics since the mid-twentieth century (and arguably even before). For an overview of the technical requirements in Expected Utility Theory, as well as a framework generalizing it to a world of cognitive biases, see Camerer & Talley (2007).

¹⁴ Risk-neutral people rank choices based on their expected value, which equals the probability of the outcome multiplied by its magnitude. Risk-adverse people are assumed to rank choices based on their expected value and the variance in potential outcomes; they prefer lower variance in outcomes to higher variance, all else equal.

make choices under uncertainty based on the “true” objective probabilities of the potential outcomes, as predicted by Bayesian reasoning. Instead, they overweigh small probabilities and undervalue large ones. In addition, they do not exhibit universally concave utility functions, but instead have utility functions that kink and become convex at a reference point that subjects identify with the status quo. Moreover, most accounts of Prospect Theory (though not Kahnemann and Tversky’s original) also posit that individuals exhibit dynamically inconsistent preferences, giving excessive weight to current costs and benefits.

As noted above, it is probably a fair reading of the literature to conclude that both the devotees of Prospect Theory and those of Expected Utility Theory tend (at least implicitly) to view them as unitary-process theories, in that each theory posits a decision-making algorithm that transcends all circumstances governed by the theory. (It is worth noting, however, that some have advocated against this view and have worked to situate both approaches within a more generalized framework (Camerer and Talley, 2007)).

Unitary-process theories are particularly attractive candidates for experimental analysis because unitary-process theories, by definition, assume that people employ a particular decision-making program in all circumstances, both familiar and unfamiliar. Given this, an experiment can refute the validity of any unitary-process theory by demonstrating that circumstances exist where people’s decision-making does not conform to the predictions of the theory -- even if the laboratory choice differs from the types of choices people make in the real world. Although such a negative experimental finding would not imply that people *never* employ that decision-making program, it would invalidate any theory which holds that people employ that decision-making program as a unitary process.¹⁵

Experiments that test unitary-process theories thus have particularly broad potential implications for legal policy. A single valid experiment showing that people do not make decisions consistent with a unitary-process theory – e.g., Expected Utility Theory -- invalidates any analysis that assumes people always make decisions according to that unitary process. Similarly, experimental confirmation of a unitary-process theory would imply that scholars could employ that theory of decision-making to predict all decisions within the scope of the theory. For example, experimental confirmation of Expected Utility Theory as a unitary-process theory would imply that scholars could analyze all legal rules assuming that people respond to them as predicted by Expected Utility Theory. Of course, it is more difficult to provide experimental confirmation of a unitary-process theory, than it is to refute it. A single valid experiment can show that decision-making does not always conform to the theory – thereby invalidating the theories’ claim to describe the one process used to make all decisions; yet a single experiment confirming the theory cannot show that decision-making *always* conforms to the theory. To confirm the theory, many more results are needed that are consistent with the theory in order to show that the theory always predicts behavior.

¹⁵ Experiments can be used more effectively to refute a unitary-process theory than to prove it, because a single experiment showing that people sometimes behave consistent with the theory does not support the hypothesis that they always do.

While it is true that a single experiment refuting a unitary-process theory is enough to invalidate it as a theory, the practical implications of this claim are likely narrower than it may at first appear. The reason for this is that unitary-process theories are not the only plausible theories of decision-making. Indeed, there is an emerging consensus among neuroscientists, neuroeconomists and psychologists that people do not use any one decision-making program to make all decisions, but instead use different decision-making programs in different circumstances. The existence of this alternative theory of behavior has important implications for those seeking to use experimental results as the basis for legal policy because it implies that people may behave inconsistently with a particular theory in some circumstances but consistent with it in others. In this case, experimental evidence that fails to confirm that people use a given program – for example, Expected Utility Theory – would imply that this theory is not a valid unitary-process theory, but does not necessarily imply that people never employ the tested decision-making program to make decisions. Multiple-process theories thus narrow the potential policy implications which can be drawn from experimental evidence and also raise the hurdles that experimenters face in designing an experiment which generates valid, controlled, and yet generalizable results.

II. EXPERIMENTAL TESTS OF UNITARY- VERSUS MULTIPLE-PROCESS THEORIES

The implications of experimental research for legal policy depend to a considerable degree on whether the experimental results should be interpreted with respect to the assumption that people employ one decision-making program to make all decisions or against the alternative assumption that people employ multiple-programs to make decisions, often simultaneously. This Section briefly discusses experiments relevant to this issue. We begin with a discussion of experimental results within Experimental Law and Economics. We then turn to a discussion of the results of experiments generated by the emerging field of Neuroeconomics. We conclude that the results of both types of experiments are inconsistent with the core premise of all unitary-process theories that people employ a single decision-making algorithm to make choices (either in general or about losses). Instead, the evidence suggests that people employ multiple decision-making programs, concurrently.

Specifically, evidence suggests that decisions are produced by the interaction of conscious decision-making processes – including deliberation that may share many features of Rational Choice Theory – and unconscious intuitive processes that often intervene in advance of deliberation. These unconscious processes can intervene in a variety of ways, affecting the information that reaches our deliberative processes, the weight we give to various pieces of information, the time and attention devoted to choosing through deliberation, and our willingness to choose based on the outcome of deliberation instead of an “intuition” about what is right. One hypothesis is that deliberative decision-making processes may conform to Rational Choice Theory (perhaps exhibiting loss aversion), while intuitive processes do not. The challenge for decision-making theory is to determine whether this is the case, the precise contours of the intuitive processes, and what factors determine whether choices are influenced by deliberative processes or non-conscious ones.

A. TESTING UNITARY-PROCESS THEORIES

Experimental Law and Economics generally has focused on testing the validity of the two leading unitary-process theories of decision-making: Expected Utility Theory and Prospect Theory. Many experiments simply test the validity of Expected Utility Theory, and examine whether circumstances exist where people do not behave consistently with Expected Utility Theory or instead systematically displayed *particular* heuristics or biases (e.g., the self-serving bias; other-regarding preferences; endowment effects; status quo biases) that caused them to behave in a manner systematically inconsistent with Expected Utility Theory.¹⁶ Other experiments seek to examine whether people behave consistently with either Prospect Theory or Expected Utility Theory.

The subset of the experimental literature that is particularly preoccupied with this “battle of the bands” between Expected Utility Theory and Prospect Theory is the literature on the so-called “endowment effect.” The endowment effect experiments were the first to test the central tenet of Expected Utility Theory that people’s preferences are constant and are not contingent on a perceived status quo ante. The endowment effect experiments – which we discuss in detail below – purport to show that the value that people attach to an object is not constant. Instead, experimenters found evidence that people *can* value objects they own and feel entitled to possess more than objects owned by others, even when possession does not in any obvious way increase the value of the object (Tversky & Kahneman, 1991; Kahneman, Knetsch and Thaler, 1991). Assuming the validity of the experiments, this evidence is sufficient to refute Expected Utility Theory’s claim to being a valid unitary-process theory because it establishes that people’s valuation of an entitlement can depend on their relationship to the object in ways that do not affect its actual utility to the person. This possibility is inconsistent with the predictions of Expected Utility Theory.

Similarly, there is evidence that people evaluate outcomes differently depending on whether they *perceive* it as a loss or a gain (regardless of its substantive characteristics) – a related phenomenon known as the framing effect. A person subject to framing effects perceives the promise that she will receive a bonus of \$100 if she reaches a particular achievement goal differently than if the \$100 is built into her initial wage and she is told that she will be docked \$100 if she does not hit that same goal. This evidence also is inconsistent with Expected Utility Theory because evidence that valuation depends on the subjects’ perceived status quo is inconsistent with the premise that utility functions are fixed, and do not vary with context. This is not only troubling for generating positive theories about how people will behave after a legal change, but it gives rise to significant normative challenges for economists, who tend to pursue utilitarian ends in their normative approach. In turn, it may generate a potentially

¹⁶ We have specifically mentioned the intellectual “battle of the bands” between Expected Utility Theory and Prospect Theory here, not only because it cleanly identifies a domain where two competing accounts of behavior prevail, but also because this debate among unitary-process theories has been at the very core of experimental economics, experimental psychology, and Experimental Law and Economics for nearly three decades. Many (though certainly not all) researchers probably still direct their work towards engaging this debate (or even resolving it), hoping to determine once and for all whether Expected Utility Theory or Prospect Theory is the “best” explanation of human behavior.

stronger case for legal rules to pursue paternalistic goals, associated with maximizing agents' "a-contextual" preferences.

Although the existing experimental results appear to refute the claim that Expected Utility Theory is a valid unitary-process theory, the experimental results do not support the claim that Prospect Theory can stand as an alternative valid unitary-process theory. Surveying this vast literature, one cannot escape the striking conclusion that neither of these theories is a valid unitary-process theory (unless it can be shown that all of the experiments on one side of the debate are wrong). Indeed, the evidence appears to refute the claim that these decisions are governed by either unitary process of decision-making. A large body of experimental evidence now shows that people systematically behave in a manner consistent with Prospect Theory in some circumstances, and yet in other circumstances they make choices entirely consistent with Expected Utility Theory.¹⁷ These experiments (if valid) refute the claim that either Expected Utility Theory or Prospect Theory is a valid unitary-process theory. It seems unlikely that there exists a unitary-process theory which can explain all the conflicting results.

A Case Study: The Endowment Effect

Our claim that neither Expected Utility Theory nor Prospect Theory is a valid unitary-process theory is likely to be controversial (at least within Experimental Law and Economics). Thus, it is useful to defend it with a more in-depth discussion of some of the literature on the issue. One of the richest lines of research in the battle between Expected Utility Theory and Prospect Theory is the extensive line of experiments on the endowment effect. Although we summarize this literature in more detail later, it is useful to consider briefly the relevance of this literature for the claim that a valid unitary-process theory of decision-making may not exist.

A central tenet of Expected Utility Theory is that people's preferences are constant across all states of the world (real or perceived). This assumption implies that mere ownership or a sense of entitlement to possession should not, after correcting for wealth effects and transaction costs, alter how people value an object. This implies that *ceteris paribus*, there should be no difference between the maximum amount that a person should be willing to pay to buy an object and the minimum amount she would be willing to accept to sell it. By contrast, Prospect Theory posits that valuation is state dependent: people evaluate their welfare in a given state with reference to the status quo, experiencing a given negative change in wealth with respect to the status quo more strongly than an equivalent positive change in wealth with respect to the status quo. Proponents of Prospect Theory argue that this is consistent with the hypothesis that the amount that someone values an object may depend on their perceived status quo – specifically, on whether the person feels an ownership entitlement to the object. Under

¹⁷ Similarly, recent neuroscience evidence indicates that people do evaluate uncertainty in ways consistent with Expected Utility Theory, in that people presented with choices that give rise to uncertain payoff select actions based on consideration of both the magnitude and the probability of the resulting outcomes, as Expected Utility Theory suggests. Indeed, researchers have provided evidence that even non-human primates select between uncertain outcomes based on their expected value (weighing both probability and magnitude), consistent with Expected Utility Theory; they also appear to engage in the type of strategic behavior suggested by the Nash Equilibrium (Glimcher & Rustichini, 2004). Behavior in markets also often conforms to Rational Choice Theory.

this view, when individuals feel entitled to an object their willingness to accept payment to part with the object will tend to exceed their willingness to pay to acquire it had they never possessed it. The spread between one's willingness to accept and willingness to pay is frequently called the *endowment effect*. This particular phenomenon has been of considerable interest to Law and Economics scholars, given its rather direct relationship to one of the most central theoretical precepts of the field: the Coase Theorem.

Experimental treatments of the endowment effect not only purport to test the phenomenon in its simpliciter, but also to test the underlying validity of Expected Utility Theory as a unitary-process theory. They also bear on the validity of Prospect Theory as a valid unitary-process theory of decision-making (Tversky & Kahneman, 1991; Kahneman, Knetsch and Thaler, 1991). A review of the experimental evidence on the endowment effect reveals a mixed landscape.

On the one hand, there are now countless examples in the literature of experiments where the evidence appears quite inconsistent with EU but consistent with Prospect Theory. Kahneman, Knetsch and Thaler's seminal endowment effect experiment is an example of this type of finding. There, the experimenters found that subjects valued a university coffee mug more if they initially felt a sense of ownership of it than if they did not (even correcting for transaction costs and wealth effects). This finding is inconsistent with the premise of Expected Utility Theory that valuation is independent of entitlement *per se*, but is consistent with the predictions of Prospect Theory. Numerous subsequent experiments, both here and in other countries, have found that subjects exhibit an endowment effect with respect to a variety of different objects (Henrich, et. al., 2005; Horowitz and McConnell, 2002 (surveying studies)). If these experiments are valid, this evidence would seem to invalidate the theory that people make all decisions based on Expected Utility Theory.

On the other hand, experimental evidence does not uniformly line up consistently with these findings. Many experiments in fact find no difference between Willingness-to-Accept and Willingness-to-Pay in situations where Prospect Theory would predict that an endowment effect should occur. Specifically, experiments have shown that people endow some goods but not others (Horowitz and McConnell, 2002 (surveying studies)). Subjects also do not exhibit an endowment effect with respect to "numeraire goods," such as money or tokens that have no intrinsic value but can be traded for goods (including mugs). Moreover, even with respect to goods that often produce an endowment effect, researchers have shown that subjects may not exhibit an endowment effect in some circumstances. For example, one field experiment found that people who collect trading cards exhibit an endowment effect, but this effect dissipates once they start trading the cards regularly (List, 2003). Moreover, Arlen, Spitzer, Talley (2001) found that a simple change in the framing of the experiment can eliminate the endowment effect even with respect to University coffee mugs.

All told, the existing evidence, read at face value, does not establish that people always exhibit an endowment effect, nor does it establish that they never do. These conflicting results are inconsistent with both Expected Utility Theory and Prospect Theory, at least when viewed as unitary-process accounts of human behavior.

There are numerous potential implications of this conclusion. Our preferred interpretation is that to the extent that either Expected Utility Theory or Prospect Theory predicts when subjects exhibit an endowment, the current evidence strongly

undermines the claim that either account constitutes a satisfactory unitary-process theory. We conjecture that it is likely that people do not employ one process to make all decisions, but instead employ multiple programs, a phenomenon that would explain the mixed results in the literature.

Nevertheless, before elaborating on that conjecture, we should also consider two other plausible hypotheses for the apparent disagreement in the literature. The first is that there is another yet-to-be-developed unitary-process theory that would better predict individual behavior. This seems possible, given the fact that both Prospect Theory and Expected Utility Theory themselves were the products of observed anomalous behavior inconsistent with existing explanations. To the extent that this explanation has merit, it warrants encouraging future theorists to craft alternative conceptual models to predict behavior in a way that offers a better explanation than both extant paradigms. For those interested in basing *current* policy on *current* research, however, the possibility of a future, more explanatory theory is of little practical value.

Another possible explanation for the mixed experimental results may be that some (or all) of the experiments are in some way invalid. Experimental results may be invalid for a variety of reasons. For example, experimental results will be invalid if subjects misunderstand the tasks they are asked to perform, are indifferent to their tasks, or are influenced by systematic factors that the experimenters did not control for, such as their non-anonymity with the experimenter. It is important to consider this possibility because, as we discuss in greater detail below, some experiments claiming to test for the presence of an endowment effect employed poorly specified experimental procedures which undermine the validity of their results (See, e.g., Plott & Zeiler, 2005, arguing that many studies finding an endowment effect either do not exclude other factors or poorly inform subjects of the nature of their decisions).

Although experimental design flaws always deserve serious consideration, our view is that the breadth and depth of extant studies renders implausible claims that such flaws have systematically permeated one side of the field or the other. We reach this conclusion based on the evidence from a number of quarters. First, within Experimental Law and Economics, the endowment effect is one of the most widely studied phenomena. While some studies finding the endowment effect admittedly fail to control for various important factors (e.g., incentive compatibility, strategic bargaining, wealth effects, anonymity effects, confusion, etc), even after their exclusion there are a number of solid studies that do offer such controls. Second, there are numerous experiments within the literature that test other predictions of Prospect Theory, and show that people behave consistent with Prospect Theory in some circumstances and Expected Utility Theory in others (See, e.g., Babcock and Loewenstein, 1997; Guth, et. al., 1982; Camerer, 2003; Fehr and Fishbacher, 2003). Moreover, evidence from the (smaller) number of debiasing studies appears consistent with leading multiple-process theories, in that it suggests decisions are the product of both deliberative and non-deliberative decision-making processes whose actuation can be manipulated by context.¹⁸ For example, there is evidence that subjects evaluating litigation settlement decisions display self-serving biases that distort their decisions, but this bias can be eliminated through rules or institutions that require them to consider the weakness or

¹⁸ For an interesting exploration of the normative implications of debiasing, see Christine Jolls and Cass Sunstein (2006).

their case or the reasons a judge might rule against them. (Babcock, Loewenstein, and Issacharoff, 1997).

Lastly, the literature from neuroscience suggests that there may be sound physiological reasons to believe that multiple-process theories are at play in many arenas of human behavior. We now review this literature.

B. MULTIPLE-PROCESS THEORIES OF DECISION-MAKING

A growing number of economists and neuroscientists are converging on the view the people employ multiple decision-making programs, often simultaneously, when they make decisions (E.g., Camerer, Loewenstein, and Prelec, 2005; Cosimedes & Tooby, 2006).¹⁹ If this is the case, then the quest for a valid theory of human decision-making lies in finding a more general meta-theory that identifies the structure of the different programs employed and also the factors that determine which program is employed.²⁰

One set of multiple-process theories posits that one set of decision-making programs is conscious and deliberative. These processes are involved when we reason carefully about choices and consider the consequences of our actions. These conscious decision-making programs may be consistent with Rational Choice Theory in some circumstances. Yet these conscious programs are not the only ones that affect choices. People's choices also are influenced by non-conscious, automatic decision-making programs that operate below our level of self-awareness. Examples of intuitive decision-making processes include emotional responses that determine people's

¹⁹ Among a number of experimental economists, this generalizing move is already under foot. A number of scholars have attempted to posit a "Generalized" Expected Utility (or GEU) account of behavior that folds in both conventional Expected Utility Theory and conventional Prospect Theory as special cases of a more robust theory of decision-making that inhabits a broader domain (see Camerer & Talley 2007 for a technical overview). For example, it may be the case that endowment effects emerge in situations where, say, parties are anticipating interacting with others within a bargaining setting, and they subconsciously augment their bargaining toughness by placing premia on objects they already possess while discounting those that they do not (See e.g., Heifetz et al. 2007). These sorts of theoretical inquiries are direct responses to the context-dependency of many sorts of behaviors (such as the endowment effect).

²⁰ In other words, the insistence on universal mutual exclusivity between dominant accounts of human decision may be the problem, and both approaches should be folded into more general, multiple-process theory that maps contexts to decision-making programs and ultimately to outcomes.

Such reflective, generalizing moves have occurred in other famous instances. For generations, for example, theoretical physics was dominated by two competing theories about the nature of light. One theory held that light is comprised of particles. The other held that light is comprised of waves. Despite decades of experiments, neither side could decisively win the debate. Some experiments found evidence of light behaving like waves. Others found evidence inconsistent with the wave theory, but consistent with the theory that light behaves like particles. Physics was only able to break the stalemate when it recognized that the competing theories were incorrectly conceived as unitary-process theories of light. Rather, they turned their attention to a meta-theory that attempted to characterize light as particles in some circumstances and waves in others. Ultimately, then, the theory morphed into one that attempted to understand the contexts under which the particle versus light analogy was most apt. It took physicists a long time to get there because it took them forever to figure out it wasn't a battle of bands. It took them forever to understand that neither theory was right. We only got traction when someone posited a theory to explain both results.

behavior far more than they recognize.²¹ Other non-conscious responses include intuitions that guide our reactions about what is ethical or moral. (For a review of some recent evidence from neuroeconomics see Sanfey et. al., 2006; Camerer and Fehr, 2006; Camerer, et. al., 2005). These more intuitive decision-making processes can intervene to affect our decision-making processes before our conscious decision-making has begun to operate.²² In some cases, these more intuitive processes may distort the strength of the information that reaches our deliberative processes. In other cases, they may operate as weighty intuitions that a particular choice is the “right choice.” These intuitions about what is “right” may dictate behavior even when we deliberate carefully.

The hypothesis that decision-making is affected by both conscious and non-conscious decision-making programs, which operate quite differently, is supported by a growing body of research. For example, there is evidence that the “intuitions” (including moral judgments) that guide people’s decision-making are not the result of conscious reasoning but instead are the product of emotional processes that operate outside the zone of our conscious decision-making.²³ Moreover, researchers have found that people’s intuition, emotions and moral judgments may intervene to override the choices that we would make if guided only by our conscious, deliberative decision-making processes.²⁴ In turn, it has been shown that one can alter decision-making by altering the strength of these non-conscious processes.

²¹ “Although visceral factors ...play an essential (and probably dominant) role in human behavior, people’s introspection about the causes of their own behavior lead them to under appreciate the influence of visceral factors and to exaggerate the importance of higher-level cognitive processes. Numerous studies have employed diverse methods to show that people tend to interpret their own behavior as the result of deliberative decision-making even when this is not the case.” (Loewenstein, 2000: 427). For a review of the literature on this issue see Wegner & Wheatley (1999).

²² These more automatic decision-making programs are not simply instincts, in the classic sense of the term. These non-conscious processes also include reasoning programs, decision rules, and moral intuitions that guide our views as to what is the right choice.

Some argue that these “intuitive” programs evolved over thousands of years to solve recurrent situations presented to early man, and that we continue to apply today without any awareness that we are doing so. Indeed some argue that our brains evolved quite precise decision-making programs specialized to solve particular recurrent problems – such as how to avoid predators, cooperate with others, and protect children. It is argued that these programs include one designed to solve decision-making problems associated with cooperation in small hunter gatherer societies. These programs – covering issues such as reciprocal cooperation, collective action and so forth -- reside within us still, it is argued, and are implicated automatically when we make decisions in certain social contexts notwithstanding the fact that modern social contexts often no longer involve the small group, repeat-interaction situations that our programs were designed to address. According to this view, in order to predict how people will behave, especially in contexts where their decisions will be affected by their moral intuitions, we need to understand which program the particular context will implicate and what inferences, motivations and emotions will that program produce (Cosmides and Tooby, 2).

²³ Numerous neuron-imaging studies show that tasks involving moral judgments activate brain areas known to process emotions. *E.g.*, Moll, et. al (2002); Heekeren, et. al., (2003); Greene, et. al., (2001); Greene, et. al., (2004); Luo, et. al., (2006); see generally Hauser (2006).

²⁴ Other lines of research also support the multi-process hypothesis. For example, in one study researchers examined the role of moral intuitions by comparing decision-making of normal subjects with subjects who have a lesion in part of the brain that guides particular moral choices. The lesion does not otherwise affect the subjects’ decision-making. The lesion subjects made choices consistent with “normal subjects when presented with choices that did not involve a particular type of moral dilemma. Yet they deviated when presented with the moral dilemma of whether they would kill someone if it meant saving other people. The lesion subjects answered consistent with strict rationality, they would act to kill

For example, in one experiment, researchers affected their subjects' decision-making simply by altering their hormone levels, while leaving all other features of the choice unaffected.²⁵ Building on evidence that people's brains release the neuropeptide oxytocin in response to an intentional signal of trust from a stranger (Zak et. al. 2005), researchers examined whether they could alter levels of "trusting" behavior by altering subjects' oxytocin levels. The oxytocin manipulation altered trust even when no other factors changed that would give rise to more trust. Even more striking, the experimenters found that oxytocin had a quite targeted effect, affecting only certain forms of trusting. Specifically, oxytocin did not affect other-regarding behavior in general. For example, it did not make subjects more trustworthy of themselves. Nor did it affect a willingness to take risk in general. Instead it appeared to affect subjects' willingness to take particular social risks arising through social interactions – like the risk of exploitation (Kosfeld, et. al., 2005). This study thus highlights the effect of non-conscious emotional responses on decision-making (Damasio, 2005, Fehr et. al., 2007b) and also shows that these responses can be quite domain specific. Accordingly, evidence that the response is triggered in one context would not imply that it would be triggered in another context.

Additional experiments suggest that people employ different decision-making programs when making choices involving social relationships than they do in other contexts. For example, one recent and particularly interesting brain-imaging study showed that people use a different part of their brains to reason about a situation where a person violated a social contract than they use when reasoning about an equivalent violation of a precautionary safety rule (Ermer, et. al., 2006). Moreover, responses to violations of social relationships appear to include the kind of neurological responses that promote the development of reciprocal cooperation between people. For example, there is evidence that subjects who decide to punish a person who has violated a trust norm activate brain areas related to the processing of rewards. This research suggests that people who punish those who abuse trust experience positive satisfaction, which

one person if it meant saving a great many more. The normal subjects generally refused to do so, concluding that it was wrong to sacrifice the innocent person even if many would perish as a result. These authors concluded that their evidence supports the conclusion that people's actions in these situations normally are guided by an interaction between intuitive reactions about right and wrong and conscious deliberation, with the latter often being quite powerful. As for the lesion patients, the lesion in question eliminated this intuitive program, enabling their conscious, deliberative processes to rule the decision unchecked (Koenigs, et. al, 2007). Another interesting feature of this study is that it provides evidence that decision-making programs may be quite problem specific. The lesion patients in this study exhibited non-standard moral judgments only for a specific class of problem; their moral judgments were normal for other types of moral problems (Koenigs, et. al, 2007).

²⁵ Another recent study supports the ideas that decision-making is the result of a complex interaction between deliberative processes (which include a concern for self-interest) and more intuitive or emotional processes that operate outside of the domain of our attention and that people may exhibit automatic processes that are targeted to particular classes of problems, and thus are domain specific. In one experiment, researchers discovered that they could alter the behavior of subjects playing the Ultimatum game by disrupting a particular area of the brain (the right dorsolateral prefrontal cortex, DLPFC). Subjects whose right DLPFC was disrupted were more willing than control subjects to accept an unfair offer in an Ultimatum game. These subjects retained their ability to determine that the offer was unfair, but no longer were willing to reject it. The researchers hypothesize that the DLPFC operates to override people's selfish inclinations to accept positive offers in order to reject offers perceived to be unfair (Knoch, et. al., 2006).

operates as a kind of a neurological reward for actions that promote trust by punishing trust abusers (de Quervain, et. al. 2004).

This evidence, while preliminary, suggests that experimental analysis of human decision-making should attend to the possibility that people may employ multiple programs when making decisions. These programs may be consistent with Expected Utility Theory or Prospect Theory (as non-unitary programs). This possibility has important implications for experimental analysis. It implies that experiments should test both unitary-process and multi-process theories. It also implies that those interpreting experiments designed as a “battle of the bands” between two unitary-process theories should examine not only whether the experiment establishes the dominance of one unitary-process theory over the other, but also whether the experimental results are sufficiently robust to suggest that the winning unitary-process theory dominates over competing multi-process theories.

While we believe that the quest for a predictive theory of human decision-making will end with the development of a multi-process theory, we acknowledge that this position is by no means universal. Many scholars, especially in Experimental Law and Economics, explicitly or implicitly hold fast to the view that human decision-making is best analyzed as a result of a unitary process. Some scholars who embrace this view may do so for many reasons. Some may have been affected (wittingly or not) by a decades long “us versus them” debate between the two dominant unitary-process theories -- Expected Utility Theory and Prospect Theory. Not only has the rhetoric of this debate encouraged such dichotomous views, but scholars steeped in one or the other may have become habituated to viewing decision-making in a unitary-process fashion. Others may resist multi-process theories for more reflective reasons, having to do with the ease with which one can test (and, in particular, falsify) a multi-process theory of human behavior. Such epistemological concerns deserve explicit attention.

C. TESTING MULTI-PROCESS THEORIES

While multi-process theories represent, in our view, a more parsimonious *descriptive* account of human cognition, they are without a doubt more difficult to test than are their unitary-process cousins. As noted above, unitary-process theories are particularly good subjects for experimental testing because they presume people to employ the same decision-making program to *all* contexts. Consequently, unitary-process theories can be falsified even in the most unrealistic of experimental settings: for if a unitary-process theory is correct, the program employed in the lab should be the same one employed in the outside world. Indeed, artificial experimental settings provide, perhaps ironically, a good test of unitary-process theories, since they permit a researcher to exercise careful control over the factors affecting subjects’ choices. A replicable experimental result that is inconsistent with a prediction from a unitary-process theory constitutes a strong refutation of the theory itself.

The opposite is true for multi-process theories. The first problem is that multi-process theories, by their very definition, inject additional analytical “degrees of freedom” into the predictive enterprise. In an important respect, multi-process theories are a type of meta-theory that subdivides the universe of human behavior into specific contextual arenas, effectively assigning that arena to some subsidiary cognitive

algorithm (such as EU, Prospect Theory, or something else). For example, there is evidence that emotional environments can cause people to deviate from the predictions of Expected Utility Theory and toward those of Prospect Theory. However, there also appears to be evidence that the ability of emotions to effect such a transformation may turn on other attributes of the context in which the decision is made. A robust and predictive multi-process theory of decision-making, then, must not only postulate the set of different processes that people might employ, but also must formulate a complete set of predictions about how these processes interact in different contexts. In short, multi-process theories must combine theories of behavior with theories of context, *and do so in a systematic way that can be tested*. And herein lies the rub: For although virtually anyone can posit some arbitrary formulation of a multi-process theory (e.g., “People employ EU on weekdays and Prospect Theory on weekends”), evidence inconsistent with that formulation does not similarly falsify multi-process theory *per se*; it may merely suggest that the posited formulation of the theory was misspecified.

The greater imperviousness of multi-process theories to falsification, in turn, raises a second challenge for researchers and policy makers: multi-process theories are currently not fully specified. Instead, the evidence favoring multi-process accounts largely consists of an assorted collection of examples from the research laboratory. Researchers have yet to match up *inductively* the assorted set of examples with a *deductive* theoretical account of context. Even more fundamentally, the literature is still largely at the stages of development / early calibration of theory. As noted above, this current state may be of limited use to policy makers, who have to make decisions now, and accordingly are limited to using current research results to inform those decisions.

Nevertheless, we think that the existing evidence for multi-process theories carries valuable lessons for scholars and policy makers of all stripes. First, for multi-process theories to be testable, they must focus on decisions that are likely to be affected by a few easily identified and intuitively attractive contextual factors.²⁶ Second, researchers interested in drawing short-term policy applications from their work might do well to develop experimental protocols that endeavor to capture some of the important elements of context that pervade the “real-world” arena of decision-making that one is attempting to explore.²⁷ A recent trend towards field experimentation, where one cedes some of the experimenter’s control over environment in exchange for real-world context, may provide a promising middle avenue for such work. Finally, before employing experimental results, policy makers should take special care to confirm whether the experimental setting is a relatively good match with the real world. Such confirmations can act as a responsible check on whether other factors that may affect context (and in turn, cognitive processes) are adequately captured in the laboratory environment. Nevertheless, notwithstanding these difficulties, we expect that it is likely that the quest for a predictive theory of human behavior will end with a multi-process theory. Experimental Law and Economics can

²⁶ Viewed from the negative perspective, researchers should avoid extravagant (or ad hoc) theories of context that themselves require troves of experimental data to test, so long as the testability of simple theories about context are not overly restricting.

²⁷ Although experimental psychologists have long been comfortable with contextually rich experimental instruments, their counterparts in economics have traditionally preferred highly artificial and a-contextual environments.

contribute to this venture by designing experiments that contribute to our understanding of how these multiple processes interact.

D. A WORKING TYPOLOGY

The preceding discussion reveals that experiments can be classified along two quite different dimensions: (1) the purpose of the experiment (to test a theory or develop one) and (2) the type of theory being tested (or implicated) by the experiment. This theme is reflected in Table 1:

Table 1: Goals of Experiments (in columns) versus Domain of Theory (in rows)

	Test Theory	Develop/Calibrate Theory
Unitary-Process Theory	A	B
Multiple-Process Theory	C	D

As was previously noted, many experiments may be read and evaluated in a contingent fashion, according to each of the cells in the above table. Experiments have the broadest immediate policy implications when they are testing an existing theory (left column), because the theory helps determine the breadth of the implications for legal policy. In considering these implications, however, it is important to examine whether the experiment is designed as a test of a unitary-process theory (cell A) or a multiple-process theory (cell C). Even when the experiment tests a unitary-process theory, it is important to consider whether the implications for legal policy would be different if instead decision-making were governed by multiple processes. This possibility potentially narrows the implications of many experiments, since the development of such theories is still underway (and thus the number of contributions in cell C may be small or even non-existent). As we will discuss in the next section, the ingredients of a “good” experimental design turn on how the study is seen to fit into the above organizational matrix.

III. ELEMENTS OF GOOD EXPERIMENTAL DESIGN

Understanding the multiple purposes and possible implications of Experimental Law and Economics reveals that there is no “one size fits all” set of desiderata for an experiment within Law and Economics. Rather, the elements required for a good experiment depend on the purpose of the experiment and the domain of the theory that motivates the study.

Nevertheless, all experiments must satisfy at least six criteria, although the importance of these criteria varies across experimental contexts. These six criteria are: Control, Internal Validity, Falsifiability of Theory, Replicability, External Validity and Contextual Attentiveness. We examine each of these criteria below, discussing the importance of each relative to where the experiment falls within the cells of Table 1. As we shall see, some desiderata (such as control and replicability) are vitally important across all cells. Others (such as falsifiability of theory) are important for all cells, but

more so for some in particular. Yet others (such as contextual attentiveness) may be manifestly unimportant for some cells while being vital for others.

A. Control

“Control” is often thought to be the *sine qua non* of experiments. It is particularly important for experiments seeking to test existing theories of human behavior. A study is said to have achieved control if the experimenter controls the factors that affect choices. For experiments that test a theory of behavior by comparing the behavior of a “control group” of subjects against a “treatment group,” this implies that the experimenter must control the differences between those two groups. The experimenter thus must control both the variation in the subject pools and also the differences in the factors relevant to the subjects’ choices.

Control enables experimenters to make causal inferences about how different factors do or do not affect choice. Indeed, the ability to control the factors operating on choice is one of the key potential strengths of experimental methods relative to empirical methods. Consequently, good experiments often take great pains to ensure that the differences between the “treatment” and “control” groups are fully pre-programmed. Rigorous adherence to protocols and consistency among experimenters are all key ingredients of maintaining experimental control.

Beyond this, control requires that any differences in outcomes between the control group and the subject group be driven by only those differences in context that the experimenter intended to introduce and not extraneous factors. This goal is easier to attain the better the subjects understand the choice, the more carefully subjects attend to the choice, and the fewer the potential extraneous considerations that might be relevant. For this reason, experimentalists often endeavor to keep their methodologies simple.

Experimenters often also attempt to impose control by granting their subjects practice rounds to familiarize themselves with the nature of the task that they will perform. This can increase control by reducing the likelihood that choices are the product of unintended subject confusion. Nevertheless, caution is needed in employing such mechanisms in some circumstances because practice rounds can have implications for the “internal consistency” of the experiment, as discussed below.

To induce subjects to focus on the factors of interest to the experiment, and not extraneous factors, experimenters employ mechanisms to induce subjects to attend carefully to the decisions they are being asked to make. Sometimes the experimental protocol may be sufficiently interesting to grab some or most subjects’ attention. The leading way to induce subjects to focus on the choice in question is to design the experiment so that the subjects’ choices affect how much money they receive at the end of the experiment.²⁸

In addition, researchers must consider whether subjects’ decisions may be affected by reputational considerations in ways that the researcher has not taken into account. There is evidence that subjects adjust their behavior (perhaps non-consciously) when the experimenter will know their choices, perhaps out of a desire to

²⁸ This claim that experiments should grant subjects payoffs that conform to their choices is accepted by experimental economists, but not by experimental psychologists. Indeed, it is one feature that distinguishes experimental economics from experimental psychology.

ensure that the experimenter (often a professor) thinks well of them. Subjects who respond to reputational concerns might make a decision that appears to deviate from rational choice, and yet which nevertheless is rational once the subjects' reputational concerns are taken into account. Accordingly, in situations where subjects might expect the researcher to prefer one choice over another, experimenters should control for this effect through a blind, or double-blind, protocol.²⁹

In addition to maintaining control over what happens to subjects in the laboratory, the experimenter must meticulously assign subjects to experimental pools in a way that ensures comparability of the treatment and control groups. Thus, for example, it would not be appropriate for an experimenter to use students from one university course as a treatment group, and those in another course as a control. It would similarly be inappropriate for an experimenter to moderate one pool while her research assistant moderated the other. While random assignment of subjects between the control and treatment groups is commonly thought to be a preferred means for controlling subject pool characteristics, with small experiments it is sometimes defensible to "engineer" some of the assignment to ensure balance of demographic, ethnographic, and economic characteristics among subjects.

Because control is such a central aspect of all experimental work, it is clearly important regardless of one's reasons for conducting an experiment. However, the importance of control is perhaps sharpest when one wishes to test specific (and often causal) theories of behavior. Control may be the most difficult to attain when one is seeking to test or develop a multi-process theory of decision-making because these theories give more explicit attention to factors likely to influence non-conscious decision-making processes (including emotions).

B. Internal Consistency

Another fundamental condition for a good experiment is that the factors that subjects actually focus on in making their choices should be the factors whose influence the experimenter is interested in testing. Decisions should not be affected by other factors not considered in the experimenter's tested hypothesis. Internal consistency is particularly important for testing theories of behavior. However, it is far from irrelevant for developing theories as well, because it is most helpful to develop a new theory against the backdrop of some existing hypothesis, often through testing (and falsifying) some domain of that existing hypothesis.

Internal consistency implies that the experimenter must ensure that the choices faced by the control and treatment groups (1) in fact differ in the ways the experimenter hypothesizes and (2) do not differ in some other ways as well. For instance, to test the endowment effect, it will be important to establish in the endowed subject pool some sense of entitlement to the item, while establishing clear non-entitlement in the unendowed group. It also is important to ensure that the endowed and unendowed groups do not differ in other ways – for example, it is important to ensure that the endowed subjects do not also feel that the experimenters want them to keep the object,

²⁹ Specifically, there is evidence that subjects' behavior differs in experiments where the experimenter knows the identity of subjects from situations where the experiment is double-blind (Hoffman, et al, 1994; Plott & Zeiler, 2005).

since in this case the experiment would test more than the effect of possession on valuation (see Plott and Zeiler, 2005).

The internal consistency criterion implies that it is important to ensure that the payoffs subjects receive when they make choices in the experiment mirror the payoffs of the choices that the researcher is seeking to examine. In the case of experiments seeking to examine decisions with financial consequences in the real world (e.g., purchase and sale decisions or gifts of money), this implies that experimental subjects should be given choices with actual, and not imagined, financial consequences for them. “Incentive compatibility” between subjects’ choices and their payoffs at the end of the experiment is important to internal consistency because there is no assurance that people evaluate choices with imagined or hypothetical payoffs the same way that they evaluate choices with real payoffs.

Incentive compatibility is particularly important when the experimental choice involves both financial and non-financial considerations, such as the subjects’ desire to be “moral” or “good.” For example, lack of incentive compatibility may undermine the internal consistency of an experiment designed to examine subjects’ reaction to a choice between an option that would enhance the subject’s financial welfare and one that the subject would view as involving more socially desirable behavior because subjects can more readily decide to be “good” if the financial consequences are purely imagined than they would in a situation involving real stakes (see generally Camerer and Hogarth, 1999). The likelihood that decision-making employs multiple processes heightens the importance of ensuring incentive compatibility in these situations.

In the case of experiments testing decisions which have no financial impact (such as a jury’s decision about guilt in a capital murder case), there is still some unresolved debate about whether monetary compensation is appropriate. The challenge with such experiments is that the real-world choice one is examining has an emotional consequence but not a financial consequence for the actors, whereas the experimental choice may have neither to the degree to which subjects know they are making a purely hypothetical choice. In such a circumstance, it can be argued that financial rewards may improve internal consistency if the financial gain operates as a valid proxy for a socially desirable gain, or some form of emotional consumption. Yet financial considerations also may distort choices by distracting subjects’ attention away from the moral considerations which would be salient in the real world. Evidence suggests, for example, that subjects are more likely to conform to rational choice predictions when their choices affect them financially or with easily commodified compensation (Camerer and Hogarth, 1999). Accordingly, while there is relatively sound justification for compensating subjects in experiments involving financially salient decisions,³⁰ with

³⁰ The proposition that robust experimental tests of decision-making require that experimental subjects be presented with experiments where the financial and non-financial rewards are similar in type and *relative* magnitude to the context being considered finds support in recent evidence on decision-making emerging out of Neuroeconomics. As previously discussed, these experiments suggest that people do not employ a single decision-making program to make all decisions. Instead, people use multiple processes that interact with each other. Moreover, the intuitive processes often are triggered automatically and are outside of our control. The actual outcome of the decision depends on how the deliberative and intuitive programs interact. This in turn depends on the presence of factors that may mute or trigger our intuitive decision-making programs. When a decision triggers an intuitive program, a subject may follow it completely – even when it is contrary to his hypothetical financial interests – if it is

non-financial decisions experimenters should take some care to consider how to motivate subjects in the least distortionary way.

Internal consistency concerns also are relevant when experimenters decide whether to run practice rounds or not. If the experiment involves testing or developing a theory about decisions that people make repeatedly – and thus after learning has occurred – then practice rounds may be warranted. But some decisions occur as one-shot transactions. In this context, practice rounds may enhance control but undermine internal validity to the extent that the choices people make are altered by the learning process. For example, Professor Charles Plott has suggested that people often do not know their own preferences initially, but learn what their own preferences are as they make decisions and experience their own reactions to the outcomes. Subjects may expect to feel one way in the beginning, but learn that they have a quite different set of preferences. For example, they may learn that they are not as averse to losses as they expected to be (Plott, 1996). In addition, under a multi-process theory of decision-making, learning could affect decision-making if subjects are more likely to be affected by deliberative as opposed to intuitive decision-making programs the more experience they have with making a particular choice. These theories about the role of learning suggest that an experiment run at the end of a learning process would provide valuable data on decision-making after learning, but might not provide valid results about how people would make similar choices for the first time. This suggests that additional experimental attention is needed to understand the ways in which the process of decision-making may affect preferences and thus outcomes.

The experimenters' choice of subject pool may present internal consistency (and control) concerns. For example, an experiment designed to test settlement decisions by people who have no information about the relevant legal rules may in fact fail to test that choice if the experimenter selects experienced lawyers as subjects and considers a dispute where the law is not completely uncertain. The subjects may be unable to ignore their private information about the likely outcome, and this may result in their making decisions based on a different factual premise than assumed in formulating the experimental hypothesis to be tested, even if told to assume that the outcome is uncertain. This problem is particularly likely to occur in experiments run on lawyers or law students that are designed to test settlement behavior with respect to a particular legal dispute where law student or lawyer subjects may have sufficient knowledge of the law in question to have private views about the likely outcome that diverge systematically from the likelihood the experimenter assumed in designing the experiment.

not contrary to his actual financial interests. Even when paid, he may be more likely to follow the intuitive program if the consequences of his choice are not salient to him – for example, if he plays the experiment only once for actual pay and he does not have the experience of following it and experiencing a financial loss. The balance between deliberative decision-making and intuitive processes may differ when subjects engage in paid practice rounds that allow them to experience the losses associated with certain actions (For evidence that paid practice rounds can affect outcomes, see Plott & Zeiler, 2005).

C. Falsifiability of Underlying Theory

Because experiments are a principal way to test/develop predictive theories, a theory's falsifiability is an important consideration in designing an experiment. A universal statement is said to be *falsifiable* if it can be shown to be false by either empirical or experimental observation. Long a precept of the physical sciences (albeit not free from controversy), falsifiability has also been embraced by both economists and psychologists, whose numbers are well represented in the pages of this introduction. As was discussed above, unitary-process accounts of human behavior are more likely to satisfy this criterion, since they make categorical predictions about human cognition across contexts. If one is purporting to use an experiment to test a theory, then it is absolutely critical that the theory itself be falsifiable – for if it is not, then the results of the experiment cannot serve the purpose to which they are ascribed. In contrast, if one is primarily interested in developing or beginning to calibrate a new theory, then the demands of falsifiability recede somewhat (though not completely). Here, the appropriate goal is largely to fill out the details of a theoretical account rather than to test it. However, the task of filling out those details is most helpful if it is done in a way that adds to the theory's prospective testability by others. An experiment that is intended to explore the contours of a multi-process theory of cognition, then, should (perhaps among other things) focus on factors that will help future researchers merge inductive observation with deductive theories.

D. Replicability

In a similar vein, experimental approaches in Law and Economics must be replicable, in that other experimenters should be able to employ the same techniques, the same protocols and the same inducements to similar subject pools to attain comparable results. Experiments whose protocols are unclear, not followed, or too dependent upon experimenters' idiosyncrasies not only raise serious concerns about replicability, but also cast doubt on whether the experiment is one that had sufficient control.

A key protocol for maximizing replicability involves assiduous and painstaking record keeping by the experimenter. It also is important that the experimenter develop a script for the experiment and follow it to the letter without casual conversations or clarifications made to subjects. In addition, one would also be well advised to keep track of specific ways in which the experiment is carried out, including the nature of the laboratory surroundings, the seating arrangement, contact with the experimenter, and how much time was spent on each aspect of the experimental protocol. Experimenter gender may even be relevant. Because replicability is central to all scientific inquiry, it is a critical desideratum regardless of what one's purpose is in experimenting.

E. External Validity/ Generalizability

A fifth desideratum for Experimental Law and Economics is that an experiment be "externally valid" (or "generalizable"). This desideratum focuses on whether the experimental design tests a choice that matches up with the real-world settings to which

the experiment is meant to extrapolate. External validity differs from internal consistency in that the latter focuses on whether the experiment as designed tests the hypothesis the experimenter claims to be testing. Internal validity thus focuses on whether the experiment in fact presents subjects with the choices (and manipulations thereof) that are intended by the experimenter. External validity, by contrast, focuses on whether the decision that the research tried to (and presumably did) examine truly can be said to provide insight into the real-world choices that the researcher claims to have provided evidence about.

Whether the researcher must be concerned with external validity depends, of course, on the purpose of the experiment. For example, external validity may not be a real desiderata for experiments designed to test existing unitary-process theories of decision-making. A finding that a subject does not employ the tested process in the lab would undermine the validity of the unitary-process theory as applied in the real world, even if the experiment does not match real-world decisions. (Of course, while an experimenter testing unitary-process theories need not worry about external validity, those evaluating such experiments should to the extent that they believe people may employ multiple processes).

In addition, to the extent that the experiment is not designed to test the validity of a unitary-process theory, but instead is designed to provide insight into how people in fact do make decisions in particular settings – as in the cases of tests of settlement behavior and jury decision-making – then the experimenter must attend to external validity and ensure that the factors dictating choice in the experiment are analogous to those of real-world settings. This concern has implications both for experimental design and subject selection. For example, external validity (and internal consistency) concerns are implicated when experimenters seek to examine how juries make decisions based on population surveys which ask random people how they would decide a particular case. This methodology raises external validity concerns because hypothetical juries may weigh competing concerns – for example, between doing justice for the plaintiff and protecting a potentially innocent defendant from having to pay – differently than do real juries faced with both a real plaintiff and the responsibility of imposing a real financial cost on a real defendant. To establish the external validity of this methodology, the researcher must show that people decide purely hypothetical cases (outside the context of a courtroom) the same way that they would decide real cases.

External validity concerns also are relevant when student subjects are used to examine decision-making by people who may have different experiences and preferences than the student subjects. These “subject pool” concerns are less important for simple tests of the validity of the classic unitary-process models absent a reason to believe that the student subject pool would employ a different decision-making program from the general population. But these concerns are relevant if experience or context is likely to affect the decision. In this case, the student subject pool may behave differently. For example, an experiment run on law students may not provide a good test of how Chief Executive Officers will approach a particular decision-making problem. (At the very least, business students might make a better subject pool). An experiment run on the other-regarding behavior of economics majors certainly cannot be relied upon to provide general results about the other-regarding behavior of the

populace at large, as economic majors appear to reason differently (by selection or training) than the general population (Frank, et al. 1993).

In the end, external validity is likely to be more of a concern for Experimental Law and Economics than for other experiments. The more normative one's research goals are, the more important it is to attempt to draw such analogies to the outside world. Because Law and Economics tends to be more normatively focused than much of economics proper, this is an important consideration.

F. Contextual Attentiveness

Particularly for experiments designed to develop or calibrate a theory of decision-making (in particular multi-process accounts), external validity may necessitate an even stricter requirement: that the experimental instrument itself attend carefully to the context of the experimental choice to ensure that the factors operating on the subjects are indeed the factors the experiment claims to be examining, and that these match the factors that would operate in the real-world choice of interest. In particular, the existing evidence that decisions are affected by factors likely to influence whether subjects follow intuitive (or emotion) programs or consciously deliberative ones suggests that great care must be taken to examine how the experimental design affects this trade-off.

On first blush, contextual attentiveness is in many ways related to external validity, but we treat it separately since its focus is somewhat different and more specialized. One can frequently overcome questions of external validity through appeals to reason and common sense. For example, the frequent use of student subjects is often a target for criticism on the grounds of external validity. However, researchers can often offer practical rationales for the practice (including the fact that students are more attentive than professionals to low-stakes compensation, thereby allowing an experiment to scale down compensation and collect more data at a more reasonable cost). Contextual attentiveness, in contrast, is more closely related to the underlying richness of the theory one is attempting to develop / test. For those attempting to generate richer, context-dependent theories of behavior, it is necessary not only to remain mindful of the nature of the experimental choice, but also the context in which it is presented. Here, it may also be important, for example, that an experimental instrument convince student subjects to imagine themselves as the very professionals the study is meant to analyze, and perhaps to situate their choice in a situation that is actually drawn from that context.

G. Synthesis

The discussion above reveals that there are clear desiderata for experiments, but the relative and absolute importance of each may depend on the purpose of an experiment within Law and Economics and the domain of the theory behind it (or implicated by it). Our overall assessment is illustrated by Table 2 below. Each column of the table represents a different cell from Table 1 above.

Table 2

	A Unitary: Test	B Unitary: Develop/Calibrate	C Multi-process: Test	D Multi-process: Develop/Calibrate
Control	V.I.	V.I.	V.I.	V.I.
Internal Consistency	V.I.	V.I.	V.I.	V.I.
Falsifiability of Theory	V.I.	I.	V.I.	I.
Replicability	V.I.	V.I.	V.I.	V.I.
External Validity	N.I.	I.	V.I.	V.I.
Contextual Attentiveness	N.I.	N.I.	V.I.	V.I.

Table 2: Relative Importance of Desiderata as a Function of the Purpose Behind an Experiment. "N.I." designates "Not Important"; "I" designates "Important"; "V.I." Designates "Very Important."

In our view, control, internal validity and replicability are important across all experiments, regardless of their purpose or their animating theory's articulated domain. Simply put, experiments cannot either test or contribute to a theory if they have not sufficiently satisfied these desiderata (at a minimum).

Falsifiability of the underlying theory is also very important (almost by definition) for any study purporting to test an existing theory. While this is almost too self-evident, it is important to remain mindful of whether an experimental research claim in categories A or C accurately engages a falsifiable dimension of a theory. For example, some endowment effect studies may not cleanly test a unitary-process account of EU because they fail to control for income effects, transaction costs, or strategic effects, each of which could explain the existence of behavior that looks like an endowment effect. Note in particular that this requirement is probably the most demanding in Column C, where multi-process theories themselves are likely to be difficult to test. In contrast, experiments meant to develop and/or calibrate new theories also must worry about falsifiability, but not as directly. Rather, they should aim (as discussed above) to generate observations that are likely to contribute to the articulation of a theory that is itself falsifiable.

The desideratum of external validity is, in our view, not as important for experiments designed to test the validity of unitary-process theories. Once an experiment establishes that a unitary process is invalid, that implies that that unitary-process theory is invalid in general. There is no need to consider whether the experiment matches how decisions are made in the real world. Of course, the implications of an experiment invalidating such a theory do turn on whether one accepts the premise that decision-making conforms to a unitary process. In contrast, external validity is very important for experiments designed to test or develop multi-process theories of decision-making. It also is important to experiments that were designed to test a unitary-process theory, but are being evaluated by those who believe that decision-making is governed by multiple processes. External validity is important if decision-making is governed by multiple processes because most multi-process theories suggest that external factors can affect how people make decisions. This suggests that laboratory results do not necessarily tell us how people make decisions in the real

world. Laboratory results only provide insights on behavior when the experiment has external validity. For this reason, our final desideratum of contextual awareness is important for multi-process theories, because these theories assume that external context affects decision-making and thus variations in context can alter experimental results.³¹

IV. EXPERIMENTAL LAW AND ECONOMICS EXAMINED

With this framework in mind, we can now examine where the literature on Experimental Law and Economics stands to date. This will also serve as a preview for the constituent portions of this book. We draw two important conclusions from this literature.

First, this literature suggests that the best hope for developing a predictive theory of human decision-making lies with multi-process theories that are sufficiently rich to capture the various decision-making programs that people use (and the factors that affect their relative power) but are sufficiently parsimonious to permit predictions about how people will make decisions in particular contexts. In particular, the existing Experimental Law and Economics literature appears to support the claim that people do not always exhibit rational choice, yet it also shows that people do not always exhibit any particular deviation from rational choice. Particularly relevant are experiments showing that experimenters can alter people's choices by altering decision-making contexts in ways likely to affect the interplay between deliberative and non-conscious (intuitive) decision-making processes. This interplay can be affected by learning and training (e.g., Plott and Zeiler, 2005; Frank, et. al, 1993; Frank, et. al. 1996), the emotional or moral context of the decision (Hoffman, et. al., 1994; Hoffman, et. al, 1996). It also can be affected by external mechanisms designed to induce superior deliberation (Babcock, et. al., 1998).

Second, this literature demonstrates the importance of attending carefully to experimental design, as well-established results have been shown to disappear when greater attention is paid to controls and internal consistency. The challenge for future research is to assist in the development of a parsimonious theory that allows us to better understand the multiple decision-making processes, including deliberative reasoning, and to predict the types of circumstances where decisions can be assumed to be deliberative (and thus perhaps rational) or not.

A. Contracting, Legal Entitlement and the Coase Theorem

The Coase Theorem is a central cornerstone of the Economic Analysis of Law (Coase (1960)). The Coase Theorem holds that legal rules governing the initial allocation of a property right will not affect the ultimate allocation of that right as long as transactions costs are low, the parties are fully informed, there are no external effects on other parties, and there are no wealth effects. Indeed, in such circumstances, parties

³¹ By contrast, the criterion of contextual awareness is really not particularly relevant to unitary-process theories but is vitally important to multi-process theories. Provided that the experiment satisfies the criterion of internal consistency and control, the results of an experiment that validates or invalidates a unitary-process theory should apply across contexts.

will bargain to reallocate the right whenever it is welfare enhancing to do so. Property rights thus will be allocated efficiently regardless of the legal rule governing their initial allocation.

Tests of the Coase Theorem

Early experiments in Law and Economics sought to test the validity of the Coase Theorem on its own terms: in a simple bargaining experiment with low transactions costs, informed parties and no wealth effects. In a pioneering experiment, Hoffman and Spitzer (1982) examined whether parties to a simple two-party bargain would select the bargain that maximized their joint welfare.

Endowment Effects

These tests of the Coase Theorem all share a central feature: the parties bargain over money. These results raise the question of whether the Coase Theorem holds when subjects instead bargain over property. The Coase Theorem assumes that the initial decision of where to allocate a good does not affect the value that parties to a potential trade attach to the good. Thus, whoever values the good the most under one initial allocation also values it the most under any alternative allocation. Bargaining naturally will lead to the situation where that party ends up with the good. The central result of the Coase Theorem -- that the initial allocation of goods will not affect outcomes when (among other things) transactions costs are low -- turns on this assumption that the amount that each person values a good is not affected by whether she is entitled to it or not. In other words, the Coase Theorem assumes that (absent wealth effects) each individual's maximum Willingness-to-Pay to purchase a good equals her minimum Willingness-to-Accept to part with it.

There are both empirical and theoretical reasons to question this assumption. Professors Aaron Tversky and Daniel Kahneman developed an alternative theory of decision-making called Prospect Theory, under which people make decisions with reference to a specific baseline reference point, and evaluate gains or losses vis-à-vis that reference point. Kahneman and Tversky also contend that people exhibit loss aversion with respect to that reference point, experiencing a substantially greater change in utility from any given reduction in wealth than they experience from an equivalent increase in wealth. Consistent with this view, scholars hypothesized that people exhibit an endowment effect, under which they value goods more when they feel entitled to them than when they do not. This endowment effect would result in people exhibiting a Willingness-to-Accept for endowed goods that exceeds their Willing-to-Pay to purchase it.

The seminal article testing for the endowment effect is Kahneman, Knetsch and Thaler (1990). The experimenters conducted an experiment designed to test willingness to trade in situations that closely approximate the assumptions of the Coase Theorem: specifically, that the good does not materially affect subjects' wealth and transactions costs are low. Experimenters randomly distributed to half of their subjects a university coffee mug that sold at the nearby university bookstore for approximately \$6 (as indicated by the price tag on the bottom of the mug). Each participant was then told that she could sell (or buy) a mug by stating her sale/purchase price. The experimenters would determine the market clearing price and all those who made offers equal to or

above that price would participate in the exchange. The experiment was designed to conform to the requirements for the Coase Theorem to hold. The object – here the mug -- had little value and thus ownership did not alter valuation by materially affecting the subject's wealth. The mug did not have any "uniqueness" value – independent of that arising from endowment since subjects could easily obtain a substitute mug at the nearby bookstore for a known price (the price was affixed to the bottom). The exchange would not affect any third parties.

The experimenters hypothesized that, given the initial random distribution of the mugs between subjects who value the mug more highly and those who do not, the experiment would confirm the Coase Theorem (and the related classic economic theory) if half the mugs traded.³² Yet in fact few mugs traded. Moreover, and more importantly, the median minimum Willingness-to-Accept by mug owners was approximately twice as high as the median Willingness-to-Pay for a mug by a non-mug owner. The experimenters interpreted this disparity between Willingness-to-Accept and Willingness-to-Pay as evidence of an endowment effect (Kahneman, Knetch and Thaler, 1990).

The endowment effect undermines a central tenet of neoclassical economics that mere ownership or possession of an asset does not affect its value. It also undermines the claim of the Coase Theorem that fully informed, unimpeded, private bargaining necessarily results in commodities flowing to the people who value them most. When value depends on initial "endowment," then bargaining will not shift property to the person who would value it most *if* he possessed it, because the prospective buyer will only offer an amount equal to his Willingness-to-Pay for the good he does *not* possess, an amount that is less than the value he would attach to it if he possessed it. Viewed in this light, an entitlement will tend to stay with its initial owner to a greater degree than may be optimal, implying that the initial allocation of property rights will affect aggregate social welfare. Indeed, the endowment effect complicates the concept of social welfare. It is difficult to determine what allocation maximizes total social welfare when the amount each person values each object depends on whether she is entitled to it or not (See generally, Arlen, Spitzer, Talley, 2001).

Researchers have found apparent experimental evidence of an endowment effect in numerous experiments, using a variety of items -- such as coffee mugs, chocolate bars and pens -- as well as subjects from different countries (Horowitz and McConnell (2002)). The perceived effect persists when subjects are repeat players in the experiment (Shogren, et. al. (1994)). Moreover, the effect is not one that subjects appear to counteract on their own. Evidence suggests that people do not anticipate experiencing an endowment effect, and thus have no reason to take steps to counteract it (Loewenstein and Adler, 1995). Moreover, learning about the effect appears to be quite localized. One study found that subjects did learn to anticipate an endowment effect in others with respect to a particular commodity after repeated exchanges, but this did not

³² The number of trades is an indirect measure of the gap. The direct measure is to compare the minimum price endowed subjects will accept to part with the mug with the maximum amount non-endowed subjects are willing to pay to obtain it. There is evidence that the latter is the better way to measure the gap and the former may not accurately determine whether a significant gap exists in the data (Franciosi et al, 1996; see also Plott & Zeiler, 2005P n. 6).

cause them to anticipate an endowment effect when dealing with a seller of a superficially different good (Van Boven, et. al. 2003).

These experiments provide important evidence about human decision-making. Yet the endowment effect literature also illustrates the importance of attending carefully to both experimental design and the possibility of multiple-process decision-making. Consequently, for reasons we shall discuss, legal policy makers should not design legal policy on the assumption that the endowment effect is present.

Legal policy making cannot rely on the existence of the endowment effect because evidence reveals that people do not exhibit an endowment effect in many circumstances. Moreover, the existing theories employed to explain the endowment effect do not predict the experimental results finding that the effect is not manifested. Indeed, the existing results provide some basis to be concerned that many positive endowment effect findings may have been the result of “subject misconceptions” (including subjects making decisions based on factors, such as reputation, that might rationally cause endowed subjects to be less willing to sell than the experimenter’s hypothesis suggested).

Specifically, experiments have shown that people endow some goods but not others. For example, people do not exhibit an endowment effect with respect to money or tokens that can be traded for goods (including mugs). Moreover, even with respect to goods that often produce an endowment effect, researchers have shown that subjects exhibit an endowment effect in some circumstances but not in others. For example, as noted above, one field experiment found that people who collect trading cards exhibit an endowment effect, but this effect dissipates once they start trading the cards regularly (List, 2003).

Beyond this, Arlen, Spitzer, Talley (2001) show that a simple change in the framing of the experiment eliminates the endowment effect even with respect to university coffee mugs. Arlen, Spitzer, Talley (2001) located the mug sale/purchase decision in an employment context in which the mug exchange was part of a wage negotiation with an imaginary employer, Acme Inc. They found that endowed subjects did not exhibit a significantly higher valuation for the mug once told to imagine themselves in this employment context. This initial evidence could be interpreted consistent with the claim that people generally exhibit loss aversion if there is a basis to predict that subjects do not feel entitled to the mug in the principal-agent context. For example, in the initial experiment, endowed subjects obtained the mug before agreeing to work for the principal but made the choice of whether to part with it afterwards. The experimenters thus considered the possibility that subjects might have no longer felt entitled to the mug once they agreed to work for a principal who they believed wanted the mug. To examine this possibility, the experimenters ran a version designed to avoid any suggestion that the company might benefit from the mug. In this version, the company presented each subject with an initial offer of employment that presented subjects with two potential wage offers that the company was apparently indifferent between: one resulted in the employee receiving a higher wage and no mug (with endowed subjects exchanging their mugs to get this offer) and the other offered a lower wage and the subject would obtain/retain the mug at the end of the experiment. Subjects in this version did not exhibit an endowment effect. This evidence, at a minimum, is inconsistent with the claim that people generally exhibit an endowment effect. It thus

also appears to be inconsistent with the claim that people generally exhibit loss aversion (at least with any straightforward explanation for what causes people to feel entitled to a good).

There are several possible explanations for these conflicting results. One is that people's valuations are affected by multiple decision-making programs, whose relative influence varies across contexts. Another potential explanation is that the variation in results stems from flaws in experiments conducted by one side of the debate or the other. It is likely that both possibilities are valid.

In recent research, Plott and Zeiler considered the second possibility, revisiting the fundamental endowment effect experiments to more fully interrogate their experimental protocols (Plott and Zeiler, 2005). Here, the object of interest was whether subjects in these experiments were confused about how the allocation mechanism was likely to work (because of inadequate explanation, lack of practice, etc). If subjects were confused in this fashion, they might adapt by adopting simple heuristics (e.g., "Sell High, Buy Low") for reasons having little to do with their sense of endowment. To address such concerns, Plott and Zeiler designed protocols that included extensive training of the subjects in the experiment through paid practice rounds, guaranteeing subjects' anonymity, and ensuring that each round of the experiment was incentive-compatible in that subjects' choices in each round directly influenced their remuneration. They found that upon introducing these protocols, subjects no longer exhibited endowment effects, even though subjects participating in an experiment without these controls did exhibit the effect. Their resulting conclusion was that the endowment effect phenomenon was the result of poorly designed protocols rather than the operation of Prospect Theory.

This line of research potentially suggests that the dichotomous results in the endowment effect literature (present in some situations while absent in others) could be an artifact of poorly designed studies from within Prospect Theory. This would suggest that people *never* exhibit an endowment effect. While this explanation is plausible, we resist this conclusion for a few reasons. First, even the earlier studies that found no endowment effect arguably suffered from similar defects in subject training and understanding. Second, one distinct change in the Plott and Zeiler studies is that they utilized much more exhaustive experimental protocols, which themselves may have muted the endowment effect by changing the context of the experiment. In particular, the Plott and Zeiler protocol required subjects to repeat the experiment. This is relevant to the issue of whether people exhibit an endowment effect in some circumstances, but not in others, because it has been conjectured that learning can affect valuation. It is argued that subjects do not initially know their own preferences and use the decision-making process itself to discover their preferences. Along one stage in this discovery process subjects may exhibit loss aversion; as subjects gain greater experience with trading the good (and experiencing the loss)³³ they may discover that the loss is less painful than anticipated, and eventually may not exhibit either loss aversion or an

³³ In the Plott and Zeiler experiment, subjects experienced 2 paid practice rounds and 14-15 actual rounds.

associated Willingness-to-Accept/Willingness-to-Pay gap (Plott, 1996).³⁴ This latter possibility suggests that additional experiments are needed to test the ways in which the process of decision-making may affect preferences, and thus outcomes.

Whatever one's ultimate conclusion about whether subjects ever exhibit an endowment effect, it is important to recognize the broader implications of the Plott and Zeiler experiments (and those in a similar vein). These experiments show that the experimental protocol – including features such as whether the subjects are paid for each round, are given paid practice rounds, and are guaranteed anonymity – can significantly affect experimental outcomes. Researchers in Experimental Law and Economics should consider utilizing such experimental protocols, or (more creatively) consider how such protocols may factor into a theory of context that would itself bear on a multi-process account of human cognition (for example, a theory of behavior might turn on whether real world experience may affect one's comfort with a decision, in turn affecting her choices).³⁵

B. Self-Serving Bias and Settlement

Other biases can cause negotiations to deviate from the predictions of Rational Choice Theory. Often people negotiate in situations where future outcome are uncertain. The question is whether this uncertainty is likely to cause parties to fail to reach welfare-enhancing agreements.

Classic economic theory suggests that uncertainty alone should not undermine negotiations, provided the two parties have the same information. Negotiations should produce welfare-enhancing deals so long as both parties understand that there are gains from the deal, notwithstanding the uncertainty, when in fact those gains exist. Classic economic theory assumes that parties evaluate future expected outcomes based on an objective assessment of the probability and magnitude of possible future events, given the information available to them. This implies that, even in the presence of uncertainty, people should be willing to negotiate a deal when it is welfare enhancing that they do so, so long as they each possess the same information about the probability and magnitude of the possible future outcomes and the information provides the correct expected value of future events. In this situation, parties will conclude that a deal is warranted whenever an objective observer would conclude that a given deal between them would be welfare enhancing (given the information available to the parties). Uncertainty itself thus does not cause negotiation failure. Negotiations fail (when they should not) only when uncertainty is coupled with asymmetric information, which causes the parties' expectations about future outcomes to diverge from each other.

By contrast, psychologists have long argued that people do not evaluate uncertain information objectively. Rather, they tend to analyze the information through a mental filter that results in them evaluating the information in a way that is favorable

³⁴ This is consistent with List (2003) who conducted a field experiment of trading card exchanges and found that subjects endow when they do not exchange cards regularly, but cease to exhibit an endowment effect once they begin to trade on a regular basis.

³⁵ This implies that the failure to pay subjects can affect experimental results even when the experiment involves a comparison of how different types of subjects behave, if these subjects when deciding in the "real world" are likely to respond differently to the financial consequences of real-world choices.

to them. Thus, people over-estimate their own skill and luck, and produce more positive estimates of future outcomes than is supportable by an objective analysis of the available information. The existence of self-serving biases has important implications for the analysis of negotiation and bargaining because it suggests that parties negotiating under uncertainty may fail to enter into welfare-enhancing negotiations, even when they possess the same information, because they evaluate the information in a self-serving way, which may cause them to over-estimate their expected outcomes should they fail to close the deal.

A team of researchers studied the effect of self-serving biases on negotiation through a long-running series of experiments that examined whether self-serving biases impede parties from settling legal disputes. The classic economic account of negotiation under uncertainty implies that parties to a legal dispute should be willing to settle it, provided that they possess the same information about the merits of each side's position. Parties with common expectations about expected outcomes at trial can only gain from settlement, as this reduces their joint litigation costs. Settlement should break down only if parties have private information about their positions that causes their subjective beliefs about expected outcomes to diverge in ways that preclude settlement (Priest and Klein). This suggests that trials should be very rare events, given that trials impose substantial costs on parties that can be avoided by settling and that discovery and other mechanisms substantially reduce the parties' information disparity. Yet, empirical evidence suggests that, although the vast majority of cases do indeed settle, a substantial number (about 5%) nevertheless go to trial. This raises the question of whether the cases that do not settle are those in which asymmetric information problems are particularly great (or litigation costs are particularly low), or do these cases provide evidence that other factors, such as self-serving bias, are affecting settlement.³⁶

To examine this issue, a team of researchers developed an experimental test to determine whether self-serving biases affect settlement negotiations (Babcock, et. al. 1993; Babcock, et. al. 1995b, Babcock et. al., 1997; and Babcock and Loewenstein, 1997). In the initial experiments, subjects were given the facts of a Texas tort case based on a trial that occurred in Texas in which a motorcyclist sued the driver of the automobile that collided with him for \$100,000. Subjects were randomly assigned to be either a plaintiff or defendant, given the facts of the case, and then told to attempt to negotiate a settlement. They were told that if they failed to negotiate, an actual judge in Texas would read these facts and decide who should prevail and what damages should be awarded (if any). The materials given to the subjects was very detailed (27 pages) and included witness testimony, police reports, maps, and the testimony of the parties taken from the original litigation. They were then asked to assess both the fair settlement value and what the judge would select (and were offered a bonus for accurately predicting the judge's award (within \$5,000 of his award)). They were then told they could negotiate over how much if anything, the defendant would in fact pay the plaintiff. This negotiation would result in an actual payment by the defendant to the plaintiff (with each \$10,000 of negotiated payment translating into an actual payment

³⁶ Evidence of other factors affecting trial settlement would be relevant not only for our understanding of the settlement process itself, but also could have implications for analysis of other areas involving negotiations (including the Coase Theorem itself).

from the defendant subject to the plaintiff subject of \$1).³⁷ If the parties did not reach a settlement in 30 minutes, the judge's decision was imposed upon them. (The judge's actual decision was that the defendant owed the plaintiff \$30,560, resulting in a \$3.06 payment from defendant subjects to plaintiffs) (Babcock, et. al., 1995). The experimenters examined the effect of self-serving bias on negotiation by manipulating whether subjects learned their roles before or after reading the facts of the case and assessing its merits. One group of subjects learned their roles before reading the facts. A second group, read the materials, predicted the judge's award and assessed fairness before learning what role they would play.

Consistent with the hypothesis that self-serving biases affect how people interpret information, the researchers found that subjects who learned their roles before reading the information were much less likely to settle than those who learned them afterwards. The lower settlement rate likely results from the fact that the difference between the defendant and the plaintiff's assessment of a fair settlement (and the difference in their assessments of what the judge would award) was substantially larger for subjects who learned their roles prior to analyzing the case than it was for subjects who learned their roles after analyzing the case. It seems that parties assessed information through a self-serving filter even when this, in the end, operated to their detriment by reducing settlement (Babcock, et. al., 1995).

This research suggests that self-serving biases affect subject's interpretation of information when the information is subject to multiple interpretations. This suggests that giving subjects more information may not promote settlement if the information is susceptible to self-serving interpretations that cause the parties' expectations of the fair outcome to diverge even more. This is contrary to the findings of classical economic analysis of settlement -- that parties are more likely to settle when they are given additional information because the additional information can only reduce the divergence in their expectations about the merits of the case.³⁸

To analyze this issue, Loewenstein and Moore (2004) examined the impact on bargaining of providing both parties with common information. Subjects were told that they were bargaining over the sale of a widget on behalf of either a buyer or seller and were informed about the value of the widget to the buyer and the cost to the seller. The zone between these constituted the direct gains from trade. Each side would retain the difference between the price he negotiated and the value/cost of the widget to his principal (buyer/seller). In addition, however, each subject was told he would receive a bonus for negotiating the deal within a certain time (and that the other side would too). In a rational actor model, this bonus also would be interpreted by both sides as a gain from trade, potentially subject to division between the parties. Yet the parties might well view the bonus as their property, independent of the bargain. Loewenstein and Moore hypothesized that subjects were likely to adopt a self-serving view of whether the bonus should be treated as a gain from trade to be divided in the bargaining -- viewing it that way if the other party's bonus was larger, but not otherwise. This

³⁷ At the beginning of the experiment, each subject was paid \$4 to participate and then defendant subjects were given an additional \$10.

³⁸ Self-serving biases also have been found outside the lab. Babcock, Wang and Loewenstein, (1996) conducted a field experiment of labor-management negotiations of public school teachers and found evidence that bargaining impasse could be attributable to self-serving biases.

difference in interpretation creates a potential wedge between the parties' assessment of the fair price that could undermine negotiation. In this situation, informing parties about each other's bonuses could be expected to improve or delay negotiation depending on whether subjects learned that their bonuses were the same (thereby diminishing the zone of disagreement) or different.

To test this, Loewenstein and Moore divided their subjects into four treatment groups. In two groups the subjects had the same bonuses: one of these groups knew this and the other thought there was a 50% chance the other side was getting a six-fold larger bonus. In the other group, the subjects were receiving different bonuses: one group believed that this was only 50% likely and the other group knew the bonuses with certainty. They then examined how long it took subjects to negotiate a deal. They found that subjects in the symmetric bonus treatments reached settlement faster when both were informed that the bonuses were the same. Information in this situation reduced the expectations' wedge created by self-serving views of the fair division of the bonuses. By contrast, the provision of information to parties in the asymmetric bonus situation significantly delayed settlement. Indeed, 18% of these subject pairs failed to reach any settlement at all (in contrast to only one pair in the other treatments). This finding is consistent with the hypothesis that the information provided was interpreted in a self-serving way by the subjects, and thus only served to exacerbate the divergence between their views of what constituted a fair exchange. This experiment not only supports the prior evidence of the existence of self-serving biases, but also reveals that these biases cannot be eliminated through the mere provision of information alone, since that information can be subject to self-serving interpretations.

Of course, this analysis raises the question of whether these biases do indeed affect actual settlement negotiations (and commercial bargains) given that such bargains involve learned intermediaries. Preliminary evidence suggests that we cannot be confident that the mere presence of learned intermediaries will eliminate such biases.³⁹ There is evidence that attorneys and experienced negotiators exhibit the same variety of self-serving biases, albeit sometimes in a smaller magnitude (Babcock and Loewenstein 1997).⁴⁰ Nevertheless, certain interventions (for example by learned intermediaries) can have the effect of debiasing subjects. Babcock et al. (1997) found that they could debias subjects by instructing them to consider either the weaknesses of their own case or the real possibility that the judge may rule for the other side.⁴¹ Debiasing dramatically increased both the probability and speed of settlement. Accordingly, this evidence suggests that lawyers may potentially form a debiasing function should they

³⁹ Experts are as likely to exhibit certain biases as non-experts because decision-making experience alone does not necessarily convey decision-making skills. For experience to help debias it must provide direct feedback about the quality of the decision. This often does not occur in practice. For example, a lawyer who recommends settlement never learns what would have happened had he gone to trial. Richard H. Thaler (1987) ("Accurate learning takes place only when the individual receives timely and organized feedback").

⁴⁰ See also Babcock, et al. (1995a: 296-97) (finding that lawyer subjects in a negotiation experiment were as affected by framing biases as were student subjects).

⁴¹ Interestingly, urging subjects to consider the strengths of the other side's case did not work (and sometimes backfired). (Babcock et al., 1997). Similarly, one cannot debias subjects simply by informing them of the existence of the self-serving bias. Subjects informed about the bias expect their counterparts to exhibit the bias but seem to believe that they themselves are immune (Babcock and Loewenstein, 1997).

require their clients to discuss the strengths and weaknesses of their case as part of the process of discussing a settlement offer.

In addition, these results raise issues about whether tort reform may have unintended effects on settlement by affecting the degree of divergence between parties' potentially self-serving assessments of a case. Babcock and Pogarsky (1999, 2001) find that a strict damage cap that significantly constrains potential trial outcomes promotes settlement by reducing the gap between plaintiffs' and defendants' expectations about likely outcomes at trial. By contrast, however, an exceedingly generous cap (that exceeds the expected outcome at trial) has the reverse effect. This cap apparently operates to anchor plaintiff-side optimism, thereby increasing the gap between the parties' expectations and trial (Babcock and Pogarsky, 2001).

Fairness and Crowding Out:

Another area that experimentalists interested in Law and Economics have begun to investigate is often lumped under the general context of "fairness." Here, experimental subjects appear to act as though they account not only for their own welfare, but also for that of their experimental counterparts and even adversaries. Technically, and contrary to many popular accounts, preferences for fairness alone are not anathema to Rational Choice Theory; indeed, conventional expected utility hypotheses place no requirements on whether one's preferences are entirely egoistic or more interpersonal in nature. That said, it is at least conventional within this literature to presume that actors care only about their own consumption, wealth, and income. In this sense, episodic demonstrations of other-regarding behavior have been interpreted as inconsistent with expected utility hypotheses, at least as popularly conceived.

Much of the foundational work in this field emanates from experiments involving zero-sum games, where subjects must divide a fixed monetary sum between them. In the (so-called) "Dictator" game, two subjects divide a fixed sum using a very simple protocol: one subject (labeled as the "Dictator") simply announces how much she and her counterpart will each receive, with the only constraint that the announced amounts must be non-negative and sum to the total amount of money available. The "Ultimatum" game, in contrast, begins with the same protocol, but adds to it the ability of the offeree to either accept or reject the proposed division, with the consequence of rejection being that both subjects receive nothing.

In the traditional rational choice paradigm, these two games might be seen as "strategic equivalents" of one another. Although the protocols they prescribe are different, both games would be expected to yield nearly identical equilibria when played by rational actors possessing complete information: The offeror proposes (and, in the case of the Ultimatum game, the offeree accepts) a limiting one-sided-division that pays the former virtually the entire pie (with an arbitrarily small sum – such as a penny – awarded to the offeree to break indifference). Interestingly, however, the experimental evidence is at odds with this prediction. Not only do the Dictator and Ultimatum games yield results that diverge from the rational choice baseline case, but they also yield results that diverge from one another.

In Dictator game experiments, various experimentalists have observed a substantial variance on divisions, ranging anywhere from the predicted, one-sided

division to a half-share for each player. The range of these behaviors suggests that this is a domain in which context and prediction are maximally reactive with one another. In particular, they raise a question about whether observed behavior is a function of native altruism or rather a more complex function of a sense of connection or “distance” from the community in which one interacts. Hoffman et al (1994; 1996) explores some of the implications of this interaction. In the Hoffman et al experiments, subjects played out the Dictator game through a carefully manipulated environment of “thick” anonymity, performing their tasks in separate rooms and/or time periods, with the experimenters taking pains to ensure anonymity not only of the subject dyad from one another, but also from other subjects, and experimenters themselves. When social distance is maximized in such a fashion, the authors find that observed behavior converges most closely with corresponding rational choice predictions (but still around 40 percent deviate from those predictions).

This result, of course, begs a debate about whether it “confirms” Rational Choice Theory or instead “confirms” native altruism as a theoretical precept of human behavior. We suggest that it probably can be read to support both claims, but by far the more interesting question concerns the interaction (on top of any ambient altruism) between context and *altruistic behavior*. Understanding and calibrating these interaction effects is obviously important for the design of legal rules.

For example, Bohnet & Frye (1999) focus on social distance as their principal experimental control variable. In their experiments, dictators and counterparts sat in the same room, but the experimental protocols ensured their ability to protect subject anonymity when desired. Perhaps because of some limited social connectedness between subjects, they observe more altruistic behavior even with fully protected anonymity. They then relax anonymity between dictators and counterparts, identifying them and (in some treatments) giving the dictator some more personal information about the counterpart. As the amount of anonymity decreases, dictators increasingly adopt more altruistic behavior.

When one moves from the Dictator game to the Ultimatum game, at least two important control factors change. First, the counterpart has now become an offeree, and has real (albeit limited) bargaining power simply to reject the offeror’s proposal. This means that the offeror must now consider not only what she herself thinks is an appropriate split of the pie, but also what the offeree will think. Second, since the extensive form of the Ultimatum game opens up the possibility of two-way communication, the offeree may embed within her response expressive manifestations (about disappointment, anger, sadness, etc.) about the offeror’s strategy, a form of communication not possible in the Dictator game. This communication effect may even be more pronounced in a repeat-play context.

Such communication appears to be present in many such experiments. In most Ultimatum game experiments, offerees tend to reject any offer giving them less than approximately 20 percent of the pie, regardless of whether the game is repeated only once, the stakes are raised somewhat, or the experiment is carried to subjects of other / different cultures. Anticipating this behavior, offerors accommodate by making higher offers than they do in the Dictator game.

As with the Dictator game, context and outcomes appear to interact heavily. For example, in Hoffman et al (1994), the experimenters vary the degree to which the role

of “offeror” is earned versus randomly assigned; and in addition they vary the extent to which the Ultimatum game is situated within a market exchange environment. Both forms of treatment appear (but on the margin and in concert) to induce proposers to reduce the distributions of their offers, while at the same time not substantially increase the likelihood that offerees reject the proposed split.

The findings of the Ultimatum and Dictator game experiments are quite foundational, and do not, in and of themselves, carry over to legal applications. However, they raise important implications for the design of legal institutions. Assume, for the moment, that policy makers wish to auger at least certain types of (efficient) altruistic behavior through their choice of legal institutions. If expressed altruism is endogenous to its contextual environment, then lawmakers would want to know which institutions tend to be substitutes for such altruistic behavior, and which appear to be complements.

It is in this spirit that we also reproduce two additional studies, one by Bohnet, Frey & Huck (2001) and the other by Fehr, Klein & Schmidt (2007). Bohnet et al. consider the (so called) “Trust” game, in which an initial mover must decide whether to entrust a later decision-maker with a decision that has both productive and distributional consequences. If the first mover entrusts, the second mover can choose whether to remain faithful to that trust (in which case each split a known surplus), or breach it. Acting in a trustworthy fashion permits both parties to realize a monetary gain. Breaching, on the other hand, if not subject to enforcement, causes the breaching party to realize a larger gain and the first mover to realize an even larger loss (so that breaching is inefficient). The experimenters then allow legal enforcement to enter the picture, permitting (with some probability) recovery in a lawsuit against breaching parties. The authors find evidence of trusting and trustworthiness when legal institutions are extremely weak (e.g., there is a very low probability of a successful lawsuit). On the other end of the spectrum, strong enforcement ensures trusting behavior, but through institutional means. Legal institutions of intermediate strength tend to perform the worst; for they both crowd out norms of trusting and altruism, and they are insufficiently strong to make trusting and trustworthiness a part of a rational expectations equilibrium. In sum, then, this set of experiments tends to suggest that law and trusting are largely substitutes within the relevant domain.

One interesting extension of this finding comes from Fehr et al. (2007), who essentially endogenize the legal crowding-out phenomenon, considering the implications that preferences for fairness have on *ex ante* contracting. In their experiments, principals and agents must decide how to design a contract in an environment where the agent’s subsequent “effort” can affect the principal’s welfare – a classic moral hazard situation. For one set of subjects, principals were given a choice between (1) making a costly investment in an imperfect third-party state verification technology – which would enable them to write explicit incentive contracts with agents; and (2) offering to make an unenforceable “bonus” payment without any enforcement by a court if the principal deems the agent’s effort to be satisfactory. Because explicit incentive contracts are costly to invoke, the unenforceable bonus contracts might dominate if they inculcated sufficient trust among the players. However, the experimenters also set up the game specifically as a one-shot game with anonymous identities of players, and thus other-regarding behavior would not be part of an

equilibrium among self-interested parties. Within this design, the bonus contract paid off, on average, more than the express incentive contract, suggesting that preferences for fairness were sufficiently prevalent to alter the principal's favored contract to the one without courts. At the same time, however, the authors found that the power of trusting behavior tended to turn on the way it was manifested in the contractual instrument. In a variation on their baseline test, the experimenters allowed subjects to enter either the contract described in (1) above, or (2) a "trust" contract involving not a bonus payment *ex post*, but rather a single high wage payment made *ex ante*, essentially pre-paying the bonus under the assumption that the agent would work hard. Here they found a substantial increase in the prevalence of shirking among agents, and a strong preference among principals for the incentive contract. The authors posit that one of the reasons that the "bonus" contract tends to fare much better than the "trust" contract is that the former makes trusting a far less risky action.

Finally, Cain et al (2005) consider a different way in which law may "crowd out" other-regarding behavior: by removing the specter of guilt from behaving in non-trustworthy ways. Here, consider the effect of disclosure rules that require an interested party (such as a stock broker) to disclose her conflict of interest before offering advice to a client. Specifically, they consider a setting where some subjects ("estimators") were asked to estimate the value of money inside a transparent jar. Estimators received compensation that was directly proportional to the accuracy of their estimates. Prior to issuing their estimates, however estimators received advice from another set of subjects ("advisors"). While some advisors received the same form of compensation as estimators, other advisors were compensated not on accuracy but rather on how high the estimator's ultimate conjecture was. Thus, these conflicted advisors had an incentive to push estimator's bids too high. Finally, the experimenters varied the conditions, at times requiring interested advisors to disclose their conflicts of interest, while other times allowing them to remain silent about it. They found, not surprisingly, that all conflicted advisors tended to render advice above what disinterested advisors gave. But more interesting, they also found that the conflicted advisors who disclosed that fact tended to give the *highest* conjectured value, effectively behaving as if the disclosure had given them license to exaggerate. The conflicted advisors that did not have to disclose may, in contrast, have been constrained by a preference for fairness that dampened their strategic incentives. Moreover, because disclosing conflicted advisors gave more distorted advice, evaluators tended to fare most poorly in the conditions where a conflict of interest was required to be disclosed.

CONCLUSION

The publication of this book serves a number of important purposes. The first is to recognize the significant and powerful contributions that researchers in Experimental Law and Economics have made during the last two decades to our knowledge of human behavior, and how it interacts with legal and regulatory environments. These contributions transcend the legal setting, and provide contributions to the analysis of human decision-making which are relevant to economists, psychologists, and policy makers interested in human behavior. The articles reproduced below are but a sampling of a number of excellent studies in a field that claims contributions from economists,

psychologists, traditional legal scholars, and permutations thereof. It is a field that is already so large that it is impossible to include all worthy contributions in a single volume.

At the same time, the analysis of this introduction suggests that Experimental Law and Economics has much further to go. Neither predominant account of human cognition and behavior within legal environments – Expected Utility or Prospect Theory – has proven to be successful as a unitary theory across all contextual arenas. Moreover, in our view, the dichotomous debate among advocates about which unitary account is valid is both misplaced and counterproductive. It frequently induces policy makers and scholars to identify with one “camp” or the other, and in so doing invites policy reforms that are likely to be inappropriately extrapolated from experimental (or theoretical) settings that are excessively artificial. Research occurring outside of Experimental Law and Economics suggests that the resolution of the debate between these competing theories likely lies with the realization that human decisions cannot be explained by any unitary-process theory. To predict choices, we need to better understand both the multiple programs which influence decision-making and the external factors that affect the relative influences of these programs.

Viewed in this light, this book represents two types of invitation for those who either wish to join the fray or already find themselves in it. The first is for theorists to consider developing alternative accounts that are better able to explain experimental observations in a meaningful, predictive and testable way than are existing theories. The second is for experimentalists to take great care to articulate both the purpose of their experimentation and the domain of the theory(ies) they wish to test/develop. Greater attention to relating how behavior depends on context may not only create benefits for the development of new theories, but it would also help to spawn policy choices that are based on responsible extrapolations from that research, even before those new theories fully emerge.

Bibliography

Jennifer Arlen (1998), 'The Future of Behavioral Economic Analysis of Law', *Vanderbilt Law Review*, 51, 1765-1788.

Jennifer Arlen, Matthew Spitzer and Eric Talley (2002), 'Endowment Effects Within Corporate Agency Relationships', *Journal of Legal Studies*, 31, 1-37.

Nava Ashraf, Colin Camerer and George Loewenstein (2005), 'Adam Smith, Behavioral Economist', *Journal of Economic Perspectives*, 19, 131-145.

Linda Babcock, Henry Farber, Cynthia Fobian and Eldar Shafir (1995a), 'Forming Beliefs About Adjudicated Outcomes: Perceptions of Risk and Reservation Values', *International Review of Law and Economics*, 15, 289-303.

Linda Babcock, George Loewenstein, Samuel Issacharoff and Colin Camerer (1995b), 'Biased Judgments of Fairness in Bargaining', *American Economic Review*, 85, 1337-1343.

Linda Babcock, George Loewenstein and Xianghong Wang (1995c), 'The Relationship Between Uncertainty, the Contract Zone, and Efficiency in a Bargaining Experiment', *Journal of Economic Behavior and Organization*, 27, 475-485.

Linda Babcock, Xianghong Wang and George Loewenstein (1996), 'Choosing the Wrong Pond: Social Comparisons in Negotiations that Reflect a Self-Serving Bias', *Quarterly Journal of Economics*, 111, 1-19.

Linda Babcock and George Loewenstein (1997), 'Explaining Bargaining Impasse: The Role of Self-Serving Biases', *Journal of Economic Perspectives*, 11, 109-126.

Linda Babcock, George Loewenstein, and Samuel Issacharoff (1997), 'Creating Convergence: Debiasing Biased Litigants', *Law and Social Inquiry* 22: 913-925.

Linda Babcock, and Greg Pogarsky (1999), 'Damage Caps and Settlement: A Behavioral Approach', *Journal of Legal Studies*, 28, 341-370.

Linda Babcock, and Greg Pogarsky (2001), 'Damage Caps, Motivated Anchoring and Bargaining Impasse', *Journal of Legal Studies*, 30, 143-159.

Linda Babcock and Claudia M. Landeo (2004), 'Settlement Escrows: An Experimental Study of a Bilateral Bargaining Game', *Journal of Economic Behavior and Organization*, 53, 401-417.

Thomas Baker, Alon Harel, and Tamar Kugler (1994), 'The Virtues of Uncertainty in Law: An Experimental Approach', *Iowa Law Review*, 89, 443 - .

Marianne Bertrand and Sendhil Mullainathan (2004), 'Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment in Labor Market Discrimination', *American Economic Review*, 94, 991-1013.

Iris Bohnet and Bruno Frey (1999), 'The Sound of Silence in Prisoner's Dilemma and Dictator Games', *Journal of Economic Behavior and Organization*, 38, 43-57.

Iris Bohnet, Bruno S. Frey, Steffen Huck (2001), 'More Order with Less Law: On Contract Enforcement, Trust, and Crowding', *American Political Science Review*, 95, 131-144.

Daylian M. Cain, George Loewenstein, and Don A. Moore (2005), 'The Dirt on Coming Clean: Perverse Effects of Disclosing Conflicts of Interest', *Journal of Legal Studies*, 34, 1-24.

Colin F. Camerer, George Loewenstein, and Martin Weber (1989), 'The Curse of Knowledge in Economic Settings: An Experimental Analysis', *Journal of Political Economy*, 97, 1232-1254.

Colin F. Camerer & Richard H. Thaler (1995), 'Anomalies: Ultimatums, Dictators and Manners', *Journal of Economic Perspectives*, 9, 209- .

Colin F. Camerer and Robin M. Hogarth (1999), 'The Effect of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework', *Journal of Risk and Uncertainty*, 18, 7-42.

Colin F. Camerer, Joseph Henrich, Robert Boyd, Samuel Bowles, Ernst Fehr, Herbert Gintis, and Richard McElreath (2001), 'Cooperation, Reciprocity and Punishment in Fifteen Small-Scale Societies', *American Economic Review*, 91, 73-78.

Colin F. Camerer (2003), *Behavioral Game Theory – Experiments in Strategic Interaction* (Princeton, New Jersey: Princeton University Press).

(2005), 'Economic Man' in Cross-Cultural Perspective: Ethnography and Experiments from 15 Small-Scale Societies', *Behavioral and Brain Sciences*, 28, 795-855.

Colin F. Camerer, George Loewenstein and Drazen Prelec (2005), "'Neuroeconomics' How Neuroscience Can Inform Economics", *Journal of Economic Perspectives*, 43, 9-64.

Colin F. Camerer and Ernst Fehr (2006), 'When Does "Economic Man" Dominate Social Behavior?', *Science*, 311, 47-52.

Ronald H. Coase (1960), 'The Problem of Social Costs', *Journal of Law and Economics*, 3, 1-44.

Leda Cosmides and John Tooby (1996), 'Are Humans Good Intuitive Statisticians After All? Rethinking Some Conclusions from the Literature on Judgment Under Uncertainty', *Cognition*, 58, 1-73.

Leda Cosmides and John Tooby (2006), 'Evolutionary Psychology, Moral Heuristics and the Law', Chapter 9 in G. Gigerenzer & Christoph Engel (Eds.), *Heuristics and the Law* (Dahlem Workshop Report 94). Cambridge, MA: MIT Press.

Rachel Croson (2000), 'Why and How to Experiment: Methodologies from Experimental Economics', *University of Illinois Law Review*, 921-945.

Rachel Croson and Jason S. Johnston (2000), 'Experimental Results on Bargaining Under Alternative Property Rights Regimes', *Journal of Law, Economics and Organization*, 16, 50-73.

Antonio Damasio (2005), 'Brain Trust', *Nature*, 435, 571-572.

Elsa Ermer, Scott A. Guerin, Leda Cosimedes, John Tooby, and Michael B. Miller (2006), 'Theory of Mind Broad and Narrow: Reasoning About Social Exchange Engages ToM Areas, Precautionary Reasoning Does Not', *Social Neuroscience*, 1, 196-219.

Ernst Fehr, Georg Kirchsteiger and Arno Riedl (1993), 'Does Fairness Prevent Market Clearing: An Experimental Investigation', *Quarterly Journal of Economics*, 108, 437- .

Ernst Fehr and Simon Gächter (2002), 'Altruistic Punishment in Humans', *Nature*, 415, 137-140.

Ernst Fehr and Urs Fischbacher (2003), 'The Nature of Human Altruism,' *Nature*, 425, 785-791.

Ernst Fehr and Bettina Rockenbach (2003), 'Detrimental Effect of Sanctions on Human Altruism', *Nature*, 422, 137-140.

Ernst Fehr and Urs Fischbacher (2004), 'Social Norms and Human Cooperation', *TRENDS in Cognitive Science*, 8, 185-190.

Ernst Fehr and John A. List (2004), 'The Hidden Costs and Returns of Incentives – Trust and Trustworthiness Among CEOs', *Journal of the European Economic Association*,

Ernest Fehr, Alexander Klein, and Klaus M. Schmidt (2007a), 'Fairness and Contract Design', *Econometrica*, 75, 121-154.

Ernst Fehr, Urs Fischbacher, and Michael Kosfeld (forthcoming 2007b), 'Neuroeconomic Foundations of Trust and Social Preferences', *American Economic Review*.

Baruch Fischhoff (1975), 'Hindsight \neq Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty', *Journal of Experimental Psychology*, 1, 288-299.

Baruch Fischhoff (1982), 'Debiasing', 422-444, in *Judgement Under Uncertainty: Heuristics and Biases* (Daniel Kahneman, Paul Slovic and Amos Tversky, ed.) (New York: Cambridge University Press).

Robert Franciosi, Praven Kujal, Roland Michelitsch, Vernon L. Smith, and Deng Gang (1996), 'Experimental Tests of the Endowment Effect', *Journal of Economic Behavior and Organization*, 30, 213-226.

Robert H. Frank, Thomas Gilovich, and Dennis Regan (1993), 'Does Studying Economics Inhibit Cooperation?', *Journal of Economic Perspectives*, 7, 159-171.

Gerd Gigerenzer (1991), 'How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases"', *European Review of Social Psychology*, 2, 83-115.

William M. Goldstein and Elke U. Weber (1995), 'Content and Discontent: Indications and Implications of Domain Specificity in Preferential Decision Making', in Busemeyer, J. R., R. Hastie, and D. L. Medlin, eds., *The Psychology of Learning and Motivation*. Vol. 32. (New York, Academic Press), 83-136.

Paul Glimcher and Aldo Rustichini (2004), 'Neuroeconomics: The Consilience of Brain and Decision', *Science*, 306, 447-452.

Uri Gneezy and Aldo Rustichini (2000), 'A Fine Is a Price', *The Journal of Legal Studies*, 29, 1-17.

Serena Guarnaschelli, Richard McKelvey and Thomas Palfrey (2000), 'An Experimental Study of Jury Decision Rules', *American Political Science Review*, 94, 407-23.

Werner Güth, Rolf Schmittberger, and Bernd Schwarze (1982), 'An Experimental Analysis of Ultimatum Bargaining', *Journal of Economic Behavior and Organization*, 3, 367-388.

Chris Guthrie, Jeffrey Rachlinski and Andrew Wistrich (2001), 'Inside the Judicial Mind', *Cornell Law Review*, 86, 777-830.

Reid Hastie, David A. Schkade and John W. Payne (1999), 'Juror Judgments in Civil Cases: Effects of Plaintiff's Requests and Plaintiff's Identity on Punitive Damage Awards', *Law and Human Behavior*, 23, 445-470.

Elizabeth Hoffman and Matthew Spitzer (1982), 'The Coase Theorem: Some Experimental Tests', *Journal of Law and Economics* 25: 73-93.

Elizabeth Hoffman and Matthew Spitzer (1985), 'Entitlements, Rights and Fairness: An Experimental Examination of Subjects' Conceptions of Distributive Justice', *Journal of Legal Studies*, 14, 259-297.

Elizabeth Hoffman and Matthew Spitzer (1986), 'Experimental Tests of the Coase Theorem with Large Bargaining Groups', *Journal of Legal Studies*, 15, 149-171.

Elizabeth Hoffman, Kevin McCabe, Keith Shachat and Vernon Smith (1994), 'Preferences, Property Rights, and Anonymity in Bargaining Games', *Games and Economic Behavior*, 7, 346-380.

Elizabeth Hoffman, Kevin McCabe, and V. Smith (1996), 'Social Distance in Dictator Game', *American Economic Review*, 86, 653-660.

Jonathan K. Horowitz and Kenneth E. McConnell (2002), 'A Review of WTA/WTP Studies', *Journal of Environmental Economics and Management*, 44, 426-444.

Christine Jolls, Cass R. Sunstein and Richard Thaler (1998), 'A Behavioral Approach to Law and Economics', *Stanford Law Review*, 50, 1471-1550.

Christine Jolls and Cass R. Sunstein (2006), 'Debiasing Through Law', *Journal of Legal Studies*, 35, 199-241.

Daniel Kahneman and Amos Tversky (1979), 'Prospect Theory: An Analysis of Decision Under Risk', *Econometrica*, 47, 263-291.

Daniel Kahneman, Jack L. Knetsch and Richard Thaler (1986), 'Fairness as a Constraint on Profit Seeking: Entitlements in the Market', *American Economic Review*, 76, 728-741.

Daniel Kahneman, Jack L. Knetsch and Richard Thaler (1990), 'Experimental Tests of the Endowment Effect and the Coase Theorem', *Journal of Political Economics*, 98, 1325-1348.

Daniel Kahneman, Jack L. Knetsch and Richard Thaler (1991), 'The Endowment Effects, Loss Aversion and Status Quo Bias', *Journal of Economic Perspectives*, 5, 193-206.

Daniel Kahneman, Edward McCaffery and Matthew Spitzer (1995), 'Framing the Jury: Cognitive Perspectives on Pain and Suffering Damages', *Virginia Law Review*, 81, 1341-1420.

Daniel Kahneman, David Schkade and Cass R. Sunstein (1998), 'Shared Outrage and Erratic Awards: The Psychology of Punitive Damages', *Journal of Risk and Uncertainty*, 16, 49-86.

Tatsuya Kameda, Masanori Takezawa, R. Scott Tindale and Christine M. Smith (2002), 'Social Sharing and Risk Reduction: Exploring a Computational Algorithm for the Psychology of Windfall Gains', *Evolutionary Human Behavior*, 23, 11-33.

Kim A. Kamin and Jeffrey J. Rachlinski (1995), 'Ex post \neq Ex Ante: Determining Liability in Hindsight', *Journal of Law and Human Behavior* 19: 89-104.

Mark Kelman, David Fallas, and Hilary Folger (1998), 'Decomposing Hindsight Bias', *Journal of Risk and Uncertainty*, 16, 251-269.

Daria Knoch, Alvaro Pascual-Leone, Kaspar Meyer, Valerie Treyer and Ernst Fehr (2006), 'Diminishing Reciprocal Fairness by Disrupting the Right Prefrontal Cortex', *Science*, 314, 829-832.

Michael Koenigs, Liane Young, Ralph Adolphs, Daniel Tranel, Fiery Cushman, Marc Hauser and Antonia Damasio (2007), 'Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgments', *Nature*, 446, 908-911.

Lewis Kornhauser and Andrew Schotter (1990), 'An Experimental Study of Single-Actor Accidents', *Journal of Legal Studies*, 19, 203-233.

Russell Korobkin and Christopher Guthrie (1994), 'Psychological Barriers to Litigation Settlement: An Experimental Approach', *Michigan Law Review*, 93, 107-92.

Russell Korobkin and Christopher Guthrie (1997), 'Psychology, Economics, and Settlement: A New Look at the Role of the Lawyer', *Texas Law Review*, 76, 77-141.

Russell Korobkin (1998), 'The Status Quo Bias and Contract Default Rules', *Cornell Law Review*, 83, 608-687.

Russell Korobkin and Thomas Ulen (2000), 'Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics', *California Law Review*, 88, 1051-1144.

Russell Korobkin (2003), 'The Endowment Effect and Legal Analysis', *Northwestern University Law Review*, 97, 1227-1293.

Michael Kosfeld, Markus Heinrichs, Paul Zak, Urs Fischbacher and Ernst Fehr (2005), 'Oxytocin Increases Trust in Humans', *Nature*, 435, 673-676.

Susan J. LaBine & Gary LaBine (1996), 'Determinations of Negligence and the Hindsight Bias', *Law & Human Behavior*, 20, 501-516.

David Laibson (1997), 'Golden Eggs and Hyperbolic Discounting', *Quarterly Journal of Economics*, 112, 443-478.

Donald C. Langevoort (1997), 'Organized Illusions: A Behavioral Theory of Why Corporations Mislead Stock Market Investors (and Cause Other Social Harms)', *University of Pennsylvania Law Review*, 146, 101-172.

John List (2003), 'Does Market Experience Eliminate Market Anomalies', *Quarterly Journal of Economics*, 118, 41-71.

George Loewenstein and Richard H. Thaler (1989), 'Anomalies: Intertemporal Choice', *Journal of Economic Perspectives*, 3, 181-193.

George Loewenstein and Drazen Prelec, (1992), 'Anomalies in Intertemporal Choice: Evidence and an Interpretation', *Quarterly Journal of Economics*, 107, 573-597.

George Loewenstein, Samuel Issacharoff, Colin Camerer & Linda Babcock, (1993), 'Self-serving Assessments of Fairness and Pretrial Bargaining', *Journal of Legal Studies*, 22, 135-159.

George Loewenstein and Samuel Issacharoff (1994), 'Source Dependence in the Valuation of Objects', *Behavioral Decision Making*, 7, 157-168.

George Loewenstein and Daniel Adler, (1995), 'A Bias in the Prediction of Tastes', *Economic Journal*, 105, 929-937.

George Loewenstein and Don A. Moore (2004), 'When Ignorance is Bliss: Information Exchange and Inefficiency in Bargaining', *Journal of Legal Studies*, 33, 37-58.

George Loewenstein, Don A. Moore, and Roberto A. Weber (2006), 'Misperceiving the Value of Information in Predicting the Performance of Others', *Experimental Economics*, 9, 281-295.

Richard McAdams and Janice Nadler (2005), 'Testing the Focal Point Theory of Legal Compliance: Expressive Influence in an Experimental Hawk/Dove Game', *Journal of Empirical Legal Studies*, 2, 87-123.

Edward McCaffery and Jonathan Baron (2006), 'Thinking About Tax', *Psychology, Public Policy, and Law*, 12, 106-135.

Samuel M. McClure, David I. Laibson, George Loewenstein, and Jonathan D. Cohen (2004), 'Separate Neural Systems Value Immediate and Delayed Monetary Rewards', *Science*, 306, 503-507.

Richard McKelvey and Talbot Page (2000), 'An Experimental Study of the Effect of Private Information in the Coase Theorem', *Experimental Economics*, 3, 187-213.

Roger G. Noll and James E. Krier (1990), 'Some Implications of Cognitive Psychology for Risk Regulation', *Journal of Legal Studies*, 19, 747-779.

Martin A. Nowak, Karen M. Page, Karl Sigmund (2000), 'Fairness Versus Reason in the Ultimatum Game', *Science*, 289, 1773-1775.

Kevin N. Ochsner and Elizabeth Phelps (2007), 'Emerging Perspectives on Emotion-Cognition Interactions', *TRENDS in Cognitive Sciences*, 11, 317-318.

Elizabeth Phelps (2006), 'Emotion and Cognition: Insights from Studies of the Human Amygdala', *Annual Review of Psychology*, 57, 27-53.

Charles R. Plott (1996), 'Rational Individual Behavior in Markets and Social Choice Processes: The Discovered Preference Hypothesis', 225-250, in Kenneth J. Arrow, Enrico Colombatto, Mark Perlman, and Christian Schmidt, eds., *The Rational Foundations of Economic Behavior* (London: Macmillan).

Charles R. Plott and Kathryn Zeiler (2005), 'The Willingness to Pay-Willingness to Accept Gap, the "Endowment Effect," Subject Misconceptions and Experimental Procedures for Eliciting Valuations', *American Economic Review*, 95, 530-545.

George L. Priest and Benjamin Klein (1984), 'The Selection of Disputes for Litigation', *Journal of Legal Studies*, 13, 1-55.

Jeffrey Rachlinski and Forrest Jourdan (1998), 'Remedies and the Psychology of Ownership', *Vanderbilt Law Review*, 51, 1541-1582.

Jeffrey Rachlinski (2003), 'The Uncertain Psychological Case for Paternalism', *Northwestern University Law Review*, 97, 1165-1226.

James K. Rilling, David A. Gutman, Thorsten R. Zeh, Giuseppe Pagnoni, Gregory S. Berns, and Clinton D. Kilts (2002), 'A Neural Basis for Social Cooperation', *Neuron*, 35, 395-405.

Jennifer K. Robbennolt and Christina A. Studebaker (1999), 'Anchoring in the Courtroom: The Effect of Caps on Punitive Damages', *Journal of Law and Human Behavior*, 23, 353-373.

Roberta Romano (1986), 'A Comment on Information Overload, Cognitive Illusions, and Their Implications for Public Policy', *Southern California Law Review*, 59, 313-327.

Dominique J.- F. de Quervain, Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr (2004), 'The Neural Basis of Altruistic Punishment', *Science*, 305, 1254-1258.

William Samuelson and Richard Zeckhauser (1988), 'Status Quo Bias in Decision Making', *Journal of Risk and Uncertainty*, 1, 7-59.

Alan G. Sanfey, James K. Rilling, Jessica A. Aronson, Leigh E. Nystrom, and Jonathan D. Cohen (2003), 'The Neural Basis of Economic Decision-Making in the Ultimatum Game', *Science*, 300, 1755-1758.

Alan G. Sanfey, George Lowenstein, Samuel M. McClure, and Jonathan D. Cohen (2006), 'Neuroeconomics: Cross-Currents in Research on Decision-making', *TRENDS in Cognitive Sciences*, 10, 108-116.

David A. Schkade, Cass R. Sunstein, and Daniel Kahneman (2000), 'Deliberating About Dollars: The Severity Shift', *Columbia Law Review*, 100, 1139-1175.

Stewart Schwab (1988), 'A Coasean Experiment on Contract Presumptions,' *Journal of Legal Studies*, 17, 237-268.

Karl Sigmund, Ernst Fehr, and Martin Nowak (2002), 'The Economics of Fair Play', *Scientific American* 83-87.

J. Shogren, S. Shin, J. Kliebenstein, and D. Hayes (1994), 'Resolving Differences in Willingness to Pay and Willingness to Accept', *American Economic Review*, 84, 259-264.

Paul Slovic and Baruch Fischhoff (1977), 'On the Psychology of Experimental Surprises', *Journal of Experimental Psychology: Human Perception and Performance*, 3, 544-551.

Alison C. Smith and Edith Greene (2005), 'Conduct and its Consequences: Attempts at Debiasing Jury Judgments', *Law and Human Behavior*, 29, 505-526.

M. J. Stallard and D.L. Worthington (1998), 'Reducing the Hindsight Bias Utilizing Attorney Closing Arguments', *Journal of Law and Human Behavior*, 22, 671-81.

Michael Strahilevitz and George Loewenstein, (1998), 'The Effects of Ownership History on the Valuation of Objects', *Journal of Consumer Research*, 25, 276-289.

Cass R. Sunstein, David A. Schkade, and Daniel Kahneman (2000), 'Do People Want Optimal Deterrence', *Journal of Legal Studies*, 29, 237-254.

Cass R. Sunstein, Daniel Kahneman, David A. Schkade, and Ilana Ritov (2002), 'Predictably Incoherent Judgements', *Stanford Law Review*, 54, 1153-1215.

Taiki Takahashi, Kikue Sakaguchi, Mariko Oki, Seijiro Homma, and Toshikazu Hasegawa (2006), 'Testosterone Levels and Discounting Delayed Monetary Gains and Losses in Male Humans', *Neuroendocrinology Letters*, 27, 439-444.

Sabrina M. Tom, Craig R. Fox, Christopher Trepel, and Russell A. Poldrack (2007), 'The Neural Basis of Loss Aversion in Decision-Making Under Risk', *Science*, 315, 515-518.

Eric Talley and Colin Camerer (forthcoming 2007), 'Experimental Law and Economics', in *Handbook of Law and Economics* (M. Polinsky and S. Shavell, eds.).

Amos Tversky and Daniel Kahneman (1974), 'Judgment Under Uncertainty: Heuristics and Biases', *Science*, 185, 1124-1131.

Amos Tversky and Daniel Kahneman (1981), 'The Framing of Decisions and the Psychology of Choice', *Science*, 211, 453-458.

Amos Tversky and Daniel Kahneman (1983), 'Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment', *Psychology Review*, 90, 293-315.

Amos Tversky and Daniel Kahneman (1991), 'Loss Aversion in Riskless Choice: A Reference-Dependent Model', *Quarterly Journal of Economics*, 106, 1039-1061.

Leaf Van Boven, George Loewenstein, and David Dunning (2003), 'Mispredicting the Endowment Effect: Underestimation of Owners' Selling Prices by Buyer's Agents', *Journal of Economic Behavior and Organization*, 51, 351-365.

W. Kip Viscusi (1999), 'How Do Judges Think About Risk', *American Law and Economics Review*, 1, 26-62.

W. Kip Viscusi (2001), 'Jurors, Judges, and the Mistreatment of Risk by the Courts', *Journal of Legal Studies*, 30, 107-142.

W. Kip Viscusi and Richard Zeckhauser (2004), 'The Denominator Blindness Effect: Accident Frequencies and the Misjudgement of Recklessness', *American Law and Economics Review*, 6, 72-94.

Roberto Weber and Colin Camerer (2003), 'Cultural Conflict and Merger Failure: An Experimental Approach', *Management Science*, 49, 400-415.

Neil D. Weinstein (1980), 'Unrealistic Optimism About Future Events', *Journal of Personality and Social Psychology*, 39, 806-820.

Neil D. Weinstein and William Klein (2002), 'Resistance of Personal Risk Perceptions to Debiasing Interventions', 313-323, in *Heuristics and Biases: The Psychology of*

Intuitive Judgment (Thomas Gilovich, Dale Griffin, and Daniel Kahneman, ed.) (New York: Cambridge University Press).

Andrew Wistrich, Christopher Guthrie and Jeffrey Rachlinski (2005), 'Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding', *University of Pennsylvania Law Review*, 153, 1251-1345.

Paul J. Zak, Robert Kurzban, William T. Matzner (2005), 'Oxytocin is Associated with Human Trustworthiness', *Hormones and Behavior*, 48, 522-527.

Paul Zak and Ahlam Fakhar (2006), 'Neuroactive Hormones and Interpersonal Trust: International Evidence', *Economics and Human Biology*, 4, 412-429.