PAPERS on Economics & Evolution

0609

Democratic Epistemics: An Experiment on How to Improve Forensic Science

by

Roger G. Koppl

The *Papers on Economics and Evolution* are edited by the Evolutionary Economics Group, MPI Jena. For editorial correspondence, please contact: evopapers@econ.mpg.de

ISSN 1430-4716

Max Planck Institute of Economics
Evolutionary Economics Group
Kahlaische Str. 10
07745 Jena, Germany
Fax: ++49-3641-686868

Democratic Epistemics: An Experiment on How to Improve Forensic Science*

Roger G. Koppl
Institute for Forensic Science Administration
Fairleigh Dickinson University
Madison, NJ 07940, USA
koppl@fdu.edu

ABSTRACT

In "monopoly epistemics," one privileged actor is asked to identify the truth. In "democratic epistemics," several independent parties are asked. In an experiment contrasting them, democratic epistemics reduced the systemic error rate by two-thirds, supporting the claim that replacing monopoly epistemics with democratic epistemics would reduce error rates in forensic science and other areas. It also suggests first, the potential of "epistemic systems design," which employs the techniques of economic systems design to address issues of veracity, rather than efficiency, and second, the value of "experimental epistemology," which employs experimental techniques in the study of science. Research of the sort described here puts evolutionary epistemology into practice by seeking to find the proper design principles for error-correcting social institutions.

July 2006

* Many persons helped me to produce this paper and I thank them all. I got helpful comments from, Alvin Goldman, Kathleen Haspel, Dan Houser, Jeffrey Hsu, Richard Langlois, Heather Loenser, Maria Minniti, Paul Rubin, John Schiemann, Laurie Treleven, and Xia "Lily" Wang. Michael Adams, Bill Schwartz, and Maria Minniti provided lots of needed encouragement. The study was funded jointly by two colleges of Fairleigh Dickinson University, namely, the Silberman College of Business and the Petrocelli College of Continuing Studies. For that funding, I thank David Steele, former Silberman dean, and Kenneth Vehrkens, the Petrocelli dean. Dan Houser provided vital organization and spiritual support at an early stage in the project. The experiments were run in the Pennsylvania Laboratory for Experimental Evolutionary Psychology. I thank the PLEEP Director, Robert Kurzban, for allowing me to use his lab, for vital help with the experimental design, for teaching me how to run an experiment, and for suggesting the term "democratic epistemics." If the present study has any merit, it is because of Rob's help. I thank Kelly Asao for exceptionally good and energetic work as a research assistant. None of these people can be blamed for any errors or infirmities contained herein, except Rob Kurzban, who should have let me list him as a co-author.

Introduction and Summary

Economists typically employ the normative criterion of efficiency. They have given less attention to truth. In many contexts, however, veracity may be more important than efficiency. Police forensics, pure science, espionage, auditing, clinical medical testing, drug screening, judicial proceedings, private arbitrage, government investigations, and academic economics are all social processes that generate, in one way or another, judgments of truth and falsity. In such processes it is worth knowing which institutional arrangements will tend to produce true judgments and which will tend to produce falsehood and error. Research addressed to this issue puts the social epistemology of Alvin Goldman (1999) into practice by seeking to find the proper design principles for error-correcting social institutions. The same could be said of evolutionary epistemology.

Radnitzsky and Bartley (1987) is the classic statement of evolutionary epistemology. The leading figure in this field is Karl Popper. Other important names associated with it include W. W. Bartley III, Donald T. Campbell, Antony Flew, F. A. Hayek, Konrad Lorenz, and Gerard Radnitzsky.

Bradie and Harms (2004) gives a useful overview with an updated bibliography. As Bradie and Harms explain evolutionary epistemology is "evolutionary" in two senses. First, natural selection "is the generator and maintainer of the reliability of our senses and cognitive mechanisms." Second, "trial and error learning and the evolution of scientific theories are construed as selection processes." These two insights of evolutionary epistemology help to explain how our knowledge could be fallible and grossly imperfect while nevertheless providing serviceable guidance to human action.

Bartley (1987) contrasts sociology of knowledge and evolutionary epistemology. Both recognize that social interests influence science. Typically, however, proponents of the sociology of knowledge uncritically assume that "social distortion cannot be corrected, modulated, compensated for" in something like the way we attempt to filter the noise out of a radio signal. Evolutionary epistemology, by contrast, recognizes the importance of "a new social account of knowledge: an account dealing with the question of how to optimize the rules and practices of the community so as to diminish distortion" (p. 447). Bartley identifies a mechanism design problem. Unfortunately, there seems to be little or no work on this problem in the tradition of evolutionary epistemology. This paper takes Bartley's project seriously by reporting on the first results of an experimental program addressing the problem of designing error-minimizing mechanisms in forensic science and other areas.

Like Goldman's social epistemology, the research project described here extends beyond the realm of pure science, to which evolutionary epistemologists give near exclusive attention, to consider all social processes from the epistemic point of view. In Koppl (2005a) I paraphrase Goldman (1999) by defining "epistemic systems" as "social processes viewed from the perspective of their tendency to help or frustrate the production of truth" (91). In that paper I describe some game-theory models of epistemic systems that apply widely. In this paper I report on an experiment that applies the theory of epistemic systems to an important issue in forensic science, namely the possibility of reducing error rates through a reorganization of the network structure of forensic science.

The current institutional structure of forensics typically grants to an individual lab a kind of monopoly on the analysis of any evidence sent to it: Once a given lab has received and analyzed a body of evidence, it is unlikely that the evidence will be examined by any other lab. In Koppl (2005b) I suggest breaking this monopoly by periodically and randomly sending evidence to more than one lab for analysis. The experiment discussed below provides an initial test of the likely consequences of implementing such "democratic epistemics."

As I will explain more carefully below, subjects were assigned to the role of either a Sender – an analog for a forensic scientist – or a Receiver – an analog of a decision-maker such as a judge or jury. The Senders were shown a figure and asked to provide a "report" to Receivers. The payoff structure gives Senders a bias in favor of reports that may be inaccurate. The experiment compares what happens when you have one Sender per Receiver to the results of having three Senders per Receiver. The first case represents the current practice of giving forensic labs a kind of monopoly on the analysis of any evidence they receive. The second case creates a kind of democracy in decision making. The relevant systemic error rate fell from 75% to 25% when moving from the monopoly regime of one Sender per Receiver to the democratic regime of three Senders per Receiver.

The experimental result just previewed supports the claim that replacing monopoly epistemics with democratic epistemics would reduce error rates in forensic science and other areas. It also suggests the potential value of developing "epistemic systems design," which employs the techniques of economic systems design (Smith 2003) to address issues of veracity, rather than efficiency. It may also suggest the value of employing experimental techniques in the study of science itself, thus creating a field that might be called "experimental science studies" or, perhaps, "experimental epistemology."

I begin with a discussion of forensic science and its institutional structure. I then review the theory of epistemic systems, apply the theory to forensic science, and draw out some testable implications of the analysis. Much of this section uses language borrowed form Koppl 2005a. Next, I describe the experimental design and report my results. I close with a discussion (the last paragraph of which draws on the conclusion of Koppl 2005a) that includes comments on forensic science and epistemic systems design.

I. Forensic Science as the Problem Context

The proper function of forensic science is to extract the truth. "As it is practiced today," note Saks et al. (2001), "forensic science does not extract the truth reliably. Forensic science expert evidence that is erroneous (that is, honest mistakes) and fraudulent (deliberate misrepresentation) has been found to be one of the major causes, and perhaps the leading cause, of erroneous convictions of innocent persons" (Saks et al. 2001, p. 28). Two cases of mistaken fingerprint identification illustrate problems that seem to be characteristic of forensic science today. Cole reports on these two and another 20 cases of misattribution in the US and UK (2005, pp. 1001-1016).

In 2004, the FBI arrested Brandon Mayfield as a material witness in the Madrid train bombing of March 2004. He had been identified as the source of a latent print found on a bag of detonators found near the crime scene. After assigning three of its top fingerprint examiners to the case, the FBI declared a "100 percent match" to Mayfield (Kershaw 2004). The Spanish National Police objected to this match, however, and declared for Ouhnane Daoud. The Spanish authorities seem to have been correct and the FBI withdrew its identification and released Mayfield (Office of the Inspector General 2006).

Marion Ross was murdered in her home in Kilmarnock, Scotland in 1997. Detective constable Shirley McKie, among others, was assigned to the case. Her prints were taken "for elimination purposes" (McBeth 2004). Top experts at the Scottish Fingerprint Service (SFS), a branch of the Scottish Criminal Records Office (SCRO), declared match of a latent print left at the scene to McKie, whose was then presumed to have entered the crime scene *ex post*. McKie denied having entered the house and insisted that the latent print could not be hers. She testified to this effect at the trial of David Asbury, who was convicted of the Ross murder. Recalcitrant in her position, McKie was arrested for perjury and tried. She was found innocent when the Scottish fingerprint evidence was challenged by outside experts testifying at her trial. Later she received an apology from the Scottish Justice Minister and, eventually, £750,000 from the Scottish Executive in an out of court settlement. As of February 2006, the "management at SCRO 'acknowledges and accepts' that a mistake was made, but the six experts who made the identification do not" (McDougall 2006).

After serving 5 years in prison, David Asbury was released and his conviction voided because the SCRO fingerprint evidence used against him was found to be unreliable (Innocent 2002). Either a killer has been released or an innocent man was jailed for five years for a crime he did not commit. *Tertium non datur.*

SCRO made a presentation of the McKie evidence to a group including Arie Zeelenberg, the head of the National Fingerprint Service of the Netherlands, and Torger Rudrud, the assistant chief of police in Norway. The presentation claimed 45 points of similarity between the latent and known prints. SCRO procedures require only 16 points for a match. In a report on this presentation and evidence, Rudrud and Zeelenberg "criticise the SCRO experts' quality of analysis and accuse them of 'ignoring' points which show the prints could not have been made by the same person, and wrongly 'promoting' points to make it appear that they matched." They say, "Without valid reason in this presentation we see that 'noise' is upgraded to similarities if convenient." They were fitting the noise. "Point seven and point eight are marked by SCRO as similar characteristics. It is clear at first sight that they are located outside the contour of the latent and beyond the outskirts of it ... it is most likely the pattern of the wood we are looking at. In fact, the same applies to point 11" (McDougall 2006). It is a serious error to mistake wood grain for a ridge impression.

There is no universally agreed number of points for a match. Different jurisdictions within the US require different numbers and some have no particular number, relying only on the judgment and discretion of the individual examiner.

As I note in Koppl (2005b), DNA exonerations in the US and other events have induced an extensive literature on the limited reliability of forensic testimony in court. Examples include, Giannelli 1997, Illinois 2002, Jonakait 1991, Kaufman 1998, Kelly and Wearne 1998, McDougall 2006, McRoberts *et al*. 2004, Moenssens 1993, Office of the Inspector General 1997, 2004, 2006, Risinger *et al*. 2002, Saks 1998, Saks and Koehler 2005, and Saks *et al*. 2001. The appendix to Koppl 2005b reviews evidence that forensic testimony is not reliable. Proficiency tests reported in Peterson *et al*. (1995a, 1995b), for example, seem to suggest a rate of false positives for fingerprints of at least 2%. One test from the period after the study of Peterson and his co-authors produced a 20% rate of false positives (Grieve 1996). "Whether the rate of false identifications for fingerprints is closer to 2% or 20%," I have remarked, "it is (literally) infinitely higher than zero rate sometimes claimed by fingerprint examiners" (2005b, p. 278). I estimate that a 2% rate of false positives would correspond to about 1,400 false convictions per year in the US (2005b, p. 278). Whatever the precision of this estimate, the problem seems serious enough to justify public attention.

In Koppl (2005b) I review several more types of evidence on the reliability of forensics, including cases in which botched DNA analyses have led false convictions (Raziq and Werner 2004) and, in at least one documented case, the false exoneration of a child rapist (Teichroeb 2004). Here, I will mention only one more type of evidence, namely, the documented existence of rogue forensic scientists such as Fred Salem Zain, Ralph R. Erdman, Loise Robbins, Michael West, and Thomas N. Curran (Kelly and Wearne 1998, Giannelli 1997). In this group, Fred Zain's 15 years of "rampant falsification" (Kelly and Wearne 1998, p. 13) probably represents the longest career of any forensic scientist identified as a rogue. Ralph Erdman, however, may be the most flamboyant example. He "faked more than 100 autopsies on unexamined bodies, and falsified dozens of toxicology and blood reports. Dozens of other autopsies were botched. In one case he lost a head" (Kelly andWearne1998, p. 13).

It seems that all of the forensic sciences, including DNA typing, are less reliable than we might have imagined. Indeed, "The formal and anecdotal evidence available to us point to an ongoing crisis of forensic science" (Koppl 2005b, 281). Many scholars, journalists, activists, and others have recognized the need to improve forensic science. No consensus exists, however, on the best strategy for improvement. In Koppl (2005b) I identify eight remediable features of the current institutional structure of forensic science, each of which reduces reliability. Corresponding to each flaw (as I see it) in the institutional structure is my suggestion for amending current institutions. Table 1 summarizes the argument of Koppl (2005b).

Place Table 1 about here

The most important of my eight proposals for change is "rivalrous redundancy," which would eliminate the monopoly position now enjoyed by most forensic laboratories. Rivalrous redundancy would produce several competing forensic labs in any jurisdiction. Subject to the constraints of feasibility, some evidence should be chosen at random for duplicate testing at other labs. The same DNA evidence, for example, might be sent to more than one lab for analysis. The forensic worker need not know whether the evidence is examined by another lab. He will know that there could be another lab, and sometimes is.

Ironically, rivalrous redundancy would reduce the costs of forensic testing (Koppl 2005b, pp. 274-275). The key point is that no increase in infrastructure is required. In the current system, if I may ignore some inessential complications, each jurisdiction is served by one lab and each lab serves one jurisdiction. Under rivalrous redundancy, by contrast, a given lab might serve six jurisdictions, while each of those six jurisdictions is served by six separate labs. The reorganization does not require any new facilities. There would be some increase in the number of tests performed, but the number of redundant tests would be a fraction of the total. If reliability increased, the slight cost of running some redundant tests would be fully compensated by reductions in other costs, as explained in Koppl 2005b. Even in the direct costs of forensics would likely fall, especially if the fee system proposed by Saks *et al*. (2001) is adopted.

There is a crisis in forensic science. Rivalrous redundancy is important component of a policy to improve forensics. In the next section, I examine rivalrous redundancy using the theory of epistemic systems. I begin the section with a review of epistemic systems together with some comments on the relevant literature. I then make the application to forensics and consider the strategic implications of rivalrous redundancy. This discussion will complete the context of my experiment, to which I will then turn.

II. Epistemic Systems as the Theoretical Framework

Epistemic systems, as I have indicated above, are "social processes viewed from the perspective of their tendency to help or frustrate the production of truth. In "Epistemic Systems" (Koppl 2005a) I presented some game-theory models of epistemic systems that may be useful in "comparative institutional epistemics," which is the study of the epistemic consequences of alternative institutions. All these models have one or more "Senders" who search a "message space," and deliver a message to one or more "Receivers." If there are two or more Senders, they are in a position of strategic interdependence with respect to the messages they send. One may often be interested in knowing which message vectors are Nash equilibria. Thus, a forensic scientist searches a message space with the messages "match," "no match," and "inconclusive." He chooses "match," say, and sends the message by testifying in open court. After receiving his message or messages, the Receiver produces a "judgment." The jury, in our forensics example, decides whether the fingerprint left at the crime scene belongs to the suspect. This particular "judgment" is an input into the jury's larger deliberations. I am mostly interested in the truth value of these judgments, such as the jury's judgment that the print came from the suspect. Some arrangements induce more truthful judgments than others. For example, an arbiter for whom one party is a repeat customer is likely to be biased. An independent arbiter is more likely to give a truthful account of who is at fault in a dispute.

Somewhat more formally, an ***epistemic system*** is a set of ***Senders***, $\mathscr{S}$, a set of ***Receivers***, $\mathscr{R}$, and a set of ***messages***, $\mathscr{M}$. The Senders may have a probability distribution over messages, showing the subjective probability that each message is true. The Senders send messages to the Receivers, who somehow select one message from the message set and declare it "true." This is the ***judgment***, of the Receiver(s). For example, we might have a system with one Sender and one Receiver and in which the Receiver always nominates the message he gets from the Sender. Typically, the Senders are experts advising the Receivers. The Receivers may or may not be experts. In science, the set of Receivers is in principle identical to the set of Senders. In expert testimony, the set of Senders is disjoint from the set of Receivers. In these models we generally assume that Senders have a limited interest in the truth. They may disregard the truth or allow veracity to compete with other values such as fame and fortune.

By assuming that Senders have a limited interest in the truth, I am following the example of David Hume. "Political writers," wrote Hume, "have established it as a maxim, that, in contriving any system of government, and fixing the several checks and controuls of the constitution, every man ought to be supposed a knave, and to have no other end, in all his actions, than private interest. By this interest we must govern him, and, by means of it, make him, notwithstanding his insatiable avarice and ambition, co-operate to public good" (Hume 1777, Part I, Essay VI, "Of the Independency of Parliament," in paragraph I.VI.1). Hume here expresses the idea behind "checks and balances." In considering ways to amend our institutions so as to improve epistemic performance, it seems prudent to estimate the relative merits of different arrangements under pessimistic assumptions about human motives. Although not all of us are knaves, virtue is a scarce resource. We should therefore consider such amendments to our institutions as economize on human virtue.

The models given below are related to a large body of past work. It is difficult even to construct a reasonably complete list of related fields, which include information economics, sociology of knowledge, evolutionary epistemology, judgment aggregation, science studies, philosophy of science, epistemology, economics of science, informatics, cognitive science, and social psychology. The function of an economic system is to produce utility, not veracity. But economic decisions depend on judgments of truth and falsity on many topics including the likely returns from different investments. Thus, the economist F. A. Hayek (1935) and others argued that rational economic calculation is not possible under socialism, whereas J. M. Keynes (1936) argued (in effect) that rational economic calculation is not possible under capitalism. G.L.S. Shackle (1972) first brought the term "epistemics" to economic theory, defining "epistemics" as "the theory of thoughts" (1972, p. *xx*). In philosophy, Alvin Goldman (1978, 1999, 2001) used the terms "epistemics" and "social epistemology" to refer to studies of the sort I am attempting here. Recall that my definition of epistemic systems is a close paraphrase of Goldman's definition of social epistemology (Goldman 1999). Goldman's veritistic social epistemology, in which some claims to truth are better than other, contrasts sharply with theories pretending to be neutral on truth, for example that of the sociologist David Bloor (Goldman 2001). Bloor (1978) carries on the tradition of the "sociology of knowledge," which is generally traced back to Karl Mannheim (1936[1985]). The Mannheim tradition contrasts with that of Robert Merton who recognized that the reliability of science was a product of the social institutions of science, not the merit of individual scientists (Merton 1937). Psychology, especially social psychology, addresses many of the issues raised by research into epistemic systems. The leading study in this area is probably the famous conformity study of Solomon Asch (1951), who showed that most people will (in a certain laboratory setting at least) adjust their opinions away from the obvious truth in order to conform to majority opinion.

In economics and game theory, the models of this paper might be most closely related to two sets of results. First, there is the recent literature on "judgment aggregation." Examples include Kornhauser and Sager 1986, Kornhauser 1992, and List and Pettit 2002. Second, there is the large body of work on asymmetric information, signaling games, and sender-receiver games. Examples include Akerlof (1970), Spence (1973), Blume et al. 1998, Green and Stokey 1980, and Crawford and Sobel (1982). In Koppl 2005a I explain some points of difference between my models of epistemic systems and earlier game theory treatments of similar topics (93).

Finally, two groups of scholars deserve mention for work that not only relates to epistemic systems as here defined, but also employ experimental techniques. Larry Miller (1987) conducted a pioneering study. It applied experimental methods to the problem of reducing error rates in forensic science through a reorganization of forensic work. He had a group of trained forensic scientists do hair analysis under two conditions. In the first condition, the hair from the crime scene and one suspect hair were compared. In the second condition, the hair was compared to several suspect hairs. In other words, the second condition created an "evidence lineup." The evidence lineup reduced the error rate by about 90%. The current study builds on Miller's neglected work by applying the techniques of experimental economics to similar questions regarding the organization of forensic work.

More recently, Itiel Dror and David Charlton, together with their co-authors have been doing important experimental work relating to fingerprint experts. Dror *et al*. (2005) show that when fingerprint examiners are given emotionally charged contextual information on their cases, they are more likely to find a match where none exists. Dror *et al*. (2006) used suggestion to induce fingerprint experts to unwittingly reverse earlier and correct judgments of a match. The experimental design induced the experts to believe they were reviewing, not their own earlier cases, but the Brandon Mayfield misidentification discussed above. Dror and his colleagues are using the concepts and techniques of cognitive psychology, which give less attention to social structure than does the current study. Although this is a legitimate distinction, it is one of emphasis. Cognition occurs within a social context; the consequences of a given social structure depend on the agents' cognition.

These preliminary comments should have prepared us for my last and most formal (least informal) explanation of epistemic systems. An epistemic system is an ordered triple, $<\mathcal{S}, \mathcal{R}, \mathcal{M}>$. The set $\mathcal{S}$ (of Senders) is indexed by $i \in I^*$. A member of $\mathcal{S}$ is represented by $s_i \in \mathcal{S}$. The set $\mathcal{R}$ (of Receivers) is indexed by $j \in J^*$. A member of $\mathcal{R}$ is represented by $r_j \in \mathcal{R}$. The set $\mathcal{M}$ (of messages) is indexed by $h \in H^*$. A member of $\mathcal{M}$ is represented by $m_h \in \mathcal{M}$. Typically, there will be a finite number of Senders and Receivers. It may often be convenient to assume the message space is infinite. If I^* has a largest element, denote that element I. Define J and H similarly.

Senders and Receivers have value functions over messages. These might be utility functions or payoff functions. For Receivers, $V_{r_j} = f_{r_j}(m_{s_1}, m_{s_2}, ..., m_{s_I})$. For Senders, $V_{s_i} = f_{s_i}(V_{r_1}, V_{r_2}, ..., V_{r_J}; m_{s_1}, m_{s_2}, ..., m_{s_I})$. For example, suppose we have a law suit in which the plaintiff's attorney has hired an expert witness to estimate the money value of the claimed harm. Here, the Sender is an expert witness hired by the Receiver, the plaintiff's lawyer. Up to some limit of plausibility, the Receiver prefers higher estimates to lower estimates. This may induce a similar preference in the Sender, who wants the plaintiff's lawyer to become a repeat customer. If there are two or more Senders, they are in a position of strategic interdependence with respect to the messages they send. As I have indicated already, one may often be interested in knowing which message vectors are Nash equilibria.

Models of epistemic systems will probably be most useful when one must assume that Senders, or Receivers, or both place little or no value on truth. We want a world in which people prefer to send and receive true messages. In many contexts, however, people prefer to send or receive false messages. Most criminals prefer to deny their crimes. Many scientists prefer their own theories to competing theories that come closer to the truth. In such contexts, models of epistemic systems may help us to amend our social institutions so as to produce more truth.

A simple example

Let $\mathcal{M}=\{0,1\}$. The Receiver always nominates the sent message flawlessly. The message is sent over a noiseless channel. Figure 1 illustrates.

Place Figure 1 about here

The dashed arrow represents the Sender choosing from the message set. The solid arrows represent the transmission to the Receiver and the selection of a message by the Receiver.

Notice that Figure 1 looks like it came from Shannon's information theory. This is no coincidence. The basic idea of epistemic systems is that we eliminate from information theory the exogenously given distribution over messages and replace it with search of the message space and strategic choice of messages.

In the context of Figure 1, assume the Receiver's value function is

$$U(x) = \begin{cases} 1 \text{ if } x = 0 \\ 0 \text{ if } x = 1 \end{cases}$$

In this case the Receiver is interested in the content of the message (whether it is 1 or 0) but not in its truth. Assume further that the Sender estimates the probability that 1 is true to be 0.75. His subjective probability that 0 is true is, therefore, 0.25. Assume the Sender's value function is

$$V(x) = P(x \text{ is true}) E[U(x)],$$

where $P(\bullet)$ denotes probability and $E(\bullet)$ denotes expected value. E[$U(x)$] denotes the Sender's expectation of the Receiver's value function, $U(x)$. The Sender, in this example, values the truth, but also wishes to please the Receiver. Assume, finally, the Sender knows $U(x)$. Then E[$U(x)$]=U(x), $V(1) = 0$, and $V(0) = 0.25$. In this case, the Sender sends 0 and the Receiver's judgment is 0 even though the Sender thinks 1 is three times more likely. This model is purposefully quite simple. In a rough and ready sort of way, however, it might apply to many command and control situations. If the Receiver is in a position of dominance over the Sender, the Sender may craft his message to please the Receiver rather than reveal the truth. It is a commonplace that dictators and Hollywood stars receive nothing but praise and celebration even under desperate circumstances.

A small variation in the simple model of this section reveals the importance of the principle of "information hiding." Borrowing a term from computer science (Parnas 1972), Richard Langlois (2002) has introduced the concept of "information hiding" to economics. Information hiding is "enforced ignorance among the parts" of a system (Koppl and Langlois 2001, p. 294). For example, information shared with an attorney is hidden from the jury of a criminal trial.

Imagine that the Sender in the previous model does not know the Receiver's utility function, $U(x)$. The Sender wants to please the Receiver, but does not know which message is preferred. In this situation, E[$U(x)$] is 0.5 regardless of x. Under these assumptions, $V(1) = 0.375$ and $V(0) = 0.125$. The Sender sends 1 and the Receiver's judgment is 1.

Recall that the Sender in this model estimated the probability of 1 being true to be 0.75. Imagine this is true on average over time. In other words, the relative frequency of 1 being true is 0.75. Then information hiding raises the epistemic efficiency of this system from 0.25 to 0.75. This result illustrates the great importance of information hiding in epistemic systems. A well-designed epistemic system will typically have a modular structure with information hiding. It may seem counter-intuitive to say that we wish to hide information in order to generate better judgments about the truth. But this is done all the time. In science we have double-blind testing in which the scientist arranges to have certain information hidden from him. The law courts hide information from the jury.

In Koppl 2005a I apply this simple framework to analyze forensics. I also develop the framework in directions that I will not address here. Thus, I am ignoring the important distinctions between Delphic and discursive systems and between open and closed epistemic systems.

Application to forensics

Bias seems to be an important source of error in police forensics (Koppl 2005b). There are two ways to handle bias: eliminate it or compensate for it. I adopt the second strategy in the experiment discussed below.

Under current institutions in the common-law countries, crime labs are typically organized under the police. For example, the Scottish Fingerprint Service (SFS) is a branch of the Scottish Criminal Records Office (SCRO) and the leading crime lab in the US is the FBI. This mode of organization creates a pro-police bias (Risinger *et al*. 2001). The bias need not be conscious. Conscientious employees of the police may wish to please their superiors and do a good job. But these very desires, laudable in themselves, may induce an unconscious bias in favor of the police theory in each case. The police generally ask for a test when they believe a match will identify their suspect. Thus, forensic workers tend to have a bias in favor of finding matches. For this analysis, therefore, I will assume that forensic labs prefer to send the message 1, "match." This bias produces different results depending on whether forensic labs have monopoly power.

Figure 1 represents the current situation in the United States and elsewhere. The message space is given by $\mathscr{M}$={0,1}, where 0 represents "no match" and represents "match." (As far as I can tell this simplification of the message space does not matter for the implications I will draw. As we will see, the experimental design uses a three-element message space and differs in other inessential particulars as well.) The Sender is a forensics lab and the Receiver is a judge or jury. Figure 1 reflects the current monopoly situation in forensics whereby, as I have said, evidence going to any one lab and will probably not be examined by any other lab.

In the simple situation of Figure 1, the Sender has a dominant strategy, namely 1, "match." The lab always sends the message 1 and the crime lab adds no new information to the system. It is not reducing error rate in the system. In this case, forensic science does not increase the epistemic reliability of the criminal justice system. (Koppl 2005c explains why a pro-police bias may increase the system-level error rate.)

Notice that I am counting all false reports as "errors" whether they are lies or honest mistakes. There are no bright lines as we pass from willful falsehood, to conscious bias, to unconscious bias, to perfect objectivity. Thus, it may be fruitful to sometimes lump all false reports together without attempting to discriminate between the fraudulent and non-fraudulent cases.

Now assume that three independent forensic scientists examine the evidence as illustrated by Figure 2. This redundancy will not help if it is "mere redundancy." As explained in Koppl (2005b) mere redundancy means that there is more than one Sender examining evidence, but no specific incentives for one Sender to discover the errors of the other(s). The redundant bits of the system just lie side by side. If each Sender sends the same message, the Receiver's judgment is the common message. In this model, redundancy does not alter the behavior of the Senders. Each of the three Senders in this model sends 1. The epistemic efficiency of the system is low because the system produces the right judgment only when the message "1" happens, coincidentally, to be true. In this case also, forensic science does not increase the epistemic reliability of the criminal justice system.

Place Figure 2 about here

Now imagine that each Sender's payoff depends on the choices of the other Senders and of the Receiver. The Receiver, we will imagine, always accepts the majority opinion among Senders. Each Sender receives a positive payoff only if his selection is nominated by the Receiver. The minority Sender looses; he receives nothing or, perhaps, is fined. This payoff structure creates "rivalrous redundancy."

In this game, each Sender wants to be in the majority and there are two Nash equilibria, (0,0,0) and (1,1,1). Where there are multiple Nash equilibria, standard game theory cannot decide which, if any, will prevail. It is at least possible, however, that the truth may be more salient (Schelling 1960) than other messages. In that case, the epistemic efficiency of this system will be relatively high and forensic science will increase the epistemic efficiency of the criminal justice system.

Increasing the number of Senders from one to three introduces a Nash equilibrium in which honesty is the best policy. Our question is whether that Nash equilibrium is a Schelling point. In other words, is the truth more salient than answers serving a pro-police bias? The experimental design of this study was meant to find out.

III. The Experiment

Objectives

The experiment examines three questions:

1. Everything else equal, will people transmit false information to a decision-maker if they have an incentive to do so?
2. Can a simple institutional mechanism, democratic epistemics, be implemented to reduce or eliminate the distorting effects of the bias in (1)?
3. How does the size of a bias (1) interact with the mechanism in (2)?

Answering question 2 means testing democratic epistemics as an institutional mechanism, tying together the research question with the experimental method. Question 3 addresses robustness.

The experiment gives Senders an incentive to produce biased reports on an observed event. The Receivers guess what event happened after reading the report(s) of one or more Senders. In the baseline, "monopoly," condition, each Receiver gets one report from one Sender. In the "democratic" condition, each Receiver gets three reports from three Senders. In the baseline condition, the Sender's dominant strategy is to lie whenever the truth does not conform to the Sender's bias. In the democratic condition, each Sender has a greater incentive to be in the majority than to act on his or her bias. Thus, the Senders have a coordination problem. In the democratic condition, truth telling is one of many Nash equilibria. The experimental study of Bacharach and Bernasconi (1997) supports the conjecture that Senders may choose truth telling because it is the most salient strategy in the sense of Schelling (1960). They provide evidence on salience, but do not consider whether truthfulness creates salience. The chance that Senders will choose truth telling depends on several variables such as the size of the bias and number of Senders per Receiver. These dimensions are explored in different experimental treatments as described below.

Experimental Design

A new task, dubbed "The Science Game" is introduced to mirror the essential structure of forensic science. The appendix contains two examples of the experimental instructions used in the study. In this game, the Senders represent forensic labs and the Receivers represent a judge or jury. There are N Senders and one Receiver, each describable, in the behavioral game theory tradition, as having a set of information, a set of decisions, and payoffs associated with all possible outcomes. In each round, Nature (i.e., a randomizing device) chooses a "correct Object" from a set of 3 objects or shapes: circle, triangle, square. This information is known to Senders but not the Receivers. Senders send a message consisting of one of the objects to Receivers, who then must submit a decision about what they believe the correct Object to be. (In the context of forensic science, this message corresponds to a lab indicating "match," "no match," or "inconclusive.") The Receivers' information consists only of the communications they will receive from Senders, their own payoffs for correct versus incorrect guesses, and limited information about Senders' payoffs (see below).

Senders get a payment, P_C, if the Receiver indicates the correct Object. Senders, however, also get an additional payment depending on which object the

Receiver chooses *independent of the correct Object*. This payment corresponds to the bias on the part of the lab personnel. This payoff information is private (i.e., unknown to the Receivers).

All experimental sessions run in the Penn Laboratory for Experimental Evolutionary Psychology (PLEEP) using pen and paper data collection. The PLEEP lab has 16 subject stations with partitions, which prevent players from seeing one another or each others' actions or materials. Anonymity was preserved by assigning participants identification numbers at the start of the experiment. Participants were paid by taking an envelope labeled with their unique numeric identifier filled with their earnings. Through "experiments@Penn," the subject pool for these experiments was the general University of Pennsylvania community, including staff and students.

Except for age, there are no obvious, radical differences between the demographics of the participants in this study and the demographics of forensic scientists in the US. The demographics of our participants may not matter, however, for the legitimacy of our results as applied to forensic science or any other lab-based social process. It is a cliché that experimental studies tend to use college students as participants. In an important recent study, Henrich *et al*. (2005) explicitly addressed the possible limits of this practice. They examined the role of culture in generating certain results from experiments with human subjects. They found that a substantial portion of variation in the behavior of participants in ultimatum, public goods, and dictator games was explained by two factors, namely, the degree of market integration and the level of the payoffs to cooperation in everyday life. Significantly for the value of our study, they found that "individual-level economic and demographic variables do not consistently explain behavior within or across groups" (pp. 797-798). Our experimental participants live in the same sort of society as participants in modern lab-based social processes in the US and many other countries. Thus, although the context of the study of Henrich *et al*. was different than the current study, it suggests that the subject pool may give us an appropriate representation of the behavior of forensic scientists.

Experimental Treatment I: High & Low Bias

In the baseline condition, there was one Sender and one Receiver. Senders received a payment of $3 for leading the Receiver to the correct Object (a "hit") if the Sender sends the correct Object and the Receiver chooses this Object. Senders may have received an additional payment of either $1.00 or $5.00 for a “supplemental” object. The prospect of this payment induced a bias in the preferences of Senders. The Sender received this payment or not as a function of the object chosen by the Receiver independent of what the correct Object is. Thus, if the correct Object is square, the supplemental object is circle, and the Receiver chooses circle, the Sender gets $5. The Sender thus has an incentive to report the supplemental object (circle in our example) regardless of what the correct Object is. This incentive is the Sender’s bias.

The size of the bias is manipulated as a within-subjects treatment. In the “High Bias” condition, the bias is greater than the payoff to the Sender for the Receiver making the correct decision. If the Receiver chooses the correct Object, the Receiver receives $5, otherwise $2.

Senders were informed that 1) Receivers must guess the correct Object based only on the information that the Sender sends 2) Receivers earn more money if they guess the correct Object, and 3) that this is the *only* means by which Receivers earn money. Senders are not given quantitative information about Receiver’s payoffs. This minimizes problems associated with other-regarding preferences observed in games such as the Dictator Game (Hoffman *et al*. 1996) because Senders do not know how much they can benefit the Receiver by sending correct information.

In each experimental session, 10 rounds of this game were played. The correct Object and supplemental shape for each round was determined randomly (before the experiment), and given to Senders at the beginning of each round. Five High and Five Low bias rounds were played, with their order being randomly pre-set. No feedback was given to players regarding the Receiver’s guess about the object in each individual round.

Participants were told that one round of play will be selected randomly, and that the outcome for that round would determine their monetary payoff for the experiment.

Experimental Treatment II: Number of Senders

A second experimental treatment was run with 3 Senders. Everything was identical, including the High and Low Bias treatments, except that there were three Senders per Receiver. Each Sender was be told that the Receiver will have access to the information (shape) sent by each of the three Senders, and then will guess the correct shape based on this information. Crucially, Senders were not informed of the biases of the other Senders, which were not necessarily the same in each round.

Hypotheses

My hypotheses are that

1) the biases of the baseline condition will induce inaccurate reporting on observed events,
2) larger biases will produce larger number of inaccurate reports, and
3) democratic epistemics will reduce or eliminate these inaccuracies.

Let us review each of the hypotheses:

First, “the biases of the baseline condition will induce inaccurate reporting on observed events” and “larger biases will produce larger number of inaccurate reports.” These hypotheses are principally a matter of confirming the legitimacy of the experimental method. If either hypothesis were falsified by the experiment, it would probably be rash to conclude that people do not respond to incentives. It would be more plausible to question the experimental design or execution.

Second, “democratic epistemics will reduce or eliminate these inaccuracies.” If this hypothesis is sustained, then it is likely that decision makers can understand and respond to the incentives of democratic epistemics in lab-based social processes. In this case we will have evidence that the epistemic performance of the system depends as much on its institutional structure as the skill or integrity of the individuals in the system. Such a result would tend to strengthen the view that institutional change may be an effective method of reducing error rates in forensic science and other areas.

Results

With the help of PLEEP personnel I ran 54 subjects, all recruited from the general population at the University of Pennsylvania as described above. I ran a mix of sessions with one Sender and three Senders. Tables 2 through 7 report the findings. These tables report chi-square values. I used the Pearson uncorrected chi-square test, which is the standard version available in spreadsheets and statistical software programs. Similar results on statistical significance emerged from the Fisher exact test.

Place Table 2 about here

Table 2 combines the 1-Sender and 3-Sender conditions. The left two columns show the overall number of correct and incorrect messages sent under the high-bias and low-bias conditions. People send the correct information 96% of the time with low bias, and 50% of the time with high bias. The right two columns restrict attention to the cases in which the bias disagrees with the truth. In these cases, people send the correct information 95% of the time with low bias, and 23% of the time with high bias. As measured by the standard chi-square test, these results were significant far beyond the 1% level and cannot reasonably be attributed to small-sample bias. In the high-bias condition, a larger fraction of incorrect messages were sent.

Table 3 compares the 1 and 3-Sender conditions considering both cases in which Sender bias agrees with the truth and cases in which it does not. Considering both high and low-bias conditions together, 70% of messages were correct in the 1-Sender condition and 79% correct in the 3-Sender condition. This result is significant only at the 7% level and may thus reflect nothing more than small-sample bias. Considering only the high-bias condition, 41% of messages were correct in the 1-Sender condition and 59% correct in the 3-Sender condition. This result is significant at the 2% level and may thus reflect something more than small-sample bias. A larger portion of correct messages were sent under the 3-Sender condition.

Table 4 compares the 1 and 3-Sender conditions considering only cases in which Sender bias disagrees with the truth. Considering both high and low-bias conditions together, 55% of messages were correct in the 1-Sender condition and 68% correct in the 3-Sender condition. This result is significant only at the 4% level. Considering only the high-bias condition, 14% of messages were correct in the 1-Sender condition and 33% correct in the 3-Sender condition. This result is significant at the 2% level. This last result implies that the error rate in messages was 86% in the

1-Sender condition, but only 67% in the 3-Sender condition. The rate of errors in messages was lower in the 3-Sender condition.

Place Tables 3 & 4 about here

It is important that Senders send correct messages, but only because correct messages from Senders increase the chance that Receivers will select the correct message when making their judgments. Thus, the more interesting error rates are the error rates for Receivers, or "systemic error rates." Table 5 combines the cases in which Senders had low bias and high bias. Recall that in the low-bias condition, a Sender has no incentive to send an erroneous message. Thus, the values in Table 5 produce a conservative test of the benefit of democratic epistemics. They reveal, nevertheless, an error rate of 36% in the 1-Sender condition and only 15% in the 3-Sender condition. This result is significant beyond the 1% level and cannot reasonably be attributed to small-sample bias. Table 6 restricts attention to the more interesting cases in which Sender bias disagreed with the truth. It reveals an error rate of 47% in the 1-Sender condition and only 21% in the 3-Sender condition. This result is significant far beyond the 1% level and cannot reasonably be attributed to small-sample bias. Table 7 further restricts attention to the most interesting cases in which Sender bias disagreed with the truth *and* at least one Sender is in the high-bias condition. It reveals an error rate of 75% in the 1-Sender condition and only 25% in the 3-Sender condition. Figure 3 illustrates. The result from Table 7 is significant far beyond the 1% level and cannot reasonably be attributed to small-sample bias.

Place Tables 5, 6, & 7 about here

Place Figure 3 about here

The logic of the results on systemic error rates can be seen by considering the imaginary case in which each Receiver always accepts the majority opinion and in each group of three Senders, one sends a false message and the other two send a truthful message. The error rate in messages will be 33%, but the systemic error rate will be zero. Through strategic redundancy, democratic epistemics reduces systemic error rates by more than the reductions in error rates in messages.

The results just reviewed show that 1) experimental participants can be induced to send true information when the induced bias is small, 2) participants can be induced to send biased information when the induced bias is large, 3) democratic epistemics reduce the error rate in messages, and 4) democratic epistemics produce and even more dramatic reduction in systemic error rates.

IV. Discussion

While the current study supports the claim that democratic epistemics should replace monopoly epistemics, more experimental work is needed. To cite only two obvious examples, future work should make further variations in the number of Senders per Receiver and in the relative size of the Senders' bias. These are but two of a large number of variations required in just one relatively narrow line of experiments. There are, moreover, many similar lines of experimentation in the same general area, wherein the normative criterion applied is not efficiency, but veracity. Forensics is one of several areas of application. Others include pure science, espionage, auditing, clinical medical testing, drug screening, judicial proceedings, and private arbitrage.

All such experiments would form part of the field of epistemic systems design, which I have defined above as the application of the techniques of economic systems design to questions of veracity, rather than efficiency. Economic systems design is a branch of experimental economics, which uses money payments to create "induced preferences" (Smith 1976). Economic systems design has produced a major change in how researchers design economic institutions. It uses "the lab as a test bed to examine the performance of proposed new institutions, and modifies their rules and implementation features in the light of the test results" (Smith 2003). Economic systems design combines game theory with laboratory experiments in the search for improvements in the institutions of exchange (McCabe, Rassenti, and Smith 1991).

This study has adapted the techniques of economic systems design to aid in discovering institutional changes that will improve not the efficiency, but the veracity of forensic science and social processes, including other lab-based social processes. Results will help us to understand whether democratic epistemics is feasible and desirable in contexts such as forensic science.

In epistemic systems design, as in economic systems design, researchers adjust their institutional design until it produces the desired results in laboratory settings. Each experiment may produce unexpected and, perhaps, undesired results. These surprises induce the researcher to modify his institutional design and test the new design in the laboratory, where new surprises may induce another round of adjustments. Once this process has converged on a design that seems to produce maximal efficiency or maximal veracity, the researcher is ready to implement his design in the social world where experience will suggest the need for further adjustments, large or small.

Epistemic systems design is possible because we *construct* the truth in an experimental economics laboratory. We are in the godlike position of saying unambiguously what the truth is and how close to it our experimental subjects come. We construct the truth, the preferences, and the institutional environment of choice. We construct, in other words, the world in which in which we place our subjects. From this godlike perspective we are in a position to compare the epistemic properties of different institutional arrangements. When we return from our constructed world to the real world, we lose our privileged access to the truth and return to the groping ignorance under which all humanity toils. But we carry with us a knowledge of which institutional structures promote the discovery and elimination of error and which institutional structures promote error and ignorance. This knowledge can be carried from the constructed world of the laboratory to the natural world of social life because of the common element in both worlds, namely, the human mind. The one vital element of the experimental world that is *not* constructed is the human mind, which makes choices within the institutional context of the laboratory experiment. It is this same element that makes choices in the institutional structures of the natural world of social life. Thus, the sort of laboratory experiment described in this study cannot tell us which particular expert judgments are correct and which incorrect, but they can tell us that the monopoly structure of forensics today produces a needlessly high error rate.

The techniques of epistemic systems design can be applied to a variety of problems including forensic science and pure science. The application to forensic science is an important part of "forensic science administration," the study of how to reduce error rates in forensic science by a reorganization of forensic work. I call for work in this field in Koppl 2005b. As Lawrence Kobilinsky and I have said, "Forensic science administration is a branch of social science. It studies the organization of forensics labor in the criminal justice system, using the tools of social science and business administration. Forensic science administration studies forensic science within its legal and political context" (Koppl and Kobilinsky 2005).

When applied to pure science, the techniques of epistemic systems design produce experimental science studies or, if one prefers the label, experimental epistemology. (I do not imply that science studies and epistemology are the same, however.) Science and technology studies can be very roughly divided into two traditions, that of Robert Merton (Merton 1937, Kitcher 1993, Goldman 1978,1999, 2001, Koppl 2005b) and that of Karl Mannheim (Mannheim 1936, Bloor 1976, Fuller 1988). Epistemic systems design as here conceived falls squarely within the broadly Mertonian tradition. In the past, disputes in this field could be addressed only through historical research and field studies. It now seems possible to address a significant fraction of them with the tools of epistemic systems design. In particular, the network structure relating one lab to another can be manipulated in the laboratory. Thus, it seems possible to address the role of the network structure of pure science in producing reliable knowledge. Indeed, one may describe my proposed move from monopoly to democratic epistemics in forensic science as a move to a network structure for forensic science that more nearly resembles the network structure of pure science.

Epistemic systems design translates evolutionary epistemology into an empirical research program aimed at minimizing error rates. It thus takes seriously W. W. Bartley III's call for an "ecology of knowledge" (1987, p. 447). Bartley says, "hopelessness about objective standards of truth" characterizes the conclusions of the sociology of knowledge. But "such conclusions would never have been reached had the investigation of academic institutions been initiated by economists rather than sociologists" (p. 447). Perhaps, then, it is no accident that my attempt to engage Bartley's project in the ecology of knowledge is rooted in my training in economics.

Epistemic systems design may help us to understand which social institutions produce truth and which do not. Such knowledge permits us to use an indirect strategy for the discovery of truth and the elimination of error. Rather than attempting to instruct people in how to form true opinions, we might reform our social institutions in ways that tend to induce people to find and speak the truth. Comparing the epistemic properties of alternative social institutions is "comparative institutional epistemics." At the margin it may be more effective to give people an interest in discovering the truth than to invoke the value of honesty or teach people the fallacies they should avoid. When we rely on experts such as forensic scientists to tell us the truth, it seems especially likely that institutional reforms will have a higher marginal value than exhortations to be good or rational. If virtue and rationality are scarce goods, we should design our epistemic institutions to economize on them.

References

Akerlof, G. A., 1970. The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism. Quarterly Journal of Economics 84(3): 488-500.

Asch, S. E., 1951. Effects of group pressure upon the modification and distortion of judgement. In Guetzkow, H., ed., Groups, leadership and men. Pittsburgh, PA: Carnegie Press.

Bacharach, M. and Bernasconi, M., 1997. The Variable Frame Theory of Focal Points: An Experimental Study. Games and Economic Behavior 19(1): 1-45.

Bartley, W. W. III, 1987. Alienation Alienated: The Economics of Knowledge *versus* the Psychology and Sociology of Knowledge. In Radnitzsky, G. and Bartley, W. W. III, edited, Evolutionary Epistemology, Rationality, and the Sociology of Knowledge. La Salle, Illinois: Open Court, 1987.

Bloor, D., 1976. Knowledge and Social Imagery. London: Routledge & Kegan Paul.

Blume, A., DeJong, D., Kim, Y., and Sprinkle, G., 1998. Experimental Evidence on the Evolution of Meaning of Messages in Sender-Reciver Games. The American Economic Review, 88(5): 1323-1340.

Bradie, M. and Harms, W., 2004. Evolutionary Epistemology. The Stanford Encyclopedia of Philosophy (January Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/entries/epistemology-evolutionary/>. The url for a revised version of Fall 2006 is expected to be <http://plato.stanford.edu/archives/fall2006/entries/epistemology-evolutionary/>.

Cole, S., 2005. More Than Zero: Accounting for Error in Latent Fingerprint Identification. The Journal of Criminal Law & Criminology 95(3): 985- 1078.

Crawford, V. P. and Sobel, J., 1982. Strategic Information Transmission. Econometrica 50(6): 1431-1451.

Dror, I. E., Charlton D., and Péron A. E., 2006. Contextual information renders experts vulnerable to making erroneous identifications. Forensic Science International 156: 74-78.

Dror, I. E., Péron A. E., Hind, S., and Charlton, D., 2005. When Emotions Get the Better of Us: The Effects of Contextual Top-down Processing on Matching Fingerprints. Applied Cognitive Psychology 19: 799-809.

Fuller, S., 1988. Social Epistemology, Bloomington: Indiana University Press.

Giannelli, P. C., 1997. The Abuse of Evidence in Criminal Cases: The Need for Independent Crime Laboratories. Virginia Journal of Social Policy & the Law, 4: 439-478.

Goldman, A., 2001. Social Epistemology. In Zalta, E. N., edited, The Stanford Encyclopedia of Philosophy (Spring 2001 Edition) URL = <http://plato.stanford.edu/archives/spr2001/entries/epistemology-social/>.

Goldman, A., 1999. Knowledge in a Social World. Oxford: Oxford University Press.

Goldman, A., 1978. Epistemics: The Regulative Theory of Cognition. The Journal of Philosophy 75: 509-523.

Green, J. R. and Stokey, N. L., 1980. A Two-Person Game of Information Transmission. Discussion paper 418, Center for Managerial Economics and Decision Sciences at Northwestern University.

Grieve, D., 1996. Possession of Truth. Journal of Forensic Identification. 46(5): 521–528.

Hayek, F.A., 1935. Collectivist economic planning: Critical studies on the possibilities of socialism. London: George Routledge & Sons, Ltd.

Henrich, J. and 16 other authors, 2005. "Economic Man" in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-Scale Societies. Behavior and Brain Sciences 28: 795-855.

Hoffman, E., McCabe, K., and Smith, V., 1996. Social distance and other-regarding behavior in dictator games. American Economic Review 86: 653–660.

Hume, D., 1777. Essays Political, Moral, and Literary. Available online at http://www.econlib.org/library/LFBooks/Hume/hmMPL.html.

Illinois, State of, 2002. *Report of the Governor's Commission on Capital Punishment* State of Illinois, April 15, 2002.

Innocent, 2002. David Asbury: Three and a half years, downloaded from http://www.innocent.org.uk, 30 June 2006.

Jonakait, R. N., 1991. Forensic Science: The Need for Regulation. Harvard Journal of Law and Technology 4: 109- 191.

Kaufman, F., 1998. *Commission on Proceedings Involving Guy Paul Morin*. Queen's Printer for Ontario.

Kelly, J. F. and Wearne, P., 1998. *Tainting Evidence: Inside the Scandals at the FBI Crime Lab*, New York: The Free Press.

Kershaw, S., 2004. Spain and U.S. at Odds on Mistaken Terror Arrest. The New York Times, 5 June 2004.

Keynes, J. M., 1936. The General Theory of Employment, Interest and Money. London: Macmillan. Reprinted in Collected Writings, Vol. VII. 1973, New York: St. Martin's Press.

Kitcher, P., 1993. The Advancement of Science. Oxford: Oxford University Press.

Koppl, R., 2005a. Epistemic Systems. Episteme: Journal of Social Epistemology 2(2): 91-106.

Koppl, R., 2005b. How to Improve Forensic Science. European Journal of Law and Economics 20(3): 255-286.

Koppl, R., 2005c. Diversity and Forensics: Diversity in Hiring is Not Enough. Medicine, Science and the Law 46(2).

Koppl, R. and Kobilinski, L., 2005. Forensic Science Administation: Toward a New Discipline, manuscript.

Koppl, R. and Langlois, R. N., 2001. Organizations and Language Games. Journal of Management and Governance 5: 287-305.

Kornhauser, L. A., and Sager, L. G., 1986. Unpacking the Court. Yale Law Journal 96: 82-117.

Kornhauser, L. A., 1992. Modelling Collegial Courts II: Legal Doctrine. Journal of Law, Economics and Organization 8: 441-470.

Langlois, R. N., 2002. Modularity in Technology and Organization. Journal of Economic Behavior and Organization 49(1): 19-37.

List, C., and Pettit, P., 2002. Aggregating Sets of Judgments: An Impossibility Result. Economics and Philosophy 18: 89-110.

Mannheim, K., 1936[1985]. Ideology and Utopia: An Introduction to the Sociology of Knowledge. New York: Harper Books.

McBeth, J., 2004. Fingerprint row detective cleared by US forensics. Scotsman.com 21 September 2004.

McCabe, K. A., Rassenti, S. J. and Smith, V.L., 1991. Smart Computer Assisted Markets. Science 254(5031): 534–38.

McDougall, L., 2006. Shirley McKie: Was it Really an Honest Mistake? Sunday Herald, 12 February 2006.

McRoberts, F., Mills S., and Possley, M., 2004. Forensics Under the Microscope: Unproven techniques sway courts, erode justice. Chicago Tribune 17 October 2004.

Merton, R. K., 1937[1957]. Science and the Social Order. Reproduced in his Social Theory and Social Structure, revised edition. New York: The Free Press.

Miller, L. S. 1987. Procedural Bias in Forensic Science Examinations of Human Hair. Law and Human Behavior 11:157-163.

Moenssens, A. A., 1993. Novel Scientific Evidence in Criminal Cases: Some Words of Caution. Journal of Crime Law and Criminology 84: 1-?.

Office of the Inspector General, United States Department of Justice, 1997. The FBI Laboratory: An Investigation into Laboratory Practices and Alleged Misconduct in Explosives-Related and Other Cases. http://www.usdoj.gov/oig/special/97-04a/index.htm.

Office of the Inspector General, United States Department of Justice, 2004. The FBI DNA Laboratory: A Review of Protocol and Practice.

Office of the Inspector General, United States Department of Justice, 2006. A Review of the FBI's Handling of the Brandon Mayfield Case, Unclassified Executive Summary.

Parnas, D.L., 1972. On the Criteria for Decomposing Systems into Modules. Communications of the ACM, 15(12): 1053–1058.

Peterson, J. L., Crim D., and Markham, P. N., 1995a. Crime Lab Proficiency Testing Results, 1978–1991, I: Identification and Classification of Physical Evidence. Journal of Forensic Sciences 40(6): 994–1008.

Peterson, J. L., Crim D., and Markham, P. N., 1995b. Crime Lab Proficiency Testing Results, 1978–1991, II: Resolving Questions of Common Origin. Journal of Forensic Sciences 40(6): 1009–1029.

Radnitzsky, G. and Bartley, W. W. III, edited, 1987. Evolutionary Epistemology, Rationality, and the Sociology of Knowledge. La Salle, Illinois: Open Court.

Raziq, D. and Werner A., 2004. DNA Testing: Study Calls into Question Long-Trusted Lab Results. The IRE Journal 27(1): ,13-14 & 30.

Risinger, M., Saks, M. J., Thompson, W. C., and Rosenthal, R., 2002. The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion. California Law Review 90: 1-56.

Saks, M. J., 1998. Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science. Hastings Law Journal 49: 1069 -1141.

Saks, M. J. and Koehler, J.J., 2005.The Coming Paradigm Shift in Forensic Identification Science. Science 309: 892-895.

Saks, M. J. *et al*., 2001. Model Prevention and Remedy of Erroneous Convictions Act. Arizona State Law Journal 33: 665-718.

Schelling, T.C., 1960. The Strategy of Conflict. Cambridge, MA: Harvard University Press.

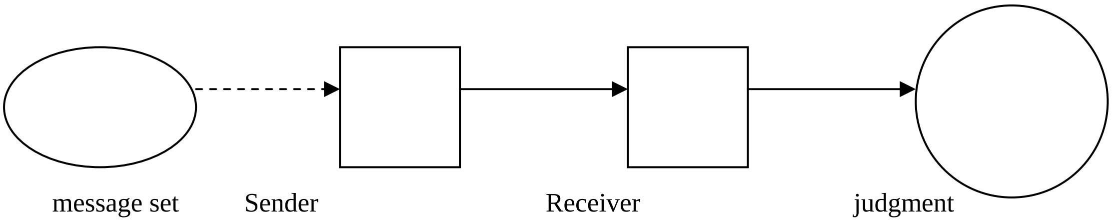
Shackle, G.L.S., 1972. Epistemics and Economics. London: Cambridge University Press.

Smith, V.L., 2003. Constructivist and Ecological Rationality in Economics. American Economic Review 93: 465- 508.

Smith, V.L., 1976. Experimental Economics: Induced Value Theory. American Economic Review 66: 274-279.

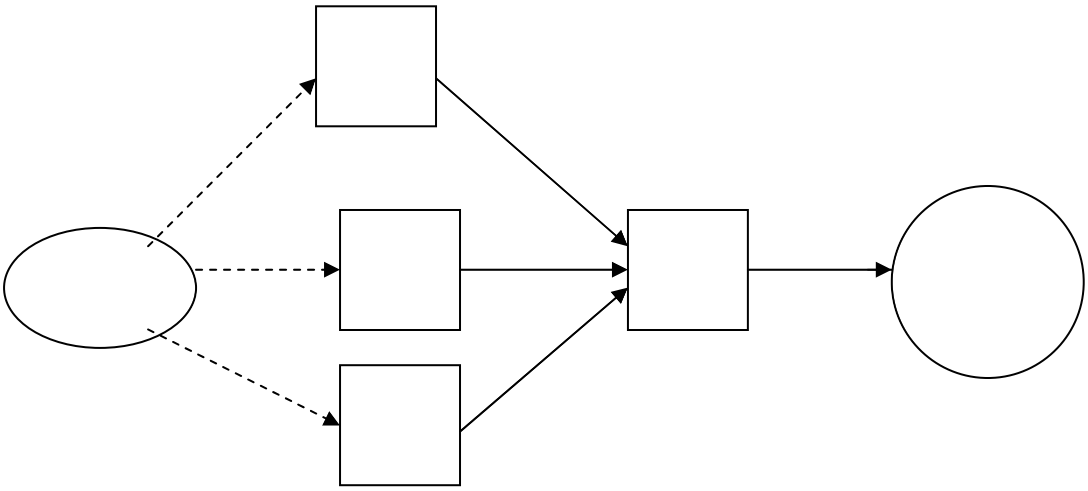
Spence, M., 1973. Job Market Signaling. The Quarterly Journal of Economics 87(3): 355-374.

Teichroeb, R., 2004. Rare Look Inside State Crime Labs Reveals Recurring DNA Test Problems. Seattle Post-Intelligencer 22 July 2004.



message set Sender Receiver judgment

Figure 1



message set Senders Receiver judgment

Figure 2

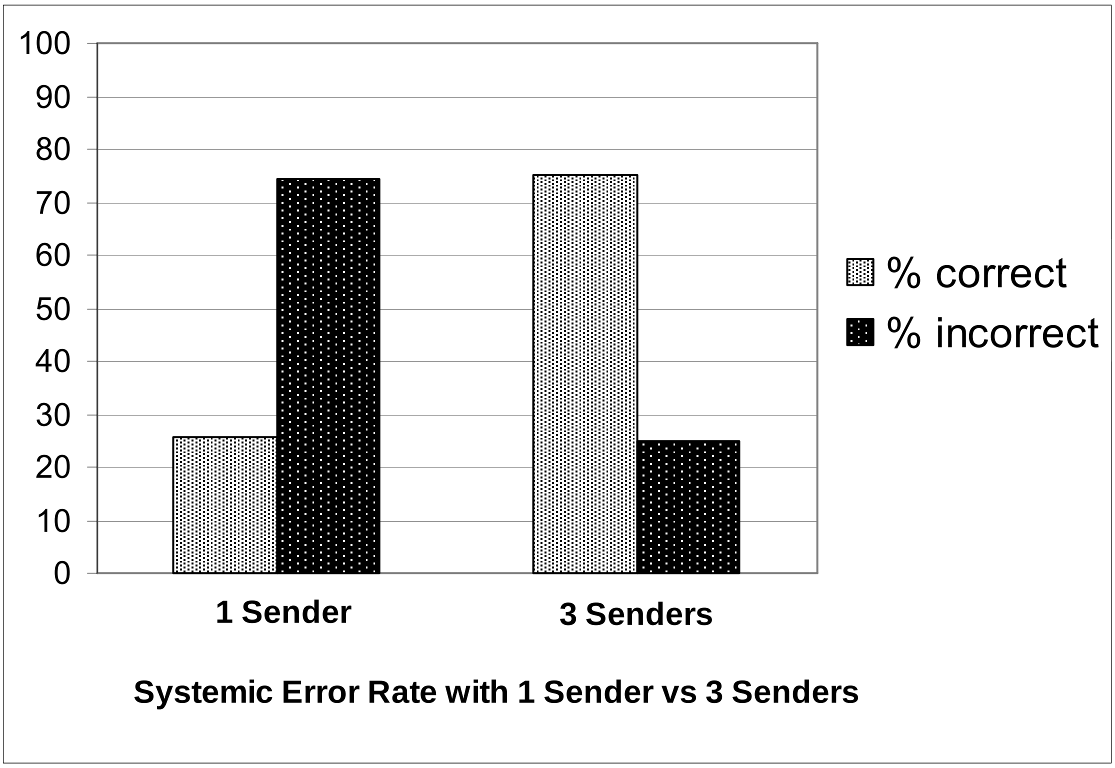


Figure 3

With one Sender per Receiver, the Receivers guessed the wrong shape 75% of the time and the correct shape only 25% of the time. With three Senders per Receiver, the Receivers guessed the wrong shape only 25% of the time and the correct shape 75% of the time. This result supports the claim that replacing monopoly epistemics with democratic epistemics will reduce error rates in forensic science and other areas.

Current system	Resulting problem	Proposed institutional change	Explanation or comment
Monopoly	Sloppy, biased, and sometimes fraudulent work	Rivalrous redundancy	There should be several competing forensic labs in any jurisdiction. Subject to the constraints of feasibility, some evidence should be chosen at random for duplicate testing at other labs. The same DNA evidence, for example, might be sent to more than one lab for analysis. The forensic worker need not know whether the evidence is examined by another lab. He will know that there could be another lab, and sometimes is.
Dependence	Bias	Independence	Crime labs should be independent of police and prosecutors.
Poor quality control	Persistently poor work	Statistical review	Statistical review would support improved quality control. For example, if a given lab produces an unusually large number of inconclusive findings, its procedures and practices should be examined
Information sharing	Conscious and unconscious bias	Information hiding	Evidence should be prepared for testing so as to shield the lab doing a test from all extraneous knowledge of the case particulars.
No division of labor between forensic analysis and interpretation	Error from false interpretations of legitimate results.	Division of labor between forensic analysis and interpretation	When this measure is combined with the provision of forensic counsel for the defense, errors of interpretation are less likely to go unchallenged.
Lack of forensic counsel	False convictions	Forensic counsel for the indigent	Forensic science decides many criminal cases and yet we do not have a right to forensic counsel similar to our right to legal counsel.
Lack of competition among forensic counselors	Poor quality forensic counsel	Forensic vouchers	A voucher system would give forensic counselors to the indigent an incentive to provide high-quality services to their clients.
Public ownership	Weak financial incentives to provide high-quality work	Privatization	Unlike public labs, private labs would be subject to meaningful fines and civil liability. In the US, the federalist structure of government means federal regulation and oversight are easier when labs are private.

Table 1

The proposals of Koppl (2005b) in tabular form.

	Sent correct, overall	Sent incorrect, overall	Sent correct, when bias disagrees with the truth	Sent incorrect, when bias disagrees with the truth
Low Bias condition	166	7	111	6
High Bias condition	77	76	23	76
	$\chi^2 = 89$ d.f.=1 p < 0.0005		$\chi^2 = 117$ d.f.=1 p < 0.0005	

Table 2

Messages sent, one and three Sender conditions combined

	Sent correct, high and low bias combined	Sent incorrect, high and low bias combined	Sent correct, high-bias condtion	Sent incorrect, high-bias condition
One-Sender condition	104	45	30	44
Three-Sender condition	139	38	47	32
	$\chi^2 = 3$ d.f.=1 p = 0.0714		$\chi^2 = 5$ d.f.=1 p = 0.0191	

Table 3

Messages sent, combining cases in which bias does and does not agrees with truth

	Sent correct, high and low bias combined	Sent incorrect, high and low bias combined	Sent correct, high-bias condtion	Sent incorrect, high-bias condition
One-Sender condition	54	45	7	44
Three-Sender condition	80	37	16	32
	$\chi^2 = 4$ d.f.=1 p = 0.0369		$\chi^2 = 5$ d.f.=1 p = 0.0210	

Table 4

Messages sent, when bias does not agrees with truth

	Receiver guessed correctly	Receiver guessed incorrectly
One-Sender condition	95	54
Three-Sender condition	50	9
	$\chi^2 = 10$ d.f.=1 p = 0.0030	

Table 5

Receiver guesses, comparing the one and three Sender conditions regardless of the direction of Sender bias

	Receiver guessed correctly	Receiver guessed incorrectly
One-Sender condition	52	99
Three-Sender condition	31	8
	$\chi^2 = 26$ d.f.=1 p < 0.0005	

Table 6

Receiver guesses, comparing the one and three Sender conditions when the truth disagrees with Sender bias

	Receiver guessed correctly	Receiver guessed incorrectly
One-Sender condition	13	38
Three-Sender condition	24	8
	$\chi^2 = 20$ d.f.=1 p < 0.0005	

Table 7

Receiver guesses, comparing the one and three Sender conditions when the truth disagrees with Sender bias and at least one Sender is in the high-bias condition.

Appendix: Experimental Materials Used.

Intructions for Senders in the 1-Sender conditions

INSTRUCTIONS

Please read all the instructions before making any decisions.

Thank you for your participation. You have been asked to participate in a decision-making experiment. Your decisions today are anonymous. No one, including the experimenters, will know which person made which decision. Because your decisions are private, we ask that you do not tell anyone your decisions either during or after the experiment. By showing up for today's experiment you have already earned $5. This is yours to keep regardless of any decisions that you subsequently make during the experiment. If you have any questions, please raise your hand and an experimenter will come by to answer them.

Overview

This experiment involves three shapes: triangle, circle, and square. We will pick one and show it to you. The shape you are shown is the "correct shape." Someone else has to guess what shape you were shown. You will provide that person with information by making a "report" on what shape you saw. You are not required to report the correct shape. The person who gets this report will then use it to guess the correct shape. You get money two ways. First, you get money if the person guesses correctly. Second, you get money if the person guesses **a supplementary shape**, whether or not that's the correct shape. The process of reporting shapes will be repeated ten times. The details are explained below.

Decision Makers

The participants in this experiment have been divided into two types of decision maker, Decision Maker 1s (DM1) and Decision Maker 2s (DM2). **You are a Decision Maker 1**. You will be participating in ten interactions with other decision makers. You will have the opportunity to earn additional money through these interactions, as explained below. In each interaction, you are randomly matched with other Decision Makers. You will not be told who you have been matched with, either during or after the experiment. All interactions are anonymous.

Payment for Decision-Making Tasks

At the conclusion of the session, one of the ten interactions will be randomly selected according to a pre-determined procedure and you will receive the money from the outcome of that one interaction.

Coding System to Maintain Anonymity

There is an index card in your envelope containing ten different numbers, one for each interaction. Each one of these codes will match a set of numbers on the top right corner of the ten interactions that you have in this packet. You will present this index card to the experimenter after all 10 interactions have taken place. Your payment will be marked by one of these secret codes.

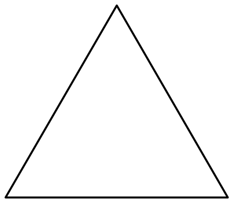
What happens in each interaction

Each interaction begins when the experimenter announces it. "Interaction one begins now." "Interaction two begins now." And so on.

You will be shown one of three shapes on a piece of paper:
1) circle, 2) triangle, 3) square.

The next page is an example. In the example, the correct shape is a triangle.

The correct shape for this interaction is:

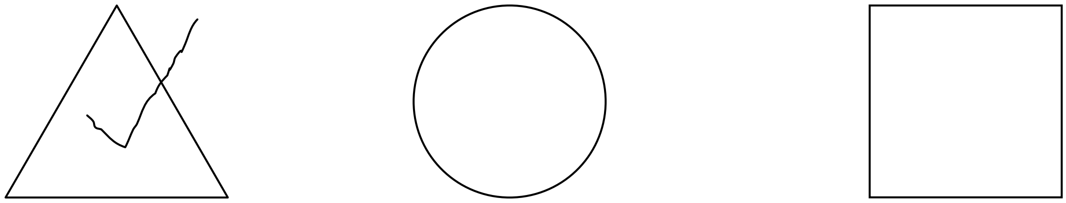


After viewing the correct shape, you will "report" which shape appeared:
1) circle, 2) triangle, 3) square.

You will make your "report" by completing a Report Form. You complete the form by putting a check on the shape of your choice. Please use the marker provided. The next page is an example of a completed Report Form. In the example, the reported shape is "triangle." In each interaction you may report any one of the three shapes found in the Report Form. This Report Form will be given to DM2.

Report Form

DM1 has reported the shape with the check mark on it.



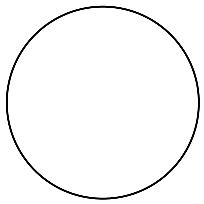
After seeing your report, DM2 will guess what the correct shape is. He/she has three choices: 1) circle, 2) triangle, 3) square.

DM2 will try to guess the correct shape based only on the report you send him/her. DM2 earns money in an interaction if and only if he/she correctly guesses the correct shape.

Your payment will depend on DM2's choice and will have two parts. **First part:** You will get $3.00 if your DM2 guesses the correct shape ***and*** you reported the correct shape. Otherwise you will get nothing for this part of your payment. **Second part:** Independently of the first part, if DM2 guesses **the supplementary shape** you will get $1.00 or $5.00, depending on the value of that supplementary shape to you. You will be told the value of the supplementary shape at the beginning of each interaction by being shown a piece of paper indicating the value. The next page is an example.

(**Remember**: While a dollar payment will be calculated for each of the ten interactions, you will get an actual payment only for one randomly chosen interaction.)

The supplementary shape is:



The value of circle is:

$1.00

Here are three examples of how payments are calculated. The examples assume that circle is the supplemental shape.

Correct Shape	Value of a Circle to You	Your Report	DM2's Guess	Your Payoff
Triangle	$1.00	Triangle	Triangle	$3.00
Triangle	$1.00	Circle	Circle	$1.00
Triangle	$5.00	Circle	Circle	$5.00

How to Make a Report

You have a packet for each interaction. The first page of the packet tells you the correct shape for that interaction and the value of a circle to you, for that interaction. The second page of the packet is a Report Form. You will complete the Report Form, which will be sent to DM2 by an experimenter. The next page gives an example of a completed Report Form. In the example, the reported shape is "triangle." In each interaction you may report any one of the three shapes found in the Report Form.

When you are Done

When you have completed all ten interactions, the experimenter will ask you to complete a questionnaire asking for some additional information. This information remains confidential and anonymous. When you have completed this questionnaire, please submit it to the experimenter who will hand you a debriefing form and your payment. Please do not discuss this study with others, as it is important that future participants do not have prior knowledge of the content of this study.

You now have the opportunity to participate in ten interactions with DM2s we have randomly matched to you.

Intructions for Recievers in the 3-Sender condition

INSTRUCTIONS

Please read all the instructions before making any decisions.

Thank you for your participation. You have been asked to participate in a decision-making experiment. Your decisions today are anonymous. No one, including the experimenters, will know which person made which decision. Because your decisions are private, we ask that you do not tell anyone your decisions either during or after the experiment. By showing up for today's experiment you have already earned $5. This is yours to keep regardless of any decisions that you subsequently make during the experiment. If you have any questions, please raise your hand and an experimenter will come by to answer them.

Overview

This experiment involves three shapes: triangle, circle, and square. We will pick one and show it to three other people. The shape those people are shown is the "correct shape." Your job is to guess what shape those people were shown. To provide you with information for your guess, each person will make a "report" on what shape he/she saw. Those people are not required to report the correct shape. You will receive a total of three reports. You get money by guessing the correct shape. The process of guessing shapes will be repeated ten times. The details are explained below.

Decision Makers

The participants in this experiment have been divided into two types of decision maker, Decision Maker 1s (DM1) and Decision Maker 2s (DM2). **You are a Decision Maker 2**. You will be participating in ten interactions with other decision makers. In each interaction, you will be randomly matched with three DM1s who are in this room. You will have the opportunity to earn additional money through these interactions, as explained below. In each interaction, you are randomly matched with other Decision Makers. You will not be told who you have been matched with, either during or after the experiment. All interactions are anonymous.

Payment for Decision-Making Tasks

At the conclusion of the session, one of the ten interactions will be randomly selected according to a pre-determined procedure and you will receive the money from the outcome of that one interaction.

Coding System to Maintain Anonymity

There is an index card in your envelope containing ten different numbers, one for each interaction. Each one of these codes will match a set of numbers on the top right corner of the ten interactions that you have in this packet. You will present this index card to the experimenter after all 10 interactions have taken place. Your payment will be marked by one of these secret codes.

What happens in each interaction

Each interaction begins when the experimenter announces it. "Interaction one begins now." "Interaction two begins now." And so on.

For each interaction, the DM1s will be shown one of three shapes:
1) circle, 2) triangle, 3) square.
All three DM1s see the same shape.

Each DM1 will then be prompted to "report" to you which shape he/she saw. No DM1 knows what any other DM1 reports. Each DM1 must give you one of the following three reports:
1) circle, 2) triangle, 3) square.

After seeing the reports of all three DM1s, you will make one of three guesses about what the "correct" shape is: 1) circle, 2) triangle, 3) square.

Your payment depends on whether you guess correctly. If you guess correctly, you get $5.00. If you guess incorrectly, you get $2.00. Although your guess may influenced by the reports of the DM1s, your personal payoff does not directly depend on their reports.

(**Remember**: While a dollar payment will be calculated for each of the ten interactions, you will get an actual payment only for one randomly chosen interaction.)

Here are three examples of how payments are calculated.

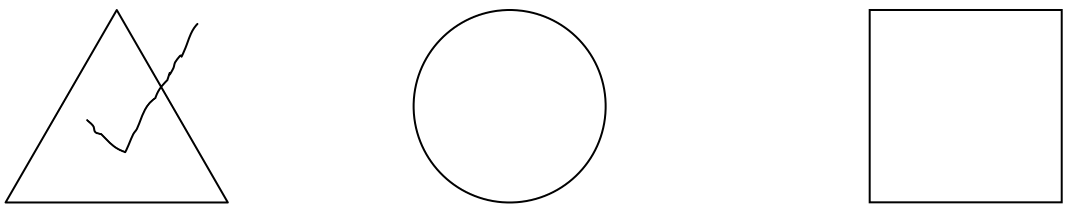
Correct Shape	First DM1's Report	Second DM1's Report	Third DM1's Report	Your Guess	Your Payoff
Triangle	Triangle	Triangle	Triangle	Triangle	$5.00
Circle	Triangle	Square	Circle	Triangle	$2.00
Circle	Triangle	Square	Circle	Circle	$5.00

How a DM1 Makes a Report

For each interaction, you will receive three Report Forms from the three DM1s you have been matched with. Each Report Form has been completed by a different DM1. The next page gives an example of a completed Report Form. In the example, the reported shape is “triangle.” In each interaction, each DM1 may report any one of the three shapes found in the Report Form.

Report Form

DM1 has reported the shape with the check mark on it.



How to Record Your Guess

You have one “DM2 Decision Form” for each interaction. You will complete the Decision Form and turn it in with the other materials. The last page of these instructions is an example of a completed Decision Form. In the example, the DM2 guessed shape is “square.” In each interaction you may guess any one of the three shapes found in the Decision Form.

When you are Done

When you have completed all ten interactions, the experimenter will ask you to complete a questionnaire asking for some additional information. This information remains confidential and anonymous. When you have completed this questionnaire, please submit it to the experimenter who will hand you a debriefing form and your payment. Please do not discuss this study with others, as it is important that future participants do not have prior knowledge of the content of this study.

You now have the opportunity to participate in ten interactions with DM1s we have randomly matched to you.

DM2 Decision Form

Place a check mark on the shape you guess is the correct one. Use the marker provided.