CHAPTER

13

c0013

Behavioral Game Theory and the Neural Basis of Strategic Choice

Colin F. Camerer

OUTLINE

s0010

GAME THEORY

p0010 Game theory is a very general language for modeling choices by purposive agents in which the actions of other agents can affect each player's outcomes. The elements of a game are players, strategies, information, a structure or "game form" (e.g., who chooses when), outcomes which result from all players' strategy choices and information, and preferences over those outcomes. The generality of this language makes it applicable to many levels of analysis (from biology to international politics). Examples include evolutionary adaptation of genes competing on the basis of fitness, face-to-face bargaining, competition and collusion among firms, behavior in rule-governed games like chess and poker which are computationally demanding, inferring something about a person's preferences or future intentions from their actions, delicate diplomatic bargaining in which language has multiple meaning to different audiences, and designing rules to raise the most revenue from government auctions of scarce resources.

The aim of this chapter is to introduce some of the essential components of game theory to neuroscientists (for more details, see Chapter 5 in this volume), and to summarize some emerging regularities from psychological and neural analysis which suggest ways to put the theory of behavior in games on a biological basis. p0020

Game theory is useful in two ways. Simply specifying the details of a strategic interaction and matching p0030

GLIMCHER 978-0-12-374176-9 00013

CH013.indd 193 5/12/2008 6:19:38 PM

it to familiar categories of games (such as the prisoners' dilemma game) is helpful as a language to describe a situation and point out its crucial features (Aumann, 1985). Classification of this sort can be helpful even if there is no deep mathematical analysis of what players are likely to do.

p0040 Most of game theory, however, analyzes what players are likely to do once a game is fully specified mathematically. Analytical game theory assumes players choose strategies which maximize utility of game outcomes given their beliefs about what others players will do. This means that the most challenging question is often how beliefs are formed. Most theories assume that beliefs are derived from some kind of analysis of what other players are likely to do, given the economic structure of the game. In equilibrium analysis, these beliefs about others are assumed to be correct, which solves the problem of how to specify reasonable beliefs by equating them with choices. While analytical game theory has proved enormously powerful, there are two shortcomings of its tools that limit its use as a complete model of behavior by people (and other levels of players).

p0050 First, many of the games that occur naturally in social life are so complex that it is unlikely that players instantaneously form accurate beliefs about what others would do and therefore can choose equilibrium strategies. It is therefore useful to consider what strategies might be chosen by players with bounded rationality, or when there is learning from repeated play. (A different approach, "evolutionary game theory", assumes that agents in a population play fixed strategies, but population pressure adjusts the statistical mixture of strategies across the population – i.e. the "market share" of each strategy in the population – so that successful strategies are reproduced more frequently.)

p0060 Second, in empirical work, only received (or anticipated) payoffs are easily measured (e.g. prices and valuations in auctions, the outcome of a union–management wage bargain, or currency paid in an experiment). But game theory takes as its primitives the *preferences* players have for the received payoffs of all players (utilities), and preferences are generally taken to be most clearly revealed by actual choices (see Chapters 3 and 4 in this volume). Inferring from strategic choices alone both the beliefs players have about choices of others, and their preferences for outcomes which result from mutual choices, is therefore especially challenging. One shortcut is to have a theory of theory of social preferences – how measured payoffs for all players from an outcome determine players' utility evaluations of that outcome – in order to make predictions. Emerging concepts of social preference and their neural correlates are reviewed by Fehr (Chapter 15) and Camerer (Chapter 13).

Hundreds of experiments show that analytical p0070 game theory sometimes explains behavior surprisingly well, and is sometimes badly rejected by behavioral and process data (Camerer, 2003). This wide range of data – when game theory works well and badly – can be used to create a more general theory which approximately matches the standard theory when it is accurate, and can explain the cases in which it is badly rejected. This chapter describes an emerging approach called "behavioral game theory," which generalizes analytical game theory to explain experimentally-observed violations by incorporating bounds on rationality in a formal way.

Like analytical game theory, behavioral game p0080 theory is efficiently honed by laboratory regularity because the structure of the game and resulting payoffs can be carefully controlled in the lab (in field applications it is usually hard to know what game the players think they are playing). However, behavioral game theory is ultimately aimed at practical questions like how workers react to employment terms, the evolution of Internet market institutions for centralized trading (including reputational systems), the design of auctions and contracts, explaining animal behavior, and players "teaching" other players who learn what to expect (such as firms intimidating competitors or building trust in strategic alliances, or diplomats threatening and cajoling).

BEHAVIORAL GAME THEORY s0020

Behavioral game theory is explicitly meant to pre- p0090 dict how humans (and perhaps firms and other collective entities) behave. It has four components: representation, social preferences over outcomes, initial conditions, and learning.

Representation s0030

How is a game perceived or mentally repre- p0100 sented? Often the game players perceive may be an incomplete representation of the true game, or some elements of the game may be ignored to reduce computationally complexity. This topic has been studied very little, however (Camerer, 1998).

One example is multi-stage alternating-offer bar- p0110 gaining. In this game, agents bargain over a sum of money, and alternate offers about how to divide the sum. If an offer is rejected, the available money shrinks (representing the loss of value from delay).

The game ends when an offer is accepted. One version of the game that has been studied experimentally has three stages, with sums varied randomly around $5, $2.50, and $1.25 (if the last offer is rejected, both players get nothing). If players are self-interested and plan ahead, the prediction of game theory is that the player who makes the first offer should offer $1.26 and the other player will accept it. (Assuming players only care about their own payoffs, the prediction of game theory comes from forecasting what would happen at every future "subgame" and working backward ("backward induction"). In the third stage, player 1 should expect that an offer of $.01 will be accepted, leaving $1.24 for himself. Player 2 should anticipate this, and offer $1.25 to player 1 out of the total of $2.50 in the second stage (just a penny more than player 1 expects to get in the third stage), leaving $1.25 for himself (player 2) in the second stage. In the first round, player 1 should anticipate that player 2 expects to earn $1.25 in the second stage and offer $1.26.) Empirically, players offer more than predicted, around $2.10, and much lower offers are often rejected.

p0120 One possible explanation is that players care about fairness. Beliefs are in equilibrium but they reject low offers because they prefer to get a larger share of a smaller sum of money. Another explanation is that players do not plan ahead. If they act as though the game will last only two periods, for example, then the equilibrium offer is $2.50; so the empirical average offer of $2.10 might reflect some mixture of playing the three-period game and playing a truncated two-period game.

p0130 Camerer *et al*. (1993) and Johnson *et al*. (2002) compared these two explanations by using a "Mouselab" system which masks the three varying dollar amounts that are available in the three rounds in opaque boxes (like in the game show *Jeopardy*!). Information is revealed when a computer mouse is moved into the box (and the box closes when the mouse is moved outside of it). They found that in about 10–20% of the trials subjects did not even bother to open the box showing the sum of money that would be available in the second or third stage. Their information look-up patterns are also correlated with offers they make (subjects who looked ahead further made lower offers). By directly measuring whether players are opening the value boxes, and how long those boxes are open, they could conclude that subjects were computing based on an attentionally limited representation of the game. Keep in mind that these games are simple and players are capable of perceiving the entire game. (In one condition, subjects play against a computer algorithm which they know is optimized to earn the highest payoff, and expects that human players will do the same. At first, subjects' looking strategies and offers are similar to those when they play human opponents. However, when they are gently told that it might be useful to look at all three amounts available for bargaining and work backward ("backward induction"), subjects learn rapidly to play the optimal strategy and open all boxes.) Further work on limited representations could study more complicated games where truncation of representations is likely to be even more dramatic and insightful.

SOCIAL PREFERENCES OVER OUTCOMES s0040

As noted above, when the payoffs in a game are measured, a theory of preferences over payoff distributions is needed to fully specify the game. This is a rich area of research discussed by Camerer (2003: Chapter 2), Fehr and Camerer (2007); see also Chapter 15 of this volume). p0140

Initial Conditions or One-shot Play s0050

Many games are only played once (a "one-shot" game), and in other cases an identical game is played repeatedly (a "repeated game"). In many games, it is not plausible that beliefs will be correct immediately in one-shot games or in the first period of a repeated game (as assumed by equilibrium models), without pre-play communication or some other special condition. Two types of theories of initial conditions have emerged: cognitive hierarchy (CH) theories of limits on strategic thinking; and theories which retain the equilibrium assumption of equilibrium beliefs but assume players make stochastic mistakes. p0150

Cognitive hierarchy theories start with the presumption that iterated reasoning is limited in the human mind, and heterogeneous across players. Limits arise from evolutionary constraint in promoting high-level thinking, limits on working memory (Devetag and Warglien (2003) show a correlation across subjects between working memory, as measured by digit span, and choices linked to the number of steps of thinking), and adaptive motives for overconfidence in judging one's relative skill (i.e., people may stop after some iteration because they think others must not have thought any further than they did). p0160

Denote the probability of a step-k player i choosing strategy *j* by $P_k(s_i^j)$. The payoffs for player *i* if the other player (denoted $-i$) chooses s_{-i}^h are given by a payoff function of both strategies denoted by $p_i(s_i^j, s_{-i}^h)$. p0170

Assume a distribution $f(k)$ of k-step types. Zero-step players choose randomly (i.e., $P_0(s_i^j) = 1/n$ if there are n strategies). k-step players form a conditional belief $g_k(t)$ about the percentage of opponents who do t steps of thinking. One specification is that k-step players guess the relative proportions of how much thinking other players do correctly, but they do not realize others might be doing k or more steps of thinking, that is

$$g_k(t) = f(t)/(S_{m=0}^{k-1} f(m))\ for\ t = 0\ to\ k-1$$

(The simpler specification $g_k(h) = 1$ for $h = k-1$ (k-steppers believe all others do exactly one less step) is often more tractable and is also widely used; see Camerer *et al.*, 2004; see also Nagel, 1995; Stahl and Wilson, 1995; Costa-Gomes *et al.*, 2001; Costa-Gomes and Crawford, 2006; Wang *et al.*, 2006.)

p0180 Given their beliefs, k-step players figure out what all lower-step types will do and form an expected payoff for strategy s_i^j

$$E_k(s_i^j) = S_t g_k(t) S_k P_t(s_j^h) p_i(s_i^j, s_j^h)$$

A level-k player responds with a logit (softmax) choice function

$$p_k(s_j^j) = \exp(?E_k(s_i^j))\,/S_n \exp(?E_k(s_i^n))$$

This model is easy to compute numerically because it uses a simple loop. Behavior of level-k players depends only on the behavior of lower-level players, which is computed earlier in the looping procedure. In contrast, equilibrium computation is often more difficult because it requires solving for a fixed-point vector of strategies which is a best response to itself.

p0190 A useful illustration of how the cognitive hierarchy approach can explain deviations from equilibrium analysis is the "p-beauty contest" game (Nagel, 1995; Ho *et al.*, 1998). In this game, several players choose a number in the interval [0, 100]. The average of the numbers is computed, and multiplied by a value p (in many studies, $p = 2/3$). The player whose number is closest to p times the average wins a fixed prize.

p0200 In equilibrium, by definition, players are never surprised what other players do. In the p-beauty contest game, this equilibrium condition implies that all players must be picking p times what others are choosing. This equilibrium condition only holds if everyone chooses 0 (the Nash Equilibrium, consistent with iterated dominance).

p0210 Figure 13.1 shows data from a game with $p = .7$ and compares the Nash prediction (choosing 0) and the fit of a cognitive hierarchy model (Camerer *et al.*, 2004).

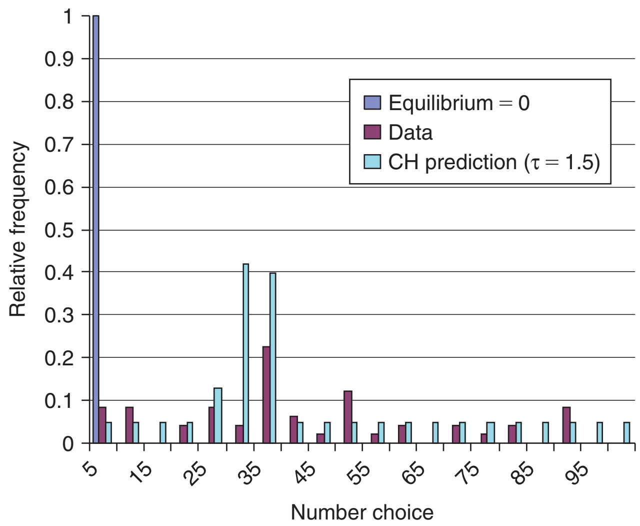


FIGURE 13.1 Data and predictions for .7 times the average game. Players choose numbers 0–100 simultaneously; the player closest to .7 times the average wins a fixed monetary prize. Data are closer to the prediction of a cognitive hierarchy (CH) model than to the equilibrium prediction of 0. Reproduced from Camerer and Fehr (2006), with permission. f0010

In this game, some players choose numbers scattered from 0 to 100, many others choose p times 50 (the average if others are expected to choose randomly), and others choose p^2 times 50. Groups of subjects with high analytical skill and training in game theory do choose lower numbers (although 0 is still rarely chosen). When the game is played repeatedly with a fixed group of players who learn the winning number, number choices do converge toward zero – a reminder that equilibrium concepts can reliably predict where an adaptive learning process leads. Costa-Gomes *et al.* (2001), Camerer *et al.* (2004), Costa-Gomes and Crawford (2006) and earlier studies show how these cognitive hierarchy theories can fit experimental data from a wide variety of games, with similar thinking-step parameters across games. Hedden and Zhang (2002) show similar evidence of limited thinking from cognitive experiments.

The cognitive hierarchy theories deliberately allow for the possibility that some players do not correctly guess what others will do. Another approach, called "quantal response" equilibrium (QRE), retains the assumption that each player's beliefs are statistically correct, but uses a softmax choice function so that choices are not always payoff-maximizing best-responses. (That is, players can make mistakes in strategy choice, but large mistakes are rarer than small mistakes.) The expected payoff of player i's strategy s_i^h is $E(s_i^h) = S_k P_{-i}(s_j^k) p_i(s_i^h, s_{-i}^k)$, and choice probabilities are given by a softmax function (see above). QRE fits a wide variety of data better than do Nash p0220

GLIMCHER 978-0-12-374176-9 00013

CH013.indd 196 5/12/2008 6:19:39 PM

t0010 TABLE 13.1 "Work–shirk" game payoffs (Dorris and Glimcher, 2004)

		Employer	
		Inspect	Don't inspect
Worker	Work	.5, 2–I	.5, 2
	Shirk	0, 1—I	1, 0

Note: Employer mixed-strategy equilibrium probability is (.5, inspect; .5, don't inspect). Worker mixed-strategy equilibrium probability is (I, shirk; 1–I, work).

TABLE 13.2 Variation in equilibrium shirking rates, cognitive hierarchy prediction, and actual human and monkey shirk rates in the work–shirk game t0020

	Inspection cost I			
	.1	.3	.7	.9
Equilibrium p(shirk)	.10	.30	.70	.90
CH prediction (τ = 1.5)	.28	.28	.72	.72
Human data	.29	.48	.69	.83
Monkey data	.30	.42	.64	.78

Data from Dorris and Glimcher (2004) (Table 13.1). CH predictions from online calculator at http://groups.haas.berkeley.edu/simulations/ch/default.asp.

predictions (McKelvey and Palfrey, 1995, 1998; Goeree and Holt, 2001). QRE also circumvents some technical limits of Nash Equilibrium. In Nash Equilibrium, players can put zero weight on the chance of a mistake or "tremble" by other players, which can lead to equilibria which are implausible because they rely on threats that would not be carried out at future steps of the game. In QRE, players always tremble and the degree of trembling in strategies is linked to expected payoff differences (*cf*. Myerson, 1986).

p0230 The essential elements of CH and QRE can also be synthesized into a more general approach (Camerer *et al*., 2008), although each of the simpler components fits a wide range of games about as well as a does more general hybrid model.

p0240 One goal of CH and QRE is to explain within a single model why behavior is far from equilibrium in some games (like the p-beauty contest) and remarkably close to equilibrium in others. An example is games with mixed equilibrium. In a mixed equilibrium, a player's equilibrium strategy mixes probability across different strategies (that is, there is no combination of strategies played for sure – "pure strategies"– which is an equilibrium).

p0250 One game that has been studied relatively frequently in neuroeconomics, which only has a mixed equilibrium, is the "work or shirk" inspection game shown in Table 13.1 (Dorris and Glimcher, 2004). The economic story surrounding the game is that a lazy worker prefers not to work, but an employer knows this and sometimes "inspects" the worker. There is no pure equilibrium, because the worker only works because of the fear of inspection, and the employer does not inspect all the time if the worker is expected to work. Instead, both players mix their strategies. For the Table 13.1 game payoffs, employers inspect half the time and workers shirk I% of the time (where I is the cost of inspection). This game is in a class called "asymmetric matching pennies," because the worker prefers to match strategies on the diagonal (working if they are inspected and shirking if they aren't) and the employer prefers to mismatch.

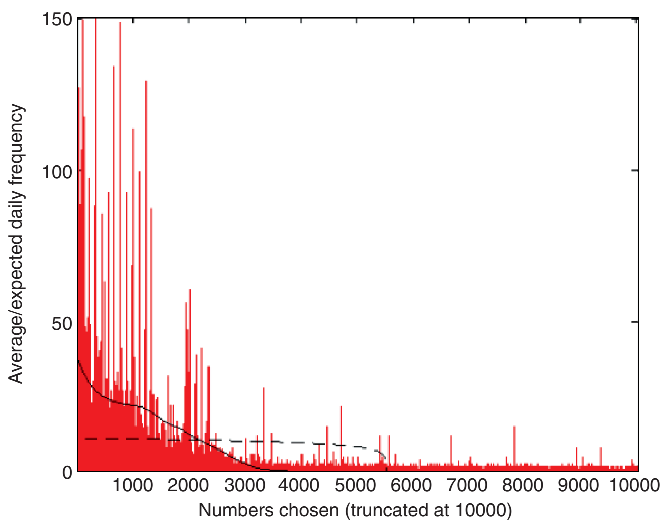
Empirically, in games with mixed equilibria the relative frequencies of strategies chosen in the first period are actually remarkably close to the predicted frequencies (see Camerer, 2003: Chapter 3) although they are regressive: That is, actual play of strategies predicted to be rare (common) is too high (too low). Table 13.2 illustrates results from human and monkeys in Dorris and Glimcher (2004). The monkey and human data are very close. The CH prediction fits the data much better than the Nash prediction for I = .1, and is equally close for other values of I. p0260

The CH model explains why choices are so close to the mixed equilibrium probabilities through the heterogeneity of players. Low-level players randomize, but higher-level players respond to expected randomization. The mixture across those level types tends to be close to the first-period data, but is closer to equal mixing than the equilibrium predictions (which typically fits data better). p0270

CH models have also been applied in two field settings. In a Swedish lottery game called LUPI (Östling *et al*., 2007), players chose integers from 1 to 99,999 and the *l*owest *u*nique *p*ositive *i*nteger wins (hence the name LUPI). About 50,000 people played the lottery each day. The (symmetric) Nash Equilibrium prediction is approximately equal choice of numbers from 1 to 5000, a sharp drop-off in choice from 5000 to 5500, and very few choices above 5500. This prediction is derived from extremely complicated algebra, using only the rules of the game, the range of numbers, and the number of players as inputs (with zero free parameters). Figure 13.2 shows that actual behavior in the first 7 days of play is surprisingly close to this prediction (shown by the dotted line). However, there is a clear tendency to choose too many low numbers, p0280

GLIMCHER 978-0-12-374176-9 00013

CH013.indd 197 5/12/2008 6:19:39 PM



f0020 **FIGURE 13.2** LUPI game results from the first week of Swedish lotteries. n = 53,000 players choose integers 1–99,999. The lowest unique integer wins a large prize (100 Swedish Kroner, ~10,000 €). The symmetric equilibrium prediction is shown by the dotted line and CH best-fitting model is shown by the solid line (t = 2.89).

too few numbers from 2500–5000, and too many numbers above the drop-off at 5500. The cognitive hierarchy model (the solid line) explains these deviations reasonably well with a best-fitting value of t = 2.98, comparable to values from 1–2 which fit experimental data well. Brown *et al.* (2007) also use the CH model to explain why moviegoers seem to ignore the fact that movies which are not reviewed before they are released tend to be low in quality. This strategic naïveté leads to a box-office premium from withholding poor movies for review. Their analysis estimates a best-fitting t = 2.89 for moviegoer behavior, close to the LUPI game estimate and earlier lab estimates.

s0060 Learning

p0290 When a game is played repeatedly, agents can learn from the payoffs they get and from the strategies other players choose, and can also learn about what other players are likely to do. Many models of these learning processes have been proposed and tested on a wide variety of experimental games (see Camerer, 2003: Chapter 6). The general structure is that strategies have numerical attractions that are updated based on observation of payoffs and actions of other players. Attractions determine choice probabilities using a logit or comparable rule. The difference across models is how attractions are updated.

p0300 There are several important differences across models. Denote a strategy j's attraction for player i after period t by $A_i^j(t)$. In reinforcement learning, the attraction of the chosen strategy is updated by the received payoff

$$A_i^j(t) = fA_i^j(t-1) + (1-f)p_i(s_i^j, s_{-i}(t))$$

where $s_{-i}(t)$ is the strategy actually chosen by opponent $-i$ in period t and f is a geometric decay. Note that this can be written as

$$A_i^j(t) = A_i^j(t-1) + (1-f)[p_i(s_i^j, s_{-i}(t)) - A_i^j(t-1)]$$

The payoff surprise $p_i(s_i^j, s_{-i}(t)) - A_i^j(t-1)$ is a prediction error – the difference between the received payoff and the previous attraction – so the learning rule is a form of temporal-difference learning rule (see Chapter 22 of this volume). A different approach is to update beliefs about what other players will choose, then use those new beliefs to update attractions – as in "fictitious play" learning, which keeps track of the fraction of previous choices by other players of each strategy (possibly geometrically-weighted to incorporate forgetting or perception of non-stationarity in opponent play).

Camerer and Ho (1999) noted that the reinforcement rule written above and fictitious play are both special cases of a more general "experience-weighted attraction" (EWA) family in which p0310

$$A_i^j(t) = \{fN(t-1)A_i^j(t-1) + d(s_i^j, s_i(t)) * p_i(s_i^j, s_{-i}(t))\}/N(t)$$

where

$$N(t) = f(1 - ?)N(t-1) + 1$$

is a cumulated weight of experience.

When ? = 0, the rule is a TD-like averaging rule. The weight on new payoff information is $1/(fN(t-1)+1)$ which falls over time t as $N(t-1)$ grows, so that learning slows down. This algebraic form expresses a time-adjusted learning rate. (However, unlike the standard temporal difference rule, when ? = 0 the rule cumulates payoffs rather than averages them. This allows attractions to grow outside the bounds of payoffs which, in the softmax rule, means that probabilities can lock in sharply at extreme values of 0 or 1.) The key term is $d(s_i^j, s_i(t)) = d + (1-d)I(s_i^j, s_i(t))$, where $I(x, y)$ is an identity function which equals 1 if $x = y$ and 0 otherwise. This "imagination" weight is 1 for the chosen strategy and d for unchosen strategies. When $d = 0$, the model reduces to reinforcement of the strategy that is actually played. When $d = 1$, it is mathematically equivalent to fictitious play; both payoffs from strategies that are actually played and "fictive" p0320

GLIMCHER 978-0-12-374176-9 00013

CH013.indd 198 5/12/2008 6:19:39 PM

payoffs from unplayed strategies influence learning equally strongly. The insight here is that learning by updating beliefs about other players' choices (using fictitious play) is exactly the same, mathematically, as generalized reinforcement in which unchosen strategies are updated by the payoffs they would have created. In computer science terms, EWA represents a hybrid of model-free learning from choices and model-based learning (which uses information about unchosen strategy payoffs through a "model" which is the structure of the game).

p0330 Ho *et al*. (2007) propose and estimate a "self-tuning" version of EWA in which *f* and *d* are functions of experience (with N(0) = 1 and ? = 0 for simplicity), so that there are no free parameters except for the response sensitivity?. The function f is interpreted as a "change-detector" which adjusts the learning rate to environmental uncertainty. When another player's behavior is highly variable, or changes suddenly, *f* falls so that more relative weight is placed on new payoff information. Behrens *et al*. (2007) have found neural evidence for such a learning-adjustment process in decision problems with non-stationary payoffs. Soltani *et al*. (2006) have simulated behavior of a similar "meta-learning" model which explores learning model parameters (Schweighofer and Doya, 2003) and shown that it fits some aspects of monkey behavior.

p0340 All these models are adaptive because they use only previous payoffs in the updating equation. However, eyetracking experiments show that players do look at payoffs of other players (Wang *et al*., 2007) and are responsive to them. In the empirical game learning literature, learning rules that anticipate how other players might be learning are termed *sophisticated*. Stahl
AUQ1 (2003) and Camerer *et al*. (2002) proposed a sophisticated rule in which players believe that others are learning according to EWA and respond to expected payoffs based on that belief.

p0350 If players are playing together repeatedly, sophisticated players could also take account of their current actions when considering what other players will do in the future, a process called *strategic teaching*. Chong *et al*. (2006) showed evidence of strategic teaching in games based on on trust and entry deterrence.

s0070 PSYCHOLOGICAL AND NEURAL EVIDENCE

p0360 Behavioral game theory analyses of experimental data have proceeded along a parallel track with other psychological and neural studies, but the tracks have rarely met. This section mentions some types of neural activity which might be linked to behavioral game theory constructs in future research.

Theory of Mind (TOM) and Human-computer Differences s0080

Theory-of-mind (TOM) refers to the capacity to make accurate judgments about the beliefs, desires, and intentions of other people, which are crucial inputs for appropriate social judgment and for social success (see Chapters 18 and 19 of this volume for further discussion of these issues across species). TOM is thought to be impaired in autism. It is widely thought that neural components of TOM include anterior and posterior cingulate, medial frontal cortex (Frith and Frith, 2006), paracingulate cortex, superior temporal sulcus (STS), and the temporal-parietal junction (TPJ). There is lively empirical debate about which of these regions are involved in different kinds of social reasoning and attribution. For example, Saxe and Powell (2006) argue that bilateral TPJ is unique for understanding another person's thoughts, and develops later in life, while mPFC is more useful for more general social understanding (e.g., sensations that other people feel). p0370

If TOM is indeed a separate faculty, it certainly is necessary to reason strategically about likely actions of other players in games. Despite this obvious link, there is only a modest number of studies searching for activity in areas thought to be part of TOM in strategic games. p0380

The first example is McCabe *et al*. (2001). They studied two-player trust games. (Trust games are a sequential form of the well-known prisoners' dilemma (PD), with the modification that a defection by the first player always creates a defection by the second player. The sequentiality allows separation of trustingness and trustworthiness, which are confounded in the PD.) In a typical trust game, one player can end the game (giving both players 10, for example), or can trust a second player. Trust creates a larger collective gain (40), but the second player can share it equally (both get 20) or can keep it all (see Chapter 5 of this volume). Contrasting behavior with human partners and that with computer partners, they found that high-trust players had more activity in the paracingulate cortex and speculated that trust requires careful consideration of likely behavior of other players. Activity in the same general region is reported by Gallagher *et al*. (2002) in a PET "rock, paper, scissors" game when playing an experimenter rather than a computer opponent. p0390

Given the link between autism and TOM, it is natural to use games to ask whether autists play differently p0400

GLIMCHER 978-0-12-374176-9 00013

CH013.indd 199 5/12/2008 6:19:40 PM

from control players. In the widely-researched ultimatum game, one player offers a share of money to another player, who can accept it or reject it. In these games, players typically offer 30–50% of the money, and offers that are too low are often rejected. Sally and Hill (2006) found that autists are much more likely to offer zero, apparently neglecting or misjudging the second player's move. Importantly, autistic children who offer positive amounts make a wide variety of offers, while positive offers by autistic adults consolidate around an equal split (similar to typical offers by normal adults). This consolidation of offers in adulthood around "normal" behavior is consistent with many reports that adult autists cope by learning explicit strategies for socially appropriate behavior.

s0090 Thinking Steps and Iterated Beliefs

p0410 The TOM evidence suggests that people are doing *some* strategic thinking, since playing humans versus computers activates TOM areas. The question raised by CH models, and their empirical success in explaining experimental data, is how *much* strategic thinking players do, what neural areas implement more strategic thinking, and related questions that arise.

p0420 Bhatt and Camerer (2005) compared player A's choices in games, A's expressed belief about B's choices, and A's "second-order" belief about B's belief about A's choice. Second-order beliefs are important in maintaining deception, because a successful deception requires A to make a certain choice and simultaneously believe that B believes he (A) will make a different choice. Second-order beliefs are also important in models of social image, in which a players' beliefs about what another player believes about his intentions or moral "type" influence utility (that is, players like to believe others believe they are good). (See Dufwenberg and Gneezy, 2002; Andreoni and Bernheim, 2007; Dillenberger and Sadowski, 2007; Ellingsen and Johannesson (2007) note the implications of this view for worker motivation in firms.)

p0430 One finding from Bhatt and Camerer's study is that in games in which a players' beliefs are all in equilibrium, there is little difference in neural activity when the player is making a strategy choice and expressing a belief. This suggests that the mathematical state of equilibrium (a correspondence of one player's beliefs with another player's actual choices) is also manifested by a "state of mind" – an overlap in the brain regions involved in choosing and guessing what others choose. They also find that second-order beliefs tend to err on the side of predicting that other players know what you will do, better than they actually do. (This bias is related to psychological research on the "curse of knowledge" – the tendency of experts to think that novices know what they know (see any computer manual for evidence, or Camerer *et al*., 1989) and the "illusion of transparency;" Gilovich *et al*., 1998.) That is, players who planned to choose strategy S guessed that other players thought they would play S more often than the other players actually did. There is differential activity during the second-order belief task and first-order beliefs task in the insula, which has been implicated in the sensation of "agency" and self-causation and may help account for the self-referential bias in second-order beliefs.

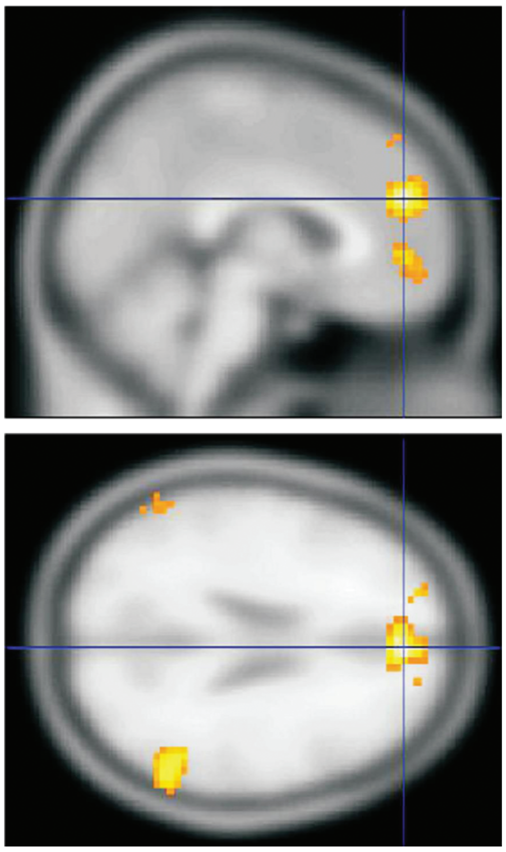
So far, only one fMRI study has looked directly for p0440
neural correlates of the steps of thinking posited by the cognitive hierarchy model described earlier in the chapter. Coricelli and Nagel (2007) used a series of "beauty contest" number-choosing games (as shown in Figure 13.1). Players chose numbers 0–100 and have a target equal to p times the average number (for various p). Playing human opponents versus computers showed differential activity in medial paracingulate cortex and bilateral STS, as in other TOM studies. They classified players, using their choices, into low strategic reasoners (one step of reasoning, choices around p*50) and high strategic reasoners (two steps of reasoning, choosing around p^2*50). The high-step reasoners showed very strong differential activity (playing humans versus computers) in paracingulate, medial OFC, and bilateral STS (see Figure 13.3).

Since game theory presents many different tools to p0450
tap different aspects of strategic thinking, many more studies of this type could be done. The eventual goal is a mapping between the types of strategic thinking involved in games, components of theory of mind, and an understanding of neural circuitry specialized to each type of thinking.

Learning s0100

Studies of the neuroscientific basis of learning in p0460
games fall into two categories. One category consists of attempts to see whether behavior exhibits some of the properties of reinforcement learning.

Seo and Lee (2007) recorded from monkey neu- p0470
rons in dorsal anterior cingulate cortex (ACC) during a matching-pennies game (the work–shirk game with equal payoffs in all cells) played against various computer algorithms. They found neurons with firing rates that are sensitive to reward and to some higher-order interactions with past choices and rewards. Behaviorally, the monkeys also play a little closer to the mixed equilibrium when the computer algorithms



f0030 FIGURE 13.3 Differential activity in playing human vs computer opponents in "p-beauty contest" game for high-reasoning players. Players choose numbers 0–100 and winner earns p*(average number) for various values of p. High reasoners are those exhibiting "level 2" thinking (choosing closer to p^2 times 50 than p times 50). Enhanced activity for these subjects for human versus computer opponents is in paracingulate and medial paracingulate PFC (top), and bilateral STS (bottom). Reproduced from Coricelli and Nagel (2007), with permission.

are designed to exploit temporal dependence in the monkeys' play. Using the work–shirk game in Table 13.1, Dorris and Glimcher (2004) also found that monkeys play close to the mixed-equilibrium proportions, and adjust their strategy mixtures surprisingly rapidly, within 10–20 trials, when the game payoff parameters change. However, they note that neural firing rates in lateral intraparietal sulcus (LIP) do not change when strategies change, as long as the relative expected utility of strategies is the same. The LIP neurons are clearly encoding relative value, not choice rates.

p0480 The second category of neuroscientific studies explores generalizations of reinforcement which posit that learning can be driven by forces other than simply immediate reward. Lohrenz *et al*. (2007) define "fictive learning" as learning from counterfactual or imagined rewards (the d term in the EWA model). In an investment game (based on actual stock market prices), they show that a fictive learning signal is evident in caudate, close to a caudate region that encodes prediction error (the difference between outcome and expectation). The fictive signal also predicts changes in investment behavior.

Another interesting kind of learning arises when players engage in a repeated game. King-Casas *et al*. (2005) studied a repeated trust game. In the one-shot game, an "investor" player can invest an amount X from a stake of 20 which triples in value to 3X. The second, "trustee," player repays an amount Y, so the investor earns $(20 - X) + Y$ and the trustee earns $(3X - Y)$. Notice that the total payment is $20 + 2X$, so the collective payoff is maximized if everything is invested … but the investor cannot count on the trustee repaying anything. (Economists call this a game of investment with "moral hazard" and no enforcement of contracts, like investing in a country with poor legal protection.) King-Casas repeated the game 10 times with a fixed pair of players to study dynamics and learning, and scanned both investor and trustee brains simultaneously. Trustees tend to exhibit two kinds of behavior – they either reciprocate an uptick in investment from period $t - 1$ to t by repaying a larger percentage ("benevolent"), or reciprocate an uptick by investing less ("malevolent"). Figure 13.4 shows regions which are activated in the trustee choice period t by (later) benevolent trustee actions. The interesting finding, from a learning point of view, is that anticipation of benevolent "intention to trust" moves up by about 14 seconds from early rounds of the 10-period game (rounds 3–4) to later rounds 7–8. There is also both a within-brain correlation of this signal (trustee anterior cingulate and caudate) in anticipation of the later choice, and a cross-brain correlation (investor middle cingulate MCC and trustee caudate). That is, just as trustees are anticipating their own later benevolent action reward value, investors are anticipating it as well in medial cingulate cortex. This is a dramatic sign of synchronized anticipation due to learning, which could only be seen clearly by scanning both brains at the same time. p0490

Hampton *et al*. (2007) used the work–shirk game (Table 13.1) to investigate neural correlates of "sophisticated" learning. Suppose an employer player, for example, has some inkling that others are learning from their own (employer) choices. Then if the employer chooses Inspect in one period, that choice has an immediate expected payoff (based on the employer's beliefs about what the worker will do) and also has a *future* influence because it is predicted to change the worker's beliefs and hence to change the worker's future play. Hampton *et al*. include this "influence value" as a regressor and correlate its numerical value p0500

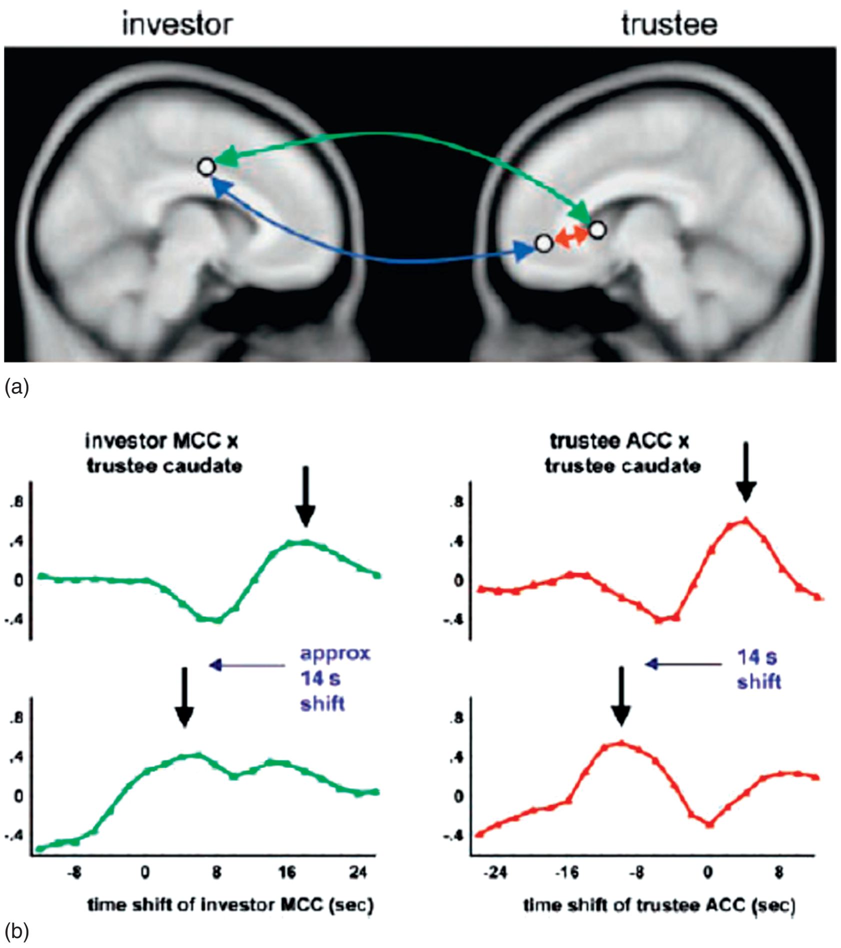
f0040


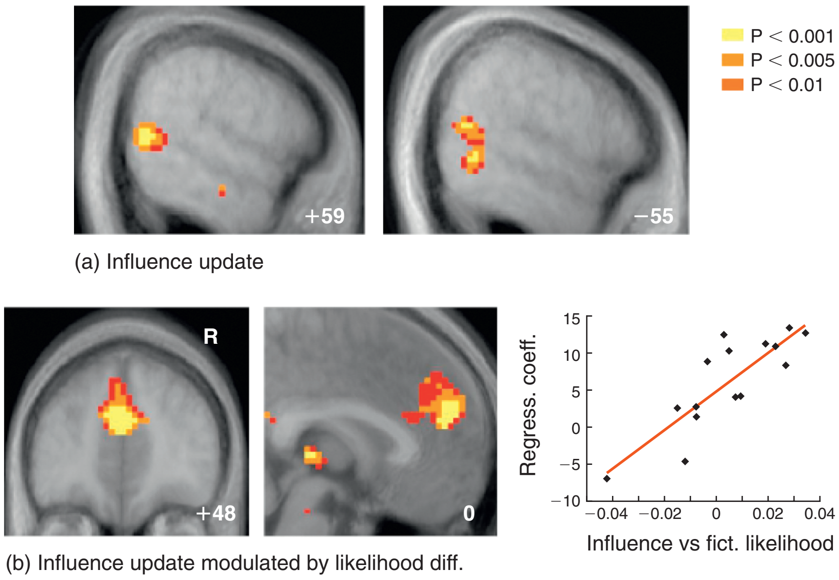
FIGURE 13.4 Trust-game activity in investor (first-moving player) and trustee (second-moving player) brains in 10-period trust game. (a) Regions activated in period *t* by "intention to trust" behavioral signal (reciprocal response to activity in the previous period) during trustee repayment phase. (b) Graphs show a time series of correlation between brain signals at different points in time. Positive correlations indicate two types of brain activity are highly correlated at the point in time indicated on the *y*-axis (0 is choice onset). Top time series is early rounds 3–4; bottom time series is later rounds 7–8. Correlations shift forward in time (~14 s) from early to late rounds, for both the cross-brain correlation of investor middle cingulate (MCC) and trustee caudate (left graphs), and within brain correlation of trustee ACC and trustee caudate (right graphs). The forward shift indicates that learning creates anticipation of likely behavior within the trustee's own brain (right graph) and between the two players' brains (left graph). Reproduced from King-Casas *et al*. (2005), with permission.

with activity in the brain. The influence value (teaching) component activates posterior STS (Figure 13.5a) on a trial-by-trial basis. Furthermore, subjects can be categorized, purely from their behavioral choices, by how much better the influence model fits their choices than does a purely adaptive fictitious play model. There is a strong cross-subject correlation between the improvement in predicting behavior from including influence (Figure 13.5b *x*-axis) and activity in medial paracingulate in response to trial-by-trial influence value (Figure 13.5b *y*-axis). Along with the behavioral and eyetracking evidence, this finding provides direct fMRI evidence that human learning in games sometimes includes some degree of sophistication (as proposed and shown in behavioral data by Camerer *et al*., 2002, and Stahl, 2003).

s0110
Deception

p0510 Deception is an important topic that the combination of game theory and neuroscience may help illuminate. Game theory offers a rich set of games which characterize when deception is expected to occur, and potential field applications of these games. A useful aspect of the game theory is that it considers jointly the actions of a deceptive player and a player who can anticipate deception.

A useful mathematical example is "strategic information transmission" between an informed "sender" p0520
and a "receiver" (Crawford and Sobel, 1982). Consider a security analyst who has good information about the value V of a company, and a client who wants to know what V is. The analyst sends a message M to the client, who then chooses an action number A (a perception of V which triggers an investment policy). Often there is a conflict of interest in which the analyst wants the client to choose a number which is the true value plus a positive bias B. That is, the analyst's payoffs are highest if the client chooses an action A = V + B, and the client's payoff is highest if A = V. Notice that the message M is "cheaptalk:" in game theory language, this means that the message is costless (so the willingness to pay for a message is not a calculated strategy



f0050 **FIGURE 13.5** Correlation of the "influence value" of current action on future reward (through the influence on an opponent's future choice) with BOLD signal in fMRI. (a) Numerical influence values correlate with activity in bilateral STS trial-by-trial. (b) Cross-subject correlation of the extent to which choices are fit better by including an influence value term (*x*-axis) and strength of the influence value regressor in paracingulate cortex (*y*-axis). Reproduced from Hampton *et al*. (2007), with permission.

that is informative) and does not bind behavior (so if the sender promises to do something and does not, there is no explicit penalty). The message number M can only influence payoffs through any information the client infers from the message which changes his action. This model has broad application to corporate advertising, expert services (fixing cars to fixing human bodies), political and personal promises, and bargaining.

p0530 When there is a large bias B, it can be proved mathematically that little truthful information is transmitted through the message M, if players are assumed to be self-interested and strategically sophisticated. To illustrate, suppose the values are integers 1–5 and the bias is B = 1. Little information is transmitted, because any message about V that is truthful and believed will also be sent if the analyst has a lower value than V. For example, suppose when V = 3 the analyst truthfully announces M = 3, and the client believes it and chooses A = 3. Then an analyst who knows V = 2 will send the same message of 3, since he wants the client to choose V + B, which is 2 + 1 = 3. As a result, the equilibrium which conveys the most information is for the analyst to say M = 1 when V = 1 (admitting bad news) and to mix across 2–5 otherwise. When the bias is larger, B = 2, there is no information conveyed at all in equilibrium. These equilibria are inefficient because if players could somehow agree to be truthful – or if a third party could certify information or penalize the analyst for exaggerating – then the players together would have a higher collective payoff.

p0540 Despite this strong prediction, several experiments have shown that in games with this structure, a substantial amount of information is communicated truthfully (and believed). There are two possible sources of this "overcommunication." One source is a feeling of guilt the sender has from exaggerating the truth and sending a message M which is higher than the value of V. Adding a negative utility from guilt into the analyst's payoffs can lead to an equilibrium with more truthful revelation. A second source is cognitive difficulty; it may just be hard to figure out how much to exaggerate, or how another player will react, so that heuristics like telling the truth or sending the message V + B (and hoping the client accepts it) are chosen. Wang *et al*. (2006) explored these causes in a sender-receiver game using eyetracking and pupillometry. They found that analysts do not look at the client's payoffs very often (although the equilibrium analysis predicts they have to, in order to figure out how receivers will react to different messages). Those who look at the client payoffs more often are not less deceptive, as a simple guilt theory predicts. Furthermore, the looking patterns and choices are mostly consistent with doing one or two steps of strategic thinking. Interestingly, using a combination of looking at information and pupil dilation when the analyst makes his choice gives a statistical prediction of the true state V which is sufficiently accurate to improve the analyst's experimental profits (on paper) by 10–20%.

Bhatt *et al*. (2007) have studied a closely-related game they call "yard sale bargaining," using fMRI. p0550 In this game, a seller has an item with value zero (so he would accept any positive price for it). A buyer has a value for the item V from 1–10, which the buyer knows but the seller does not. The buyer learns his value and suggests a price (a kind of cheaptalk, as in

GLIMCHER 978-0-12-374176-9 00013
CH013.indd 203 5/12/2008 6:19:44 PM

the analyst–client game), S. The seller sees S and then states a final take-it-or-leave-it price P. If the buyer's value is above the price $V \geq P$, the object is sold at the price P. If players are self-interested and strategic, there is no suggested price function S(V) which conveys any information. The reason is that any suggestion S which is believed to credibly convey the idea that the value is V would also be used by savvy buyers with values higher than V. So, in theory, the seller should completely ignore the suggestion S and state a price of 5–6 (which maximizes expected profits knowing nothing about V).

p0560 As in the analyst–client game, Bhatt *et al*. see that there is substantial revelation of information about value (contrary to the theory). A typical suggestion function is $S = V/2$ and a typical pricing function is $P = S + 2$. That is, the buyers often say they can pay half of what they can actually afford, and the sellers seem to guess this and pick a price which is the suggested price marked up by 2 units. (Mathematically, these strategies imply that trade takes place when $V \geq P$, or $V \geq (V/2) + 2$ which implies $V \geq 4$, so that more trades take place than would in equilibrium. The fMRI shows that buyers who are less truthful about their values have greater activity in bilateral dorsal striatum (as if they were expecting larger profits). Sellers who are more sensitive to suggested price have greater activity in right anterior temporal sulcus, and less activity in the anterior cingulate. These regions are consistent with the hypothesis that believing messages is the default mode for sellers: Since the anterior cingulate is often involved in response conflict, lowered activity means the sellers who respond to suggested price are not registering a conflict between the suggested price and the likely buyer value. Responding to suggested prices activates the TOM temporal sulcus area (trying to infer the buyer's intention or state of mind), and ignoring those suggestions recruits ACC in order to resolve cognitive conflict (Kerns *et al*., 2004).

s0120 **Reputation**

p0570 An important concept in repeated game theory with private information is *reputation*. Private information is usefully characterized by a *type* a player has, which is randomly determined at the start of a repeated game. In theory, a player's actions are designed to satisfy short-term goals and also to either convey (in a cooperative) game or hide (in a competitive game) the player's type. Player A's reputation is the belief, in the eyes of other players, about player A's type.

p0580 For example, an important analysis of the repeated prisoners' dilemma (PD) starts with the idea that some players always cooperate. Then, even players who are selfish will choose to cooperate, in order to maintain a reputation as the kind of player who cooperates, because having such a reputation encourages cooperation from other players in the future.

Two neural studies indirectly tap aspects of repu- p0590
tation. Singer *et al*. (2004) found that showing faces of people who had previously cooperated activated the nucleus accumbens. This is the first direct evidence that a game-theoretic reputation generates a neural value signal. Delgado *et al*. (2005) used fMRI to explore neural reactions to behavior in a repeated cooperation game when the scanned subject's opponent begins with a good, neutral, or bad reputation created by a picture and short blurb about an opponent's behavior (in Bayesian terms, the blurb creates a prior belief that the opponent will behave cooperatively or not). They found that during the outcome phase, if the partner behaves cooperatively, compared to uncooperatively, there is differential activity in the caudate nucleus (and several other areas). However, there is no such difference in this contrast if the partner had a good reputation to begin with. The time course of activity is consistent with the idea that bad behavior is "forgiven" (in neural terms, does not generate as much reward or prediction error signal) if the partner is a good person.

CONCLUSIONS AND FUTURE RESEARCH s0130

Game theory is useful for creating a precise math- p0600
ematical model linking strategy combinations to payoffs, a kind of periodic table of the elements of social life. Predictions are made using various behavioral assumptions about how deeply people reason and how they react to observed behavior. Hundreds of experiments suggest that players do not always reason very strategically, evaluation of payoffs often includes social elements beyond pure self-interest, and players learn from experience.

So far, there has been only limited use of game the- p0610
ory and neuroscientific tools to link strategic thinking to neural activity. This limited contact is probably due to the fact that psychologists have not used the major tools in game theory, which may in turn be due to skepticism that the rationality-based analyses in game theory are psychologically accurate.

One promising point of contact is between theo- p0620
ries of strategic thinking and "theory of mind" (TOM) regions thought to be necessary for understanding beliefs, desires, and thoughts of other people. The few available studies tend to indicate that TOM areas are

GLIMCHER 978-0-12-374176-9 00013

CH013.indd 204 5/12/2008 6:19:49 PM

activated in playing mathematical games, but a closer link would be very useful for both fields.

p0630 Game theory could also be useful in understanding disorders. Some psychiatric disorders could be understood as disorders in normal social evaluation and prediction. For example, anti-social personality disorder seems to disrupt normal valuation of the consequences of one's actions on others. Paranoia in psychosis and schizophrenia could be defined symptomatically as overpredicting a hostile (payoff-reducing) reaction of others to one's own choices. Autism can also be seen as a disorder in evaluating expected social behavior. Using a battery of games involving altruism, fair sharing, and trust, Krajbich *et al*. (2008) have found that patients with ventromedial prefrontal cortical damage act as if they exhibit less parametric guilt – giving less and acting in a less trustworthy fashion – than do normal controls and control patients with damage in other regions.

p0640 Game theory is also a tool for understanding expertise and increasing skill. In a game, there is usually a clear performance metric – who makes the most money? Understanding extraordinary skill in bargaining, poker, and diplomacy may illuminate the everyday neural bases of these skills and permit effective training.

References

Aumann, R. (1985). What is game theory trying to accomplish? In: K. Arrow and S. Honkaphoja (eds), *Frontiers of Economics*. Oxford: Basil Blackwell, pp. 28–76.

Andreoni, J., and Bernheim, B.D. (2007). Social image and the 50–50 norm: a theoretical and experimental analysis of audience effects. August, http://econ.ucsd.edu/~jandreon/WorkingPapers/socialimage.pdf

Behrens, T.E.J., Woolrich, M.W., Walton, M.E., and Rushworth, M.F.S. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.

Bhatt, M. and Camerer, C.F. (2005). Self-referential thinking and equilibrium as states of mind in games: fMRI evidence. *Games Econ. Behav.* 52, 424–459.

Bhatt, M., Lohrenz, T., Montague, R.M., and Camerer, C.F. (2007). Neural correlates of lowballing and gullibility in "yard-sale bargaining". Working Paper, Caltech.

Brown, A.L., Camerer, C.F., and Lovallo, D. (2007). *To review or not review? Limited strategic thinking at the box office*. Pasadena, CA: California Institute of Technology.

Camerer, C.F. (1998). *Mental Representations of Games*. Princeton, NJ: Princeton University Press.

Camerer, C.F. (2003). *Behavioral Game Theory: Experiments on Strategic Interaction*. Princeton, NJ: Princeton University Press.

Camerer, C.F. and Fehr, E. (2006). When does "Economic Man" dominate social behavior? *Science* 311, 47–52.

Camerer, C. and Ho, T.H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica* 67, 827–874.

Camerer, C., Loewenstein, G., and Weber, M. (1989). The curse of knowledge in economic settings – an experimental analysis. *J. Political Econ.* 97, 1232–1254.

Camerer, C.F., Johnson, E., Rymon, T., and Sen, S. (1993). Cognition and framing in sequential bargaining for gains and losses. In: K.G. Binmore, A.P. Kirman, and P. Tani (eds), *Frontiers of Game Theory*. Cambridge: MIT Press, pp. 27–47.

Camerer, C.F., Ho, T.-H., and Chong, J.-K. (2004). A cognitive hierarchy model of games. *Q. J. Economics* 119, 861–898.

Camerer, C.F., Rogers, B., and Palfrey, T. (2008). Heterogeneous quantal response equilibrium and cognitive hierarchies. *J. Econ. Theory*, (in press).

Chong, J.-K., Camerer, C., and Ho, T.-H. (2006). A learning-based model of repeated games with incomplete information. *Games Econ. Behav.* 55, 340–371.

Coricelli, G. and Nagel, R. (2007). *Guessing in the Brain: An fMRI Study of Depth of Reasoning*. Working Paper, Lyon University.

Costa-Gomes, M.A. and Crawford, V.P. (2006). *Cognition and Behavior in Two-Person Guessing Games: An Experimental Study*. London: UCLA, Department of Economics.

Costa-Gomes, M.A., Crawford, V.P., and Broseta, B. (2001). Cognition and behavior in normal-form games: an experimental study. *Econometrica* 69, 1193–1235.

Crawford, V.P. and Sobel, J. (1982). Strategic information transmission. *Econometrica* 50, 1431–1451.

Delgado, M.R., Frank, R.H., and Phelps, E.A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618.

Devetag, G. and Warglien, M. (2003). Games and phone numbers: do short-term memory bounds affect strategic behavior? *J. Econ. Psychol.* 24, 189–202.

Dillenberger, D. and Sadowski, P. (2007). *Ashamed to Be Selfish*. Princeton, NJ: Princeton University.

Dorris, M.C. and Glimcher, P.W. (2004). Activity in posterior parietal cortex is correlated with the subjective desirability of an action. *Neuron* 44, 365–378.

Dufwenberg, M. and Gneezy, U. (2000). Measuring beliefs in an experimental lost wallet game. *Games Econ. Behav.* 30, 163–182.

Ellingsen, T. and Johannesson, M. (2007). Paying respect. *J. Econ. Persp.* 21, 135–149.

Fehr, E. and Camerer, C.F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* 11, 419–427.

Frith, C.D. and Frith, U. (2006). The neural basis of mentalizing. *Neuron* 50, 531–534.

Gallagher, H.L., Jack, A.I., Poepstorff, A., and Frith, C.D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage* 16, 814–821.

Gilovich, T., Savitsky, K., and Medvec, V.H. (1998). The illusion of transparency: giased assessments of others' ability to read our emotional states. *J. Pers. Social Psychol.* 75, 332–346.

Goeree, J.K. and Holt, C.A. (2001). Ten little treasures of game theory and ten intuitive contradictions. *Am. Econ. Rev.* 91, 1402–1422.

Hampton, A., Bossaerts, P., and O'Doherty, J. (2007). Neural correlates of mentalizing-related computations during strategic interactions in humans. Working Paper, Caltech.

Hedden, T. and Zhang, J. (2002). What do you think I think you think? Strategic reasoning in matrix games. *Cognition* 85, 1–36.

Ho, T.H., Camerer, C.F., and Weigelt, K. (1998). Iterated dominance and iterated best response in experimental "p-beauty contests". *Am. Econ. Rev.* 88, 947–969.

Ho, T.H., Camerer, C.F., and Chong, J.-K. (2007). Self-tuning experience weighted attraction learning in games. *J. Econ. Theory* 127, 177–198.

Johnson, E.J., Camerer, C., Sen, S., and Rymon, T. (2002). Detecting failures of backward induction: monitoring information search in sequential bargaining. *J. Econ. Theory* 104, 16–47.

GLIMCHER 978-0-12-374176-9 00013

CH013.indd 205 5/12/2008 6:19:49 PM

Kerns, J.G., Cohen, J.D., MacDonald, A.W., III *et al*. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science* 303, 1023–1026.

King-Casas, B., Tomlin, D., Anen, C. *et al*. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.

Krajbich, I., Adolphs, R., Tranel, D. *et al*. (2008). Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. Working Paper, Caltech.

Lohrenz, T., McCabe, K., Camerer, C.F., and Montague, P.R. (2007). Neural signature of fictive learning signals in a sequential investment task. *PNAS* 104, 9493–9498.

McCabe, K., Houser, D., Ryan, L. *et al*., (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA,* 98, 11832–11835.

McKelvey, R.D. and Palfrey, T.R. (1995). Quantal response equilibria for normal form games. *Games Econ. Behav.* 10, 6–38.

McKelvey, R.D. and Palfrey, T.R. (1998). Quantal Response equilibria for extensive form games. *Exp. Economics* 1, 9–41.

Myerson, R.B. (1986). Acceptable and predominant correlated equilibria. *Intl J. Game Theory* 15, 133–154.

Nagel, R. (1995). Unraveling in guessing games: an experimental study. *Am. Econ. Rev.* 85, 1313–1326.

Nietzsche, F. (1996). *Human, All Too Human: A Book for Free Spirits*. Cambridge: Cambridge University Press.

Östling, R., Wang, J.T.-y., Chou, E., and Camerer, C.F. (2007). Field and lab convergence in Poisson Lupi games. Working Paper Series in Economics and Finance. Stockholm: Stockholm School of Economics.

Sally, D. and Hill, E. (2006). The development of interpersonal strategy: autism, theory-of-mind, cooperation and fairness. *J. Econ. Psychol.* 27, 73–97.

Saxe, R. and Powell, L.J. (2006). It's the thought that counts. *Psychological Sci.* 17, 692–699.

Schweighofer, N. and Doya, K. (2003). Meta-learning in reinforcement learning. *Neural Networks* 16, 5–9.

Seo, H. and Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27, 8366–8377.

Singer, T., Kiebel, S.J., Winston, J.S. *et al*. (2004). Brain responses to the acquired moral status of faces. *Neuron* 41, 653–662.

Soltani, A., Lee, D., and Wang, X.-J. (2006). Neural mechanism for stochastic behaviour during a competitive game. *Neural Networks* 19, 1075–1090.

Stahl, D.O. (2003). Sophisticated learning and learning sophistication. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=410921

Stahl, D.O. and Wilson, P.W. (1995). On players' models of other players: theory and experimental evidence. *Games Econ. Behav.* 10, 218–254.

Wang, J.T.-y., Spezio, M., and Camerer, C.F. (2006). Pinocchio's pupil: using eyetracking and pupil dilation to understand truth-telling and deception in biased transmission games. Pasadena, CA: Caltech.

Wang, J. T.-y., Knoepfle, D., and Camerer, C.F. (2007). Using eyetracking data to test models of learning in games. Working Paper, Caltech.

Author Query
{AUQ1} Not in ref list; pls. supply