The Challenge of Flexible Intelligence for Models of Human Behavior

Mathew D. McCubbins
University of Southern California
Marshall School of Business
Los Angeles, CA
mmccubbins@marshall.usc.edu

Mark Turner
Case Western Reserve University
Department of Cognitive Science
Cleveland, OH
turner@case.edu

Nicholas Weller
University of Southern California
Department of Political Science
Los Angeles, CA
nweller@usc.edu

Abstract

Game theoretic predictions about equilibrium behavior depend upon assumptions of inflexibility of belief, of accord between belief and choice, and of choice across situations that share a game-theoretic structure. However, researchers rarely possess any knowledge of the actual beliefs of subjects, and rarely compare how a subject behaves in settings that share game-theoretic structure but that differ in other respects. Our within-subject experiments utilize a belief elicitation mechanism, roughly similar to a prediction market, in a laboratory setting to identify subjects' beliefs about other subjects' choices and beliefs. These experiments additionally allow us to compare choices in different settings that have similar game-theoretic structure. We find first, as have others,that subjects' choices in the Trust and related games are significantly different from the strategies that derive from subgame perfect Nash equilibrium principles. We show that, for individual subjects, there is considerable flexibility of choice and belief across similar tasks and that the relationship between belief and choice is similarly flexible. To improve our ability to predict human behavior, we must take account of the flexible nature of human belief and choice.

The Assumption that Actions Follow Beliefs

Game theoretic models are utilized across a variety of domains to address important problems such as allocation of security forces (Pita et al. 2011), allocation of health care services (Roth 1990), and the design of institutions (Kagel and Roth 1997). Even a survey of surveys would be beyond the scope of this paper. (For a start, see Fudenberg and Tirole 1991, Ordeshook 1986, Nisan et. al. 2007, Tirole 1988.) Despite the prominent, and often quite successful, applications of game theory in these settings, we also observe many situations in which behavior does not accord with the predictions derived from game theory (for a survey see Camerer 2003).

To address the discrepancy between predicted and actual behavior, scholars have taken various approaches. Three of the most common approaches are to propose that discrepancies arise from (1) cognitive biases and dysfunctions in the decision-making of players (Kahneman and Tversky 1979; Rabin and Thaler 2001; Ainslie 2001, Elster 1999); (2) mismatches between a game's payoffs and an individual's utility (Hoffman et. al. 1994, Rabin 1993); and (3) the effects of uncertainty, bounded search ability, or limits in thinking about others' likely behavior (Simon 1957, Gigerenzer and Selten 2002, Stahl and Wilson 1994; Crawford and Costa-Gomes 2006).

Although experimental subjects regularly make choices that do not comport with Nash equilibrium strategies (or even von Neumann-Morgenstern utility maximization), this does not imply that human reasoning is flawed. Rather, human intelligence is flexible, creating enormous diversity of beliefs and choices, the challenge is that the models to which we put them to the test are not flexible. Humans are able to solve many tasks that are quite difficult (Gigerenzer 2000, 2008: Turner 2009). To build a better theory of human behavior, we must start with an appreciation for how we actually reason. As cognitive science has shown, intuitive notions of how the mind works (vision, language, memory, etc.) may be very useful for the human being to hold as scaffolding for consciousness, but they are comprehensively wrong and simplistic. Intuitive notions of how we reason are not a basis for science. How we reason must be discovered, not assumed, and certainly not borrowed from intuition

As is well known in game theory, Nash equilibrium requires players to have correct and consistent beliefs (Rasmusen 2006). To have "correct beliefs" is to regard other players as "Nash players" and to predict that they follow Nash equilibrium (NE) strategies. It is also required that players have "common knowledge" that they are all Nash players, that is, that they know that other players know that they themselves are following Nash equilibrium strategies, and so on, ad infinitum. Lupia et. al. (2010) point out that "Common Nash refinements have similar attributes. Although these refinements differ in what they allow players to know and believe, they continue to require that actors share identical conjectures of other players' strategies" (p. 106)

If players do not believe that other players will play consistent with Nash equilibria, then it is no longer true, relative to their beliefs, that their own best response is to follow a Nash equilibrium strategy. Recent experimental work has shown that subjects' beliefs frequently do not match our theoretical assumptions and that their behavior can be reasonable, given their beliefs (McKenzie and Mikkelsen 2007). Some prior work on subjects' beliefs in experimental games also suggests that subjects often have non-equilibrium beliefs (Kuhlman and Wimberely 1976; Croson 2007). In what follows, using a within-subjects design, we investigate choices in a large battery of games, and we elicit subjects' beliefs about other subjects' choices *and beliefs* in these games, using an analog of prediction markets.

Experimental Design

We report on a portion of our battery of tasks here related to the well-known Trust game (as developed by Berg et al. 1995). In our experiments, subjects know that their choices are always private and anonymous, even to the experimenters at the time of the experiment. Subjects receive no feedback, during the course of the experiment, about the consequences of their choices, except for quizzes related to our narratives/manipulations (subjects may, for some of our tasks, be able to infer the consequences of their choices). For each task, subjects are randomly matched to another subject. Thus, to the extent possible, given they were in narratives that describe games, every task is a single shot, separate from the prior and future choices. We also ensure that no subject knows anyone else in either of the two rooms of the experiment.

The Trust game involves two players. Each player begins with a $5 endowment. The first player chooses how many dollars, if any, to pass to an anonymous second player. In our experimental protocols, we use no labels other than "the other person(s)." To avoid suggesting an investment or reciprocity frame we label actions as "transfer." The first player keeps any money he does not pass. The money that is passed is tripled in value and the second player receives the tripled amount. The second player at that point retains the original $5 plus three times the amount the first player passed, and decides how much, if any, of that total amount to return to the first player. The second player at the moment of choice in the Trust game is in a role that is equivalent to the role of Dictator in the Dictator game. The subgame perfect Nash equilibrium (SPNE) is that Player 1 will send $0 and Player 2 will send $0. This is also a dominant strategy equilibrium.

These equilibrium strategies derive from assumed beliefs: the assumption is that all players maximize economic payoff and believe that all other players do the same. In the Trust game, a Player 1 with these beliefs concludes that Player 2 will return nothing and so, as a maximizer, Player 1 sends nothing. The beliefs that players hold about other players lead to the belief at every level of recursion that all players will send $0, will guess that others will send $0, will guess that others will predict that everyone will send $0, and so on ad infinitum.

But what happens if a subject with these Nash beliefs finds himself off the equilibrium path? In the Trust game, only Player 2 could make a choice after finding himself or herself presented with an off-the-equilibrium-path choice. If Player 2 is gifted with anything more than his or her $5 endowment, the subgame perfect Nash equilibrium strategy is still to send $0 back.

We add elements to the basic Trust game to tap into subjects' beliefs. Our belief elicitation mechanism borrows from the idea of a prediction market (Wolfers and Zitzewitz 2004). For the Trust game, we ask Player 1 to make two additional decisions and Player 2 to make one additional decision. We do not ask subjects to report their expectations or beliefs, because asking for a report might have normative implications. Rather, we ask them to "guess" other subjects' choices, or to guess other subjects' "predictions." In general, we try to provide little or no framing of the experimental tasks offered to our subjects. After Player 1 makes his choice about how much to pass, we ask him to guess how much Player 2 will return. Before Player 2 learns Player 1's choice, we ask Player 2 to guess how much money Player 1 passed. We also ask Player 2 to guess how much Player 1 predicted she would transfer. After Player 2 learns Player 1's choice, we ask Player 2 to guess how much Player 1 predicted she would return. All players know that all players earn $3 for each correct guess and earn nothing for a guess that is wrong.

The questions we ask vary slightly for each task, but as an example, here is the exact question we ask Player 2: "How much money do you guess the other person transferred to you? If you guess correctly, you will earn $3. If not, you will neither earn nor lose money." We add similar incentivized prediction tasks to various experimental tasks. Players do not learn whether their predictions were right or wrong and subjects never have any information about other subjects' guesses.

Players in the Trust game know that they are randomly paired with another subject in a different room. Later in the experiment, all subjects also make choices as Player 2, randomly assigned to the player in the other room who was Player 1. Accordingly, all subjects first make choices as Player 1 and then, roughly 90 minutes later, make choices as Player 2 (randomly assigned to a different Player 1). They thus play Trust twice, but in different roles. Player 1 never learns the consequences of any of his or her choices in the Trust game. Player 2 can of course infer the consequences of his or her own choices.

Subjects also make decisions in a variety of other games, including a Dictator game and what we call the Donation game. In both these games, each subject is randomly paired with another subject in another room. In the Dictator game, The Dictator (Player 1) and the Receiver (Player 2) have endowments identical to those the subject in the role of Dictator faced when he or she was in the role of Player 2 faced in Trust. Accordingly, the Dictator game was identical right down to the specific endowments to the second half of the Trust game. In effect, each subject replayed the second half of the Trust game, but now without the reciprocity frame. The SPNE is for the Dictator to send $0 to the Receiver. The Donation game is identical, except that each player begins with a $5 endowment and the amount Player 1 chooses to send is quadrupled before it is given to Player 2 (making it roughly similar to the choice faced by Player 1 in the Trust game). The SPNE is again for the Donor to send $0.

Our subjects also play, among other things, a unanimous Public Goods game with nine other players, randomly assigned. Each of the ten players in the group is endowed with $5. In these games, players must decide whether to keep their $5 or contribute all $5 to a "pot." In this task, if all players contribute their $5 endowment to the "pot," then the money is tripled, and the money is distributed equally to all players, in which case each player gets $15. There are two pure strategy Nash equilibria to this game. In one, no one contributes, and each subject keeps the $5 endowment. In the second, everyone contributes, and everybody earns $15. Subjects' behavior in this step-level unanimity Public-goods game is conditioned on their beliefs about which equilibrium will arise. If a player believes all other nine players are going to contribute, then he or she should contribute; if not, then not.

At the end of the experiment, we present the subjects with the few tasks that would allow them to learn something about the choices made by subjects in the other room. So, for example, they are presented the tasks for Player 2's choice in the Trust game as one of their final tasks. In this last stage, we have no choice but to provide subjects with feedback in the form of information about what other subjects have done. For example, Player 2 in Trust must know what Player 1 chose to send. Only in this last stage, then, is there any chance for learning or development of individual or group reputations. The order of experimental tasks is identical for all subjects in the experiment except during the public goods game, where by design we systematically manipulate the order of two tasks: (1) making choices and (2) guessing the choices made by others.

The subjects in our experiment completed the tasks using pen and paper in a controlled classroom environment. Subjects were recruited using flyers and email messages distributed across a large public California university and were not compelled to participate in the experiment, although they were given $5 in cash when they showed up. A total of 180 subjects participated in this experiment. The experiment lasted approximately two hours, and subjects received on average about $41 in cash. The experiment was followed sometime later with a post-experiment questionnaire, for which subjects were also paid.

Uncovering Subjects' Beliefs

In many of our tasks, we ask subjects to make guesses about other players' actions and predictions. We pay the subjects $3 for correct guesses. All subjects in our experiments know this. Do subjects believe what game theory assumes they believe? The answer is, mostly, no.

The subgame perfect Nash equilibria (SPNE) in the Trust Game is that neither Player 1 nor Player 2 will send any money to the other. All should believe that all others will predict that no one will send money, and all such beliefs should be recursive, so that Player A believes Player B believes Player A believes Player B will send no money, and so on for any number of steps and for any subject in any role.

But we see quite the contrary in our experiments: Figure 1 shows that only 68 of 180 subjects as Player 2 believe that Player 1 will send nothing. In other words, 62% of subjects have "incorrect" beliefs, this is beliefs contrary to those that support SPNE strategies.

Figure 2 shows guesses made by Player 1 of the amount Player 2 will return. We include even the Player 1s who sent nothing. (Since Player 2 begins with a $5 endowment, Player 2 can transfer money even if Player 1 sent nothing.) Ninety-two of the 180 subjects guess that Player 2 will return $0, but 88, or 49%, believe that Player 2 will return some money. This means that 49% of these subjects have "incorrect" beliefs. Their beliefs diverge broadly from SPNE, across a large span of possible returns.

We also investigate beliefs in other games. We have 160 of our 180 subjects participate in a unanimous Public Goods game (in one session, one room received flawed instructions for this task and thus the behavior for all 20 subjects is dropped from our analysis of this task). In this task, subjects have two choices: to contribute their $5 or not. There are two pure strategy equilibria—one in which no one contributes and one in which all subjects contribute. Therefore, we are particularly interested in whether beliefs are consistent with one or the other of those equilibria. Subjects play the Public Goods game with nine other subjects, and we ask subjects to predict how many of the other nine subjects will contribute to the Public Good. We find that 47 of the 160 subjects (29.4%) guess that zero

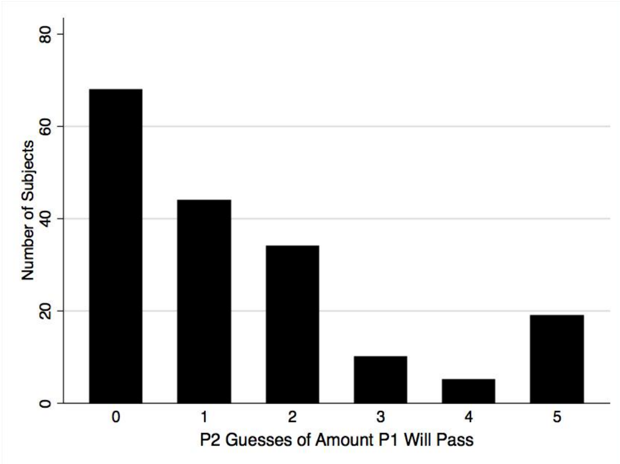

Figure 1: Player 2 guesses of amount Player 1 sent

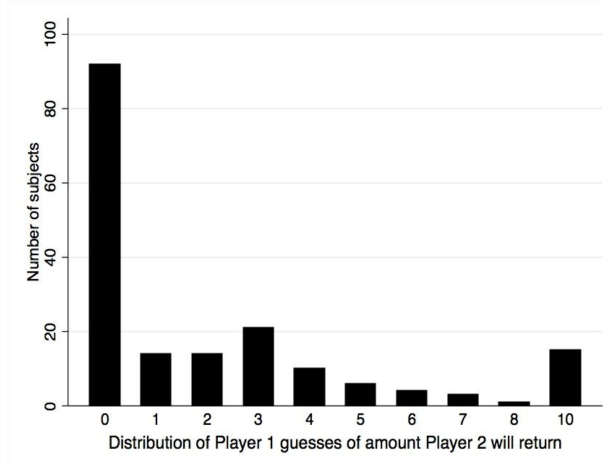

Figure 2: Many Player 1s expect money to be returned

other subjects will contribute, and 42 subjects (26.2%) guess that all nine of the other subjects will contribute. The rest of the subjects, however, hold beliefs at variance to pure strategy NE: their guesses span the range of participation levels and are neither simple nor uniform.

One important telltale in this game comes when we balance the order of the tasks. In some cases, we ask the subjects first to choose whether to contribute and second to guess the other subjects' choices. In other cases, we present the tasks in the reverse order. For the theory of games, the order of these two tasks cannot affect subjects' strategies or choices. But our experiments show that subjects who choose first guess on average that 3.3 other players will contribute, while subjects who guess first guess on average that 4.6 other players will contribute ($p=0.03$ in a Kolmogorov-Smirnov equality of distributions test), with 80 subjects in each group. Further, in an equality of proportions tests 25% of subject choose to contribute when making their choice before their prediction, whereas 43% choose to put their money in the pot when prompted about their beliefs before they made their choice ($p<.03$). This result suggests that changing the order of belief elicitation and choice significantly affects subjects' beliefs. This simple change in task order does not accord with Nash equilibrium expectations. Are the subjects who guess after they choose simply winging it first and rationalizing later, or are the others simply winging their guesses first and then choosing according to something else later?

While we don't know where beliefs come from, we can compare subjects' beliefs about others in one part of the Trust game with their choices in that same part of the Trust game. For example, we can examine the difference between what a subject choose to do as Player 1 in the Trust game, and what they believe as Player 2 that Player 1 will do. The modal category is subjects believe that other subjects will play like them: 109 of the 180 subjects guess that the choice of the Player 1 with whom they are randomly matched will be the same as their own choice when they were Player 1. For these subjects, theory of mind might equal theory of self, or this may simply represent the "false consensus" effect in which people think others are more like them than they actually are (Ross et al. 1977), or it might be akin to the curse of knowledge, but we can't really tell. Perhaps most surprising, there is a large variance, with 71 subjects (39%) making guesses that differ from their own choices.

Consistently Inconsistent

The standard approaches to explaining departures from NE strategies (other-regarding preferences, cognitive constraints, or decision-making biases) implicitly assume that players deviate from game-theoretic expectations in consistent ways. For example, if players prefer to reduce inequality, that preference should be stable across all manner of economic games. Or, if players cannot perform backward deletion of dominated sub-games, as game theory requires, then this handicap should operate in all game environments of equal difficulty. To date, there has been little focus on identifying the extent to which players have consistent beliefs or behavior across games.

Cognitive science gives us considerable reason to doubt that players will behave identically across different environments, because changes in environment lead to changes in mental activation, which affects beliefs and behavior. As Sherrington famously wrote, the state of the brain is always shifting, "a dissolving pattern, always a meaningful pattern, though never an abiding one" (Sherrington [1941], 1964). If the particular tasks, and order of those tasks, induce different mental activations, then belief and behavior should vary accordingly. Our experiment is designed to shed light on whether subjects have consistent beliefs and make consistent choices.

Our first cut at this question is simply to examine the number of subjects who have beliefs consistent with NE across a variety of tasks. In the Trust game, subjects make predictions as Player 1 about the behavior of Player 2 and as Player 2 about the behavior of Player 1. We already demonstrated that in either single task, a great many subjects do not have SPNE beliefs. In Table 1, we show the number of subjects with Nash beliefs and non-Nash beliefs as both Player 1 and Player 2 in the Trust game. If Player 1 has Nash beliefs, it means that this subject guessed that Player 2 would return nothing. If Player 2 has Nash beliefs, it means that the subject guessed that Player 1 would send nothing. Overall, out of 180 subjects in our analysis, only 63 subjects made guesses as both Player 1 and 2 that were consistent with Nash beliefs. In other words, only 35% of our subjects have consistently "Nash beliefs" *even inside this one game*.

		Player 2 Nash beliefs	Player 2 Non-Nash beliefs
Player 1	Nash beliefs	63	29
	Non-Nash beliefs	5	83

Table 1: Subjects' beliefs in Trust game (N=180)

There were 83 subjects who lacked Nash beliefs in both part of the Trust game, 29 subjects who possessed "Nash beliefs" as Player 1 but not as Player 2, and only 5 subjects who possessed "Nash beliefs" as Player 2 but not as Player 1. Our experiment does not allow us to identify why players' beliefs diverge from the NE beliefs, but it is clear that most subjects deviate from "Nash beliefs" during at least one of the experimental tasks.

We can also examine the actions/choices of individual subjects across a number of similar tasks to see if individual subjects behave consistently. In particular, we look at subject behavior across a set of tasks, all of which involve choosing how much money to transfer to another person, and in which the outcome of that decision is not contingent on the other person. In the Trust game, subjects play the role of Player 1 and 2 during the course of the experiment. There were 60 subjects who were "fully Nash actors" throughout the game; that is, they chose SPNE strategies (i.e., $0) as both Player 1 and Player 2.

Another way to investigate consistency of behavior is to examine the choices of subjects who as Player 2 in Trust received money from Player 1. Of the 100 subjects who received money as Player 2 in Trust, only 62 returned any of the money to Player 1. Additionally, of those 62, only 40 sent money in the Dictator game. This shows that many subjects do not behave consistently in these two tasks, in which their actions could reduce inequality. Further, of the 40 who sent money in Dictator, only 29 also send money in the Donation game. This means that of the 100 subjects who received money as Player 2 in Trust, only 29 sent money to the other player in all three related tasks. This shows that subjects do not behave consistently even in their violations of SPNE.

We have seen, and the literature broadly documents, that subjects deviate remarkably from Nash equilibrium strategies. We have reported how subjects' beliefs deviate from those necessary to support equilibrium strategies. We have also shown that these deviations are not consistent. Accordingly, it is doubtful that proposals to explain deviation from Nash strategies and beliefs by attributing to subjects a particular consistent mental or behavioral signature will succeed.

Does behavior accord with beliefs?

Now, we ask whether actions are minimally rational, that is, do subjects' actions accord with their beliefs? To begin, we investigate whether action and belief accord in the Trust game. Figure 3 shows the decisions of subjects as Player 1 about the amount of money to send to Player 2 in the Trust game. Over one-half of subjects (100 out of 180) pass money, which is inconsistent with a SPNE strategy, and on average subjects pass $1.43. Of the 100 subjects who receive money as Player 2, 62 of them return some money to Player 1. On net, Player 1 loses money. This result—that players choose to pass some money in the Trust game—has been well-documented (Berg et al. 1995).

Figure 4 displays the difference between the amount Player 1 sends to Player 2 and the amount Player 1 guesses Player 2 will return. Recall that any money sent by Player 1 is tripled before it is sent to Player 2, (e.g., if Player 1 sends all $5, then Player 2 has $20, and if Player 2 splits that money, then Player 1 and Player 2 end with $10 each, and we would say that each has "earned" $5 through their actions). Figure 4 shows that there are only a few players

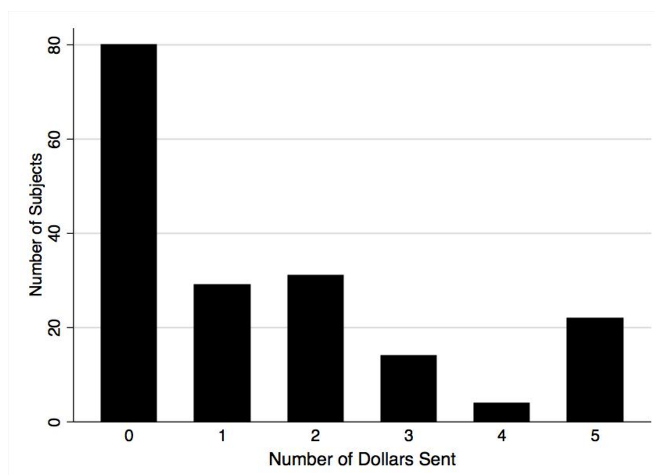

Figure 3: Distribution of amount Player1s pass in Trust game

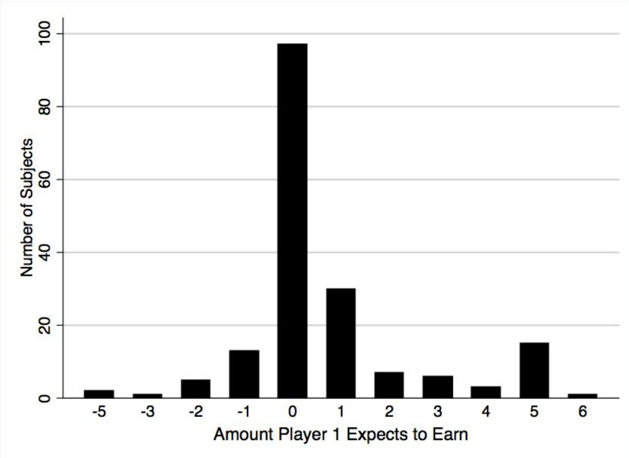

Figure 4: Guesses by Player 1 of profit from choice

who guess that they will lose money by sending money to the other player. Mostly, players expect to benefit from their decision. The beliefs held by these players imply not only that they do not expect others to play consistently with SPNE strategies, but also that they expect, on average, to profit from their non-SPNE strategy to send money.

There are 100 subjects who as Player 1 in Trust chose to send a positive amount to Player 2, and 20 of those players guess they will not receive anything in return. These 20 players guess that Player 2 will follow a SPNE strategy. These 20 subjects cannot simultaneously be maximizing their payoffs and hold the belief that Player 2 will follow a SPNE strategy of returning $0 so it is hard to see how their choices accord with their own beliefs. We must either conclude that they are not payoff maximizers or relax the assumption that subjects act according to beliefs. One way to relax that assumption is to give up the assumption that believing, preferring, deciding, and acting are simultaneous and coordinated mental events. Perhaps subjects act without fully activating their decisions, or believe without activating the consequences of those beliefs for action, or act without activating beliefs, and so on.

Recall that earlier we identified 60 subjects who were “fully Nash actors” in the Trust game, that is, the subjects whose actions as both Player 1 and 2 were consistent with SPNE strategy. We turn now to these 60 subjects and examine whether their beliefs are “fully Nash” in the Trust game. The answer is no. First, let us consider these 60 subjects in the role of Player 1 in Trust. Of these 60 subjects, 56 guessed as Player 1 that Player 2 would return nothing, which is consistent with SPNE. They also guessed Player 2’s prediction of the amount they will pass. Only 40 of the 60 “fully Nash” Trust players (66%) guessed that Player 2 predicted that they would transfer $0. The other 20 of the 60 “fully Nash” Trust players (1/3rd) lacked that SPNE belief. They also guessed Player 2’s prediction of Player 1’s guess of what Player 2 will return. Of the 60 subjects, 49 (81%) had beliefs consistent with SPNE. These results show that even the 60 “fully Nash” Trust subjects hold beliefs whose degree of consistency with SPNE principles varies question by question even when we look at only those questions asked of them when they are in the role of Player 1. Beliefs show flexibility.

We next turn to the beliefs of those 60 “fully Nash” Trust subjects when they are in the role of Player 2 in Trust. Of the 60, 44 guess that Player 1 will transfer nothing; that is, 16 of 60 (27%) lack SPNE beliefs. Of the 60, 35 guess that Player 1 predicts that they will return nothing; that is, for this question, 42% of these 60 “fully Nash” Trust subjects have beliefs that are inconsistent with SPNE. Overall, non-SPNE beliefs are quite common even among the 60 “fully Nash” actors in the Trust Game. Beliefs show flexibility

Next, we ask whether the 60 “fully-Nash” actors in Trust are “fully Nash” in the related Donation and Dictator games. Here we find that 57 of the 60 subjects pass $0 in the Dictator game and 50 of the 60 pass $0 in the Donation game. If we focus on those 57 subjects who are “fully Nash actors” as both Player 1 and Player 2 in Trust and also as Dictator in the Dictator game, we find that 48 of the 57 (84%) pass nothing in the Donation game. Therefore, across our entire 180 subjects, only 48 (27%) have consistent Nash behavior in Trust, Donation, and Dictator. The results from our battery of experimental tasks demonstrate that subjects regularly deviate from SPNE in both their beliefs and behavior, that the deviations are themselves inconsistent, and that there is variation in the degree to which behavior accords with belief.

Discussion

Our results show, as is usually shown, that subjects deviate from game-theoretic predictions. Our findings also show that these deviations are not consistent; they depend on the specific setting and task. Our results demonstrate that beliefs are also inconsistent. These deviations are so pervasive and so various even within subject that is seems unwarranted to refer to them as “deviations.” On the contrary, consistent “Nash behavior and beliefs” appear to be remarkable deviations from human cognitive patterns and human behavioral norms. It may be that people can be trained to comply with these deviations, at least to some extent, for rare and specially-designed cultural conditions, such as strategic board games, under the additional stipulations that the other human beings in the story are somehow constrained and also trained to be deviant in the same ways.

Our results are not consistent with a view that decision and action are coordinated around inflexible beliefs and preferences. Rather, different tasks and settings appear to lead to different mental activations in subjects, and subjects respond flexibly. This flexibility has not been well-appreciated by existing approaches to modeling human action in economic settings. Research into decision-making should turn now to the goal of discovering what those cognitive patterns of decision-making actually are.

Acknowledgments. McCubbins acknowledges the support of the National Science Foundation under Grant Number 0905645. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Turner acknowledges the support of the Centre for Advanced Study at the Norwegian Academy of Science and Letters.

References

Ainslie, G. 2001. *Breakdown of will*. Cambridge: Cambridge University Press.

Aumann, R., and Brandenberger, A. 1995. Epistemic conditions for Nash equilibrium. *Econometrica* 63:1161–80.

Berg, J. E., Dickhaut, J., and McCabe, K. 1995. "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, Vol. 10, 1995, pages 122-142.

Camerer, C. 2003. *Behavioral Game Theory.* Russell Sage Foundation, New York, New York/Princeton University Press, Princeton, New Jersey

Costa-Gomes, M. A., and Crawford, V. P. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*. vol. 96(5), pages 1737-1768.

Croson, R. 2007. "Theories of Commitment, Altruism and Reciprocity: Evidence from Linear Public Goods Games." *Economic Inquiry*, Vol. 45, pp. 199-216.

Elster, J. 1999. *Strong Feelings: Emotion, Addiction, and Human Behavior*. MIT Press.

Engelmann, D., and M. Strobel, 2004. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments," *American Economic Review*, vol. 94(4), pages 857-869.

Fudenberg, D., and J. Tirole. 1991. *Game Theory*. MIT Press.

Gigerenzer, G. 2000. *Adaptive thinking: Rationality in the real world.* New York: Oxford University Press.

Gigerenzer, G. 2008. *Rationality for mortals: How people cope with uncertainty.* New York: Oxford University Press.

Gigerenzer, G., and Selten, R. 2002. *Bounded Rationality: The Adaptive Toolbox*. MIT Press

Hoffman E., McCabe, K., Shachat, K. and Smith, V. 1994. "Preferences, Property Rights, and Anonymity in Bargaining Games." *Games and Economic Behavior* 7(3): 346-380

Kagel, J.H. and Roth, A.E. eds. 1997. *The Handbook of Experimental Economics*. Princeton University Press.

Kuhlman, D. M., and Wimberley, D. L. 1976. "Expectations of choice behavior held by cooperators, competitors, and individualists across four classes of experimental game." *Journal of Personality and Social Psychology, 34*, 69-81.

Kahneman, D., and Tversky, A. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, Vol. 47, No. 2. pp. 263-292.

Lupia, A., Levine, A. S., and Zharinova, N. 2010. "Should Political Scientists Use the Self Confirming Equilibrium Concept? Benefits, Costs and an Application to the Jury Theorem." *Political Analysis* 18:103-123.

McCubbins, M. D., and Turner,M. In press. "Going Cognitive: Tools for Rebuilding the Social Sciences." In Sun, Ron, ed. *Grounding Social Sciences in Cognitive Sciences.* Cambridge MA: MIT Press. Chapter 14.

McKenzie, C. R. M., and Mikkelsen, L. A. (2007). "A Bayesian view of covariation assessment." *Cognitive Psychology*, 54, 33-61

Nisan, N., Roughgarden, T, Tardos, E., and Vazirani, V. V. 2007. *Algorithmic Game Theory*. Cambridge University Press.

Ordeshook, P. C. 1986. *Game Theory and Political Theory*. Cambridge University Press.

Pita, J., Kiekintveld, C., Tambe, M., Steigerwald, E. and Cullen, S. 2011. "GUARDS - Innovative Application of Game Theory for National Airport Security." In *Proceedings of the International Joint Conference on Artificial Intelligence* (IJCAI).

Rabin, M. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review, LXXXIII*. 1281-1302.

Rabin, M. and Thaler, R.H. 2001. "Anomalies: Risk Aversion," Journal of Economic Perspectives. vol. (1), pages 219-232

Rasmusen, E. 2006. *Games and Information.* Oxford: Blackweller Publishers, Fourth Edition.

Ross, L., Green, D., and House, P. 1977. "The 'False Consensus Effect': An Egocentric Bias in Social Perception and Attribution Processes." *Journal of Experimental Social Psychology*. 13, pp. 279-250.

Roth, A.E. 1990. "New Physicians: A Natural Experiment in Market Organization," *Science*, 250, pp. 1524-1528.

Sherrington, C.S. Sir. [1941] 1964. *Man on his Nature*. [The Gifford Lectures, Edinburgh, 1937–1938. New York: The Macmillan Co.; Cambridge: The University Press, 1941]. New York: New American Library.

Simon, H. 1957. "A Behavioral Model of Rational Choice", in *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. New York: Wiley.

Stahl, D. and Wilson, P.1994. "Experimental Evidence on Players' Models of Other Players," *Journal of Economic Behavior and Organization*, 25, 309–327.

Turner, M. 2009. "The Scope of Human Thought." http://onthehuman.org/humannature/

Tirole, J. 1988. *The Theory of Industrial Organization*. MIT Press.

Wolfers, J. and Zietzwitz, E. 2004. "Prediction Markets." *Journal of Economic Perspectives.*