



ELSEVIER

Contents lists available at ScienceDirect

## Journal of Economic Behavior &amp; Organization

journal homepage: [www.elsevier.com/locate/jebo](http://www.elsevier.com/locate/jebo)

## Theory and experiment: What are the questions?☆

Vernon L. Smith\*

ESI at Chapman University, One University Drive, Orange, CA 92866, United States

## ARTICLE INFO

## Article history:

Received 30 October 2007

Received in revised form 30 January 2009

Accepted 4 February 2009

Available online 24 February 2009

## JEL classification:

B41

C92

## Keywords:

Experimental economics

Game theory

Methodology of science

## ABSTRACT

This paper deals generally with testing questions that arise both when experimental observations are in accord with the actions we predict, and when they are not. In both cases the inference of truth from observation is inherently ambiguous, and we face the daunting challenge of using our experimental skills and imagination to reduce this ambiguity. Primarily and most difficult of all we have to constantly reevaluate everything, including ourselves, especially in examining how we talk about and interpret our data. Although I will be drawing on examples and experience from laboratory experiments, the issues I consider apply just as meaningfully to other empirical studies whether from field experiments or observations from past records of socioeconomic processes.

© 2009 Published by Elsevier B.V.

... “when we praise any actions, we regard only the motives that produced them, and consider the actions as signs or indications of certain principles in the mind or temper . . . We look within to find the moral quality . . . (but) we cannot do this directly . . . When we require an action, or blame a person for not performing it . . . we esteem it vicious in him to be regardless of it. If we find, upon enquiry, that the virtuous motive was still powerful . . . tho’ checked in its operation by some circumstances unknown to us, we retract out blame, . . . (Hume, 1739; 1985, pp. 529–30).

## 1. Introduction

In this passage, I see Hume reminding us that action is produced by motive only in conjunction with other circumstances of action, and we cannot infer motive, or any mediating condition, from action alone. Mindfully, we all know that the statements we test with observations are of the form: If  $H(M)$ , where  $H(M)$  is a hypothesis concerning motive, and given the decision circumstances,  $C$ , then action  $A$  follows from the conjunction of  $H(M)$  and  $C$ , and not the reverse; i.e., we cannot conclude anything about the implications of  $A$  for  $H(M)$  independently of  $C$ . But we cannot be continuously mindful in all that we do. In constructing formal theory we try carefully to avoid confusing if-then with if-and-only-if, statements, but when we leave our closet, we easily slip casually into this confusion, theorists and experimentalists alike. Moreover, human motivation may be so inextricably bound up with circumstances that embody previous experience that it is not even meaningful to separate them.<sup>1</sup>

☆ Based on an address, “What are the questions?” at the 2007 meetings of the Economic Science Association in Rome, Italy.

\* Tel.: +1 714 628 2830; fax: +714 628 2881.

E-mail address: [vsmith@chapman.edu](mailto:vsmith@chapman.edu).

<sup>1</sup> In elaborating on my interpretation and use of this quotation from Hume, Bart Wilson reminds me that Wittgenstein would ground motive in action, “not by identifying an inner motive but through . . . actions as woven into our ‘form of life.’” (Personal correspondence, 14 February 2007).

This paper deals generally with testing questions that arise both when experimental observations are in accord with the actions we predict, and when they are not. In both cases the inference of truth from observation is inherently ambiguous, and we face the daunting challenge of using our experimental skills and imagination to reduce this ambiguity. Primarily and most difficult of all we have to constantly reevaluate everything, including ourselves, especially in examining how we talk about and interpret our data, and, in certain sensitive experiments, how we communicate word descriptions to our subjects and solicit the same from them.<sup>2</sup> Although I will be drawing on examples and experience from laboratory experiments, the issues I consider apply just as meaningfully to other empirical studies whether from field experiments or observations from past records of socioeconomic processes.

My purpose is to promote a discussion that probes more deeply into the implicit, often hidden, multiplicity of mediating conditions or auxiliary hypotheses on which we rely in order to extract meaning from tests of formal hypotheses. The exercise is difficult because I must necessarily go beyond the words we usually find adequate in our discourses. Hence, the paper's style is to raise questions, without any presumption that many of the answers are within the range of data now available.

## 2. Why so many experiments on seemingly narrow topics?

The answer is simply *stated*: we want to reduce error and to understand its sources. 'Error' is defined as a discrepancy between what we observe and what we expect as an implication of our predictive hypotheses; i.e., it includes specification error in all its forms as well as sampling error. This discrepancy depends upon factors that arise as part of the testing process or from the hypothesis generating methods, or both.<sup>3</sup> Predictive hypotheses, as expectations about observations, are sometimes inspired initially and directly from a theoretical model. I choose the word "inspired" rather than the word "derived" advisedly; this is because of the large reservoir of personal experiential knowledge that we must draw on in order to make operational a test of *any* theoretical construct with empirical observations. That knowledge is always subject to revision in the light of new direct examinations.

In exploring error it is important to probe beyond the restricted confines of the original theory and examine the assumptions we make in interpreting test outcomes. This is because testing involves a blizzard of narrowly prescribed circumstances that are not part of the theory. Error may be dependent on those circumstances, and patterns in such dependence may harbor clues to a more comprehensive understanding of behavior than can be gleaned within the restricted framework of reasoning within the theory.

An intensively studied case where predictions fail is the dictator game, originally motivated by the conjecture that rejections in the ultimatum game – where predictions also failed – were due to "fairness," meaning, I believe it was thought, a culturally derived social standard for sharing a fungible resource (money) under specified but singularly unfamiliar (see the discussion below) choice rules.<sup>4 5</sup> Thus Forsythe et al. (1994) reasoned that if "fairness" fully accounts for the ultimatum results then removing the responder's veto power should make no difference. We learned that it made a big difference, but the dictator game took on a special life of its own because scholars – conditioned to expect domination – were puzzled as to why 80% of the dictators were giving positive if modest amounts of money to anonymous persons.

Similarly, many important questions remain when the theory accounts for observations, and predictive error is ostensibly low, but we seek to understand the theory's success:

How are decision makers who have not the academic knowledge of either the theorist or the experimentalist able to approximate the predictions of the theory?

Do they think about the problem the way we do?

If not, why do we observe the predicted outcome?

I will turn first to a celebrated case in which theory has seemed to be predictive.

<sup>2</sup> As will be evident below, I am thinking of two-person games (ultimatum, dictator, trust), and the many variations on VCM games, see Smith (2008).

<sup>3</sup> When we do test bedding exercises in Economic System Design, 'error' is any observed gap between design objectives and observations. Thus our objectives are typically maximum efficiency, reduced decision or transactions cost, simplicity of operations, and so on.

<sup>4</sup> Thus, in Wierzbicka (2006) we have:

"The ubiquity of the words *fair* and *unfair* in modern English discourse, across a wide range of registers, is all the more remarkable given that these words have no equivalents in other European languages (let alone non-European ones) and are thoroughly untranslatable (p. 141) ... the everyday word *fair* has crystallized in its meaning political and philosophical ideas that were developed in the seventeenth and eighteenth century by the thinkers of the British Enlightenment and that have become entrenched in modern Anglo consciousness (p. 152) ... there are indeed some universal moral norms and values, but to think that 'fairness' is among them is an Anglo centric illusion (p. 162) ... The earliest examples that sound 'modern' in their use of *fair* are those with the phrase *fair play* ... (that) had its opposite in *foul play*, not in *unfair play* (p. 164) ... team sport ... provides a perfect model for 'fair' interaction because the emphasis is on rules and procedures, which are blind to the individual players interests and which everyone voluntarily accepts (p. 166) ... The concept of 'fairness' is an artifact of modern Anglo culture and one of the key concepts that speakers of Anglo English live with and see as the best basis for social life and interpersonal interaction (p. 166)." If subjects use the colloquial *unfair* in reference to their experience in an ultimatum or dictator game does this reflect the rules, the outcomes, the outcomes under the rules, or other circumstances such as that the experimenter is staking the game? It might be that in these games people really have in mind *justice* not *fairness*, since there is incomplete overlap between English and other languages in the translation of the former, and the former is about substantive outcomes not procedure. The standard forms of justice (distributive, corrective, retributive, commutative) all concern the result, not the procedures that arrive at the result. Wierzbicka (2006, p. 165) Experimenters use *fair* in the substantive utilitarian sense but our only clues as to what the subjects mean is through explorations of how the circumstances (context) of decision affect behavior.

<sup>5</sup> For a summary of experiments that probe issues related to the "fairness" interpretation of ultimatum game behavior, see Schmitt (2004).

### 3. When theory passes our tests, what are the questions—the double auction example?

Static competitive market equilibrium theory has performed well under double auction (DA) rules, and this is often touted as a victory for theory. It was indeed a remarkable and surprising discovery, but the theory failed to predict the weak conditions under which these outcomes would prevail.<sup>6</sup> The participants in these experimental markets need *only* private information and moderate repetition; they do not require sophistication or deep experience, except in certain very complex multiple commodity and network market environments; nor must they have an understanding of economics. In fact knowledge of economics does not appear to carry an advantage for the typical experimental DA participant.

Moreover, these documented results are robust in that they extend to different subject pools, to quite complex multiple market supply and demand environments, and in a qualified sense to other trading institutions.

**Subjects:** Hundreds, probably thousands of experiments by now have demonstrated robustness over an incredible range of people—with students at all levels, business persons (including regulated utility managers), government officials (including regulators), citizens of Russia and China, and academic professionals including game theorists.<sup>7</sup>

**Complexity:** Starting in the 1980s, new tests in multiple interdependent markets induced Constant Elasticity of Substitution (CES) values on two commodities. A CES payoff function together with a budget constraint implies that the price an individual is expected to pay for commodity A depends on the price of B and vice versa. But this statement only has meaning contingent on the prices (and therefore knowledge of how the budget limits choice) being known in advance, whereas the DA trading task is for a group of buyers and sellers to discover the equilibrium prices and corresponding trading volumes that satisfy the static individual budget constraints that emerge out of that process. Sixteen replications yielded DA convergence in only a few more repetitions than in isolated single commodity markets.<sup>8</sup> But modeling any such dynamic adjustment process has tended to exceed the bounded rationality of economic theorists.<sup>9</sup>

**Institutions:** Sealed bid-offer call markets and posted offer markets have been compared with DA across identical single commodity environments (Smith et al., 1982; Ketcham et al., 1984) With some exceptions, these alternatives to DA also converge to the market clearing equilibrium, although DA tends to converge more rapidly and consistently to efficient allocation levels (see the discussion and references in Davis and Holt, 1993).

The challenge of the DA empirical results has not yielded game theoretic models that predict convergence to a static competitive equilibrium. This failure is not for a lack of effort. Wilson (1987), one of the best theorists in the business, modeled DA trading with the standard assumptions of what is common knowledge—number of agents, their distribution of values, capacity to compute the equilibrium, etc. The model itself generates degeneracy in the end game. He acknowledges that we have a tool problem: “The crucial deficiencies, however, are inescapable consequences of the game theoretic formulation.” (Wilson, 1987, p. 411) We appear to be up against the ultimate limitations of the complete information common knowledge game theoretic model. We have encountered a dead end. We cannot model and predict what our subjects can routinely accomplish without complete information, and in some examples with complete information convergence is worse than under incomplete (private) information. Friedman (1984) is more successful, using an unconventional no-congestion assumption to finesse the usual problems, concluding efficiency and a final clearing price but the theory does not yield testable predictions about the dynamics of convergence.

### 4. When theory fails our tests, what premises might we question?

The short answer is that *any* combination of the assumptions of the theory and its implementation in the experiment may be among the culprits. Some good examples arise in two-person game theory: under the procedures commonly used and interpreted, the theory has not fared well in tests of ultimatum, dictator, investment, and trust games. The eight item list I will consider includes some of the formal axioms of game theory as well as some of the common assumptions that experimentalists make in order to construct and implement tests.

#### 4.1. Backward Induction

**Assumption 1.** A basic formal postulate is that People will use *backward induction* to analyze their situation, to determine the consequences of one's own and the choices of others, and to choose accordingly.

<sup>6</sup> Financial asset trading environments based on rational expectations (fundamental value) theory using DA rules provide an exception to these findings in the sense that convergence is by comparison very slow. (Smith et al., 1988; Porter and Smith, 1995). I return to this topic below in discussing the issue of subject “confusion.”

<sup>7</sup> At a conference of leading international theorists, the group was about to participate in an oral double auction price discovery experiment. After passing out the values and costs to the buyers and sellers and describing the rules, a leading game theorist raised his hand, and asked, “How can I make a decision—I do not have enough information.” Answer: “Don't worry you will be able to handle this,” as indeed he did, but there is a puzzle as to the processes whereby our brains have this and other skills so deeply hidden from our calculating self-aware minds.

<sup>8</sup> Williams and Smith (1986), Williams et al. (2000); also see the convergence results reported in Smith (1991) and Plott (2001).

<sup>9</sup> Except see the significant early contribution of Reiter (1959, 1981) also, the intriguing recent results in Gjerstad (2007) who reports experiments that use the DA to implement the Hahn–Negishi adjustment process, a theoretical model which allows for disequilibrium trades.

In effect, we have assumed that subjects' behavior is equivalent to having had the course and mastered the principles; they think like game theorists and the way we experimentalists think in constructing tests.

I will discuss several examples of empirical challenges to this assumption:

In ultimatum games (framed as a bilateral trade), subjects offer more under instructional treatments in which the proposer is asked to "consider what choice you expect the buyer to make . . . (also) consider what you think the buyer expects you to choose." Presumably, if people are already applying backward induction this instruction should make no difference (Hoffman et al., 2000). We were surprised by the outcome – proposers offered more – and we explored it more completely than we had originally planned or expected. We thought the slight modification in the instructions was a benign means of encouraging thoughtfulness; *What was our error?* We interpreted the data from the perspective of game theory, which is not the perspective of the subjects—remind them to think ahead like a game theorist and they appear to focus on socially derived circumstances (what do you and other expect?). Hence, we seem to encounter heightened concern about the possibility of implicit disagreement if the offer is too low. Wilson (2008) offers this interpretation). This is one of the many ways in which our learning can be occasioned by varying instructions and procedures. Experimenters who conduct these exercises routinely encounter surprises and it behooves us to ask after their implications for changing how we interpret observations. A procedure used to test a theory that leads to correct predictions is not "right," nor is one that leads to incorrect predictions, "wrong." Rather, we need to ask what the totality of this evidence tells us about the maintained (auxiliary) assumptions of theory and its test framework.

In asset trading, I and my coauthors conjectured that price bubbles were a consequence of some subjects having home-grown expectations of prices rising and on that supposition "reasoning" as follows: capital gains can be earned by buying shares in expectation of reselling them at a higher price and similarly by selling shares in anticipation of buying them back at a lower price (Smith et al., 1988). The proposition is falsified by the following procedure: Endow some subjects with cash and the right to buy shares to hold against dividend payments. They cannot, however, resell for cash. Endow others with the right to hold shares for dividends or sell shares against cash but they cannot repurchase any shares sold. Hence, if bubbles arise because subjects are mindfully expectant of an increase in prices they cannot profit from this expectation, and likewise if they expect prices to decline. The important implication is that no profit motivated trader reasoning as above will sell at prices below or buy at prices above the fundamental value. But Lei et al. (2001) apply these procedures, and report that bubbles can be observed much as in experiments in which these restrictions are not applied to trading behavior. Their results imply that our original interpretation of subject thought processes and actions were not correct because we had slipped into thinking that – given expectations of rising prices – the subjects reason like economists.

These results are sometimes interpreted to mean that the subjects are "confused." Notice, however, that this means that they are "confused" in the sense of failing to reason about asset trading the way we do as economists, by applying backward induction based on given expectations. That is, we have erred in believing they would think as we expected. The tests tell us what subjects do *not* do, but they help not at all in understanding how subjects think and why they do what they do. Moreover, the results do not explain why subjects commonly converge to intrinsic equilibrium value after becoming twice previously experienced. Groups ranging from business persons to stock traders to high schools students are all equally "confused." Subjects do not routinely apply reason to the tasks we give them the same way we do in our constructivist models. This seems to be true both when they achieve our predicted outcomes and when they do not. Is not our task to understand behavior? Imagine either Franz de Waal or Jane Goodall concluding from their studies that their Chimp subjects are "confused." Even so Goodall and de Waal are criticized for "anthropomorphizing" their subjects.<sup>10</sup>

You might say that we "anthropotheorize" our subjects.

Similar findings have been reported in VCM games where the standard observation is that subjects start by contributing unexpectedly large amounts to the public good then reduce them over successive periods. This has been explained by the conjecture that subjects give in anticipation that others will subsequently reciprocate their early giving, but when this expectation fails to be fulfilled they become disheartened and reduce their contributions.

Houser and Kurzban (2002, hereafter HK; also see Andriani, 1995) compare contributions in two treatment conditions: one consisting of a collective of four persons; the other consisting of one person and three computer "players" or robots. Individual subjects in the latter condition were instructed that the contributions of the computer players were fixed *independently of their individual contributions*; moreover, in this condition on each round of play, the humans were reminded that "This period, regardless of what you do, the computers will contribute . . ." (HK, p. 1068). HK find that (1) "Confusion," in the sense of error in which individuals fail to act rationally on the implications of the instructions, was responsible for approximately half of all alleged "cooperation" in VCM experiments. (2) All of the decay over time in the level of contributions was attributed to reductions in "confusion." Subjects give to the public good although they cannot reap any possible future benefit and currently are giving up money by not allocating all of their endowment to the private good. In repetition many subjects appear simply to be learning by experience the implications of what was stated in the instructions, viz., that individual payoffs from the public good depend only on total giving not on whether the individual contributed to that total. Implicit in this discovery is the hazard we face as experimentalists in drawing inferences from single play observations in which

<sup>10</sup> This is why ". . . it has not been clear that being an economic agent has had any advantages in the scientific study of economic behavior." (Smith, 1991, p. 259) Economists, as economic agents, do not have access to the unconscious mental processes that govern their own behavior in market and other social exchange processes. Can you model your own real time behavior in a DA experiment?

subjects have no opportunity to adjust decisions in real time based on experience, and we have to rely on how subjects apply reason to their task.

Again, these results tell us what the subjects do *not* do, and we remain uniformed about what mental processes underpin their behavior. Many subjects do not “get the message” that we thought should be transparent in the initial verbal information. I suggest that their brains’ require the real time feedback of experiential reinforcement—the hypothesis is that people learn primarily by doing, or observing others doing, not by deliberation and abstract analysis applied to their task description.

Finally, *Johnson et al. (2002)* using “look ahead mouse lab” technology report that in three-stage declining pie alternating offer games 19% of the subjects fail to look ahead to the next stage. When playing robots the subjects frequently rejected equilibrium offers. This is consistent with the hypothesis – as in HK – that the “mouse lab” subjects do not apply what we would regard as the most elementary backward induction and economic reasoning when analyzing the consequences of their choices.<sup>11</sup> *Johnson et al. (2002)* cite many examples in which people fail to use backward induction; through repeat experience, however, subjects can be induced to behave in accordance with the implications of backward induction. This does not imply that they generalize that learning to other situations as a problem solving principle.

Earlier, a significant study by *Harrison and McCabe (1992)* reported convergence toward equilibrium in alternating offer, declining pie, game experiments in which subjects repeatedly obtained experience with the future sub games before playing the full game. Subjects needed to experience the future to take it properly into account, which suggests that people may “induct forward,” in the sense of choosing, then modifying their choice in repetitions when the outcomes turn out to be unsatisfying to them. They arrive at their play choices by adaptation and experience over time, not by reasoning ahead and taking account of both own and other choices. This finding is particularly interesting since an important function of futures markets – or so we have discovered – is to serve in enabling people to internalize a reading on each other’s expectation of a future event by making a current market in claims contingent on that future event. Thus, futures markets have been found to moderate laboratory stock market bubbles.<sup>12</sup> In this case futures’ trading is an example of a market institution that serves to facilitate looking ahead, obtaining information and performing backward induction for individuals. This market helps them to perform a function in which they seem to be deficient when left to their own unaided rationality devices based on the same information available to futures traders. Apparently, people need to see how futures’ traders aggregate “look ahead” information, and then they modify their asset trading behavior. In this sense, futures markets can serve as a specialized form of experiential aid to individual rationality. Thus, we learn that specialization is not just “limited by the extent of the market,” but is an integral part of that extension!

In the above examples of “confusion” it is natural to instruct or teach subjects more fully, and use tests to screen out those who do not reason as we do in order to be sure that we have people whose understanding of decision making corresponds to what we understand by it. This procedure is of merit: it enables us to verify whether we have identified a source of our explanatory error. Contrarily, the problem with this procedure is that it may insulate us from exploring how subjects think about the tasks we give them. In the asset market, VCM and declining pie environments, subjects get it right on their own with enough (or the right kind of) experience and it would be nice to know more about how that process works. For example, suppose we tested people’s comprehension of the instructions, then compared the behavior of those who pass with those who fail in separate experiments.

What I think we have discovered is that backward induction is not a natural tool to use in human decision processes, and is very difficult to teach using pencil and paper exercises that people absorb, generalize as a principle and automatically apply to unfamiliar circumstances. Those of us who learned the backward induction principle long ago easily forget how difficult it was to internalize, and we can still fall into error. Perhaps many people explore an environment by making an accustomed move, observe the consequence, then in repeat play make whatever adjustment over time that seems to be sustainable and serve their interest in the same or sequences of different games. This calls for caution in assuming that our subjects will come off the starting blocks in one-shot games and get that socially complex and unfamiliar circumstance right.

Theorists involved in the design of the FCC spectrum auction process did not use backward induction to examine their own auction proposals. Based on the idea that licenses had a common value for all bidders, English auction procedures were recommended since this allows the “given” dispersed private information on common value that bidders have collected to be fully revealed at auction and a higher sale price obtained—an important static equilibrium theorem. But if that were true, so called “sophisticated” bidders using consultants would conclude, in the reality of an open planning environment, that they should invest nothing to acquire information that would be revealed later at auction. No information, no variation in estimated value to be revealed and the static theorem becomes vacuous (*Banks et al., 2003*). At the foundation of common value English auction theory there is degeneracy in the end game. The designers did not find it natural to look ahead and apply backward induction to their own design proposals.

<sup>11</sup> Also see *Herz (2007)* who examines the ‘mouse lab’ acquisition of information in a “partners” vs “opponents” trust game whose structure is like that studied by *Burnham et al. (2000)*, but simplified to enable greater focus on information processing. Among the many interesting results reported are the observations that “partners” open significantly more payoff boxes, and spend significantly more looking time (double) at boxes than “opponents.” These and other observations, such as “partners” reopening more boxes than “opponents,” are consistent with the hypothesis that “partners” are engaged in explorations that show a greater orientation toward relationship building within the single sequential play than “opponents.”

<sup>12</sup> See *Porter and Smith (1995)* and *Noussair and Tucker (2006)*.

Actually, there are surely many theorems that hold only because the modeler limits the boundaries of what will be formally analyzed by “cutting the tree,” usually in the pursuit of tractability, with the desirable objective of simplifying and understanding the elements of a problem.<sup>13</sup> Since such cognitive limitations must always be the forte of real people, this gives us a hint as to why full backward induction might be unnatural, and not worth the trouble, as a general principle of life to internalize. Of course this certainly does not deny that there are significant strategic interactions in which the parties to an interaction should or would conduct a careful analysis of their situation. The standard game-theoretic framework is surely a good thing to know if you already are reasonably certain that the players’ have access to its principles, including strictly opposed interests. But one of our principal tasks is to understand the homegrown means by which subjects make decisions in every day human life, and their expression of that experience in the lab. There is time enough for us to teach certain principles in the classroom without assuming that they apply to the behavior of all economic agents. If that assumption were true we would have nothing to teach.

#### 4.2. Independence of history and future

**Assumption 2.** In game theory an important conceptual distinction is made between a “stage” game played once, and a game in which two players engage in repeat play of the same such game. As experimentalists we have assumed in our implementation of this conceptual distinction that subjects make choices in a (stage) game based on what we *understand* by it—a single play through a decision tree between ‘strangers’ that is devoid of a history and future.<sup>14</sup>

We do not have good evidence that this proposition is valid. We have known at least since [Fouraker and Siegel \(1963\)](#) that subjects behave differently in repeat play than in single play. But this tells you only that reputation building is more important in repeat play of the same game than in single play, not that it is absent in one-shot games, or that it is not important in sequences of different stage games each played once with the same or distinct partners. Hence, the question – Do subjects choose in stage games as if they understand what we understand? – is not answered by saying that subjects know the difference between a one-shot game and a repeated game. Of course they play differently, but I am talking about the one-shot base line, and whether it is devoid of reputation considerations in play across the games people experience in the world (or in the lab), whether the “same” game is repeated or not.<sup>15</sup>

The problem with [Assumption 2](#) is revealed when you contemplate the agony of attempts to banish a history and a future in the matching protocols:

- a. Random rematching of pairs in a finite sample was of course inadequate.
- b. Or, let every subject play every other exactly once; this is judged inadequate because of contagion: no one must play anyone who played with someone you played earlier.
- c. If you have each person play their matched counterpart in the population exactly once you encounter the problem that you have to have large N to get large samples. But large N collides with the finding that trusting behavior declines significantly in a trust game when group size is increased ([Burnham et al., 2000](#)). Hence, you have to have large N overall, and hold constant the smaller number who are present in a given location. But these two conditions juxtaposed may strain credibility with the subjects.
- d. As in footnotes 17 and 18, one recruits from pools that are inexperienced in previous two-person game experiments.

All this is mind-boggling. Moreover, we still do not know what this means for how the brains of our subjects process elaborate matching protocols. Yet different protocols commonly generate differences in observed behavior. The bottom line appears to be that the abstract concept of single play invokes conditions sufficiently remote from much human experience that it may be operationally difficult to penetrate.

In single play dictator, investment trust and trust games, the results are affected by double blind treatments that are difficult to reconcile with postulate 2 if these games are correctly portrayed as representing single isolated play. Context and third party knowledge reflect characteristics of subject sociality in a world of repeat play that seems to import a history and/or a future in spite of player anonymity.

An important example illustrating how easily procedures can introduce a future is found by comparing dictator game results reported by Cox: in one ([Cox, 2004](#)), the amount given is 36.3% of the stakes; in the other ([Cox, 2000](#)), dictators give

<sup>13</sup> Fifty years ago this was the common general equilibrium critique of all partial equilibrium analysis, but the latter has been resurrected with the explosion in applications of Nash’s work.

<sup>14</sup> The procedures used to implement these single play conditions were pioneered by [Siegel and Fouraker \(1960\)](#) and [Fouraker and Siegel \(1963\)](#). Subjects were matched anonymously in a large sample for a single interactive play of a specified game, paid their earnings privately and separately, never to be matched again. Sid Siegel was a master of such technique. In the 1963 work the sample was large indeed: 106 subjects participated in the five bilateral monopoly experiments; 291 participated in ten different oligopoly experiments (6 quantity adjuster and 4 price adjuster). Subjects were all randomized into treatments *before* they arrived at a session, and were notified a week in advance of the night they were scheduled to appear for a session. ([Fouraker and Siegel, 1963](#), p. 22, 114–115).

<sup>15</sup> At Arizona, in the late 1980s and 90s, I and my coauthors conducted a large number of two-person game experiments. These were spread over many semesters, and we commonly recruited from the subject data bank pool only those (mostly new additions to the pool) who had not been in previous such two-person experiments. Thus did we seek to control for “history” in the sense of experimental play—a huge challenge.

51.8%.<sup>16</sup> The difference between these two sets of observations is attributable to a single difference in the procedures: the larger amount is transferred when subjects are informed that a second undefined Task 2 will follow the first, and after they have completed the first they will receive instructions about the second. Hence, it appears that any suggestion of a “future,” even if undefined, and otherwise thought to be “independent,” substantially increases giving—an important sensitivity finding.

From the perspective of the subjects, however, the dictator game task must seem strange. You are recruited to the lab, awarded a costless right to some of the experimenter’s money, and given an opportunity to transfer any part or none of it to a second anonymous person who has nothing to do except receive the money. (It is not the standard interactive game in which payoffs jointly depend on the decisions of both players.) “The God’s must be crazy!” But if there is a second task to follow, perhaps it is not so crazy; is it better to give in case there will be a subsequent return on your generosity?

Just because people cannot be relied upon to backward induct with mindful rationality, does not mean that their socially conditioned autonomic brains are oblivious to all considerations or hints of a future, especially it seems, in the case of the dictator game. I think these results are significant in hinting as to how geared subject brains are to be cognizant (if unconsciously) of building valuable reputations in the repeat interactions of daily life—considerations that it would be convenient to believe are fully controlled in the single play of a stage game between anonymous pairs of subjects.

When a subject faces a task choice in a one-shot game what is her homeostatic “set point?” (Camerer et al., 2005) How safe are we in assuming that the set point corresponds to the experimentalist/theorist definition: a single interaction between strangers in which the players take no account of reputation issues? How is that set point affected by circumstances and history? Suppose a person’s accustomed behavioral mode is to try cooperation in most or all first encounters, because this seems to have had long term benefits, and to modify it only when it is discovered to ill serve her in a particular situation. Then when matched by the gods with another like person, this “set point” will earn her more money than if she always acted according to Assumption 2.

This interpretation is consistent with trust game data reported by Rigdon et al. (2007) in which subjects are allocated randomly between two repeat play treatments: in the control each trial rematches subjects in pairs at random; in the research treatment subjects are rank ordered on each trial by frequency of historical cooperation over the previous five decision trials, and then matched according to these “trust” and “trustworthy” scores. Subjects are *not informed of the scoring or matching rules*, and in both treatments are informed only that on each trial they will be paired with another individual. How do people respond in discovering through experience that they are in more, compared with less, cooperative behavioral environments? Any person with a cooperative “set point” in the research treatment is more likely to be matched with another such person on any trial; potentially both earn a higher payoff and, compared with the control, there is an unannounced reinforcement for cooperation to build over time. Alternatively, a subject discovering that he or she is in an environment of cooperation may feel they can profit even more by defecting. Hence, cooperation could build, decay or be indistinguishable from the control, depending on how subjects respond to their experience in the treatment condition.<sup>17</sup> From the first block of trials, 1–5, to the last, 15–20, the ratio of percent cooperation in the sorting treatment to the random control increases by 100% for Players 1 (“trust”), and by 50% for Players 2 (“trustworthy”).

In game theory, utility for the individual is the utility from his current choice,  $s_i$ , in the stage game,  $(1 - d)u_i(s) + dV_i(H(s))$ , where  $s = (s_1, \dots, s_i, \dots, s_n)$  for  $n$  players,  $d$  is a discount factor,  $H$  is the history of play, and  $dV_i(H)$  is  $i$ ’s endogenous discounted value of continuation. In the standard theoretical argument, if the game is repeated, it means that the individual has incentive to forgo currently choosing dominance to avoid spoiling his future, and use the interactive opportunity to develop a mutually beneficial exchange relationship over time (see Sobel, 2005). If the same game is not repeated with the same person ( $n = 2$  in our experiments), there is no “history,” and by hypothesis the continuation value is assumed to be identically zero—a stark contrast with repetition. But why should a real person see no continuation value across stage games with different but culturally more or less similar strangers? Can we ignore the fact that each person shares cultural elements of commonality with the history of others? Why is  $H(s)$  in actual human experience derived only from play of the same game with the same person? Is not culture about multilateral human sociality? These empirical extra theoretical questions critically affect how we interpret single play observations.

#### 4.3. Complete information and dominance

**Assumption 3.** Subjects have *complete information* on own and other payoff.

**Assumption 4. Domination.** Given a choice between any two amounts of money, subjects always choose the larger amount whatever the circumstances.

<sup>16</sup> The reported levels of dictator giving are higher than in the standard dictator game in these treatments because the transfers are tripled. This treatment allows the results to be compared with giving in the investment trust game (Berg et al., 1995).

<sup>17</sup> It is straightforward to model the case in which a small population of cooperators can invade a population of defectors through mutual recognition and deliberately engaging in forms of positive reciprocity. But Rigdon et al. (2007) hypothesized that many people have instinctive cooperative tendencies that are non-deliberative and invoked in response to what they experience.

**Table 1**

Percent who trust with and without direct punishment.

Frequency Treatment trial block	Percent cooperation conditional on play in cooperative subgame	
	Trust, no direct punishment	Trust with direct punishment
1–5	64.4	82.1
6–10	84.6	86.4
11–15	89.0	92.6
16–20	91.0	91.4

Based on McCabe et al. (1996), Tables 6 and 7.

In same-pair repeated games domination yields good predictions when subjects are matched anonymously with *private information*, but not under *complete information*.<sup>18</sup> (McCabe et al., 1996, 1998). Private information is also the condition under which market equilibrium theory yields good predictions. Hence, it appears that across *all* interactive situations, repeat play and private information are the conditions under which equilibrium theory is very effective in organizing our observations. In two-person interactions such replicable experimental results motivated the theorems in Kalai and Lehrer (1993) which capture the phenomenon of convergence but do not address or predict the properties of its dynamic path. This is an important empirical victory for economic/game theory. Where game theory fails to predict well is in the much less realistic case in which the cash value of outcomes to other(s) as well as self is known perfectly.

Domination has been thought to have failed decisively in many complete information games, and many have invoked social preference explanations and tests (see the discussion and references in Fehr and Fischbacher, 2002). But this explanation is only justified if we know that our other assumptions are not violated. Thus, if some do not backward induct or do not choose as if mindful of the meaning of single play ( $V_i(H(s)) = 0$ ), we cannot be secure in our preference conclusions. I am not claiming that there is no room for social preferences, only that the common approach is to impose that interpretation by implicitly assuming/believing that our test frameworks control for *all* the assumptions of game theory that may result in cooperative choices, *except* domination.

#### 4.4. Folk theorem

**Assumption 5.** Repeat play of a stage game favors cooperation because in “long” games subjects can rationally use punishment to induce domination opponents to play cooperatively.

Table 1 compares two trust games, each using repeat play with the same pairs: one in which defection can be punished directly at a cost in the sub game containing the cooperative outcome; the other in which defection can only be punished on the next round by playing the sub game containing the equilibrium.

The prospect of costly direct punishment increases the frequency of conditional cooperation compared with the indirect punishment case, but over time the game with the indirect punishment option (asymptotically in the average) approaches the frequency of cooperation in the game with direct punishment. Hence, in this example, the Folk Theorem has validity both when people have to devise their own forms of punishment in repetition, and when it is provided for them directly in the decision tree. Moreover this convergence to cooperation emerges “quickly” – within 20 trials – so well practiced it would appear are most subjects in their cultures of repeat social interaction. In this case a relatively small nine percent get locked into non-cooperative play in the last trial block.

In repetition of the same game, people deviate from their initial choice as they discover, along with their paired counterpart, strategies that improve on that initial outcome. Empirically, Houser has shown that subjects can be usefully “Typed” by their initial round-one choice, a characteristic which continues to have significant predictive power across individuals even as they adapt to different treatment circumstances in dynamic play over time in various two-person trust game environments (see for example, Houser, 2003). Do these “types” define social preferences? Perhaps they do, but the problem with invoking other regarding utility is that it is confounded with the observation that such “types” have been observed to reap larger average individual rewards in both single and repeat play by choosing cooperatively. For example, we have reported single play trust game data comparing faculty with undergraduates (Coricelli et al., 2000). The undergraduates take substantially less time to choose at each decision node than the faculty and earn higher average monetary rewards because they cooperate more. Are we to judge the undergraduates to be irrational? Strange, that faculty requires more time for an analysis that one expects to be quick and easy for sophisticated players. What in their brains are giving them pause in this straightforward application of game theoretic reasoning?

Consider again the trust games with and without a direct punishment option for defection on offers to cooperate: across all single play and repeat play matching protocols, average earnings from cooperative play never fall below sub game equilibrium earnings. In all but single play in the pure trust game people offering to cooperate make more money on average than if they played non-cooperatively. This is shown in Table 2.

<sup>18</sup> The recording of such cooperation has become famously commonplace going back decades to the earliest experimental games. (See e.g., Fouraker and Siegel, 1963).

**Table 2**

Average earnings per trial for players offering to cooperate: non-cooperative equilibrium earnings = 40.0; cooperative earnings = 50.0.

Treatment protocol	Trust; no direct punishment	Trust with direct punishment
Single play	40.0	46.2
Repeat single	NA	41.5
Same pairing, repeat play	46.9	46.6
Random pairing, repeat play	41.2	40.7

From McCabe et al. (1996), Table 4.

In the lab we offer people choices in a single isolated stage game; is that circumstance isolated from the person's social history? Have subjects' brains encoded experience from isolated single play games, or from sequential mixtures of single play games, repeated games, games that are repeated conditional on the outcome, and games in which it is not clear who was the first or second mover. In life whether a game is repeated or not may depend on the outcome of the initial encounter, and is not neatly defined in advance as assumed in theory and implemented by instructions; rather games are defined only by past and continuing experience. The "isolated" game is an abstract construct. It is conceptually very important to make these distinctions when we are engaged in closet theorizing, as mental control experiments, but this does not mean that subjects' brains, or even our brains, encode these distinctions as part of our autobiographical experiential knowledge.

#### 4.5. Context irrelevance

**Assumption 6.** Context does not matter, only the underlying abstract game structure.

Perhaps the most useful learning from ultimatum and dictator games has been in demonstrating the power of context – instructions and procedures – to shift decisions significantly and measurably as much as cultural and information treatments. In ultimatum games the data can be shifted from 44.1% offered to 27.8% offered (Hoffman et al., 1994). Similar context effects appear in dictator and trust games. As indicated previously, in trust games words like "opponent," "partner" and "counterpart" are associated with different patterns of play over time. And beware always of inadvertently creating an in-group or out-group "set point" context or any slight hint of a future interaction in dictator games.

Context variation may include an allegedly benign expansion of the decision space to include "cheap talk." Thus, ultimatum games do not permit responders to express their concern (displeasure, emotion) – I am going to call it "disagreement" (Wilson, 2008) – with the terms of the proposer's offer, except by rejecting the offer. Houser and Xiao (2005) study ultimatum games in which the responder, in addition to deciding whether to accept or reject the offer, has the opportunity to send any freely composed written message to the proposer. In the baseline experiments in which no auxiliary message can be sent – only "accept" or "reject" – 12 of 20 (60%) of the offers of \$4 or less (the stakes are \$20) are rejected by the responders. In the comparison experiments where an auxiliary message can be sent, 6 of 9 (32%) of the offers of \$4 or less are rejected. The contents of the messages in the latter case clearly vent negative emotions toward the proposer, and appear to provide a low cost communication of strong disagreement that relieves half the responders from incurring the high cost of rejection. This behavior cannot be understood rationally in single play games, but it can be understood in the experiential history of ongoing interactive social context, where people can credibly express their disagreement and expectation failure short of taking precipitate actions – an eye for an eye and a tooth for a tooth – that will spoil continuation and the prospect of an improved future. Why should we suppose that such histories are made irrelevant by our careful attempt to create the conditions we think should be sufficient for one-shot play?

And why are so many dimensions of context significant treatments in determining action? Following Hume (1739; 1985) we would say that each context defines elements in the rich set of circumstances other than explicit payoff motivation. I suggest that when people walk into an unfamiliar situation they bring a brain encoded with categories generated from the relative frequency of exposure to discrete past experiences in life (see Hayek, 1952). Based on the instructions and procedures, people search that data bank for personal knowledge relevant to the task they face and then choose an action according to their motivation. Under this hypothesis, each subject's personal autobiographical knowledge is filtered by contextual circumstances for relevance to the decision at hand. This conjunction determines the person's "set point," in which personal knowledge also appears to reflect important life style characteristics derived from the human career of sociality.

#### 4.6. Equivalence of the standard game forms

**Assumption 7.** The two standard game forms – extensive and normal – are equivalent.

Long ago Schelling (1960) made cogent arguments that one should not expect the two game forms to be equivalent. Schelling's main tool was not game theory, but ordinary common sense in thinking through how experience and circumstances might reasonably influence how individuals behave. There are now many experimental studies falsifying the equivalence hypothesis in a variety of different games (Schotter et al., 1994; Rapoport, 1997; McCabe et al., 2000).

**Table 3**  
Percent dictators giving nothing.

Treatment	\$10 Stakes (%)	\$40 Stakes (%)
Single blind; experimenter gives endowments	19	15
Single blind; DMs earn endowments	79	70
Double blind; DMs earn endowments	95	97

Based on Cherry et al. (2002).

The latter argue that the important principle that allows better coordination “derives from the human capacity to read another person’s thoughts or intentions by placing themselves in the position and information state of the other person.”<sup>19</sup> (McCabe et al., 2000, p. 4404). Such “mind reading” to detect intentions underlies reciprocity in extensive form games. This interpretation has found support in the brain imaging study reported by McCabe et al. (2001). It was indicated 250 years ago by Adam Smith: “As we have no immediate experience of what other men feel, we can form no idea of the manner in which they are affected, but by conceiving what we ourselves would feel in a like situation.” (Smith, 1759, 1982, p. 9). Krueger et al. (2007) report that in repeat play of trust games activation of the intention detecting modules (paracingulate cortex) occurs in the trust building phase for those choosing to trust unconditionally. In the maintenance phase activation is relegated to older brain modules (septal) as the interaction becomes routine and people settle into reciprocal coordination on the cooperative outcome. As Polanyi (1962) would say, what begins as “focal awareness” soon becomes an automatic part of our “subsidiary awareness.” Focal awareness places larger demands on the brain’s resources for reading intentionality, and it seems that our brains long ago evolved mechanisms to economize on those resources.

These results and that of others cited above imply that the extensive and normal forms are not played as if they were the same games. Players use moves to read and signal intentions that are not the same when actually experienced in sequential move form as when imagined in a mental experiment corresponding to the same sequence, but expressed in the normal form.

Cooper and Van Huyck (2003) draw a similar conclusion. In particular they suggest that players in the first position in their games are motivated to choose moves that include the second player in the decision process, and thus invite the ready development of a relationship. I think this readiness to develop a relationship in stage games hints strongly as to why subjects who are not dominance players (Tables 1 and 2) are able conditionally, in some games, to reach more profitable cooperative outcomes in repeat play with the same pairs.

#### 4.7. Own vs other people’s money

**Assumption 8.** Monetary payoffs matter, but not *who* provides the money or *how* people acquired the stakes: there is no OPM (other people’s money) problem.

In dictator games, Cherry et al. (2002) have demonstrated that this implicit (and traditionally unexamined) assumption is false. Although larger stakes increase altruism somewhat, the results are dominated by whether the money was a freebie from the experimenter or earned, and is significantly further influenced by whether third parties (single vs double blind) can know and identify individual actions.<sup>20</sup> This is shown in Table 3.

Oxoby and Spraggon (2008) have further examined the robustness of these results—100% of their dictators give nothing. Consequently, there is a substantial difference between how people allocate experimenter supplied endowments between own and other use and how they allocate money that they have earned from the experimenter based on pre game task performance. After all it is rare to see people walking around in the streets giving small amounts of their money to anonymous strangers. The data appears to send an important message: as experimenters we have learned a lot about behavior in two-person games using OPM, and have used such findings to fit mixtures of own and other preferences to the data, but these findings may perhaps tell us little about people’s behavior when using their own money. There is a large literature interpreting behavior as due to “fairness” in games using the experimenter’s money, but how “fair” are people with their own money? Cherry et al. (2002) and Oxoby and Spraggon (2008), suggest that in the dictator game, at least, some 0–3% of them are “fair” after passing through a procedure that legitimizes the funds as their own property. The bad news here is that it seems that some of us may have spent a lot of time, money and effort on a special case that does not seem to generalize. But our knowledge is growing and the good news is that much new learning – pro or con – is likely as all the old questions concerning dominance vs social preferences must be asked yet again using various procedures for implementing decision when people use their own or earned money. The theorist’s work is from sun to sun, but the experimenters work is never done. What we do not know, as yet, is the extent to which this variation in dictator results extends to other games. In this exploration we have to proceed on a case by case basis.<sup>21</sup>

<sup>19</sup> This is also central to the idea in Schelling (1960) that focal points facilitate coordination.

<sup>20</sup> It seems that Hoffman et al. (1994) did not go far enough in requiring only that subjects earn the right to the first mover position.

<sup>21</sup> Thus, Schwartz and Ang (1989) found that asset bubbles were observed when subjects were required to bring their own \$20 endowment (OPM being a possible explanation of bubbles). Alternatively, Cherry et al. (2005) found no strong OPM effects in VCM games.

## 5. Closure

Game theory does not predict when and where its principles can be applied except through the filter of our judgments as to the validity of its assumptions in our test environments. How sensitive are our test conclusions to the Assumptions 1–8? How do we come to know their separate validity given the ambiguity of composite tests and the challenge of creating more precise tests? The maintained assumption of game theory is that the players' interests are in opposition, but the theory relies on extra theoretical knowledge of when this is the case. Hence, the impromptu *ex post* complaints, when experiments produce unpredicted results, that the stakes or subject sophistication “must have been” too low—conditions that were not explicit in the original theory but were quickly invoked after seeing the anomalous test results.

As experimentalists we have come to the realization that people's decisions, consistent with their cultural experience, are often as sensitive to the specific context as to variation in the structure of the game. That is how brains work. An understandable fault we share is the error of thinking too much like economists and game theorists when interpreting the decision choices of our subjects. The standard game-theoretic assumptions are an important generator of precise theorems, but this desirable property should never be confused with the *precision of the tests the theorems motivate*. Modeling skill is essential when constructing abstract first-cut representations of socioeconomic systems and their underlying function. Moreover, these representations are often accurate predictors of the outcomes we observe – as in markets, various auction institutions and in two-person games under private information – but those skills fail to enable us to see how and why people who do not have our professional expertise, can or cannot reach our predicted outcomes across the spectrum of games we study.

In two-person games of complete payoff information, one of our tasks in coming to terms with experimental observations is to reexamine our nearly universal procedure of serving as the third party provider of the resources that motivate the decisions we observe. We cannot assume without evidence that this particular circumstance does not qualify the interpretation of our results in the games we study. The dictator investigations have shown that these circumstances matter.<sup>22</sup> This is not to say that the focus on OPM is irrelevant, only incomplete. Both government and private organizations are commonly engaged in the allocation of resources supplied by others, but this is not the framework from which traditional economic and game theoretic principles arose.

In thinking about how we can better understand these issues, I want to suggest first that we ask of any of our experiment designs the following question: how would we conduct them if people brought (or otherwise provided) their own money? Secondly, we need to move past the narrow confines of a theoretical perspective that emphasizes games of pure distribution, particularly games of redistribution of the experimenter's money.

The important property of an economy is its capacity to generate wealth from the specialization that is made possible through exchange (or other forms and norms of sharing the fruits of specialization). Game theory and its application in experiments takes as given that a game can be nonzero sum, but does not ask why or how these circumstances might occur and how they might effect the perceptions and motivation of the participants. Here is simply one illustration of how these circumstances might be modified in standard experiments:

In the ultimatum game the experimenter uses some procedure for awarding money to A but his right to it is conditional on B's agreement to the share offered by A; the dictator game removes B's right of veto. Anyone who brought their own money to these experiments would surely elect not to play either game, as they are games of pure redistribution of the money brought to the table. (It would be judged “unfair” in the Anglo sense.) I am reminded of Ellsberg's insightful paper on zero sum game theory, a theory launched by the matching of reluctant dualists when viewed within a larger opportunity space (Ellsberg, 1956).

We could give the constant positive sum ultimatum game economic content as follows: Each player provides \$M of his own money. Some procedure is used for pairing the subjects, and determining who is to be Player 1 and who Player 2; this procedure in some variations might incorporate an earned and/or investment feature. It is understood that their pairing has economic significance in the sense that there are synergistic gains from the interaction equal to some fixed sum  $y > 2M$ . The experimenter provides only the surplus above  $2M$  which represents the gains from specialization and exchange, as this is the one reliable source of a “free lunch” that converts economic systems into non-zero sum games. Hence, the total to be shared under the property right rules of the game is  $2M + y$ , making it feasible for each to receive a share of the jointly created net gain above their pooled initial contribution,  $2M$ .

In summary I think there is a sense in which the state of experimental economics is comparable to the description of knowledge in physics a century ago: “The mass of insufficiently connected experimental data was overwhelming . . .” (Einstein, 1949, p. 17). Hume (1739; 1985) points to an important source of unexplained variability in our data, the “circumstances unknown to us” in the expression of behavior; clues to the source of that variability, however, are often revealed whenever we probe some of the explicit or implicit premises of game theory and our empirical implementations of the theory.

<sup>22</sup> In the original investment trust game Berg et al. (1995) required the Player 1 investors to decide how much of their \$10 show up fee to send to (invest in) Player 2, suggesting that the authors were already thinking in terms of conveying a sense that the funds were the subjects' own. But the circumstances defining alternative sources of funding for the participants in trust games is yet to be systematically investigated.

## References

- Andrioni, J., 1995. Cooperation in public good experiments: kindness or confusion? *American Economic Review* 85, 891–904.
- Banks, J., Olson, M., Porter, D., Rassenti, S., Smith, V., 2003. Theory, experiment and the federal communications commission spectrum auctions. *Journal of Economic Behavior and Organization* 51, 303–350.
- Berg, J., Dickhaut, J., McCabe, K., 1995. Trust, reciprocity, and social history. *Games and Economic Behavior* 10, 122–142.
- Burnham, T., McCabe, K., Smith, V., 2000. Friend-or-foe intentionality priming in an extensive form game. *Journal of Economic Behavior and Organization* 44, 1–17.
- Camerer, C., Loewenstein, G., Prelec, D., 2005. Neuroeconomics: how neuroscience can inform economics. *Journal of Economic Literature* 43, 9–64.
- Cherry, T., Frykblom, P., Shogren, J., 2002. Hard nose the Dictator. *American Economic Review* 92, 1218–1221.
- Cherry, T., Kroll, S., Shogren, J., 2005. The impact of endowment heterogeneity and origin on public good contributions. *Journal of Economic Behavior and Organization* 57, 357–365.
- Cooper, D., Van Huyck, J., 2003. Evidence on the equivalence of the strategic and extensive form representation of games. *Journal of Economic Theory* 110 (2), 290–308.
- Coricelli, G., McCabe, K., Smith, V., 2000. Theory-of-mind mechanism in personal exchange. In: Hatano, G., Okada, N., Tanabe, H. (Eds.), *Affective Minds*. Elsevier, Amsterdam, pp. 249–259.
- Cox, J., 2000. Implications of Game Triads for Observations of Trust and Reciprocity. University of Arizona, Department of Economics, unpublished paper.
- Cox, J., 2004. How to identify trust and reciprocity. *Games and Economic Behavior* 46, 260–281.
- Davis, D., Holt, C., 1993. *Experimental Economics*. Princeton University Press, Princeton.
- Einstein, A., 1949. Autobiographical notes. In: Schilpp, P.A. (Ed.), *Albert Einstein Philosopher-Scientist*. Open Court, La Salle, IL.
- Ellsberg, D., 1956. Theory of the reluctant duelist. *The American Economic Review* 46 (5), 909–923.
- Fehr, E., Fischbacher, U., 2002. Why social preferences matter—the impact of non-selfish motives on competition. *Cooperation and Incentives*. *The Economic Journal* 112, C1–C33.
- Forsythe, R., Horowitz, J., Savin, N., Sefton, M., 1994. Fairness in simple bargaining experiments. *Games and Economic Behavior* 6, 347–369.
- Fouraker, L., Siegel, S., 1963. *Bargaining Behavior*. McGraw-Hill, New York.
- Friedman, D., 1984. On the efficiency of double auction markets. *American Economic Review* 74, 60–72.
- Gjerstad, S., 2007. Price Dynamics in an Exchange Economy. Purdue University, Economics Department, November 19, unpublished paper.
- Harrison, G., McCabe, K., 1992. Testing noncooperative bargaining theory in experiments. In: Issac, R. (Ed.), *Research in Experimental Economics*, vol. 5. JAI Press, Greenwich, CT, pp. 137–169.
- Hayek, F., 1952. *The Sensory Order*. University of Chicago Press, Chicago.
- Herz, B., 2007. Mouselab experiments on the effect of a label framing in an experimental trust game. Masters Thesis. University of Bonn, November.
- Hoffman, E., McCabe, K., Smith, V., 1994. Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior* 7 (3), 346–380.
- Hoffman, E., McCabe, K., Smith, V., 2000. The impact of exchange context on the activation of equity in ultimatum games. *Experimental Economics* 3 (1), 5–9.
- Houser, D., 2003. Classification of types of dynamic decision makers. In: Nadel, L. (Ed.), *Encyclopedia of Cognitive Science*, vol. 1. Nature Publishing Group, London, pp. 1020–1026.
- Houser, D., Xiao, E., 2005. Emotion expression in human punishment behavior. *Proceedings of the National Academy of Sciences of the United States of America* 102 (20), 7398–7401.
- Houser, D., Kurzban, R., 2002. Revisiting kindness and confusion in public good experiments. *American Economic Review* 92, 1062–1069.
- Hume, D., 1739; 1985. *A Treatise of Human Nature*. Penguin, London.
- Johnson, E., Camerer, C., Sen, S., Rymon, T., 2002. Detecting failures of backward induction: monitoring information search in sequential bargaining environments. *Journal of Economic Theory* 104, 16–47.
- Kalai, E., Lehrer, E., 1993. Rational learning leads to nash equilibrium. *Econometrica* 61 (5), 1019–1045.
- Ketcham, J., Smith, V., Williams, A., 1984. A comparison of posted-offer and double-auction pricing institutions. *Review of Economic Studies* 51, 595–614.
- Krueger, F., McCabe, K., Moll, J., Kriegeskorte, N., Zahn, R., Strenziok, M., Heinecke, A., Grafman, J., 2007. Neural correlates of trust. *Proceedings National Academy of Sciences* 104, 20084–20089.
- Lei, V., Noussair, C., Plott, C., 2001. Nonspeculative bubbles in experimental asset markets: lack of common knowledge of rationality vs actual irrationality. *Econometrica* 69, 831–859.
- McCabe, K., Rassenti, S., Smith, V., 1996. Game theory and reciprocity in some extensive form experimental games. *Proceedings of National Academy of Sciences* 93, 13421–13428.
- McCabe, K., Rassenti, S., Smith, V., 1998. Reciprocity, trust and payoff privacy in extensive form bargaining. *Games and Economic Behavior* 24 (1–2), 10–24.
- McCabe, K., Smith, V., LePore, M., 2000. Intentionality detection and ‘mindreading’: why does game form matter? *Proceedings of the National Academy of Sciences* 97, 4404–4409.
- McCabe, K., Houser, D., Ryan, L., Smith, V., Trouard, T., 2001. A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences* 98, 11832–11835.
- Noussair, C., Tucker, S., 2006. Futures markets and bubble formation in experimental asset markets. *Pacific Economic Review* 11 (2), 167–184.
- Oxoby, R., Spraggon, J., 2008. Mine and yours: property rights in dictator games. *Journal of Economic Behavior and Organization* 66, 703–713.
- Plott, C., 2001. Equilibrium, equilibration information and multiple markets: from basic science to institutional design. In: Paper Presented to the Nobel Symposium, Behavioral and Experimental Economics, December 4–6, Stockholm.
- Polanyi, M., 1962. *Personal Knowledge*. University of Chicago Press, Chicago.
- Porter, D., Smith, V., 1995. Futures contracting and dividend uncertainty in experimental asset markets. *Journal of Business* 68 (4), 509–541.
- Rapoport, A., 1997. Order of play in strategically equivalent games in extensive form. *International Journal of Game Theory* 26, 113–136.
- Reiter, S., 1959; 1981. A dynamic process of exchange. In: Horwich, G., Quirk, J. (Eds.), *Essays in Contemporary Fields of Economics*. Purdue University Press, West Lafayette [3–23. Revision of A Market Adjustment Mechanism. Institute Paper No. 1, School of Industrial Management, Purdue University, 1959].
- Rigdon, M., McCabe, K., Smith, V., 2007. Sustaining cooperation in trust games. *Economic Journal* 117, 991–1007.
- Schelling, T., 1960. *The Strategy of Conflict*. Harvard University Press, Cambridge.
- Schmitt, P., 2004. On perceptions of fairness: the role of valuations, outside options, and information in ultimatum bargaining games. *Experimental Economics* 7, 49–73.
- Schotter, A., Wiegelt, K., Wilson, C., 1994. A laboratory investigation of multiperson rationality and presentation effects. *Games and Economic Behavior* 6, 445–468.
- Schwartz, T., Ang, J.S., 1989. Speculative bubbles in the asset market: an experimental study. In: Paper Presented at the American Finance Association meeting, December, Atlanta.
- Siegel, S., Fouraker, L., 1960. *Bargaining and Group Decision Making*. McGraw-Hill, New York.
- Smith, A., 1759; 1982. The theory of moral sentiments. In: Raaphaet, D., Mactie, A. (Eds.), *Liberty Fund*, Indianapolis.
- Smith, V., 1991. *Papers in Experimental Economics*. Cambridge University Press, Cambridge.
- Smith, V., 2008. *Rationality in Economics Constructivist and Ecological Forms*. Cambridge University Press, Cambridge.
- Smith, V., Williams, A., Bratton, K., Vannoni, M., 1982. Competitive market institutions: double auctions vs sealed bid-offer auctions. *American Economic Review* 72, 58–77.
- Smith, V., Suchanek, G., Williams, A., 1988. Bubbles, crashes and endogenous expectations in experimental spot asset markets. *Econometrica* 56, 1119–1151.

- Sobel, J., 2005. Interdependent preferences and reciprocity. *Journal of Economic Literature* 42, 392–436.
- Wierzbicka, A., 2006. *English: Meaning and Culture*. Oxford University Press, Oxford.
- Williams, A., Smith, V., 1986. Simultaneous Trading in Two Competitive Markets. Indiana University, Department of Economics, Bloomington, Indiana, unpublished paper.
- Williams, A., Smith, V., Ledyard, J., Gjerstad, S., 2000. Concurrent trading in two experimental markets with demand interdependence. *Journal of Economic Theory* 16, 511–528.
- Wilson, B., 2008. *Contra Fairness*. Economic Science Institute, Chapman University, unpublished paper.
- Wilson, R., 1987. On Equilibria of bid-ask markets. In: Feiwel, G. (Ed.), *Arrow and the Ascent of Modern Economic Theory*. NYU Press, New York, pp. 375–414.