

ROUNDTABLE DISCUSSION: MUST WE CHOOSE BETWEEN
RATIONALITY AND IRRATIONALITY?

Chicago-Kent College of Law
November 6, 2004

MODERATOR

CLAIRE HILL, Chicago-Kent College of Law

PARTICIPANTS

PHIL CURRY, Simon Fraser University, Department of Economics

OLIVER GOODENOUGH, Vermont Law School

RICHARD MCADAMS, University of Illinois College of Law

KEVIN MCCABE, George Mason University, Department of Economics

ERIN O'HARA, Vanderbilt University Law School

TODD PREUSS, Emory University, Division of Neuroscience

RICHARD WARNER, Chicago-Kent College of Law

CLAIRE HILL: This is the last session of this conference, and I'd like to thank everyone who has participated. We're going to end with a roundtable discussing the following two questions. First: must we choose between rationality and irrationality? Second, what follows from making the choice or not?

RICHARD MCADAMS: Thank you, Claire, for inviting me to this conference. Must we choose between rationality and irrationality? I think so. I think that if you're trying to develop a social science understanding of human behavior, most social scientists of any stripe think that people are substantially rational: some amount of rationality is going to enter implicitly or explicitly into anyone's theory. When I say "substantially rational," I mean it in the simple instrumental sense that there is a connection between people's means and ends most of the time. I think everyone would agree that people are not perfectly rational.

So, I would defend rational choice theory as a research agenda. If you're going to try to move beyond speculative disagreements about vague differences in the degree to which people are rational to something more specific, then I think it's probably useful to develop an agenda that begins

at either one extreme or the other. And the extreme of irrationality is just incoherent. When you get the invitation from Claire to a conference at Chicago-Kent in November 2004, if you think that people are irrational you have no reason to assume that there is a conference, that Claire sent this to you, or that she has the authority to invite people. Basic means-end rationality is the only way to interpret human behavior, at least as a starting point: people who put coins in vending machines are actually trying to get from the vending machine the thing the vending machine sells.

So, as between the two extremes, I think it makes sense for the research agenda to start with the extreme of people being perfectly rational and see how far that can go. We've learned a lot in recent years by following that agenda—and we've prompted, and perhaps provoked, people into doing the kind of work that identifies where perfect rationality is a bad model to predict behavior. And I think we're moving towards a better model.

I read Claire's paper on categorization¹ and agree with the basic point she makes: we're not exhausting the model of human behavior, even on the process side, just by getting it right about how rational or irrational people are. We can't just think about the ends humans seek and how they seek those ends. I think categorization—how we classify phenomenon in the world—is an important part of the behavioral story, and we've heard a few talks on that subject here at this conference. Consider, for instance, Roland Fryer's piece on categorical cognition.² As Fryer discusses in that piece, we are boundedly rational and are constrained in the number of categories we have, such that we may end up categorizing coarsely in a way that has significant consequences—for instance, putting in three categories what might have been more “accurately” put in four.

I'm now going to speak briefly on my own work. I've been writing about the effort to explain legal compliance through mechanisms other than the fear of sanctions and the deference to legitimate authority. One of the ways that I hope to show that law works expressively is in articulating—clarifying and publicizing—categories that people then use to understand the social world. Now most law reflects the categories by which people understand the social world: the law creators already have these concepts, which they then articulate through law. But in ambiguous situations, as many situations are, or in situations where the way we understand reality is

1. Claire A. Hill, *Beyond Mistakes: The Next Wave of Behavioral Law and Economics*, 29 *QUEEN'S L.J.* 563 (2004).

2. Roland G. Fryer & Matthew O. Jackson, *A Categorical Model of Cognition and Biased Decision-Making* (Working Paper Oct. 2002, revised Nov. 2004), available at http://post.economics.harvard.edu/faculty/fryer/papers/fryer_jackson.pdf.

contested, as it often is in the context of social movements, one way law may work is by influencing how popular certain categories of understanding the world are.

Let me make a final point about the importance of categories in the context of a field I know a little about—evolutionary game theory. When you have models that try to explain how informal order arises, the models often are concealing the crucial step of how people categorize the world. So, for example, there are models in which property emerges from the state of nature, because people decide what their best strategy is, and the strategies they can play depend on certain features of the world they observe. So in a model, a player could be either the “possessor” of the property or the “non-possessor” of the property. And then you can choose to play an aggressive strategy or a passive strategy. So therefore you could play a strategy like “be aggressive when I am the possessor and be passive when I am the non-possessor.” Built into the model, however, as soon as you describe the strategies that way, you are already saying that the people in the model have categorized the world in the way that the strategies define them, that is, as possessors or non-possessors. And I think that in fact, it’s *that* preliminary step of explaining how they come to carve up the world the way they do that the models can’t say anything about. But some of the other things that we’ve heard here today do address that issue and that is a very interesting correction that my work will try to pursue.

KEVIN MCCABE: I thought it was interesting that Claire put the issue as: must we choose between rationality and irrationality, as opposed to must we make rational versus irrational choices. I think that however we define rational and irrational, we do have a sense of people making better or worse choices. But that isn’t what’s central here. What’s central is: do we want to be seen as somebody who’s led a rational life versus an irrational life? And then you’d have to say: how should such a thing be judged? What’s a rational life and what’s an irrational life? Do we use a market test? By a market test, Bill Gates leads a rational life. But there are other tests, such as an evolutionary test, and by such a test, Bill Gates fails. As far as we know, he’s got far too few children given the amount of resources he has for bringing up those children. So, in the end, I don’t think that’s a very appealing way to try to approach the issue. What is perhaps more interesting than thinking about individuals in this way is thinking about groups in this way. Do groups make better or worse decisions? And then you’d have to ask the purpose of a group. And you might feel a little more comfortable actually writing down what you think the purpose of, say, a profit maximizer or a firm of shareholders is, even though you’d still

have trouble because Roy Radner's already shown that if you're trying to maximize survival you're not going to try to maximize the present discounted value of the firm.³ So even there you would have problems finally deciding what you meant by rationality or irrationality.

The methodology I've used has been quite different. It starts with the brain—it recognizes that our brains have evolved over a very long period of time so as to function in the evolutionary environment. I don't think that's controversial—but what is a little more controversial is the argument that therefore the way we interrelate socially and the way we form our institutions is to fit the way our brains have been adapted over time. And that what we see is a process in society of trial and error, with the success of institutions being how well they fit the way we can conceive of, think about, and get what we want out of the world, constrained by what other people are trying to do and so on. I've taken less of a performance-based perspective⁴ on the fitness of institutions as is common in experimental economics study of institutional and market design. Instead, I've taken more of a coherence perspective—that is, something is rational if it fits together coherently. Institutions are rational if people want to use them and can use them easily, and if they solve a certain kind of fitness to the individuals who are using them. And I'm not, at this point, going to go much past that from a viewpoint of a positive theory of people or individuals. If you ask me: do I want to give up on normative theories as a consequence of this kind of strictly positive approach to study the world, I would say no. But I'm on far less clear ground on how to apply my science to normative theorizing about the world. That's a problem—but also a good thing.

PHIL CURRY: When I was thinking about what to say, I realized I had a choice to make. I could tell you how I thought about this question and then conclude with my answer, or I could give you my answer and then tell you how I thought about this question. If you're all fully rational, your opinion of what I say would be independent of the choice I make in that regard, since you're going to get the same amount of information—the only thing that will change is the order. However, I'm going to choose to give you my answer first and tell you what led me to that answer because if I do that, I think you'll generally all agree with me; but if you tell me you disagree, I can then say that there was a framing bias going on and that you

3. Prajit K. Dutta & Roy Radner, *Profit Maximization and the Market Selection Hypothesis*, 66 *REV. ECON. STUD.* 769–98 (1999).

4. A performance perspective is one in which the institution is designed to solve a specific performance criteria, such as maximizing surplus, minimizing costs, or maximizing revenue.

didn't incorporate the information properly. In this way, I can stroke my own ego and use a confirmatory bias on my own part.

Having said that, yes, we must choose between rationality and irrationality, and we must choose rationality. How did I get to that conclusion? I think all the talks at this conference have mentioned that we need to define both rationality and irrationality. I'm going to give a very simple definition of rationality that states that we are able to make decisions given the information available to us and, in particular, that we are able to see through the mechanisms that determine our outcomes. Further, I think another important issue, and I haven't heard anyone else mention it so far, is to define who "we" are and what is meant by "choose." The definition of "we" that I shall use to answer this question is very narrow and doesn't even apply to everyone in this room. Specifically, I consider "we" to include researchers whose goal it is to try to develop optimal policy. So it's a very policy-oriented and very academic-oriented definition of "we." With regards to "choose," I feel the appropriate definition is, "to use models in which there are agents that are able to see through mechanisms." Given these definitions of "we" and "choose," I think it straightforward that, of course, we must choose rational agents. If we don't choose rational agents, this means that policies rely on people not understanding the outcome.

We have an example of such a policy from the book of Kings in the Bible. King Solomon had to decide between two plaintiffs in front of him, each of whom claimed to be the mother of a baby. Each woman said that the child was hers. King Solomon had to decide who would get custody of the child. So, he told the women that he felt the only fair thing to do would be to cut the baby in half and give each of them one half of the child. In response, one of the women said, "no, that's a terrible solution. Give the baby to the other woman. I won't let this happen to the child." From this dialogue, King Solomon deduced that she was the mother. Only someone who was not the mother of the child would go along with such a solution. The child was thus awarded to the woman who had protested.

A government, or policymaker, can only use King Solomon's method once. It does not say anywhere in the Bible what happened the next time King Solomon had such a decision to make. That would have been interesting to know.

So, when determining optimal policy, it is necessary to assume that agents are going to be able to figure out what the outcomes of the policy are. And, perhaps surprisingly, there's a lot of evidence to suggest that people are pretty good at it, perhaps even better than the researchers who propose the policies.

There are many examples of policies that led to unanticipated events that, after the fact, seem obvious. In Canada, comprehensive unemployment insurance was introduced in 1972. Shortly after, unemployment rates soared. Why did this happen? Unemployment insurance decreased the cost of being unemployed. So, as a result of the policy change, people were willing to search longer for new jobs. This came as a surprise to people at the time. These days, it's quite clear why this happened. It was in fact some colleagues of mine at Simon Fraser University, Herb Grubel and Dennis Maki, now both professors emeriti, who first noticed this effect.⁵

There have been other examples of policies with unintended, but in hindsight perhaps obvious, side effects. Seat belt laws were introduced in both Canada and the U.S. in the mid-sixties, and accident rates went up.⁶ The reason for this is quite simple. When drivers wear seatbelts, the expected cost of an accident *to the drivers* decreases. As such, drivers reduce the precaution that they take to avoid accidents. Perhaps not surprisingly, the losers from the seat belt law were pedestrians. While the number of accidents went up, the average severity went down, although pedestrians got hit hard both literally and figuratively.

CLAIRE HILL: Phil, I can't restrain myself: what do you say about the recent American presidential election?

PHIL CURRY: This raises a very interesting issue because I don't think many economists would suggest that agents never make mistakes. In particular, when a new policy is introduced, it seems reasonable that some people are going to need some time before they figure out its implications. I believe the recent election to be such an example. Historically, Republicans and Democrats have both been centrists. The differences between them have traditionally been slight. George W. Bush, however, represents a divergence from this tendency. The current Republicans are actually quite extreme in their views. I believe Americans have become used to voting without paying much attention to the policies of the candidates, because there was little difference between them. If the current Republican Party is representative of the future, however, voters will start to pay more attention to the candidate's policies. I suppose time will tell on this prediction.

The past election also provides another example of how people make mistakes. Politics is, for many people, an emotional topic. Emotion has been a prevalent topic of discussion at this symposium. Most notable, I

5. Herbert G. Grubel et al., *Real and Insurance-Induced Unemployment in Canada*, 8 CANADIAN J. ECON. 174-91 (1975).

6. Sam Peltzman, *The Effects of Automobile Safety Regulation*, 83 J. POL. ECON. 677-726 (1975).

feel, was the presentation by Professor Clore. He noted that emotion can either be rational or irrational. Even if a policy has been around a long time, some people will make mistakes because their emotions get in the way. It would therefore seem undesirable to have a policy such that the consequences of a mistake are extreme. Some economic models provide that one way to give agents incentive to engage in a particular activity is to shoot them if they don't. These models assume that agents never make mistakes. This is not to say that economists never take mistakes into consideration. For example, work by Winand Emons⁷ has shown that when crimes can be committed by mistake (or when people can be falsely convicted), the optimal policy entails increasing penalties for repeat offenders. The penalty for the first offense should be low, because anybody might do it once. However, the probability that somebody does it twice is exponentially smaller, and so the penalty should increase.

To summarize, I believe that long-term policy should be designed with rational agents in mind. However, there is room for models that incorporate irrational agents, as it is important to understand the types of mistakes that people may make and how to insure against them.

But, I would say that my final answer to the question is yes, we must choose, and that policy-oriented questions should overwhelmingly rely on rational agents.

CLAIRE HILL: My answer to the symposium question, must we choose, is sometimes. And, when we do have to make a choice, we should choose rationality. I'm very much in the camp of the previous speakers, who've basically said this same thing. My answer to the next question is to tell you about some of the work I'm doing that assumes a choice between rationality and irrationality (choosing rationality), sometimes sidesteps the issue, and sometimes attempts to define rationality more broadly.

Defining rationality more broadly is a big mission of mine. I think rationality as it's been defined thus far doesn't take into account many things, including the kinds of things Herbert Simon is concerned about, such as how we process information.⁸ Nor does it take into account how we arrive at our intuitions, and what the process is by which we conclude that our intuitions are correct or incorrect.

7. Winand Emons, *Escalating Penalties for Repeat Offenders* (Univ. of Bern, Dept. of Econ. Working Paper, 2003), available at http://www-vwi.unibe.ch/theory/papers/emons/esc_p.pdf.

8. See generally Herbert Simon, *Rationality in Psychology and Economics*, in *RATIONAL CHOICE: THE CONTRAST BETWEEN ECONOMICS AND PSYCHOLOGY* 25 (Robin M. Hogarth & Melvin W. Reder eds., 1986).

One reason I like Simon's take on many things is because I'm sympathetic to what he describes as the psychologist's more process-based take on rationality, which is about means, rather than the economist's more substance-based take, which is often about ends. But, as I think even Simon would agree, something that's perfectly rational as concerns process can be quite irrational if the end makes no sense. What if Ellen thinks she's a witch and looks very efficiently for the most effective spells? This of course assumes that the society's assessment that there are no such things as witches who can use spells is correct.

My general take is, though, that we should generally choose rationally, for the sorts of reasons Richard McAdams gave, but—and I may believe in this more strongly than Richard does—there is an important class of cases in which rationality isn't at issue in any meaningful way. What becomes a norm and why? How does the government convince people that smoking isn't cool? How does an advertiser convince people that a particular car *is* cool? Where is rationality in that? It may establish some sort of bound, but there's lots of indeterminacy in the middle ground.

Work I've done recently tries to analyze the process by which people attach labels such as coolness, or other labels where there's no determinate "fact of the matter." I've also been trying to analyze the much broader question of how people make sense of the world, much discussed in other fields such as philosophy and psychology but curiously assumed away in economics.

Erin [O'Hara] and I are working on a project⁹ that considers, among other things, how people assess how trustworthy other people are. One important question is: how do people organize that information? We argue for four general boxes: residual trust, residual distrust, specific trust, and specific distrust, alone or in combination. Maybe I feel residual trust when it comes to the Pope. But with Bill Clinton, I might think "good guy, but not around my daughter (if I had one)." Similarly, I might think "mobster X is a bad guy, but he always picks up after his dog." And, of course, how we assess people influences what we'll do subsequently. To go back to Richard McAdams' earlier example, because he trusts me, I bet he didn't go to the website to make sure I was still teaching at Kent or call our mutual friend Jacob [Corré] and ask him the question.

Another project that I'm working on is about decision making under uncertainty. This is a very well-worn area, but I have a particular angle: how the decision is affected when you know you have to justify it. Exam-

9. Claire Hill & Erin O'Hara, *Optimal Trust* (forthcoming) (on file with authors).

ples include investment and hiring decisions. There's a concept from psychology that's very important: so-called "alignable metrics." My intuition is that if it's hard to justify making a decision, and you have no idea as to how it will turn out, it's really helpful if there's some scale that people accept on which the things you had to choose among were ranked. And there will be lots of pressure for such a scale to exist and to have the necessary legitimacy to serve its purpose. Think about the SAT or price/earnings ratios. Consider Roland Fryer's paper on majority and minority job candidates—he talked about the lack of information the majority group has about the minority group.¹⁰ Seems to me as well that if you're a member of the majority group, you may think you're really good at reading other members of your group (although you may be wrong on this—some research shows people very much over-trust their gut feelings), but not so good at reading members of other groups. You may hence be less apt to hire members of minority groups, unless you have something you can point to to justify the decision after the fact if it doesn't work out—"See? Of course I hired Jones. He had a terrific SAT score. How could I know he'd be a total screw up?" It may be that the alignable metric—in this case, the SAT—is a crummy predictor generally, or perhaps a crummy predictor for the minority group. But there will be lots of effort thrown at minority group acquisition of high SAT scores if majority group employers are relying on them, which may lead to all sorts of distortions.

One of those distortions is in the market for rankers—you can't have a zillion different rankers that all have significant legitimacy. So, you see just a few—in the case of the SATs, what are their competitors? I did some work recently on rating agencies, and it was interesting to consider how the kinds of dynamics I'm discussing affect the incredible power and dominance of the two main rating agencies, Moody's and Standard and Poor's, which most people think are scarcely cutting-edge clever. And how do they justify their methodologies—and how does their need to justify those methodologies affect those methodologies? I read recently about a subject near and dear to Fred Schauer's heart, fancy French restaurants rated highly by the Michelin guide. Some chefs were complaining that to get the high Michelin rating, they needed to be changing their flowers several times a day. I don't know about you, but I'd just as soon have silk flowers (no plastic, please) and pay less or get better food. But think about why Michelin might notice flowers. They need to train inspectors, presumably, and they want some measure of uniformity among their inspectors. The more benign explanation is that the flowers are a good proxy for quality, but since the

10. Fryer & Matthews, *supra* note 2.

inspector is eating the food, why does he need a proxy? It seems more likely to me that rank ordering makes you look to things like “how many times a day do they change the flowers” and value it over “just how good is this food anyway?”

Another project I’m working on now is how to fit people’s identities into their utility function. And not just self-serving identities, but more “neutral” and “hard wired” ones like gender. And other identities that might be interesting for law. The reason I’m discussing these two projects in the time I have here is that I see them as fitting into a rationality paradigm, broadly construed—I think they potentially can help tweak and elucidate what rationality might really consist of, and not just in an exotic “psychology gets the obscure nuances right” sense, but in senses that are really important to legal scholarship.

A final project that fits within this rubric isn’t mine, but I wanted to mention it—it’s George Loewenstein’s. He has a paper in which he shows the extent to which preferences aren’t just “what one likes” that, say, somebody is smart enough to figure out how to offer you at a price that seems worth it to you. In other words, through a series of ingenious experiments, he shows that preferences aren’t just discovered. How people come to like things, and how much they value them, has a huge amount to do with what’s presented to them, and specifics about the presentation. The notion of rationality or people making a “mistake” doesn’t really make sense here. Like the Tom Sawyer example George uses, where Sawyer, having the job of painting the fence, gets others to help him, but charges them money. They all say they had a great time—were they mistaken? How could they be? In general, people know when they’re enjoying themselves, no? There are certainly many consequences of this for standard economic theory. What the consequences are, “real world,” is another question and one that’s rather harder. Some of this line of thinking will presumably remind you of things advertisers have to think about all the time. Is the shampoo company better off trying to tell me in some credible way about the good things the shampoo can do for me? Or, is the company better off trying to conjure up some sort of glam image I’d like to be associated with? I think most decisions aren’t substance-based the way economists imagine—people simply don’t have that good a sense of what “inherent attributes” they value and how much they value them. Again, what do non-advertisers do with this? Tune in next year.

ERIN O’HARA: I want to start out with a confession, which is that I have no idea why we have to choose between rationality or irrationality. But I am so thrilled to see that Claire was willing to organize such a great

group of people who put their minds around this question, because it's one that I've been agonizing over for a while. I want to thank not only Claire but also the sponsors of the conference, the Chicago-Kent Law Review, the Gruter Institute, and the Society for Evolutionary Analysis in Law.

Like a lot of you, I'm interested in whether we can come up with a good theory of human behavior to use for policymaking. I started out as a strict law and economics scholar, and then I met people like Oliver [Goodenough], who wrote on evolutionary theory and behavioral biology, and the implications that these topics have for a theory of human behavior. I've spent the last couple of years trying to really explore the implications of this notion that our human brains are a product of evolution in terms of its cognition and in terms of our emotions. And one important part about our evolutionary history seems to be that humans have evolved in social groups, so it has to be that the way we process information and the way we process emotions have a lot to do with social relationships.

I've spent much recent time, some of that time with Claire, some of that time on my own, and some of that time with Doug Yarn, who's at Georgia State Law School, focusing on how relationships develop, how trust in relationships develop, how relationships break down, and how people who are having trouble in their relationships conciliate those relationships. For example, Claire and I are interested in the different cognitive functions of trust and the way that trust develops, is maintained, and breaks down differently in different kinds of relationships. If we think of the law as working on individual behavior in relationships and we think about trust as a cognitive function subject to the same limitations and biases as other cognitive functions, can we come up with a notion of optimal trust in relationships? And, can we use legal tools to get those people in those relationships to a better place or situation for developing their relationships with a sense of trust that they might not have on their own?

Similarly, with regard to relationships breaking down, I've been doing a lot of work on apology,¹¹ and how we have a taste for the ritual of apology, and how that has implications for the way that we deal with dispute resolutions as a formal matter in the courtroom. I tell you this just as a prelude to the fact that I started thinking about theory of human behavior as purely economic, began to think about evolutionary theory and behavioral biology, realized that these theories enriched the economic theory but were still inadequate, and so began to turn to anthropology and psychology,

11. Erin Ann O'Hara & Douglas Yarn, *On Apology and Consilience*, 77 WASH. L. REV. 1121 (2002); Erin Ann O'Hara, *Apology and Thick Trust: What Spouse Abusers and Negligent Doctors Might Have in Common*, 79 CHI.-KENT L. REV. 1055 (2004).

found competing theories there, went back to economics and behavioral biology to see if I could use those tools or that reasoning to sort through competing theories, and now find myself in the thick of not just interdisciplinary but multidisciplinary work. I've always assumed that I was just using a theory of rational behavior but trying to get a richer sense of what that rational behavior was and just woke up one day and asked, "am I talking about rational behavior anymore?" Or, is this theory of human behavior something that can no longer be classified as rational? I think we can define rationality so broadly or so loosely that I can put all of this in, but then what work is the rationality doing for me anymore? It's a problem I'm struggling with.

OLIVER GOODENOUGH: I am here wearing two hats. One is as somebody who does work in the field. The other is in my capacity as a representative of the Gruter Institute.

CLAIRE HILL: I want to acknowledge again with thanks the co-sponsorship of this symposium by the Gruter Institute, which you're here representing, and the Society for Evolutionary Analysis in Law, which Erin O'Hara is representing.

OLIVER GOODENOUGH: We all know the importance of symbolism and creating positive feelings. But this conference provides an example of a strategic affection, where we have positive feelings and also a great result. With my Gruter Institute hat on for a moment, I want to tell those of you who haven't heard of the Gruter Institute or been familiar with its operations that it is an organization dedicated to pursuing exactly the kind of questions that are on the table here and through the kind of methodology that we've seen used. I'll make as well a short plug for SEAL, the Society for Evolutionary Analysis in Law: it is well worth investigating for anybody who is interested in the kinds of topics we've discussed today. Both Gruter and SEAL have websites on which you can learn more, www.gruterinstitute.org and www.sealsite.org.

To put my scholarly hat back on: first of all, let me address the organizing question of this conference, do you have to choose rationality or irrationality? To begin, I argue with the choice of "irrational." I would prefer to phrase it as non-rational because I think that there's perhaps some different loading on the two choices if phrased that way. Rationality vs. irrationality is a very gross taxonomy, and while it has a certain kind of folk psychology life, which probably means there's some sense in it, I think that we're in a position to move beyond such a simple dichotomy. There are currently on the table a lot of different ways to view what the organizing question includes or might mean. Sometimes when the economists ask

it, they may be asking whether everybody really behaves as though they were pursuing some rational goal, even if it's not a sensible one. You can "rationally" arrive at really poor positions if you follow certain kinds of choice strategies that are sometimes linked with rationality.

So I took to heart some of Kevin McCabe's statements, linking rationality in a sense with competence as a way of approaching things. It's often been linked by others with transparency to the consciousness, which is actually not a good way to think about it. It's also often linked to coolness as opposed to affect, which is also not a good way to think about it. People can often use affect, meaning unconscious processes, to arrive at very good results, particularly in the social domain, where we have a number of processes that help us to get to much better cooperative solutions than we would if we were doing the "rat choice," prisoner's dilemma kind of process, through which we'd all end up in the wrong spot. We solve those social problems everyday through the kinds of intuitive processes that successfully bring us together and make us friendly and cooperative.

So I think that it is time we got tools on the table to begin to unpack how people really think, as opposed to making assumptions about it. I disagree with the notion that pushing rationality as far as we can is the most sensible way to run a research program at this point. Rather, we're at a stage where we can begin to make progress partly through classic psychology and its methods, and also through the new methods of cognitive neuroscience, and begin to tease out what it is that people are really doing when they make their choices about how to behave day to day. I think we'll find that the choices conform to different kinds of models in different cases.

My own work in this area includes pursuing a new model, which is to think of law in terms of being a culturally embedded and transmitted recruitment mechanism that allows us to make normative judgments by bringing other processes to bear beyond the sort of hyper-determined hot issues, introducing a cognitive richness that includes deductive logic centers and all kinds of things that we wouldn't necessarily use in a more "natural" state.¹² This recruitment allows us to improve our results. As a lawyer, I like to think of the law as the most competent means for social regulation that we've got, but that may not be saying very much. The results that we see when we look at developing countries are suggestive; however, they show that one of the best predictors for economic growth is the availability of rule of law. It's much more important than tax structures

12. Oliver R. Goodenough & Kristin Prehn, *A Neuroscientific Approach to Normative Judgment in Law and Justice*, 359 PHIL. TRANSACTIONS ROYAL SOC'Y LONDON SERIES B, BIOLOGICAL SCI. 1709 (2004).

or anything ephemeral like that. The other highest predictor is the level of general trust in society. I'm not sure that a rational choice person would have chosen either of these as the most likely predictors.

I am also working on how we set up reliable structures for cooperation, using brain science and other empirical approaches.

TODD PREUSS: I want to talk a little bit about what I think we know about human psychological organization and how that bears on the questions on rationality and non-rationality. It's clear to me that people can be rational under certain circumstances, especially when they have the information to evaluate the likely outcomes of their behavior. That's usually not the case and that's probably not usually how people govern their behavior. We know from neuropsychological studies of people with brain lesions—even before we had new techniques of neuro-imaging—that the human brain is a collection of behavior-generating devices, and these devices operate in parallel. We have a mechanism—probably a frontal lobe mechanism—that damps down some responses and potentiates others. I'm sure many are familiar with the observations of Tranel and Damasio, whose patients with the frontal lesions are capable of being very rational when questioned about responses to a hypothetical problem, but whose real-life behavior is maladaptive. Orbitofrontal lesions leave these individuals unable to evaluate their gut feelings, and gut reactions can provide useful information in certain circumstances. In addition, there are automatic response-preparation systems that operate outside of awareness. Particular events and stimuli in the environment imply particular actions; this is sometimes referred to as *affordance*. Normally, potential conflicts between different actions are resolved by the frontal lobe, but lesions can disable such executive control. One dramatic example of this is known as the “alien hand sign,” perhaps better known as the Dr. Strangelove syndrome. Recall that Dr. Strangelove has to use his left arm to keep his right arm from giving the Nazi salute. Similarly, patients with certain brain lesions, when presented with a familiar tool or instrument, cannot help but reach out and grasp it. Other lesions can disinhibit even much more complex behaviors. The French neurologist Francois Lhermitte has documented many cases like this, all of which involve frontal-lobe damage. Lhermitte would leave a familiar object like a comb on the desk and instruct the patient explicitly not to use it—to pay no attention to it—but thirty seconds later the patient would pick up the comb and start combing his or her hair. For another patient, a nurse, Lhermitte put a syringe out on the table and within a short time, without any explicit prompting, the nurse had him

standing up and was giving him an injection. When asked why they did these things, the patients would say that they were supposed to.

Objects and situation, then, automatically potentiate characteristic actions, even quite complex actions.

We seem to be a bundle of these relatively automatic systems that are somehow in normal circumstances kept under check. The question that we come to is, what elicits these things, are these rational choices that we make, or are we often just simply put in social situations where the demand characteristics of the situation are so strong that we perform these behaviors without even thinking about them? Much of our behavior is probably of the latter type. But what happens when off-the-shelf solutions don't suffice to get the job done? Then people have to make a conscious, reflective choice, as opposed to simply a *response*. Even in these situations, however, our choices can be influenced by our knowledge, our powers of reasoning, and our gut feelings. When there's no obvious, rational solution, at least in terms of what's best for the individual, we have default responses. Here we enter the realm of the anthropological. There is security and value in the group no matter what may happen to us as individuals; what we do under circumstances of uncertainty is much more a matter of what's culturally prescribed, culturally acceptable. The behavior of people in dangerous situations like combat is a very good example. The behavior of the participants has little to do with rational choice and personal benefit and everything to do with the group.

Having said these things, I want to emphasize just how little scientists know about the human brain—the emphasis here being on *human*. Functional imaging and lesion studies give us, for some, clues to the relationship between human brain structures and human behavior. We really don't know, however, how the systems of the human brain are organized and how those systems can be modified. Neuroscientists, I have to say, for the most part don't study human brains. They study mouse brains; they study monkey brains; they don't spend very much time studying human brains. Human brains, I believe, are different in important ways from those of rats and monkeys, and there are things we need to know about human brains that can only be learned by studying humans directly and by comparing humans to closely related animals, such as chimpanzees. I'm inclined to think that human behavioral control systems are really unusual. I suspect we're the only animals that face the dilemma of having a whole series of ready-made outputs available for a certain situation and are forced to choose from among them. This, as much as anything else, could define us as a species.

RICHARD WARNER: I want to begin by emphasizing that rationality is a normative ideal. And like all normative ideals, we just approximate—we never get there. In this case, Aristotle is probably the first to be crystal clear about this but it is also certainly present in Kant. Both Aristotle and Kant think of a person who is characterized as rational as having a capacity to control the issuance of behavior that's produced by various non-rational systems. They don't think we do a very good job of it. So from their point of view, while we may be shocked at the extent of irrationality, we certainly shouldn't be surprised.

So having said that, the next point I want to emphasize is that there are lots of different ideals of rationality. The ideal of rationality appropriate to trying to evaluate the most recent theories of quantum mechanics is quite different than the ideal of rationality appropriate to our law school dean trying to figure out how to steer this unruly faculty down the appropriate road. There are quite different ideals, and it's just a mistake to think, "well, how do we define rationality?" The question is, what ideal of rationality are you going to define? There are multiple ideals, which may conflict. There's a very interesting question that worries a lot of philosophers—well, given that we've got multiple ideas, how do you choose the one? Perhaps there are ways, perhaps there are not. But, again, there are multiple ideals of rationality, which means I can't really answer the question how long I should talk because that depends on what ideal you choose. Clearly, some people chose one; some another.

So the other question, do we have to choose rationality or non-rationality? Most of the panelists have considered it so far with a view to others—should I view others as rational or non-rational? Well I would ask it about each of us. I'll pick myself as an example. Do I have to give up the ideal of asking for the best reasons? My answer is, I can't. Insofar as we are persons—this is the tradition about persons from Plato on, with a small deviation in views. But insofar as we are persons, we are essentially beings that ask for reasons. It's the attempt to ask for reasons for what we do that gives our lives the continuity and characteristics of persons. It is the continuity that makes me the same person now that I was when I was eleven years old. I remember my eleventh birthday party; I don't remember somebody else's birthday party, even though I've changed tremendously over time. So to be a person is to essentially be a being that asks for reasons and evaluates himself or herself in light of ideals and rationality. We can't *not* do that. That's why normative theories of rationality sometimes yield correct predictions: to the extent that we succeed in our attempt to act for good reasons, normative theories of rationality have some predictive value: we

do what the rational person would. So that's what I have to say: we don't have a choice. We must persist in our rather poor attempts to approximate to ideals of rationality.

My work in this area focuses on one ideal of rationality that I've worried about quite a bit and that is this, if I'm to choose between mutually exclusive actions, is my choice rational only insofar as it can at least in principle be justified by showing that the reasons for the choice of action are at least as good if not better than reasons for the rejected alternatives? Should I always ask for the best reasons? This may seem the obvious truth. I think it's false, and I think it's importantly false. I think we fail to understand the relation between self, freedom, and indeed the relation between free selves and public choice government by the state if we don't get a better picture of this ideal. The reason I think it is false is easily illustrated by promising. I promised Claire that I'd come to this panel. And I wake up today and it's a really nice day. What am I suppose to do? "Really nice day, Claire's symposium. Really nice day. Claire's symposium. Oh, it's close, but Claire wins." If I do that, then I don't understand what promising is. What I'm supposed to do is reason this way: "I promised; the issue is already decided about going to the symposium." I exclude the nice day as a reason for not going. That doesn't get counted in. We see this kind of reason preemption or reason exclusion throughout a variety of our activities, and it's inconsistent with the conception I articulated earlier. I crossed certain reasons off the list for comparison and just ignored them.

QUESTIONS

CLAIRE HILL: I'd like to start off the question and answer session by asking a question: whether the distinction George made in his paper,¹³ between affective processes and more deliberative processes, is a broader distinction that has been observed in the brain? Has it been used, even metaphorically, in other contexts?

KEVIN MCCABE: I think that a lot of the recent spate of neurobiology papers that come out in science and nature are really playing with this idea, but they do it in a really simple way. Probably too simplistic. What they do is, they all rely on the idea of selecting between alternative patterns of behavior and they do it through creating cognitive conflicts in the brain—they set up oversimplistic conflicts: they tend to use something as simple as a

13. George Loewenstein, *Emote Control: The Substitution of Symbol for Substance in Foreign Policy and International Law*, 80 CHI.-KENT L. REV. 1045 (2005).

Stroop test¹⁴ type of cognitive conflict test. And then the argument is that conflict resolution occurs through prefrontal systems that are exerting cognitive control. Now that's the part I don't like; I think it's a metaphor for not understanding the fact that there's a cognitive trade off that is somehow arrived at, and that cognitive tradeoff requires prefrontal systems. Now I do think Jordan Grafman has a slightly better perspective on that.¹⁵ He has kind of a model that the prefrontal cortex is constantly in the process of encoding social and the life scripts for behaviors to the world and that those become reinforced and they get in conflict with each other. But they resolve through frontal systems working it out. The problem is: do you view the brain as setting up these competitions with a winner-take-all decision, or do you view the brain as coming to a consensus and thereby getting to a decision? And I don't think we know enough to know exactly the right story, but the winner-take-all story seems a little too easy.

TODD PREUSS: There's another aspect of automatic versus deliberative behavior. We are constantly interpreting our own behavior. Michael Gazzaniga's studies of split-brain patients cast an interesting light on our interpretative propensities. You can prompt a patient's right hemisphere to make some response, without informing the left hemisphere, and then ask the left hemisphere—the speaking hemisphere, in most people—why they did what they did. The left hemisphere's lack of understanding does not inhibit it from concocting, apparently in all sincerity, an elaborate and completely fictitious causal account of the behavior. It's likely that we're always interpreting our behavior in this way. Of course, our internal interpreter will usually know more about the circumstances preceding an action than the left hemisphere in Gazzaniga's studies. It will also have access to the emotions elicited by the situation. Whether inputs are sufficient to arrive at a valid causal account of our actions is another question, however. I suspect that our understanding of our own actions evolves, that we can examine the interpreter's account, and, over time, modify or discard its hypotheses—ultimately modifying the responses we make in particular situations. Thus, I think there's a complex, ongoing interaction between the automatic motoric and emotional and the more deliberative aspects of human behavior and cognition, that gives us hope that at least sometimes we can behave in rational, insightful ways.

14. Earl K. Miller & Jonathan D. Cohen, *An Integrative Theory of Prefrontal Cortex Function*, 24 ANN. REV. NEUROSCIENCE 167–202 (2001).

15. See Jacqueline N. Wood & Jordan Grafman, *Human Prefrontal Cortex: Processing and Representational Perspectives*, 4 NATURE REVIEWS NEUROSCIENCE 139–47 (2003).

RICHARD WARNER: I want to talk a bit about the distinction between affective and deliberative processes. Not about the neurophysiology, which I find fascinating.

It's important not to draw too sharp a distinction between the affective and the deliberative. This starts in Plato, and develops with Aristotle and the Medievals. They distinguish between the rational and non-rational parts of the psyche—the rational part is supposed to control the non-rational part. But think about your actual deliberations. Suppose I'm trying to decide whether to pick a job in a new city that means moving away from my daughter. What do I do? I tell myself narratives. I imagine pictures of what it would be like. I tell myself different stories, partly to see how I feel. This provides an enormous amount of information, but it also defines categories in which I conceptualize the problem, and it also assigns weights to different considerations. I can't imagine reasoned deliberation in any complex case without this. Now is this affective or deliberative? It's both. As Aristotle emphasizes in *De Anima*, "Look, all of this stuff about the parts of the psyche is just made up. You can divide up the psyche any way you want. It's just an artificial, theoretical distinction."

JACOB CORRÉ: There's no doubt that many processes and problems have both an affective and a deliberative component. But I have a simple question. I have a problem at work where I have to try to figure out what section 524(g)(4)(A)(iii)(III)(cc) of the Bankruptcy Code means. What is affective about that process?

RICHARD WARNER: Jacob, think of it as a continuum. There's some stuff that's like the Bankruptcy Code although if it were me, I couldn't think about it because the Code induces such revulsion. And then there's the stuff of the sort I described, in the middle. And then there's the stuff that's so emotion-ridden that it doesn't involve reasoning at all.

OLIVER GOODENOUGH: In practicing law, I do what I tell my students to do: I apply the smell test as well. You read the statute, you look at it, you say "is this within the letter of the law," and then you say, "how is the judge going to react?"

I think that we can recruit different cognitive pathways to get ourselves more into a deductive application, but why should that be the only kind of thing we call deliberative? It's presumed that deliberative processes are conscious, but that isn't necessarily so. Sometimes I wake up in the morning having solved a problem that has been stumping me for weeks, and I ask, "where did that come from?" It had to come from some place. It was competent but wasn't conscious—it's implicit in a lot of our discus-

sions that “deliberative” is in this cool, conscious box, but in fact it’s all over the place.

JACOB CORRÉ: I’m wondering if we have a really good idea of what’s at stake in getting the answer to the question [about rationality] right or wrong. Let me give you the one area where I know something about the competition between behavioral and traditional approaches. I’ve read my Markovitz [the traditional approach], and then I started reading Robert Schiller and Richard Thaler,¹⁶ who take quite a different approach. But when you get down to looking at how this affects how one would go about designing a portfolio, they have different predictions, but the expected results are really insignificant—they don’t even justify the transaction costs of thinking about the problem unless you really enjoy thinking about the problem.

CLAIRE HILL: Jacob this can’t be true about Thaler’s mutual fund—his whole pitch is that he’s going to get superior returns for his investors by exploiting market anomalies. If his fund isn’t doing better than more traditional funds, why would investors send him any money?

JACOB CORRÉ: His strategy is more than absolute return.

My question, though, is what are the areas where choosing between rationality and non-rationality most makes a difference?

CLAIRE HILL: My short, glib answer, is that I don’t know and I don’t care. (And besides that, my rating agency paper has a policy implication—seems to me I’ve paid my dues on the policy side for a few years) My more serious answer is that I’d hope that figuring things out, even things that at the outset don’t have the most immediate practical import, will lead to policy implications down the road. And, while that’s a self-serving thing to say, I actually believe it.

And, another more serious answer: whatever more navel-gazing high theory projects I’m pursuing, the trust project I’m working on with Erin does have concrete policy implications. We think a lot about why people might “get it wrong” in their assessments of how trustworthy others are, and how the law might help them get it right (and, interestingly, how in some cases the society is better off if people don’t get it right).

OLIVER GOODENOUGH: I’ve just been editing a symposium issue on Law and the Brain. And one of the pieces in it is entitled, “For the Law Neuroscience Changes Nothing and Everything.” The notion is that it doesn’t have to change that much partly because the law has groped its way

16. ROBERT J. SCHILLER, *IRRATIONAL EXUBERANCE* (2000); RICHARD H. THALER, *QUASI-RATIONAL ECONOMICS* (1991).

towards a number of reasonably competent mixed solutions over time. Think about the law of negligence, which combines a relatively naked appeal to intuitive notions of a “reasonable person” precautionary standard, while with laws like the bankruptcy code, you’re talking about a naked appeal to deductive word-based rule application. So the law has over time in fact groped its way towards some of these solutions. One area where I’m paying attention at the moment is property, and intellectual property in particular. The thesis is that we have some kind of property “primitives” in our brains, by which I mean cognitive equipment that enables us relatively well to conceive making property solutions about things and conceive in terms of property. This thesis supposes that property is not simply a cultural, or a legally, developed right.

And if we’ve got some basic equipment that helps us think property thoughts, is that equipment somewhat domain specific? Here’s where I get out on some limbs: that equipment may be relatively clear about tangible things, and less good in the territory of real property. And when we get to the product of imagination, intellectual property, we may not have any kind of effective link into that. So, for intellectual property we may have quite a good solution at a formal level which does access some good things in society—but which is relatively loosely attached to any of the normal normative structures of the brain. And if that’s the case, then one of the reasons why people illegally download music without a pang of regret is because they have a hard time feeling bad about it. How we structure our intellectual property system might be improved by this idea. The film and record companies are moving in this direction, attempting to personalize their products, putting their artists out front saying, “This is my stuff.”

ANNE TUCKERMAN: Can the irrational by definition push the rational into a decision so there is really no choice in it?

CLAIRE HILL: What’s the definition?

ANNE TUCKERMAN: Irrational is when it cannot be explained by rational events or trends. Irrational would be affective or instinctive.

I don’t know if anyone here has read *The Wall* by Jean Paul Sartre. Sartre’s view, existentialism, may be out of fashion these days. In hypothesizing that one can choose between rational and irrational, you eliminate completely this doctrine. Existentialism is based on the idea that there is no such thing as a principle or personal being or reaction to events. It’s just imposed from the outside.

GERALD CLORE: Part of the informational affective stuff I was talking about that we touched on is relevant. Part of the affect experience is up and down when it comes to arousal. The point of arousal is to say, “hey stupid,

this is important.” You’re forced by events to turn your attention to something new, and that’s the real point of emotion—to change your processing priorities.

ERIN O’HARA: I’m just now realizing how closely Todd’s comments relate to [Gerald’s] work. Gerald is saying that so much of our behavior is sort of predetermined by the environment we find ourselves in, which is a bit of what you’re getting to, and you seem to be suggesting that emotions are motivating factors for us that cause us to pay attention and then cause us to be motivated to behave and all of that can happen without ever having any conscious processing besides, “hey, I feel something and that’s motivating me.” But that creates a perception that we decided. But did we? Or is it “just” a perception?

CLAIRE HILL: What I don’t understand about the question is, why is it that the irrational has more tools at its disposal than the rational? Or are convincing tools somehow independent of rationality?

ANNE TUCKERMAN: It is a matter of definition. You can have all the reasons in the world to choose a certain solution to something but then the affective component is so strong that it pushes everything else away. And you can rationalize the affective part too into the equation.

KEVIN MCCABE: There’s one thing that’s really interesting from the viewpoint of what rational choice people will try to do, and it goes back to your [Anne Tuckerman’s] question and a little bit back to emotions again, which is, if you give the rational choice person the following paradigm, they’ll get mad at you. Say, look, I’m going to fill an urn with blue and red balls and associated with the blue balls is one outcome and associated with the red balls is another outcome and then I’m going to pull a ball at random from the urn. And then maybe give you two urns with different distributions involved and ask you which urn you want to accept. Rational choice here has thought a lot about how to make the best decision. They don’t know what to do when a yellow ball is pulled out of the urn. They want to be able to completely describe the universe of choices in which rational choices are made. But in fact we live in a world where the universe of choices is never so well defined, and, as you put it, is often thrust upon us. Choices we never had to think about before and never had to consider before. And it is clear that our brains are designed to live in that world, not in the world of the rational choice theorist.

FRED SCHAUER: Just a brief response to Kevin. It’s also the case that we often have the ability to control the size of the array of our choices. If I understand correctly the work about the tyranny of choice, people frequently take on the large array of choices initially and, at the first cut, cut

down to a small array of choices to make the array natural. Whether that work is accurate, I don't have the ability to decide. But it does suggest, as does other work, that we ought not to take the size of the array of choices as fixed, given that we have some control.

KEVIN MCCABE: So how is that accomplished?

FRED SCHAUER: As I understood you, you seem to be suggesting that we are frequently faced with a large array of choices, some of the options being uncertain, and we are built to deal with that. Some of this research might suggest that one way we deal with that is by immediately cutting down the range of choices so that we in the figurative sense make sure that there are no yellow balls in the urn.

KEVIN MCCABE: Another answer, and what the Iowa stock market, for instance, does: imagine you're running a candidate race in the stock market, but you're not sure you've got all the candidates, so you say, it's candidate one, candidate two, and the rest of the field. Now what's rest of field? Rest of field is just an acknowledgement that you may not have all the candidates. Now, if another candidate goes on the market, what do you do? You parlay the stocks of rest of field into the new candidate, so now you have three candidates and rest of field. And I think there's some sense in which that's what the brain does. It evaluates what it can, but it reserves the ability to make new evaluations. And how does it know when to make new evaluations? I think the emotions create the intentional mechanisms for when I have to do a stock split in the brain—for this new evaluation.

EDWARD RUBIN: First of all it seems unlikely that we would have a model of human behavior that didn't include purely rational behavior and purely irrational behavior, like the election, and that there wouldn't be a continuum between the two. It seems to me equally unlikely that the continuum is limited—that is to say, that it doesn't have different meanings of rationality, different modes of thought. The continuum would vary in terms of different modes of thought that we might or might not be willing to describe as rational or not, based on the fact that getting to the optimal solution under different conditions and different emotional states might involve very different modes of thought. It seems to me that we are just at the beginning of a process of discerning a great complex array of those options.

GERALD CLORE: I didn't address in my remarks what I thought about the questions of the conference because somewhere in the presentation I thought, "oh, everybody will have come to the same answer I did and there's no point in mentioning it." But I notice nobody did. That is, when I saw the question and the image of the brain on the invitation and the section mentioning evolution, I thought, "well, surely, the answer is no, we

don't have to choose because of course these concepts are really sort of old Aristotelian concepts, and now with evolutionary thought and neuroscience they just don't translate anymore." I spent a lot of my career arguing with other people about big clunky concepts like emotion and cognition and so forth. And by and large what's exciting about current times is what's also scary. We have to give up some of those concepts because they just don't go through the door. We need new concepts—which is, of course, easier said than done.

RICHARD WARNER: I can't resist. The big clunky ones, cognition—well, I've already expressed my skepticism of affective deliberation. The big chunky ones I'm willing to agree with you about—but I'm not willing to agree with you that thinking, pain, belief, and desire ought to be discarded. Because you might then charitably describe your hypothesis as someday we'll think that we don't think.

GERALD CLORE: I get really uncomfortable when somebody wants to take one of these old big concepts and say, "here it is in the brain." I think it is not feasible to come up with the right mapping for something so multifaceted and widely distributed in the brain as "cognition" or "affect."

RICHARD WARNER: I'm deeply uncomfortable with that. There are the specific mental ones—thinking, pain, and so on—the big chunky ones that we both share quite considerable skepticism about, especially with respect to the people saying, "now, there is cognition." What about the more normative ones like rationality?

GERALD CLORE: It's not really a concept in psychology. Psychologists never use the word.

ERIN O'HARA: I want to go back to Anne's first question: isn't it sometimes the case or maybe always the case that the irrational pushes the rational to a decision? You [Gerald] were talking about emotions as motivation—things that motivate or create in us to make a decision. I want to ask: consider the run of the mill case, where the affect comes along and pushes the deliberative into a decision. Contrast that with a situation of a person who's suffering from clinical depression. They can deliberate totally rationally about costs and benefits of the choice, but they are completely unable to make a decision, which I'll assume is because of the lack of affect. Is one of those rational and one irrational? How would you label and contrast those two situations?

GERALD CLORE: Note that [depression] may simply lead to rumination. When sad or anxious, perhaps I can't decide, because I am being impelled by affect not to decide. Each time I face a decision it isn't just that there is not enough affect, but rather that the information from the negative

affect I experience indicates that I have not reached a satisfactory decision. But, again, rational-irrational are not my categories, so our facts may not answer your question.

ERIN O'HARA: Would you call it functional or dysfunctional or would you just not even try to contrast the two situations?

GERALD CLORE: It depends on the situations.

BARBARA A. SPELLMAN: I have two big headaches. The first is about the affective-deliberative distinction; the second is about rationality.

We should know by now that there are lots of ways of being "rational" that are not deliberative in the sense of conscious. People make a lot of smart but "automatic" judgments. We come up with answers for things. And, think of the experience of the tip-of-the-tongue state—"I know the answer to this, but I don't know what it is." Is that deliberative, or is it unconscious, or is it automatic?

Regarding "rationality," the way people talk about it at this meeting has been making me crazy. First, I don't understand what you mean when you ask about choosing between rationality and irrationality. Who is making the choice—"we" as individuals or "we" as academics—and for what purpose? And, second, what is the definition of rationality—there are something like seven different definitions in play here. There is the Kahneman and Tversky literature that started the rationality question: comparing how people solve problems to normative ways of solving problems, and showing that people were not following rules of logic or statistics, and therefore people aren't rational. That's a consistency notion of rationality: to be rational we need a set of rules that are consistent and that find consistent answers to information.

There are also correspondence notions of rationality that keep coming up in the discussion—that is, do the decisions correspond with the truth of the world?

And you've got other definitions floating out there as well.

CLAIRE HILL: What I really intended for the symposium, which the discussions yesterday and today have made clear, is to consider what "work" rationality is doing in an analysis. Erin made the point very well. Are we getting any mileage out of the taxonomy? To which my answer is sometimes yes, but not always. Work I've done—what I was talking about earlier—about how it is that people reason, how it is that people experience things, how it is that people come to decisions—sometimes is best described as rational or not, and sometimes just as its own inquiry without a rationality overlay. I don't know if this helps your headache—or maybe it makes it worse?

BARBARA A. SPELLMAN: You sometimes treat rationality as a means, sometimes treat it as an end, and sometimes treat it as the difference between two outcomes . . .

CLAIRE HILL: But in terms of the Kahneman and Tversky normative stuff, my take is the same as your friend Gerd Gigerenzer's. I don't know if he's right as a matter of intellectual history, but he certainly captures the spirit of what I've read. He said at first, "rationality"—as in the rules of logic and consistency that you referred to earlier—was advanced as valuable as a concept because it was supposed to be descriptively true. But when that claim was proven wrong—when it became clear, thanks to Kahneman and Tversky, that people didn't actually proceed "rationally" as the term had been defined—the thrust and justification became normative: "well, people may not proceed this way, but they SHOULD." And of course the next wave of inquiries asks, sensibly—why "should they" be that way? Why should they proceed that way? And that's where I think we are now.

MARY ANNE CASE: I may intensify Bobbie's headache. We've talked about why we must choose rationality or irrationality. But I think the word we haven't talked about nearly as much as we should is *must*. The question is: are we taking *must* as a descriptive term—necessarily a choice as in, there is no way not to choose—or are we taking *must* as normative in terms of *should* we choose, even if we can avoid choosing, even if it's not a logical inevitability. Even if it's the latter, and my sense is that more of us are treating it as the latter than the former, then the question we have to ask is, can we, and if we can't, is it worthwhile talking about whether we must? I mean, is our ethical aspiration beyond the reach of human abilities something worth talking about?

CLAIRE HILL: Next time, I'll spend three months choosing a title for my symposium, making sure every word is unassailable and unambiguous.

RICHARD WARNER: There are several different ideas of rationality. So if you ask the question, "must I choose rationality," my first question is, which one? If the question is "should you give up trying to guide your life by reflection on reasons," my answer is you don't have a choice in that matter. One last exercise is to try to imagine actually giving that up. But if you actually, literally, gave that up, your life will cease to exhibit the continuity that's necessary. That's been the philosophical tradition forever. And you may reject the tradition, but the answer for that tradition is that you are essentially a being that asks for reasons.

MARY ANNE CASE: There is well within the philosophical tradition the notion that we are not the same person over time.

RICHARD WARNER: Yes but you don't really believe that do you? You believe you are the same person. You believe when you remember your eleventh birthday party that it's your birthday party, right?

MARY ANNE CASE: But there are those who don't.

RICHARD WARNER: Well with the possible exception of Derek Parfit, I don't know anybody else.

MARY ANNE CASE: Derrick Parfit is a pretty significant exception.

ERIN O'HARA: What I've had in mind was that people in this room are doing fascinating research. Some of us think we're still within the rational actor paradigm and some of us don't. The questions important to me are: do we have to choose sides on this one, and what's at stake if we do choose sides?

MARY ANNE CASE: Did you view that as a descriptive question or a normative question or both? That is to say, did you presume that a choice was inevitable or not?

ERIN O'HARA: Even the word "must" wasn't doing that much work for me, although it seems to be an interesting one for everyone else. I just woke up one morning and said, I think I'm working in a rational actor model. I think I'm broadening it so that it's a richer model and still a rational actor model. Can that be—does it have any meaning, and let's get together a group of people who are doing similar things. Some of them think they are still doing rational actor modeling and some people have abandoned that notion, and some people think it's wholly incoherent. Let's try to make some sense of all of this.

CLAIRE HILL: I was telling somebody the other day about my upcoming "rationality" symposium and they asked "hey, do you have anybody on the other side? Somebody who's against rationality?"

RICHARD MCADAMS: My answer to Erin is, it doesn't matter if what you're doing is rational choice. If that's the label others want to attach—we'll figure that out later. Just do the best job you can and later we'll be writing about trends in theories of human behavior and we'll decide on a category too.

CLAIRE HILL: Oliver I'm going to characterize what you're about to say as the last official comment of the symposium.

OLIVER GOODENOUGH: At the end of this symposium we finally get to what Erin wanted to talk about, which was, really a question of methodology rather than of what actually is going on in humans. The rational actor model is a bit like Newtonian physics: an acceptable oversimplification of what humans do that can be applied in certain analytic circumstances and that gets some good results. We can certainly identify some applications of

the rational actor model that have advanced our understanding of what humans are about. But we also need to recognize that the model has its limitations. When there was only Newtonian physics, empirical observation began to cast doubt on whether the model was true or just an oversimplification. When some physicists of that time, for whom it was the only form of physics and their only simplification, were faced with the fact that it's not actually the world, they said "well it should be the world, because I'm so good at it and it's so beautiful and so perfect." Instead of asking *if* you should choose rational choice theory, I think you need to ask *is this the right context* for choosing it. If it's the right tool for the problem it's terrific, but we are just wrong if we pretend that it's the only tool for the problem of describing how humans decide and do things.

CLAIRE HILL: Now I want to thank the participants, and encourage more unabashed drinking of wine so that the wine doesn't go to waste.