Psychological Review

The Dual Accumulator Model of Strategic Deliberation and Decision Making

Russell Golman, Sudeep Bhatia, and Patrick Bodilly Kane

The Dual Accumulator Model of Strategic Deliberation and Decision Making

Russell Golman
Carnegie Mellon University

Sudeep Bhatia
University of Pennsylvania

Patrick Bodilly Kane
McGill University

What are the mental operations involved in game theoretic decision making? How do players deliberate (intelligently, but perhaps imperfectly) about strategic interdependencies and ultimately decide on a strategy? We address these questions using an evidence accumulation model, with bidirectional connections between preferences for the strategies available to the decision maker and beliefs regarding the opponent's choices. Our *dual accumulator* model accounts for a variety of behavioral patterns, including limited iterated reasoning, payoff sensitivity, consideration of risk-reward tradeoffs, and salient label effects, and it provides a good quantitative fit to existing behavioral data. In a comparison with other popular behavioral game theoretic models fit at the individual subject level to choices across a set of games, the dual accumulator model makes the most accurate out-of-sample predictions. Additionally, as a cognitive-process model, it can also be used to make predictions about response time patterns, time pressure effects, and attention during deliberation. Stochastic sampling and dynamic accumulation, cognitive mechanisms foundational to decision making, play a critical role in explaining well-known behavioral patterns as well as in generating novel predictions.

Keywords: behavioral game theory, sequential sampling, preference accumulation, bidirectional processing, cognitive modeling

Supplemental materials: http://dx.doi.org/10.1037/rev0000176.supp

Strategic decision making is an important feature of human behavior. From cooperating on collaborative projects, to negotiating agreements, to competing over limited resources, inter-dependent strategic settings—in which outcomes depend on the choices of two or more decision makers—are ubiquitous in everyday life. Game theory provides a mathematical framework with which decision making in strategic settings can be formally represented and analyzed (Hart, 1992; Luce & Raiffa, 1957; von Neumann & Morgenstern, 1944). It is an important area of research in economics, political science, biology, computer science, philosophy, and psychology, where it is used to describe the types of coordination, cooperation, and conflict that can be observed in groups of decision makers.

Traditional game theory is concerned with studying the behavior of idealized decision makers, who deliberate rationally and who find themselves in an equilibrium in which they can form accurate expectations about each other's choices.[1] Not surprisingly, humans display pervasive and systematic departures from rationality, which cannot be accounted for by the traditional Nash equilibrium prediction or its various refine-

---

[1] The traditional Nash equilibrium solution concept abstracts away from any particular dynamical process that would presumably converge to it. The implicit assumption is that any reasonable process would converge to a Nash equilibrium. This is not so (see, e.g., Hofbauer & Sigmund, 2003), and actually long-run behavior can be quite sensitive to initial conditions as well as the specific assumptions about the dynamics describing how players adjust their strategies over time (Golman, 2011).

ments. This has led to the growth of behavioral game theory, which acknowledges bounded rationality (see Camerer, 2003a, 2003b; Colman, 2003).[2]

Models of limited iterated reasoning or error-prone reasoning relax the traditional assumptions of perfect rationality and sophistication to better predict the behavior of boundedly rational decision makers. According to these models, boundedly rational decision makers may have simplistic beliefs about the behavior of their opponents (or partners), but retain the capability to intelligently choose strategies that maximize their own rewards given these simplistic beliefs, as with Level-k reasoning (Nagel, 1995; Stahl & Wilson, 1994, 1995) or cognitive hierarchy theory (Camerer, Ho, & Chong, 2004). Alternatively, they may have sophisticated beliefs about their opponents, but occasionally err in their choices of strategies, as with quantal response equilibrium (McKelvey & Palfrey, 1995). Bringing both elements together, decision makers may hold simplistic beliefs as well as make occasional mistakes, as with noisy introspection (Goeree & Holt, 1999, 2004).

Although these behavioral game theoretic models have been developed by relaxing one or more of the unrealistically strong assumptions of Nash equilibrium, they nonetheless require the formation of probabilistic beliefs and the calculation of expected values. Decision makers often struggle with these tasks in simple choice settings, suggesting that they may not be employing them in complex strategic settings either (Gigerenzer & Todd, 1999). Indeed, Agranov, Potamites, Schotter, and Tergiman (2012) find that approximately half of subjects are unable to best respond to computer opponents known to be playing a uniform mixed strategy in the p-beauty contest game. Our approach is to begin with plausible cognitive mechanisms—the same cognitive mechanisms commonly thought to be at play in other domains of preferential decision making—rather than with the traditional Nash equilibrium benchmark. We aim to characterize the cognitive process of strategic deliberation and choice in one-shot, simultaneous-move, two-player games of complete information with finite strategy sets, that is, in the simplest class of games. In line with computational cognitive modeling of perception, memory, and other forms of judgment and decision making (Bartels & Johnson, 2015; Busemeyer & Johnson, 2004; Busemeyer & Rieskamp, 2014; Newell & Bröder, 2008; Oppenheimer & Kelso, 2015), our model describes the strategic deliberation process in terms of mental operations that involve the sampling and aggregation of information and the spread of activation through information-processing units. By developing a behavioral game theoretic model from a cognitive model, we hope to make more accurate predictions about strategy choices as well as new kinds of predictions about response time patterns, about how choice probabilities change with time pressure, and about attention during deliberation.[3] These predictions should allow for individual differences, and should apply portably across games.

Our model proposes that decision makers dynamically construct preferences among their own available strategies along with beliefs about the opponent's preferred strategies using a process of stochastic sampling and evidence accumulation (e.g., Busemeyer & Townsend, 1993; Gold & Shadlen, 2007; Ratcliff & McKoon, 2008). The processes of forming preferences and beliefs are linked (through a bidirectional network, as in Bhatia, 2016; Glöckner, Hilbig, & Jekel, 2014; Holyoak & Simon, 1999), so that provisional beliefs about the opponent's strategies influence a decision maker's preferences, and these emerging preferences in turn influence beliefs about the opponent's choices. Decision makers build up a sense of how much they like their own strategies as well as how much the opponent might like his, but need not form probabilistic beliefs. They each choose the strategy they like best (which is not necessarily a best response to what they believe the opponent will choose). Still, as decision makers are able to represent an opponent's actions, they can respond in an intelligent manner to what they think the opponent might do and also revise these beliefs as they deliberate (Goodie, Doshi, & Young, 2012; Hedden & Zhang, 2002).

We show that our *dual accumulator* model predicts a variety of behavioral patterns, including limited iterated reasoning, payoff sensitivity, and consideration of risk-reward tradeoffs (Capra, Goeree, Gomez, & Holt, 1999; Crawford, Costa-Gomes, & Iriberri, 2013; Goeree & Holt, 2001). Additionally, it provides a compelling, cognitively grounded account of the effects of strategy salience on strategic choice (Crawford, Gneezy, & Rottenstreich, 2008; Hargreaves Heap, Rojo Arjona, & Sugden, 2014; Mehta, Starmer, & Sugden, 1994). Existing behavioral game theory models cannot account for all of these patterns (with a parsimonious, consistent specification across games), and some of these effects lie entirely outside the scope of many of these models. Finally, we fit our dual accumulator model to individual subject-level data (assuming that each individual has stable parameter values across games) using an existing, publicly available data set (Stahl & Wilson, 1995) and compare our model's fit against Level-k reasoning, cognitive hierarchy theory, logit quantal response equilibrium, and (telescoping) noisy introspection. In a hold-one-out analysis, we find that the dual accumulator model makes the most accurate out-of-sample predictions.

Our dual accumulator model can be seen as a cognitive implementation of a specific, finite-horizon stochastic fictitious play algorithm (Fudenberg & Kreps, 1993). The original fictitious play algorithm assumed accumulation of payoffs through deterministic sampling of iterative best responses and focused on long-run behavior (Brown, 1951). We also rely on a process of accumulation of payoffs, but, inspired by existing cognitive models of preferential choice, we assume a stochastic sampling rule and a finite number of accumulation steps, after which the process terminates. These features make the model more realistic. Our dual accumulator model also bears a resemblance to action-sampling equilibrium (Selten & Chmura, 2008) and payoff-sampling equilibrium (Osborne & Rubinstein, 1998) in that actions are randomly

---

[2] Behavioral game theory also acknowledges learning and social utility, which are not our focus here. We avoid dealing with learning by restricting to one-shot interactions or initial play in games, where learning is not possible. Social utility may always be relevant, but we hope we can disentangle it from boundedly rational deliberation—we identify, and try to explain, behavioral patterns that we attribute to bounded rationality, and we neglect behavioral patterns that we attribute to social preferences. Putting together these distinct parts of behavioral game theory is a long-term goal that is beyond our present scope.

[3] Level-k reasoning and cognitive hierarchy theory also describe the process of strategic deliberation in terms of a sequence of cognitive steps, but those steps are more complex than the steps of information processing that we propose (Camerer, 2008). These theories may generate response time patterns and time pressure effects with an additional assumption that greater response time correlates with higher level reasoning, but some additional structure must be assumed to make such predictions precise.

sampled a finite number of times and players choose the strategy that looks best given this sample, but our model does not assume that players are in equilibrium and instead proposes sampling dynamics appropriate for one-shot play.

Our dual accumulator model builds on existing accumulator-based theories of nonstrategic risky choice (Busemeyer & Townsend, 1993; see also Bhatia, 2014 and Rieskamp, 2006), in which beliefs about probabilistic events, such as the opponent's choices, remain unchanged throughout the deliberation process. Most accumulator models of risky choice—as well as of multiattribute choice and intertemporal choice (Bhatia, 2013; Bhatia & Mullett, 2016; Dai & Busemeyer, 2014; Roe, Busemeyer, & Townsend, 2001; Turner, Schley, Muller, & Tsetsos, 2018; Usher & McClelland, 2004)—rely on two main ingredients: stochastic sampling and dynamic accumulation. These ingredients are critical in our model for making deliberation subject to intrinsic variability and requiring it to play out over time, and we show that stochastic sampling and dynamic accumulation play a central role in capturing the behavioral patterns observed in strategic choice. We conclude that a single framework based on stochastic sampling and dynamic accumulation can be used to understand choice behavior across a variety of nonstrategic and strategic settings, thereby providing a cohesive, unitary approach to modeling the cognitive underpinnings of preferential choice.

Preference Accumulation Models

Sampling and Accumulation

We use a model of sequential sampling and evidence accumulation to specify the cognitive basis of strategic deliberation and decision making. Growing out of evidence accumulation models capturing perception, categorization, lexical decision making, and memory (Gold & Shadlen, 2007; Nosofsky & Palmeri, 1997; Ratcliff, 1978; Ratcliff, Gomez, & McKoon, 2004; Usher & McClelland, 2001), models based on sampling and accumulation have been quite successfully applied to nonstrategic multiattribute, intertemporal, and risky choice (Bhatia, 2013, 2014, 2017; Bhatia & Mullett, 2016; Clithero, 2018; Dai & Busemeyer, 2014; Diederich, 1997; Fudenberg, Strack, & Strzalecki, 2018; Golman, Hagmann, & Miller, 2015; Krajbich, Armel, & Rangel, 2010; Noguchi & Stewart, 2018; Roe et al., 2001; Trueblood, Brown, & Heathcote, 2014; Tsetsos, Chater, & Usher, 2012; Turner et al., 2018; Usher & McClelland, 2004; Webb, 2018). Preferences (defined as propensities to choose the available choice options[4]) can be represented as activation strengths in network nodes corresponding to the choice options. Many accumulator models also assume that decision makers represent the decision attributes or events using a second set of nodes. The connection between the preference node for a choice option and the various attribute or event nodes is proportional to the utilities of the choice option from that attribute or event. Activation spreads from the attribute or event nodes to the preference nodes through stochastic sequential sampling: At each time period, one attribute or event is sampled at random, with sampling probability proportional to the decision maker's beliefs about attribute weights or the probabilities of the events occurring. Subsequently the activation states of the preference nodes are updated based on the strength of their connections to the sampled node—that is, based on their utilities from the sampled attribute or

event. If particular attributes or events are especially salient, then the nodes corresponding to these attributes or events receive exogenous inputs, and are thus more likely to be sampled (e.g., Bhatia, 2013). Finally, after an internally or externally controlled deliberation time, the choice option with the most highly activated preference node is chosen. The assumptions of stochastic sampling and evidence accumulation allow these models to explicitly describe the full time course of a dynamic, stochastic choice process.

From Risky Choice to Strategic Choice

Our model is most closely related to a previous accumulator model of risky and uncertain choice, decision field theory (Busemeyer & Townsend, 1993; see also Bhatia, 2014 and Rieskamp, 2006). As in decision field theory, we assume that decision makers use two main layers of nodes: one to accumulate preferences (that is, propensities to choose the available choice options), and one to represent the probabilistic events involved in the decision. In the strategic context, the choice options are the strategies available to the decision maker and the events are the possible strategies the opponent may use. Thus, the strength of the directed connection from the node representing a strategy j for the opponent (i.e., the event in which the opponent plays j) to the node representing a decision maker's strategy i is proportional to the utility of strategy i for the decision maker given that the opponent plays strategy j. In the context of risky choice, decision makers sample the events according to the subjective probabilities they assign to their occurrence. Thus, in strategic choice, strategies that are more likely to be played by the opponent should be sampled more frequently and should thereby play a larger role in determining the decision maker's preferences.

Accumulator models of nonstrategic risky choice, such as decision field theory, assume that decision makers' beliefs about events (and subsequently sampling probabilities for these events) are fixed. Thus, changes in preferences (i.e., propensities to choose the available choice options) do not in any way influence sampling probabilities. For the most part, fixed beliefs in nonstrategic, risky settings appear to be fairly reasonable: decision makers' preferences do not influence the actual probability with which different events occur. This assumption is less reasonable in strategic settings in which two or more players make strategy choices that collectively determine the outcomes of the game, so that each player's utility depends on the others' choices as well as on his own. Sophisticated opponents, who can anticipate decision makers' choices, will adjust their own choices to maximize their reward. In order to successfully play strategic games, decision makers should form beliefs about their opponents' choices through strategic deliberation, and thus, the probability with which they sample events (i.e., opponents' chosen strategies) should evolve as their own preference evolves.

---

[4] Throughout this article we use the term "preference" to refer to the decision maker's (cardinal) propensities for choosing strategies, rather than exclusively to binary relations between pairs of strategies revealed through choice. This is in line with the way this term is used in many prior cognitive models of preferential decision making (e.g., Busemeyer & Townsend, 1993).

The Dual Accumulator Model

Bidirectional Accumulation

We propose a bidirectional dual accumulator model to represent a decision maker's strategic deliberation in a two-player, finite-strategy, one-shot, simultaneous-move game (see, e.g., Bhatia, 2016; Glöckner et al., 2014; Holyoak & Simon, 1999 for prior applications of bidirectionality in nonstrategic judgment and decision making). At each time period, the decision maker samples one of his opponent's strategies based on the activations of the nodes corresponding to these strategies (i.e., his beliefs about how likely the opponent is to choose the different strategies) as well as the exogenous salience of these strategies. Nodes corresponding to his own strategies receive activation proportional to the utilities of his strategies in the event that the opponent plays the sampled strategy. The decision maker then samples one of his own strategies based on the activation of the nodes corresponding to these strategies (i.e., his preferences for each of his own strategies) as well as the exogenous salience of these strategies, and uses this sample to update his beliefs about the opponent's choices. After a finite number of time periods, the decision maker chooses the strategy corresponding to the most highly activated node. (Note that strategy salience affects the sampling probabilities, but plays no role in strategy selection contingent on a sampling history.) We assume each player in a game goes through this deliberative process independently. In essence, decision makers have dynamically changing mental representations for their own preferences and also for their beliefs about their opponents' preferences, allowing them to deliberate intelligently using perspective taking and a sophisticated theory of mind.

Formal Structure

We define a finite-strategy two-player game with a set of pure strategies for each player, $S_1 = \{s_{11}, \ldots, s_{1N}\}$ and $S_2 = \{s_{21}, \ldots, s_{2M}\}$, respectively, and a pair of reward functions u_1 and u_2 that give each player's utility for each profile of pure strategies (s_{1i}, s_{2j}) (see, e.g., Hart, 1992).[5] Thus, if Player 1 selects s_{1i} and Player 2 selects s_{2j}, the utility for Player 1 is $u_1(s_{1i}; s_{2j})$ and the utility for Player 2 is $u_2(s_{2j}; s_{1i})$, with $\boldsymbol{u}_{ij} = (u_1(s_{1i}; s_{2j}), u_2(s_{2j}; s_{1i}))$.

We represent the decision process for any given player using two layers of interconnected accumulating nodes. If the decision maker has to choose from the set of strategies $S_1 = \{s_{11}, \ldots, s_{1N}\}$, then the preference layer in our model consists of N nodes, with node i representing strategy s_{1i}. The activation of node i at time t, $A_{1i}(t)$ corresponds to the decision maker's preference for strategy i at time t. Correspondingly if the opponent has the set of available strategies $S_2 = \{s_{21}, \ldots, s_{2M}\}$, then the belief layer in our model consists of M nodes, with node j representing strategy s_{2j}. The activation of node j at time t, $A_{2j}(t)$ corresponds to the beliefs that the decision maker has about the opponent's propensity to choose strategy j, at time t. We also denote the salience bias of any strategy i (for the decision maker) or j (for the opponent) as σ_{1i} or σ_{2j}, respectively. These salience biases σ_{1i} and σ_{2j} are independent of the decision process and are determined by various exogenous, nondecision factors, such as the prominence of the strategy in the presentation of the game or in the memory of the decision maker.

At each time period t, the decision maker draws (with replacement) one sample of the opponent's strategies. We assume that a softmax (logit) function, with stochasticity parameter $\lambda \geq 0$, determines the effect of activation strength and the exogenous salience bias on sampling probability. Thus, the probability of sampling strategy j at time t is given by:

$$p_j = \frac{e^{\lambda(A_{2j}(t-1)+\sigma_{2j})}}{\sum_{k=1}^{M} e^{\lambda(A_{2k}(t-1)+\sigma_{2k})}}.$$

Increasing λ reduces the noise involved in the sampling of strategies, so that for very high values of λ, decision makers almost deterministically attend only to the single strategy with greatest total activation and salience. Note that the magnitude of λ is meaningful only in relation to the magnitude of the activations and the saliences.

If the opponent's strategy j is sampled, then the decision maker observes the utility for each strategy i conditional on the opponent playing this sampled strategy: $u_1(s_{1i}; s_{2j})$. The decision maker's preferences are then updated based on this calculated utility, so the activation for each strategy i becomes

$$A_{1i}(t) = A_{1i}(t-1) + u_1(s_{1i}; s_{2j}).$$

As discussed, beliefs about the opponent's strategies are themselves updated based on the utility the opponent would derive conditional on a sample of the decision maker's strategies. Thus, after updating activation states $A_{1i}(t)$, the decision maker draws one sample of his own strategies. The probability of sampling strategy i at time t is given by:

$$q_i = \frac{e^{\lambda(A_{1i}(t)+\sigma_{1i})}}{\sum_{k=1}^{N} e^{\lambda(A_{1k}(t)+\sigma_{1k})}}.$$

Then, after sampling strategy i, the updated activation for each opponent strategy j is

$$A_{2j}(t) = A_{2j}(t-1) + u_2(s_{2j}; s_{1i}).$$

The deliberation process begins with nodes having no initial activation: $A_{1i}(0) = 0$ for all i; $A_{2j}(0) = 0$ for all j. Activation accumulates according to these equations until a time $t = T$. At this time, the most preferred strategy—that is, the one whose node has the highest activation—is the strategy that is chosen by the decision maker. (If multiple strategies have equal highest activation, the choice between them is random.) The selection of the most highly activated strategy node at time T is noiseless and does not depend on strategy salience. The parameter T describing the number of steps of sampling and accumulation corresponds to an exogenous time limit on the deliberation process, and a proxy for it could be the amount of time taken by the decision maker to make his choice. (The mapping from time steps in our model [T] to physical time could depend on individual cognitive ability (Burnham, Cesarini, Johannesson, Lichtenstein, & Wallace, 2009; Carpenter, Graham, & Wolf, 2013; Gill & Prowse, 2016), so there still may be individual variation in T even when all decision makers have the same physical time constraints.) The proposed dual ac-

---

[5] Utility is assumed here to be specified on an interval scale, unique up to positive affine transformations, but its scale matters because choice probabilities may be sensitive to the magnitudes of utility differences. In our applications of our model to various games we assume that utility is identical to monetary payoff (though we do normalize utilities to ensure comparability across games).

cumulator model is summarized in Figure 1, and Appendix A provides pseudocode for simulating the model.

We intend for the model's parameter values to be specified for each individual person and to remain fixed across games. Fixing the stochasticity parameter λ across games requires a convention for normalizing utilities. The model is sensitive to the cardinal magnitude of utilities, but the model is invariant if we rescale all utilities and saliences by a constant factor κ and simultaneously rescale λ by the factor $\frac{1}{\kappa}$. In this paper we scale utilities to lie between 0 and 100.

Explaining Behavior

Overview of Key Patterns

We now review a series of behavioral patterns and then demonstrate how our dual accumulator model accounts for them. We exclude behavioral patterns driven by social preferences in games like the prisoner's dilemma or the dictator game. Our model would need to be augmented with a notion of social utility to account for them.

Stochastic choice. Empirical research on both strategic and nonstrategic decision making has documented both heterogeneity across subjects and intrinsic variability in the responses of individual decision makers: Not only do different individuals make different choices, but a given individual may also respond differently to a given game on different occasions, even without intervening feedback (Fragiadakis, Knoepfle, & Niederle, 2016; Hyndman, Terracol, & Vaksmann, 2015; see also Loomes, 2015

for a review of the empirical evidence about individual stochastic choice in the domain of risk).

Limited iterated reasoning. Common knowledge of rationality implies that players not only never play strictly dominated strategies, but also can eliminate them from consideration and then rule out iteratively dominated strategies, and can then repeat this process (known as iterated elimination of strictly dominated strategies) ad infinitum. In games in which players have incentives to slightly undercut each other, this implies that undercutting gets taken to its logical extreme and ends only when players reach the lower bound of their strategy set. In practice, this is rarely the case.

Consider, for example, the traveler's dilemma game. The traveler's dilemma is a generalization of the famous prisoner's dilemma, conceived in order to demonstrate the absurdity of unlimited iterated reasoning in a one-shot game (Basu, 1994). In the original parable, two travelers have lost identical items and must request some amount of compensation. The airline (which is responsible for the lost luggage) will accept the lower claim as valid and pay that amount to both players, and, to deter lying, will penalize the higher claimant with a fee and will reward the lower claimant with a bonus. We represent this game with the strategy sets $S_1 = S_2 = \{20, 30, \ldots, 90\}$, where x_{1i} and x_{2j} correspond to the amounts (in dollars) associated with strategies s_{1i} and s_{2j}, and we have utilities

$$u_{ij} = \begin{cases} (x_{1i} + \gamma, \ x_{1i} - \gamma) & \text{if } x_{1i} < x_{2j}; \\ (x_{1i}, \ x_{2j}) & \text{if } x_{1i} = x_{2j}; \\ (x_{2j} - \gamma, \ x_{2j} + \gamma) & \text{if } x_{1i} > x_{2j}. \end{cases}$$

Accumulation of Preferences

Strategies for Decision Maker

Stochastic Sampling

Strategies for Opponent

Accumulation of Beliefs
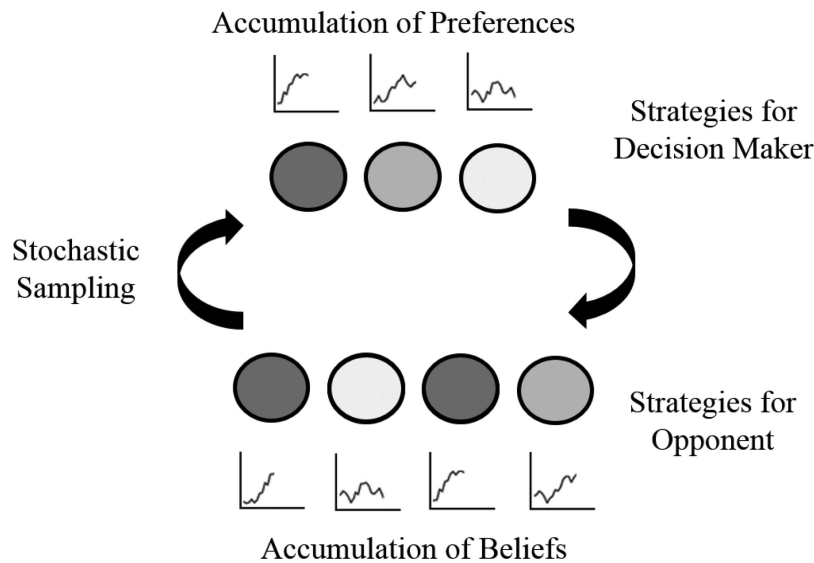
Figure 1. Illustration of the dual accumulator model. The model has two layers of accumulating nodes, one representing the decision maker's preferences and the other representing the decision maker's beliefs regarding the opponent's preferences. At each time period the decision maker samples one of his opponent's strategies and updates his propensities to choose his own strategies based on the payoffs they provide contingent on the opponent playing the sampled strategy. The decision maker then samples one of his own strategies and similarly updates his beliefs regarding the opponent's preferences. Strategy sampling probabilities depend on their activation and salience. Decisions are made in favor of the most highly activated strategy upon reaching an exogenous time limit.

The payoff parameter γ corresponds to the reward/penalty offered by the airline, and is set so that $10 < \gamma \le 20$. Table 1 presents a payoff matrix for this game.

The airline's scheme, of course, does not actually reward honesty; it rewards undercutting the other traveler. The best response is always to claim exactly 10 less than the other traveler does (if it is feasible to do so). As a result, the only Nash equilibrium strategy for both players is to claim 20. This is the lowest amount a player can claim, and players cannot make themselves better off by claiming a different amount, if their opponent also claims 20. In experiments average claims actually are well above the lower bound that Nash equilibrium predicts and are often concentrated closer to the upper end of the range of possible claims (Branas-Garza, Espinosa, & Rey-Biel, 2011; Capra et al., 1999). Similar evidence for limited iterated reasoning comes from experiments on the 11–20 game (Arad & Rubinstein, 2012), which we describe in more detail in the online supplemental materials, and the p-beauty contest, a multiplayer game in which typical decision makers appear to undercut a uniform distribution of other players' strategies once or twice rather than select the unique Nash equilibrium strategy at the lower extreme of the strategy space (Ho, Camerer, & Weigelt, 1998; Nagel, 1995; see also Costa-Gomes & Crawford, 2006 for a related two-player game).

Payoff sensitivity. Beyond showing that people do not actually race to the bottom in games like the traveler's dilemma, Capra,

Goeree, Gomez, and Holt (1999) data show that claims are higher when the reward/penalty, γ, is lower. Nash equilibrium predicts that responses in the traveler's dilemma should be independent of γ, as changing payoffs without changing best responses (e.g., changing how bad it is to play a suboptimal strategy, but not changing when a strategy is suboptimal) should have no effect on choice behavior. In practice, the finding of payoff sensitivity is not an isolated occurrence. The pattern of strategies being chosen more frequently when a varying payoff parameter gives those strategies higher payoffs also occurs in the minimum-effort coordination game (Goeree, & Holt, 2001; Van Huyck, Battalio, & Beil, 1990), the stag-hunt game (Schmidt, Shupp, Walker, & Ostrom, 2003), and the matching pennies game (Goeree & Holt, 2001), all of which we describe in more detail in the online supplemental materials. Indeed, in symmetric coordination games and extended coordination games, players are sensitive to payoff magnitudes, more frequently choosing a strategy when its payoff against a dominated strategy is increased or when the potential cost of a coordination failure is decreased (Anderson, Goeree, & Holt, 2001; Goeree & Holt, 2001).

Risk/reward tradeoffs in coordination games. Coordination games are a class of games with multiple pure-strategy Nash equilibria, in which players are incentivized to align their strategy choices with each other. Due to the presence of multiple Nash equilibria, standard game theory requires some additional criteria

Table 1

Payoff Matrix of the Traveler's Dilemma Game, With Player 1 Choosing Rows (Payoffs in Bottom Left) and Player 2 Choosing Columns (Payoffs in Top Right)

	20	30	…	80	90
20	20 / 20	$20-\gamma$ / $20+\gamma$	…	$20-\gamma$ / $20+\gamma$	$20-\gamma$ / $20+\gamma$
30	$20+\gamma$ / $20-\gamma$	30 / 30	…	$30-\gamma$ / $30+\gamma$	$30-\gamma$ / $30+\gamma$
…	… / …	… / …	…	… / …	… / …
80	$20+\gamma$ / $20-\gamma$	$30+\gamma$ / $30-\gamma$	…	80 / 80	$80-\gamma$ / $80+\gamma$
90	$20+\gamma$ / $20-\gamma$	$30+\gamma$ / $30-\gamma$	…	$80+\gamma$ / $80-\gamma$	90 / 90

to make precise predictions in these games. Schelling (1960) proposed that context can determine a focal point, a particular equilibrium strategy that most players will expect and thus choose. Harsanyi and Selten (1988) proposed two equilibrium selection criteria based on game payoffs: payoff dominance (i.e., Pareto preference among the equilibria) and risk dominance (i.e., minimizing losses in the event of coordination failure). In some games (e.g., the hi-lo game described in the online supplemental materials), these selection criteria align and actual choice behavior tends to be fairly predictable, in line with these criteria. In other games (e.g., the battle of the sexes game described in the online supplemental materials), the criteria lead to different predictions or are inapplicable. Across games the risk-dominance criterion tends to predict actual behavior somewhat better than the payoff-dominance criterion, but neither predicts as well as a behavioral rule that acknowledges risk-reward tradeoffs (Camerer, 2003a; Haruvy & Stahl, 2007; Schmidt et al., 2003; Straub, 1995).

In some games with multiple pure-strategy Nash equilibria, decision makers often do not choose any of the equilibrium strategies when the potential costs of miscoordination are too great. In fact, even when a payoff dominant pure-strategy Nash equilibrium exists, players may often choose a strategy that cannot be rationalized by any pure strategy if it provides consistently high payoffs against all of the opponent's strategies. This is the case with the Kreps game, in which decision makers (in the role of Player 2) are most likely to choose the strategy not in Nash equilibrium (Goeree & Holt, 2001; Kreps, 1995). In this game, with payoff matrix presented in Table 2, the pure-strategy equilibria are top left and bottom right, and the mixed strategy equilibrium randomizes between top and bottom and left and middle. The strategy labeled non-Nash is not part of any Nash equilibria. However, this strategy is the one that is most likely to be selected by decision makers in the role of Player 2. (Decision makers in the role of Player 1 tend to choose the top strategy, which is part of the payoff dominant equilibrium.)

Salience. Another set of behavioral patterns involves strategy salience. In many games, strategies with salient labels are more likely to be chosen, while in some other games, strategies with salient labels are actually less likely to be chosen. An example of the former scenario is a simple symmetric coordination game with multiple pure-strategy equilibria, none of them strictly risk dominant because of the payoff symmetry. In the simplest case, this game offers decision makers the choice of heads or tails and they win if they both choose the same side of the coin. A version of this game with more context involves coordinating on one out of a number of meeting points in a new city. If both decision makers go to the same meeting spot, they get to spend time together, which is a good outcome. If they go to different meeting spots, they spend time alone, which is undesirable. In coordination games with either artificial or realistic contexts, decision makers are able to achieve a fairly high rate of coordination by selecting salient strategies, such as heads in the first game, and prominent landmarks (e.g., Grand Central Station in New York City) in the second game (Schelling, 1960). Similar results are obtained in laboratory studies of these games that manipulate strategy salience using distinguishing marks or labels (Crawford et al., 2008; Hargreaves Heap et al., 2014; Mehta et al., 1994; Rubinstein, Tversky, & Heller, 1997). For example, decision makers in the laboratory can coordinate in games offering a large number of different, payoff identical, strategies, if one of these strategies is circled, underlined, or made salient using some other technique.

Salience has more subtle effects on strategic choice in games that do not involve coordination. Consider for example, the hide-and-seek game, in which the first player has to choose somewhere to hide, and the second payer has to choose somewhere to seek. Player 1 wins if they choose different locations, and Player 2 wins if they choose the same location. The utilities are simply $u_{ij} = \begin{cases} (0, 100) \text{ if } s_{1i} = s_{2j}; \\ (100, 0) \text{ if } s_{1i} \neq s_{2j}. \end{cases}$ A payoff matrix for this game is given in Table 3. The Nash equilibrium calls for uniformly mixed strategies for both players, regardless of salience. Empirically, decreasing the salience of one of the hiding spots makes decision makers in the role of Player 1 more likely to choose that spot to hide, and it also makes subjects in the role of Player 2 more likely to choose that spot to seek (Crawford & Iriberri, 2007; Rubinstein et al., 1997). A single highly salient (or "odd-one-out") strategy, on the other

Table 2
Payoff Matrix for the Kreps Game, With Player 1 Choosing Rows (Payoffs in Bottom Left) and Player 2 Choosing Columns (Payoffs in Top Right)

	Left	Middle	Non-Nash	Right
Top	70 / 100	69 / 60	66 / 62	10 / 64
Bottom	10 / 60	40 / 62	66 / 66	68 / 70

GOLMAN, BHATIA, AND KANE

Table 3
Payoff Matrix for the "ABAA" Hide-and-Seek Game, With Player 1 Choosing Rows (Payoffs in Bottom Left) and Trying to Hide and Player 2 Choosing Columns (Payoffs in Top Right) and Trying to Seek

	A	B	A	A
A	100 / 0	0 / 100	0 / 100	0 / 100
B	0 / 100	100 / 0	0 / 100	0 / 100
A	0 / 100	0 / 100	100 / 0	0 / 100
A	0 / 100	0 / 100	0 / 100	100 / 0

(Each cell shows the top-right payoff [Player 2] / bottom-left payoff [Player 1].)

hand, does not get chosen very frequently in hide-and-seek games, generally with less than uniformly random chance by hiders and with roughly uniform, or perhaps slightly higher, probability by seekers (Hargreaves Heap et al., 2014). Similarly, in a discoordination game in which both players want their choices to differ, a single highly salient strategy is not generally more likely to be chosen than any other strategy (Hargreaves Heap et al., 2014).

Predicted Behavior

Details of simulations. We use our dual accumulator model to simulate choices in a wide range of games. We describe here the predicted behavior of the model in the traveler's dilemma, the Kreps game, and the hide-and-seek game with salient labels. In the online supplemental materials we describe additional behavioral predictions for many other games using the same set of parameter values. Because the dual accumulator model relies on a stochastic sampling process, it makes stochastic choice predictions. For each of the games (and each separate player in these games, when games are asymmetric) and each set of parameter values, we simulate the model 1,000 times and report aggregate choice probabilities. When not explicitly specified, we assume that all strategies have equal exogenous salience, set to $\sigma_{1i} = \sigma_{2j} = 0$ without loss of generality.

We vary model parameters in the range $\lambda \in [.01, 10]$, $T \in [1, 30]$, and find that as long as the parameter values are not too large, the model generally produces qualitatively good predictions of behavior, consistent with the empirical findings we have reviewed, across the set of games we have considered. As we describe our model's predictions, we also discuss the shortcomings of simply assuming uniformly random sampling across all strategies, deterministic sampling of the most highly activated strategy, or unlimited decision time. Pseudocode for using the model to simulate behavior and generate predictions for these games is provided in Appendix A. We release code for simulating our model in online supplemental materials.

Traveler's dilemma. In the traveler's dilemma the dual accumulator model typically predicts behavior consistent with limited iterated reasoning as long as λ and T are not too large. This is demonstrated in Figure 2, which displays the probability distribution over strategies in the set $\{20, 30, \ldots, 90\}$ for $\gamma = 11$ and $\gamma = 19$, with varying values of λ and T. Instead of predicting that players always claim the lowest possible amount, as in Nash equilibrium, the model with $\lambda \leq .1$ and $\gamma = 11$ generates a distribution of choices that is concentrated heavily in the upper half of the range when the reward/penalty is small and shifted downward when the reward/penalty is large. The model clearly
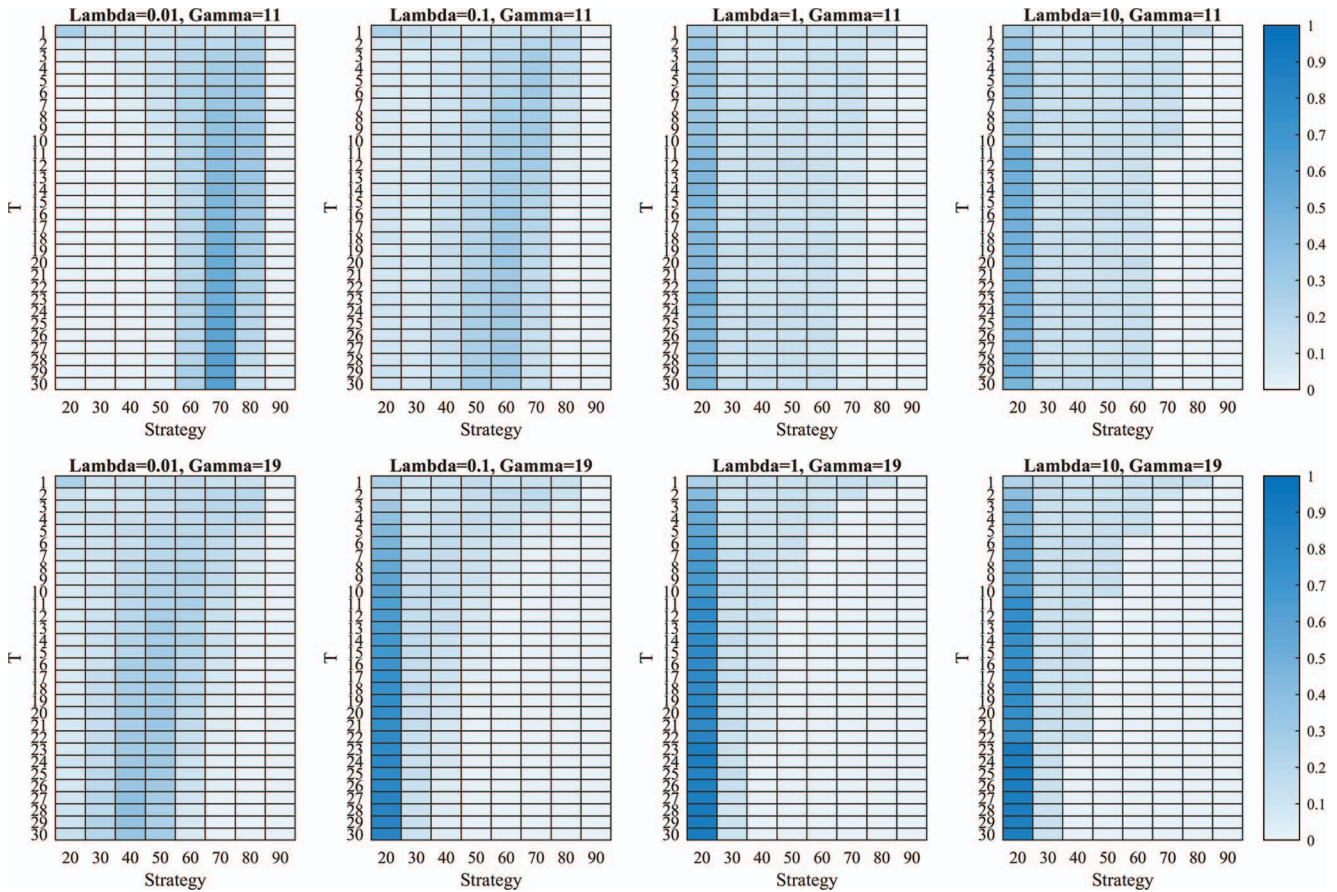
Figure 2. Simulated distribution of choices in the traveler's dilemma when the reward/penalty is small ($\gamma =$ 11) or large ($\gamma = 19$) for values of T from 1 to 30 and given values of λ. Shading indicates the probability that each strategy is chosen. See the online article for the color version of this figure.

displays payoff sensitivity. For a larger value of the reward/penalty parameter ($\gamma = 19$), the distribution of choices is uniformly smaller.

The intuition behind the model's predictions is appealing. The decision nodes in the network accumulate activation proportional to the payoffs for each strategy rather than sequentially activating only best responses. For low rewards/penalties, that is, low values of γ, the payoffs when both players make high claims are significantly higher than the payoffs when there is a low claim. The potential cost of missing out on this high payoff dwarfs the cost of making a higher claim than the opponent or the benefit of making a lower claim than the opponent. So, a few samples (or even a single sample) of the opponent playing a high claim will lead to high activation for one's own high claims. As beliefs about the opponent's strategy are updated, there will be more samples of high claims, and strategies involving an additional step of undercutting can accumulate the most utility. While high claims typically accumulate more activation, the strategy of making the highest possible claim never accumulates as much activation as the strategy involving one step of undercutting because it is dominated by it. Thus, the highest possible claim is never chosen. The number of steps of undercutting that does occur depends on payoff magnitudes.

Increasing the reward/penalty parameter γ encourages undercutting. Although it does not affect best responses (that is, the ranking of payoffs in any given sample of play), it does affect the accumulation of payoffs over time, so strategies involving more undercutting can accumulate activation more quickly.

Stochastic sampling plays an important role in the emergence of payoff sensitivity. The magnitudes of payoff differences affect the probabilities of sampling each strategy. The degree of responsiveness to the payoff parameter γ that we observe in the predicted choices for this game depends on the logit sampling parameter λ. Comparing across the columns of Figure 2, we see smaller shifts in the distribution of choices from a change in the reward/penalty parameter γ as the parameter λ increases.

The model also makes new predictions about the relationship between decision time and the strategy chosen in the traveler's dilemma. Each step of undercutting takes time to sample the strategy that already has the highest activation and then to accumulate activation to undercut. Both the decision maker's preferred claim and the beliefs about the opponent's claim should thus decrease over time. Comparing across the rows of the heat maps in Figure 2, we observe lower claims when the decision time T is larger. Indeed, experiments have revealed that decision makers take longer to choose the lowest claim than the highest claim

(though decision time patterns for intermediate responses are not so clear; Rubinstein, 2007). Similar results have been documented in the p-beauty contest game and in a market entry game (Agranov, Caplin, & Tergiman, 2015; Kocher & Sutter, 2006; Lindner, 2014). We also see in Figure 2 that responsiveness to the payoff gets stronger as the decision time T increases. Greater decision time provides more opportunity for payoffs to accumulate.

Overall, for a range of moderate parameter values, the model makes reasonable predictions. Extreme undercutting only occurs with very large values of λ or T, that is, when poorly performing strategies are rarely sampled and there are many periods of sampling and iterative updating. Assuming deterministic sampling of best responses or unlimited decision time would thus lead to poor behavioral predictions for the traveler's dilemma. Conversely, assuming uniformly random sampling would lead to unreasonably high odds of choosing 80 relative to 70, underestimating people's ability to put themselves in their opponents' shoes and think strategically about their responses.

The Kreps game. The simulated choice distributions in the Kreps game when $\lambda \leq 0.1$ also resemble actual human behavior observed in the lab. This is shown in Figure 3, which displays the choice distributions for both players in the Kreps game, for varying values of T and λ. Here Player 2 is most likely to choose the non-Nash strategy despite the availability of an alternative strategy that is part of a payoff dominant Nash equilibrium. The non-Nash strategy is the best worst-case strategy in that it never earns an extremely low payoff. Accumulation of payoffs against a fairly unpredictable player 1 thus favors the non-Nash strategy. By incorporating stochastic sampling, the dual accumulator model effectively allows decision makers to consider the potential costs of miscoordination.

Noisier sampling, that is, smaller values of the logit sampling parameter λ, and quicker decisions, that is, smaller values of the decision time T, both increase the probability of Player 1 choosing the top strategy. The top strategy, in addition to being part of the payoff dominant equilibrium, has relatively low costs when miscoordination occurs. When λ and T are both small, Player 1 chooses "top" the vast majority of the time, in agreement with the empirical record. When λ and T are both large, on the other hand,
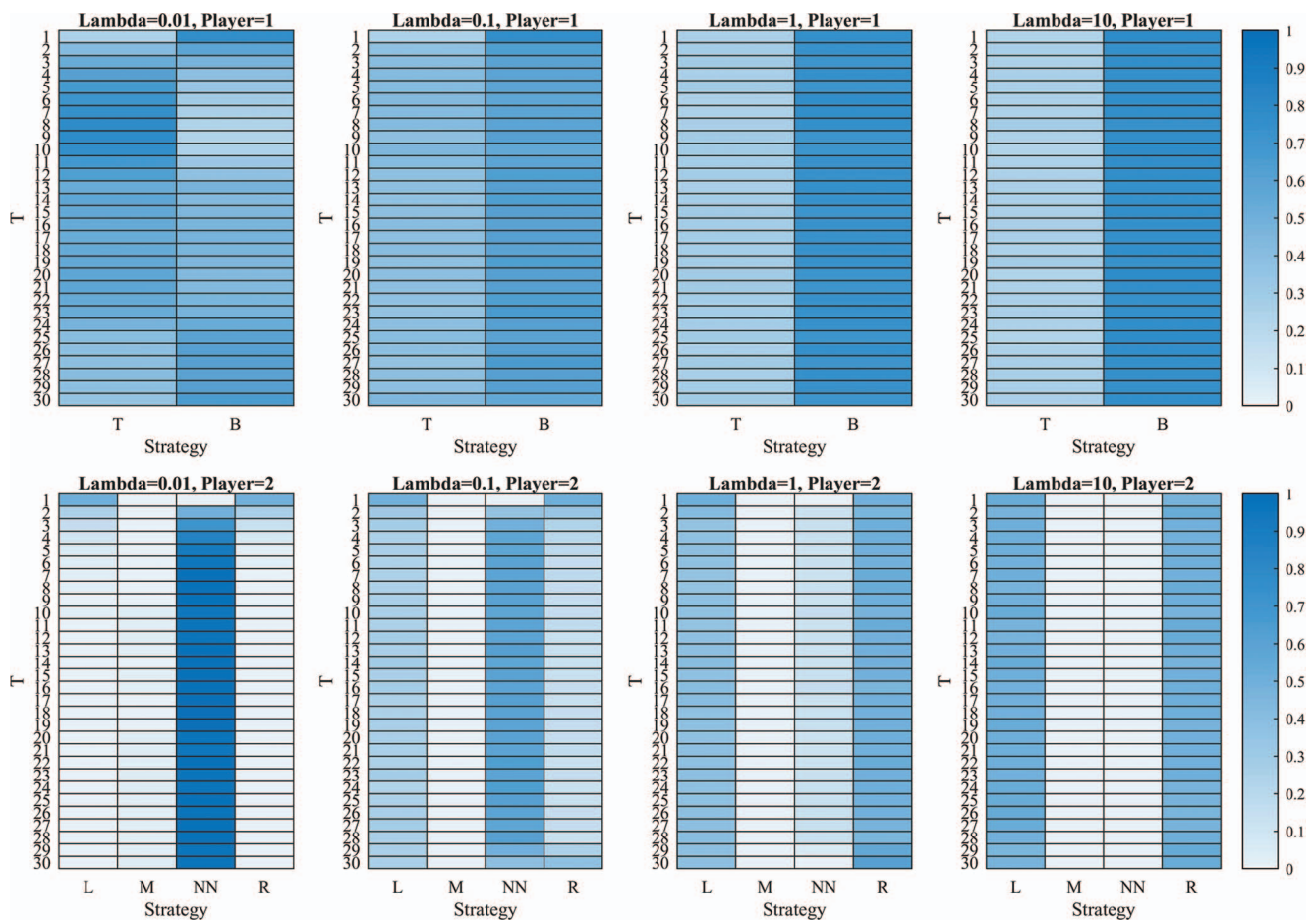
Figure 3. Simulated distribution of choices for Players 1 and 2 in the Kreps game for values of T from 1 to 30 and given values of λ. The labels T and B represent Player 1's strategies top and bottom respectively. The labels L, M, NN, and R correspond to Player 2's strategies left, middle, non-Nash, and right. Shading indicates the probability that each strategy is chosen. See the online article for the color version of this figure.

"bottom" is chosen more frequently than top because bottom accumulates more payoff every time the non-Nash strategy is sampled. With quite reliable sampling and a long decision time, Player 2's deliberation favors iterative best responses to his non-Nash strategy. In the extreme, if λ were to grow unrealistically large, Player 2 would effectively examine only best responses and would get locked into an equilibrium strategy rather than accumulating a strong propensity to choose the non-Nash strategy.

Hide and seek with salient labels. We also apply the dual accumulator model to the hide-and-seek game with four strategies labeled "ABAA." Following Bar-Hillel (2015), we consider the second strategy to be particularly salient because of its distinctive label and the first and last strategies to be somewhat salient because of their end locations, leaving the third strategy as the least salient. We represent the strategies' degrees of salience as $\sigma_1 = \sigma_2 = [50, 100, 0, 50]$. (These particular salience values are simply round numbers of comparable magnitude to the payoffs in the game. Other assumptions would also be reasonable—e.g., Crawford & Iriberri, 2007 allow for cases where the first and last strategies are more salient than the second strategy, and show that this assumption leads to better fits of

their Level-k model to the data. More generally, we could use choice data to fit a salience function alongside our model, but we would have to watch out for overfitting.)

Figure 4 shows the probabilities of Player 1 and Player 2 choosing each of the four strategies for varying values of λ and T. Consistent with Rubinstein, Tversky, and Heller's (1997) empirical evidence, summarized by Crawford and Iriberri (2007, Table 1, p. 1735), we find that for most parameter values both hiders and seekers most often choose the third (least salient) strategy, "middle A." The remaining strategies are still played occasionally, but less frequently. The precise choice probabilities vary slightly with different parameter values, but we do not identify a clear systematic relationship.

The predicted behavior is intuitive. Higher salience makes a strategy more likely to be sampled. In a coordination game, a strategy accumulates more activation when the same strategy is sampled for the opponent's play. The deliberation is self-reinforcing, and the salient strategy becomes focal. On the other hand, in a hide-and-seek game, players want to be a step ahead of their opponents. Player 1, trying to hide, will accumulate more activation for strategies that are rarely sampled. Player 1 thus quickly develops a preference for the
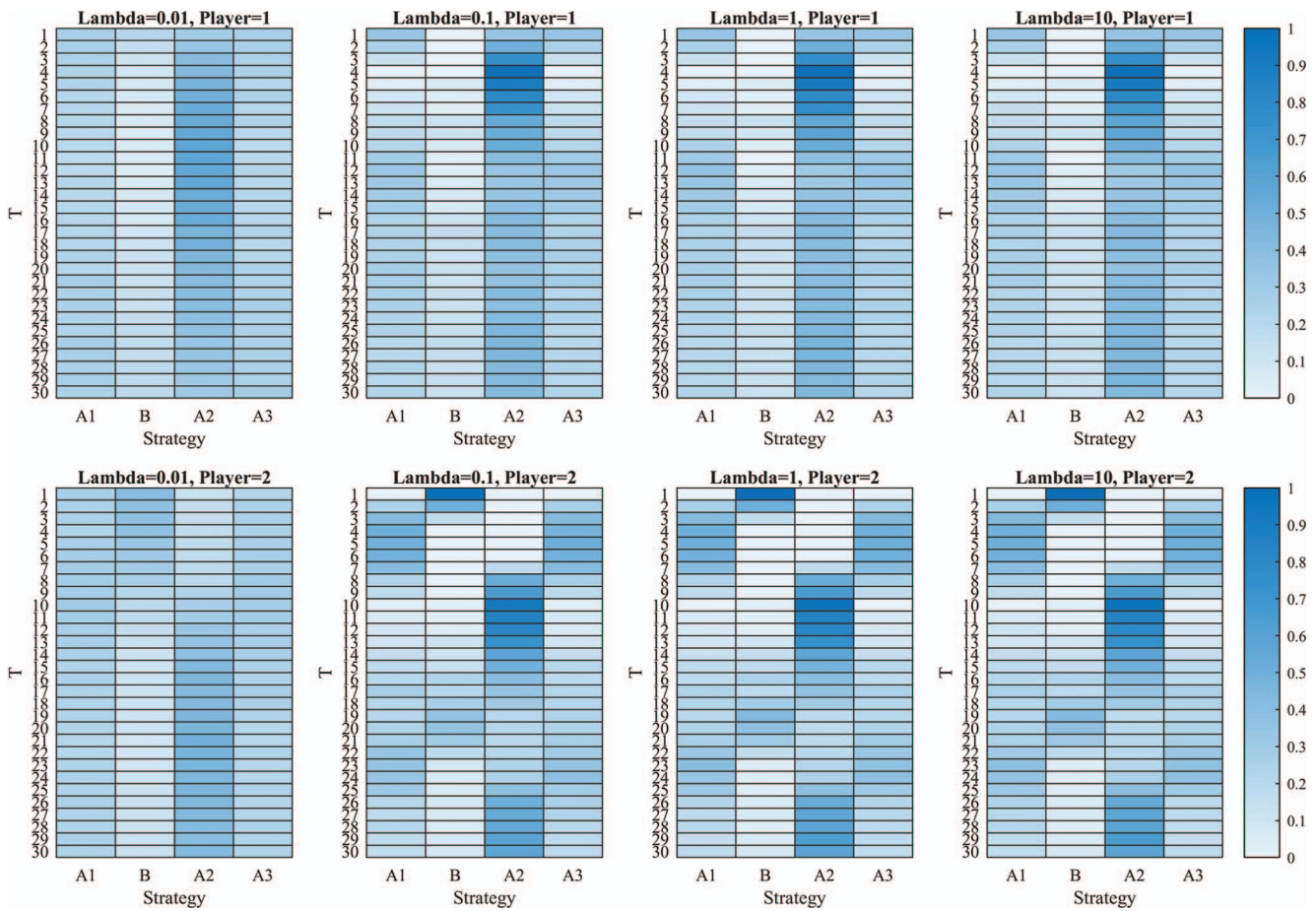

Figure 4. Simulated distribution of choices in the hide-and-seek game with strategies having salient labels "ABAA" ($\sigma_1 = \sigma_2 = [50, 100, 0, 50]$), for values of T from 1 to 30 and given values of λ. Player 1 hides and Player 2 seeks. Shading indicates the probability that each strategy is chosen. See the online article for the color version of this figure.

least salient strategy. Player 2, trying to seek, will accumulate more activation for strategies that are frequently sampled. Initially, the more salient strategies are sampled more frequently, but as the deliberation proceeds, Player 2 recognizes that Player 1 has an inclination to choose the least salient strategy, and Player 2 develops a preference for it as well. This reasoning, if allowed to go on long enough, would become circular. Thus, finite decision time plays a critical role in cutting off the deliberation and triggering a decision.

Cross-game predictions. Our goal is to make good predictions of behavior across a variety of games with the model's parameter values fixed across games for each individual subject (but allowed to vary between subjects). Allowing different parameter values for each game would give the model too much flexibility. Fixing parameter values across games also leads to new predictions about correlations across games. For example, an individual who tends to take longer to make decisions (i.e., with higher T) is more likely to choose a lower value in the traveler's dilemma and also to choose the top strategy in the role of Player 1 in the Kreps game. We are not aware of any existing dataset describing individual behavior across all the games we consider, so our predictions about cross-game correlations remain untested at present.

We can at least make aggregate (population-level) predictions across a variety of games and compare them to the existing empirical data. In principle, we should assume a distribution of parameter values to make population-level predictions. Just for simplicity, however, we will present the dual accumulator model's predictions for a single pair of parameter values, $\lambda = .01$ and $T = 10$, fixed across a variety of games for which comparable empirical data exists. Comparisons across games require consistent units of payoff, and to generate the dual accumulator model predictions, we adopt the specifications given by the payoff matrices presented here and in the online supplemental materials, so that all games have roughly comparable payoff magnitudes. These payoff matrices are normalized versions of the games that have been studied empirically. Our specification of the traveler's dilemma also departs from the version studied empirically in that we adopt a much smaller strategy set. This precludes direct comparison of our predictions and the existing empirical data, but we can still observe limited iterated reasoning in both cases.

Figure 5 presents the predicted behavior of the dual accumulator model, representative empirical data, and Nash equilibria for the traveler's dilemma, the Kreps game, the "ABAA" hide-and-seek game, the 11–20 game, a pair of stag hunt games, a pair of matching pennies games, a battle of the sexes game, the "odd-one-out" hide-and-seek game, and the "odd-one-out" discoordination game. With parameter values fixed across these games, the dual accumulator model generates limited iterated reasoning in the traveler's dilemma and 11–20 game, payoff sensitivity in the stag hunt game, responsiveness to risk-reward tradeoffs in the Kreps game and battle of the sexes, and salience effects in the hide-and-seek and discoordination games. The model's predictions are not perfect: It does not account for choices of the maximum claim in the traveler's dilemma, a weakly dominated strategy that may nevertheless be popular because of a round-number preference or because it is cooperative; it slightly underestimates propensity to choose stag in the stag hunt game, which may appeal to subjects who engage in team reasoning; and it reacts a bit too strongly to salience, likely because we do not optimally fit a salience function,

but rather arbitrarily adopt round numbers as salience values. Nevertheless, on the whole, the dual accumulator model makes qualitatively reasonable predictions.

Stochasticity and dominance. Stochastic choice is necessary to account for behavioral findings in games. Our dual accumulator model admits trial-to-trial variability in choice, using a psychologically grounded stochastic sampling mechanism that is commonly used to describe stochastic choice behavior in nonstrategic settings (see again Busemeyer & Townsend, 1993; also Bhatia, 2014; Rieskamp, 2006). Existing models have mostly relied on noise in utility or in choice responses to generate stochastic choice predictions (e.g., Goeree & Holt, 2004; Goeree, Holt, & Palfrey, 2005; McKelvey & Palfrey, 1995). This approach implies that accidental selection of strictly dominated strategies is as likely other similarly costly errors. Empirically, however, strictly dominated strategies are selected quite rarely; for example, in various games examined by Costa-Gomes, Crawford, and Broseta (2001), decision makers choose dominated strategies less than 10% of the time (see also Stahl & Haruvy, 2008 for additional data and Loomes, 2015 for a related discussion). Weakly dominated strategies are chosen with some frequency in games with many strategies, such as second-price auctions (Kagel, Harstad, & Levin, 1987) and two-person p-beauty contest games (Camerer et al., 2004; Grosskopf & Nagel, 2008).

While generating stochastic choice within the space of non-dominated strategies, the dual accumulator model predicts that strictly dominated strategies will never be chosen in any game. The deliberation itself is a stochastic process, but the selection of a strategy at the end of the deliberation is noiseless. Given any history of sampling (regardless of T and λ), a strictly dominated strategy will accumulate less activation than a strategy that dominates it, so the strictly dominated strategy cannot be chosen.[6] (This extreme prediction can be weakened by introducing an additive noise term of relatively small magnitude compared with the noise terms necessary to generate stochastic choice when stochastic sampling is neglected.). Weakly dominated strategies can persist for small T, as they may be randomly chosen when the history of sampling includes only opponent strategies against which they do equally well as the strategies that dominate them. However, weakly dominated strategies must be chosen less frequently than the strategies that dominate them, and they must be chosen less frequently as T increases.

Comparison With Existing Behavioral Game Theoretic Models

Overview of Models

Our dual accumulator model shares common elements with many other models in behavioral game theory. Some earlier models have proposed that individuals engage in an iterative process of deliberation that may terminate before reaching a point of self-consistency. Stahl and Wilson (1994, 1995) and Nagel (1995), for example, propose *Level-k reasoning*, which hypothesizes that peo-

---

[6] Stochastic choice based on sequential sampling has already been shown to avoid dominated choices in multi-attribute and risky choice (e.g. Busemeyer & Townsend, 1993; Roe et al., 2001).
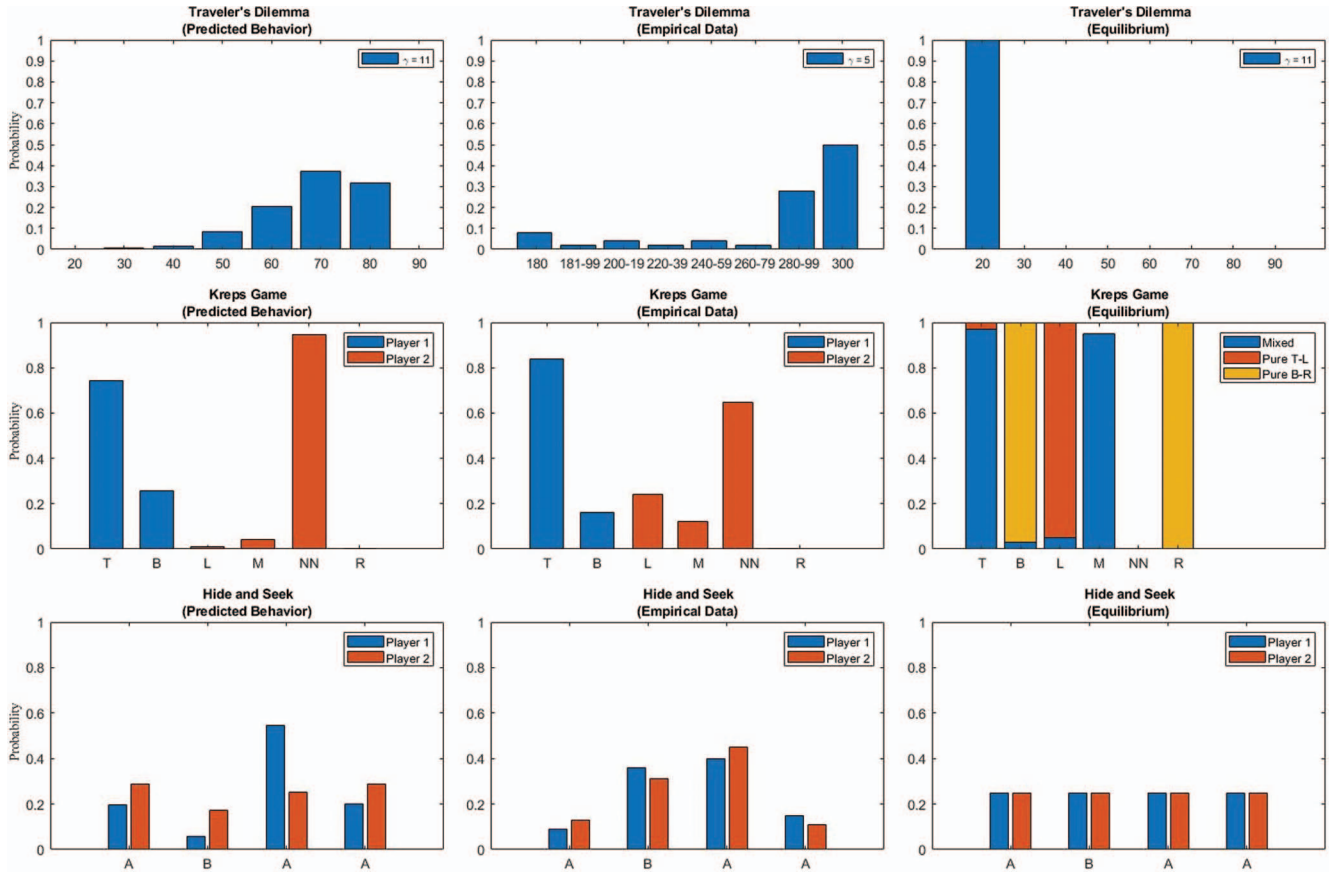
Figure 5. Comparison of predicted behavior, representative empirical data, and Nash equilibria for a variety of games (for which the empirical data exists). The predicted behavior is the result of simulating the dual accumulator model 1,000 times with fixed parameter values $\lambda = .01$ and $T = 10$, ignoring individual heterogeneity. The predictions are for the specifications of the games presented here and in the online supplementary materials; these payoff matrices are normalized versions of the games studied empirically, and the traveler's dilemma game in particular is specified with a smaller strategy set (and a different value of the payoff parameter γ) that makes direct comparison with the empirical data difficult. The "battle of the sexes" game is specified with $\gamma_1 = \gamma_2 = 33$. We assume salience values $\sigma_1 = \sigma_2 = [50, 100, 0, 50]$ in the "ABAA" hide-and-seek game, $\sigma_1 = \sigma_2 = [100, 0, 0, 0]$ in the "odd-one-out" games, and no salience effects in the other games. The original empirical data was reported by: Goeree and Holt (2001) for the traveler's dilemma game, the Kreps game, and the matching pennies games; Rubinstein et al. (1997) for the "ABAA" hide-and-seek game; Arad and Rubinstein (2012) for the 11–20 game; Schmidt et al. (2003) for the stag hunt games; Cooper, DeJong, Forsythe, and Ross (1989) for the battle of the sexes game; and Hargreaves Heap et al. (2014) for the odd-one-out games. The Kreps game has multiple equilibria (labeled Pure T-L, Mixed, and Pure B-R, respectively) shown as overlapping distributions, whereas the battle of the sexes and stag hunt games have multiple equilibria shown as separate distributions. See the online article for the color version of this figure. (*Figure continues on next page.*)

ple vary in the depth of strategic thought they are capable of or choose to engage in. The variable *k* refers to the number of steps of reasoning that an individual goes through. Level-0 types are nonstrategic and typically assumed to choose their strategy uniformly at random. Level-1 types choose a best response to the mixed strategy played by the Level-0 types. Level-2 best responds to Level-1, and so on. The distribution of types with each level of strategic sophistication can be empirically estimated, and typically Level-1 and Level-2 types are most common.

Camerer, Ho, and Chong (2004) propose a variant of Level-*k* reasoning—the *cognitive hierarchy* model—in which individuals

best respond to the mixture of all types less sophisticated than they are. To describe the mixture of types in the population, they specify a Poisson distribution over the levels of reasoning.

Level-*k* reasoning and cognitive hierarchy theory relax the equilibrium assumption of sophisticated expectations, but retain the assumption of rationality in that individuals (above Level-0) are assumed to be maximizing expected utility, given their unsophisticated strategic beliefs. Level-*k* theory and cognitive hierarchy theory match empirical data very well in many settings (Camerer et al., 2004; Crawford et al., 2013; also see Bosch-Doménech et al., 2002; Costa-Gomes et al., 2001; Costa-
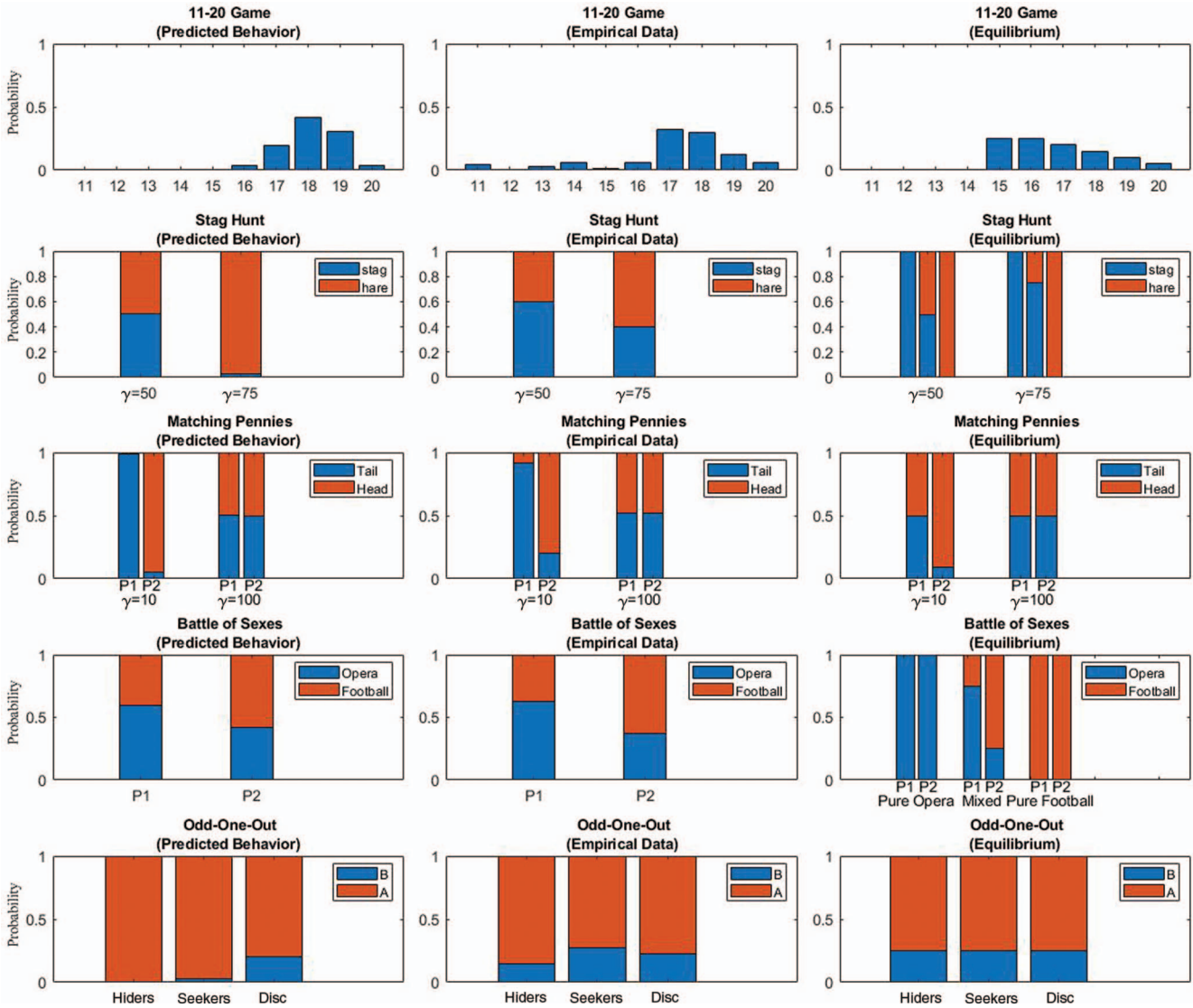
Figure 5. (continued)

Gomes & Crawford, 2006; Costa-Gomes, Crawford, & Iriberri, 2009; Ho et al., 1998). In particular, they predict choices near the upper end of the range in the traveler's dilemma and, similarly, predict modal choices in the p-beauty contest. They also predict behavior well in coordination games (Camerer, 2003a; Haruvy & Stahl, 2007). However, if best responding is not relaxed to allow stochastic responses (as in Stahl & Wilson, 1995 or Rogers, Palfrey, & Camerer, 2009), then Level-k reasoning and cognitive hierarchy theory predict that changes in payoffs that do not alter best responses to lower-level types would not affect behavior, so there are some intuitively reasonable patterns of payoff sensitivity that these theories (specified deterministically) rule out (although, to our knowledge, these predicted patterns of payoff insensitivity have not yet been experimentally tested).[7] Level-k reasoning and cognitive hierarchy theory can account for certain salience effects, for example, choice patterns in coordination and hide-and-seek games,

by assuming that Level-0 types are more likely to choose salient strategies (Crawford et al., 2008; Crawford & Iriberri, 2007). However, with a fixed distribution of types, Level-k reasoning cannot simultaneously account for discoordination game and hide-and-seek choice behavior: Averaging across discoordinators, hiders, and seekers, it predicts that a highly salient strategy is chosen too often relative to its empirical frequency (Hargreaves Heap et al., 2014). Other experiments also suggest that levels of reasoning vary across games (Burchardi & Penczynski, 2014; Camerer et al., 2004; Georganas, Healy, & Weber,

---

[7] For example, in Goeree and Holt's (2001) extended coordination game, we would expect the probability of Player 1 choosing L to increase continuously in *x*, whereas deterministic Level-k reasoning predicts a jump discontinuity at $x = 90$ and insensitivity to *x* elsewhere.

2015; Hyndman et al., 2015; Cooper, Fatas, Morales, & Qi, 2016; Fragiadakis et al., 2016).[8]

Some existing models generate stochastic choice predictions. McKelvey and Palfrey (1995) propose *logit quantal response equilibrium*. Quantal response equilibrium hypothesizes that people occasionally choose suboptimal responses to the opponent's expected play, but the frequency of these mistakes varies inversely with their cost. The logit equilibrium, in particular, specifies a conveniently tractable functional form for the distribution of responses, and a *rationality parameter* η calibrates the magnitude of the cost of a mistake to its tolerated frequency. Given the expectation that the opponent uses the mixed strategy ω_2, the decision maker responds with the mixed strategy ω_1 that calls for strategy i to be played with probability

$$\omega_{1i} = \frac{e^{\eta u_1(s_{1i}; \omega_2)}}{\sum_{k=1}^{N} e^{\eta u_1(s_{1k}; \omega_2)}},$$

and the opponent does the same, with her mixed strategy specified by

$$\omega_{2j} = \frac{e^{\eta u_2(s_{2j}; \omega_1)}}{\sum_{k=1}^{M} e^{\eta u_2(s_{2k}; \omega_1)}}.$$

Logit equilibrium relaxes the assumption of perfect rationality, but retains the assumption of sophisticated expectations in that individuals are assumed to respond to the expected play of their opponents. By allowing for mistakes, the logit equilibrium model also fits empirical data in a variety of settings (Anderson, Goeree, & Holt, 1998; Anderson et al., 2001; Capra et al., 1999; Goeree & Holt, 2005). In particular, as it has properties of monotonicity and responsiveness to payoffs (Goeree et al., 2005), it generates the commonly observed pattern of payoff sensitivity. It also approximates the risk-dominant Nash equilibrium in coordination games (Anderson et al., 2001). However, logit equilibrium only generates patterns of limited iterated reasoning when there are payoff asymmetries (as there are in the traveler's dilemma) and cannot account for modal choices in the p-beauty contest. Moreover, if the rationality parameter is assumed to be stable across games, then the best-fitting logit quantal response equilibrium is fairly noisy and would thus predict a fairly high frequency of selection of strictly dominated choices (Goeree & Holt, 2004).[9] It is also unclear how salience could be incorporated into a logit equilibrium prediction.

Elements of both of these behavioral theories—iterative belief formation and a logit response rule—have been brought together in a model of *noisy introspection* (Goeree & Holt, 1999, 2004).[10] It was conceived as an extension of logit equilibrium more suitable for one-shot games in which players have no opportunity to learn or adjust toward equilibrium. In its most parsimonious specification for a symmetric game, the noisy introspection model predicts a mixed strategy

$$\omega = \lim_{n \to \infty} \varphi_\mu(\varphi_{\tau\mu}(\ldots, \varphi_{\tau^n\mu}(\omega_0))),$$

where φ_μ is the logit response function with rationality parameter $1/\mu$, and ω_0 is an arbitrary mixed strategy (conveniently taken to be uniform because $\lim_{n \to \infty} \varphi_{\tau^n\mu}(\omega_0)$ is approximately uniform regardless of how ω_0 is specified), and μ and τ are free parameters. Noisy introspection captures behavioral patterns involving limited iter-

ated reasoning and payoff sensitivity pretty well (Breitmoser, 2012; Goeree & Holt, 2004), but it does have an extra degree of freedom giving it additional flexibility. Again it is also unclear how salience could be incorporated into noisy introspection.

A couple of existing models share our assumption that players respond to a sample of the opponent's play. Selten and Chmura's (2008) *action-sampling equilibrium* proposes that players take a finite sample of their opponent's play and best respond to it. The players are assumed to be in an equilibrium so that the distribution of sampled play of the opponent is a multinomial distribution with event probabilities equal to the probabilities that each of the opponent's strategies is a best response to a sample of one's own play. Osborne and Rubinstein's (1998) *payoff-sampling equilibrium* is very similar, except that it assumes that players take distinct samples of their opponent's play to estimate the expected payoff for each of their own strategies separately (and then choose the strategy with the highest estimated expected payoff). Again players are assumed to be in an equilibrium so each sample of opponent's play is an independent draw from a multinomial distribution, and the underlying event probabilities are the probabilities that each of the opponent's strategies earns the highest payoff against independent samples of one's own play.

The equilibrium assumption in both of these sampling equilibrium models leads to some bad predictions in one-shot games (although, in fairness, these models were initially intended to describe long-run behavior after a learning and adjustment process had time to operate). For example, they predict that players will never play iteratively dominated strategies, in contrast to the behavioral pattern of limited iterated reasoning. Strict Nash equilibria will always be equilibria of these models as well, so in coordination games these sampling equilibrium models cannot make unique predictions that depend on risk-reward tradeoffs. And as with Nash equilibria, salience may be a factor that determines which of multiple equilibria is a focal point, but there is no role for salience to play in games with a unique mixed equilibrium.

Other behavioral game theoretic models, such as Stackelberg reasoning (Colman & Bacharach, 1997; Colman, Pulford, & Lawrence, 2014), social projection theory (Acevedo & Krueger, 2005; Krueger, 2007; Krueger, DiDonato, & Freestone, 2012), team reasoning (Bacharach, Gold, & Sugden, 2006), and virtual bargaining (Misyak & Chater, 2014; Misyak, Melkonyan, Zeitoun, & Chater, 2014), specify alternative processes of strategic deliberation and choice. According to the theory of Stackelberg reasoning, decision makers in simultaneous-move games often behave as if they believe that they are first-movers in a sequential game. Social projection theory proposes that players expect their coplayers to reason similarly as they do. According

---

[8] But see also Camerer and Smith's (2012) finding of fairly high correlation between estimates of an individual's behavioral type across different sets of games when those sets include a greater number of games.

[9] Heterogeneous logit parameters do improve fit, but do not mitigate overestimation of the frequency of choices of strictly dominated strategies (Golman, 2012; McKelvey et al., 2000; Rogers et al., 2009).

[10] Rogers et al.'s (2009), truncated heterogeneous logit equilibrium also brings together iterative belief formation with a logit response rule and captures behavioral patterns involving limited iterated reasoning and payoff sensitivity. It warrants consideration in future research.

to the theory of team reasoning, decision makers do not attempt to maximize their individual payoffs, but rather choose jointly on the basis of collective preferences, such as the maximization of total payoffs. The theory of virtual bargaining builds on both Stackelberg reasoning and team reasoning, proposing that decision makers choose on the basis of what they would agree to do together while considering the possibility that the other player might always remain one step ahead. These four models of collective reasoning can in some cases account for cooperative behavior without assuming social preferences, but they fail to account for patterns of payoff sensitivity, and typically predict payoff-dominant equilibria rather than risk-dominant equilibria in coordination games.

Fast and frugal heuristics from the domain of multiattribute or risky choice, such as the one-reason heuristic "take-the-best" (Gigerenzer & Goldstein, 1996) or the priority heuristic "avoid-the-worst" (Brandstätter, Gigerenzer, & Hertwig, 2006), could be applied to strategic choice in games. These heuristics could account for limited iterated reasoning, some patterns of payoff sensitivity, and responsiveness to risk-reward tradeoffs in coordination games, but they would fail to account for strategic considerations, for example, they do not predict sensitivity to the other player's payoffs, as observed in the matching pennies game.

Model Fits

In order to assess the dual accumulator model's ability to account for empirical data relative to competing models, we fit the dual accumulator model and several competitor models to an existing experimental data set. The data set, reported by Stahl and Wilson (1995), consists of the strategy choices of 48 subjects who each played a set of 12 3×3 symmetric games once without feedback. We chose this data set because Stahl and Wilson (1995) reported each individual subject's choice in each game (and the data were publicly available) and all the games have payoffs between 0 and 100, so we can avoid dealing with issues that arise when payoff magnitudes vary across games.[11]

We include competitor models that generate many of the observed qualitative patterns of behavior (i.e., some form of limited iterated reasoning, payoff sensitivity, and risk-reward consideration in coordination games): Level-k reasoning, cognitive hierarchy theory, logit quantal response equilibrium, and noisy introspection. We fit all models to each individual subject's choice data, restricting parameters to be fixed across games (but allowing them to vary across subjects). Because standard Level-k reasoning makes deterministic individual predictions (except for the completely random Level-0 types), we also include a version of Level-k reasoning with noise, where the model's probabilistic choice predictions become $(1 - \varepsilon)$ $\mathrm{BR}^k (\omega_0) + \varepsilon \omega_0$, where ω_0 is the uniform mixed strategy, $\mathrm{BR}^k(\omega_0)$ is the Level-k best response, and ε is an individual free parameter. Because cognitive hierarchy theory is designed to make aggregate predictions for pooled data, we distinguish two ways of applying it to individual subject level data. In the *Poisson cognitive hierarchy* specification, we estimate a best-fitting Poisson distribution from the best-fitting levels of reasoning for each individual subject, and higher-level types best

respond to the mixture of lower-level types implied by this Poisson distribution. In the *empirical cognitive hierarchy* specification, the higher-level types best respond to the empirical mixture of lower-level types determined by each individual subject's best-fitting level. (In both cases, a type's best response depends on the distribution of types. Thus, to estimate an individual's type requires an estimate of the distribution of types, but the estimate of the distribution of types depends on the estimate of each individual's type. Because of this circularity, these models must be estimated recursively. Given this complication, we do not include a version of cognitive hierarchy theory with noise.) Additional details regarding model fits are provided in Appendix B. To put the dual accumulator model on equal footing with the other models, we do not make use of its ability to incorporate salience, effectively assuming here that all strategies have equal salience.

We fit the models for each subject by minimizing the mean squared error (*MSE*) between the model's predictions and the data. For each subject the *MSE* is $\frac{1}{g}\sum_{j=1}^{g} \frac{1}{c}\sum_{i=1}^{c} (y_{j,i} - \hat{y}_{j,i})^2$ where $g = 12$ is the number of games, $c = 3$ is the number of choices per game, $y_{j,i}$ is the indicator variable for whether the subject took the ith choice in the jth game, and $\hat{y}_{j,i}$ is the predicted probability of choosing the ith choice in the jth game. The overall *MSE* for each model (averaged across all 48 subjects) are presented in Table 4. The dual accumulator model has the second lowest *MSE*, behind only the noisy introspection model. The model with the third lowest *MSE* is Level-k reasoning with noise. However, these models have two free parameters, whereas (deterministic) Level-k reasoning, cognitive hierarchy theory, and logit quantal response equilibrium each only have one free parameter, so we should be concerned about overfitting. To make a fair comparison between the models, we run a hold-one-out analysis. For each subject, we compute the *MSE*s for each game after fitting the models to that subject's observed choices in the other games. The overall out-of-sample *MSE* for each model (averaged across all 12 games and all 48 subjects) is also presented in Table 4. The dual accumulator model has the lowest out-of-sample *MSE* in the hold-one-out analysis, indicating that it makes more accurate predictions than any of the competitor models on this dataset. The differences in *MSE*s between the dual accumulator model and every other model are statistically significant at $p < .001$, as evaluated by a paired t test applied to the model *MSE*s for each game and each subject. Note that the difference between the *MSE* of fits on the full dataset and the out-of-sample *MSE* suggest that there is some degree of overfitting for all models. We would expect these differences to drop (and our out-of-sample predictive accuracy rates to improve) with larger training data sets.

Finally, we examine the best-fitting individual-level parameters of the dual accumulator model (reported in Table 5). There is heterogeneity in individual parameters, with best-fitting λ in the range [0, 9.8] and the best fitting T in the range [1, 48]. (When $T = 1$, the value of λ makes no difference, as we assume no salience effects in these games, so we leave it unspecified.) We observe median values of $\lambda = 1.8$ and $T = 24$. These

---

[11] We did not fit the models to any other data set or select this particular data set on the basis of the model fits it gives.

Table 4

Full Sample MSE and Out-of-Sample MSE Averaged Across All 12 Games for Each of the Seven Models We Fit to the Stahl and Wilson (1995) Data

Model	Parameters per individual	*MSE* full sample	*MSE* out-of-sample
Dual accumulator	2	.1434	.1775 ($SE = .0067$)
Level-k	1	.1871	.2210 ($SE = .0074$) ($t = 6.04$)
Level-k with noise	2	.1653	.2000 ($SE = .0063$) ($t = 3.76$)
Poisson cognitive hierarchy	1	.1921	.2068 ($SE = .0058$) ($t = 4.42$)
Empirical cognitive hierarchy	1	.1944	.2049 ($SE = .0053$) ($t = 4.62$)
Logit quantal response equilibrium	1	.1971	.2174 ($SE = .0037$) ($t = 6.36$)
Noisy introspection	2	.1104	.2300 ($SE = .0112$) ($t = 5.04$)

Note. *MSE* = mean squared error. The out-of-sample *MSE* column also presents *t*-values from a *t*-test with the dual accumulator model. Here we use a paired *t*-test with each observation corresponding to the *MSE*s for each subject in each game ($df = 575$ in each test, and all tests are significant at $p < .001$).

parameters are larger than we expected on the basis of our simulations in the previous section. They are likely subject to overfitting, given the challenge of fitting our model to just 12 choices for each individual, and we would expect to get more reliable estimates with a larger dataset. Still, even with some overfitting in these particular parameter values, the dual accumulator model retains out-of-sample predictive power in this dataset. The best-fit parameters and *MSE* values for each of the cross-validated fits for each participant and each model are provided in online supplemental materials.

We gain some insight into the success of the dual accumulator model by comparing the best-fitting individual-level parameters of this model to the individual types specified by the Level-k reasoning model, also reported in Table 5. The existing literature on Level-k reasoning suggests that most people are Level-1 or Level-2 types, whereas here we find that most of the individuals are classified as Level-0 types, that is, as completely unpredictable. (Similarly, when we consider Level-k with noise, most of the individuals appear to be very noisy, i.e., with $\varepsilon > 0.5$.) The difference from the existing literature arises because we are classifying individuals based on their behavior across multiple games, whereas the common practice in the existing literature is to estimate population frequencies based on the aggregate behavior in a single game. If Level-1 and Level-2 type behavior are fairly prevalent in the aggregate, but individuals are not consistently employing a single level of reasoning across games, then we get this divergence. In other words, aggregate behavior may often appear as if most people were engaging in one or two levels of reasoning because alternative decision processes may with moderately high probability mimic a mix of Level-1 behavior in some games and Level-2 behavior in other games. In any case, the prevalence of Level-0 estimates at the individual subject level indicates that most individuals are not generally behaving in accordance with any fixed level of reasoning. Of course, most individuals classified as Level-0 types are not actually making completely unpredictable choices. The dual accumulator model fits these individuals much more accurately. A few subjects (a small, but non-negligible fraction of the population) do appear to consistently apply a fixed level of reasoning (mostly, but not exclusively, one level of reasoning). The dual accumulator model can mimic Level-1 reasoning when $\lambda = 0$ and T gets large, so we fit these individuals only slightly less accurately than the Level-k reasoning model.

Discussion

A Cognitive Model of Strategic Decision Making

In this article we formally specify the computational operations involved in sampling and manipulating information during strategic deliberation, and by doing so, propose a new cognitive model of strategic deliberation and decision making (for an overview of cognitive modeling of decision making, see Bartels & Johnson, 2015; Busemeyer & Johnson, 2004; Busemeyer & Rieskamp, 2014; Newell & Bröder, 2008; Oppenheimer & Kelso, 2015). Our model links empirical findings in behavioral game theory with theoretical constructs in psychology and cognitive science, which have previously been used to study memory, perception, and nonstrategic choice.

One benefit of the cognitive modeling approach is that we are able to make predictions about the process of deliberation, including about cognitive variables that determine the use of information in strategic choices. For example, our model, which specifies the formation of preference in terms of the spread of activation based on the sampling of nodes, permits a link between attentional measures and the computational steps used to generate choice. Subsequently, our model can be used to make predictions about how attention is distributed during deliberation, and, moreover, about correlations between attention to various strategies and an individual's choice frequencies. For example, it predicts a positive correlation between attention to one of the opponent strategies and choice of a best response to that strategy (even if attention to that strategy has faded by the end of the decision process). The dual accumulator model also predicts that attention changes as the deliberation progresses, with frequently chosen strategies getting more attention as time goes on. Although there may be discrepancies between external information search and internal sampling, these predictions could be tested with eye-tracking or mouse-tracking data. Indeed, in a recent eye-tracking study, Stewart, Gächter, Noguchi, and Mullett (2016) provide some empirical evidence for both of the above predictions of our model. A number of articles in psychology and economics use process-tracing measures to evaluate models of strategic choice (Brocas, Carrillo, Wang, & Camerer, 2014; Camerer, Johnson, Rymon, & Sen, 1993; Costa-Gomes et al., 2001; Devetag, Di Guida, & Polonio, 2016; Johnson, Camerer, Sen, & Rymon, 2002; Polonio, Di Guida, & Coricelli, 2015; Stewart, Gächter, Noguchi,

Table 5
Best-Fitting Parameter Values for Each Subject in Stahl and Wilson's (1995) Data Set

Subject	Dual accumulator		Level-k	Level-k with noise		Poisson CH	Empirical CH	Logit QRE	Noisy introspection	
	λ	T	k	k	ε	k	k	η	μ	τ
1	1.1	7	0	4	.67	0	0	1.8	.01	1.8
2	0	3	0	1	.50	0	0	4.6	.1	11
3	—	1	0	2	.67	0	0	.1	.03	4
4	9.6	5	0	4	.93	0	0	1.5	.7	1
5	2.6	48	0	5	.62	0	0	4.6	.09	1.5
6	—	1	0	0	1	0	0	.1	10	1
7	1	7	0	4	.73	0	0	2.9	.008	1.9
8	0	10	0	1	.50	0	0	4.8	.2	1
9	0	33	1	1	.12	1	1	4.8	.004	1000
10	0	37	1	1	.12	1	1	9.6	.001	700
11	0	49	1	1	0	1	1	8	.004	1000
12	0	49	1	1	0	1	1	8	.004	1000
13	0	17	1	1	.37	2	2	3.6	.006	51
14	0	49	1	1	0	1	1	8	.004	1000
15	0	16	1	1	.37	1	1	3.5	.08	1000
16	0	7	0	1	.62	2	0	3.6	.001	250
17	0	13	1	1	.37	1	1	5.1	.07	1000
18	0	14	1	1	.25	1	1	5	.02	21
19	.01	45	2	2	0	0	0	.1	.001	20
20	6.8	48	0	5	.75	0	0	3.5	.02	3.7
21	6.8	48	0	5	.75	0	0	3.5	.02	3.7
22	6.1	30	0	2	.53	0	0	3.5	.01	5
23	6.1	30	0	5	.62	0	0	3.5	.02	3.8
24	6.8	48	0	5	.75	0	0	3.5	.02	3.7
25	9.8	7	0	3	.68	0	0	3.6	.02	2.2
26	7	48	0	2	.67	0	0	3.6	.01	5
27	9.7	6	0	2	.80	0	0	3.6	.007	2
28	8.1	31	0	2	.67	0	0	3.6	.02	1.2
29	7.3	10	0	2	.80	0	0	3.6	.02	1.2
30	6	8	0	4	.80	0	0	3.5	.06	4.9
31	2.2	6	0	2	.87	0	0	3.5	.007	2
32	2	8	0	2	.73	0	0	3.5	.001	220
33	—	1	0	2	.67	0	0	3.5	.02	3.8
34	2.4	48	0	2	.53	0	0	3.5	.009	1.9
35	1.1	46	0	6	.51	0	0	4.4	.004	1.7
36	1.8	7	0	4	.80	0	0	3.5	.007	2
37	.9	12	6	6	.15	0	0	.1	.001	2.9
38	.01	42	2	4	.40	0	0	3.5	.001	1.6
39	2.3	46	0	6	.51	0	0	.1	.009	1.9
40	2.3	44	0	4	.60	0	0	3.5	.009	2.9
41	1.2	6	0	2	.60	0	0	3.5	.01	1.8
42	9.5	33	0	2	.80	0	0	3.5	.01	5
43	1.3	5	0	2	.53	0	0	3.5	.006	1.1
44	2.5	37	0	2	.73	0	0	3.5	.001	220
45	.9	7	0	6	.51	0	0	1.8	.009	1.8
46	8.1	45	0	3	.55	0	0	3.5	.02	4
47	1.9	48	0	2	.87	0	0	3.5	.008	23
48	1.4	45	6	6	.44	0	0	.1	.01	1.8

Note. QRE = quantal response equilibrium; CH = cognitive hierarchy.

& Mullett, 2016; Wang, Spezio, & Camerer, 2010). Data from these articles could, in future work, be used for quantitative fits of our model that utilize both choice and attentional measures.[12]

The cognitive modeling framework also allows us to provide a more complete characterization of salience effects, as we can make predictions about how salient labels affect attention during deliberation to go along with predictions about how salient labels ultimately affect choices. Our treatment of salience—assuming

that salience influences strategy sampling, which in turn determines the formation of beliefs and preferences, rather than assuming a direct effect on choice—aligns with the modeling of salience

---

[12] Note that attentional predictions such as these are not unique to our model. Level-k models, for example, make predictions about attention in strategic choice, some of which have already been tested.

effects in computational theories of memory and attention. This alignment points to a promising direction for actually measuring salience on a cardinal scale (and thus removing a degree of researcher freedom in the assignment of salience). We could measure the salience of labels from their attentional effects or memory effects in nonchoice tasks and then input these measurements into our model to predict strategic choice frequencies (see Li & Camerer, 2019).

The dual accumulator model also specifies response time as an important variable in the decision process. Thus, we can make predictions about how choice probabilities correlate with response time or change under the influence of time pressure (i.e., with a smaller upper bound on T). For example, under time pressure we would expect to observe noisier and less strategically sophisticated decisions (as decision makers would have less time to sample and aggregate strategy preferences). In the study of nonstrategic choice, testing such predictions is a common way in which cognitive models are evaluated. Agranov, Caplin, and Tergiman (2015) introduce a strategic choice process protocol for eliciting provisional choices as subjects deliberate, which could be used to perform similar tests in the strategic domain.

There are two free parameters in our dual accumulator model: the decision time parameter, T, and the stochastic sampling parameter, λ. The dual accumulator model generates realistic behavior for moderate values of T and λ, that is, for settings in which there is some finite period of dynamic accumulation and some degree of noise in sampling. Indeed, as strategic deliberation does not always proceed until it converges to equilibrium, finite-horizon dynamic accumulation is the primary mechanism generating limited iterated reasoning. Thus, commonly observed patterns of limited iterated reasoning in games such as the traveler's dilemma would no longer be reconciled with the model if T were too large. Of similar importance, stochastic sampling is a critical mechanism for generating realistic patterns of payoff sensitivity. Because any strategy always has some chance of being sampled and greater accumulated activation increases these chances, increasing a strategy's payoffs makes it more likely that the strategy will be considered and, thus, tends to make it more likely that it will ultimately be chosen. Stochastic sampling also plays a critical role in realistically navigating risk-reward tradeoffs in coordination games, as early (stochastic) samples of the opponent's strategies expose the decision maker to these risk-reward tradeoffs. Thus, with deterministic sampling (extremely large λ) we would be unable to generate realistic behavioral patterns in the minimum-effort coordination game, the stag hunt, the hi-lo game, the battle of the sexes, and the Kreps game. Likewise, if strategy sampling were uniform (extremely small λ), bidirectional connections would be irrelevant and deliberation would be nonstrategic. We would be unable to reconcile realistic strategic behavior in the matching pennies game and the battle of the sexes, to name just a couple of examples. With uniformly random sampling, choices are also uninfluenced by strategy salience, and the model would also be unable to account for any of the observed findings related to salience.

Ultimately, the model's key behavioral properties depend critically on its dynamic and stochastic processes. Many scholars have suggested that behavioral theories of decision making can, with

incorporation of these fundamental cognitive processes, describe a wide range of behavior (see Busemeyer & Townsend, 1993 for an introduction; Busemeyer & Rieskamp, 2014; provide a more recent overview of key issues). Our results reinforce these claims by demonstrating the explanatory power of stochastic sampling and dynamic accumulation in strategic choice.

Relationship With Existing Cognitive Models

Our proposed dual accumulator model resembles an existing accumulator model in the domain of nonstrategic risky choice, decision field theory (Busemeyer & Townsend, 1993; see also Bhatia, 2014 and Rieskamp, 2006). We take from decision field theory the assumption that decision makers use two separate sets of nodes to represent their preferences (i.e., propensities to choose each of their choices options) and their beliefs about the occurrence of uncertain events. The connections between these nodes correspond to the rewards generated by the available choice options if the various events occur. Event probabilities in risky choice are considered to be exogenous, unrelated to one's preferences, so activation flows one way, from belief nodes to preference nodes, in decision field theory. In strategic choice, however, the choices of the opponent are codependent on those of the decision maker, so we are forced to introduce bidirectionality. Our dual accumulator model can thus be seen as the natural extension of decision field theory to strategic decisions, thereby providing further evidence that a fundamental cognitive process—based on sequential, stochastic sampling of information and the dynamic accumulation of evidence—supports decision making across multiple domains.[13]

There is considerable evidence that decision makers are able to represent the preferences and beliefs of others separately from their own (e.g., Frith & Frith, 1999; Gallese & Goldman, 1998; Wimmer & Perner, 1983). Some experimental work shows that people form theory-of-mind representations when they play strategic games. Hedden and Zhang (2002), for example, find that players in sequential move matrix-based games have sophisticated beliefs about the opponent's preferences, and that these beliefs are dynamically modified based on the evidence presented to the decision maker during the decision process. They also find that these beliefs predict the choices of participants, suggesting that decision makers are responding to what they think their opponent will do. Goodie, Doshi, and Young (2012) also find that players' beliefs about their opponent's preferences are fairly complex, and are formed in response to the players' own preferences. Of course, perspective taking may be imperfect. Decision makers are better at predicting their opponent's choices when they are asked to imagine themselves in the role of the opponent relative to when they are asked to play against the opponent (Zhang, Hedden, & Chia, 2012). Still, we could not just model the opponent's choice as an event with uniform probabilities. Some kind of strategic deliberation is necessary. Our approach integrates strategic theory-of-mind representations into an evidence-accumulation model of decision making.

Our approach is also closely related to cognitive models based on neural networks with recurrent connectivity (Bhatia, 2016; Glöckner et al., 2014; Holyoak & Simon, 1999; Simon, Krawczyk,

---

[13] In contrast, many other behavioral game theory models would make poor predictions in the domain of nonstrategic, risky choice.

& Holyoak, 2004). Recurrence in these networks is often bidirectional; the activation of cues and decision attributes may influence and be influenced by beliefs and preferences. Recurrent neural networks have been very successful in capturing coherence effects in information processing and decision making and are able to shed light on a number of related findings regarding constraint satisfaction in high-level cognition (see Bhatia, 2016 for a summary).

The bidirectional feedback in the above models and in ours is very similar. Both kinds of networks involve recurrent connections between two different sets of nodes, with sequential updating (though our model permits asymmetric connections, which is not the case with existing bidirectional neural networks in cognitive decision modeling). This similarity implies that our dual accumulator model could be adapted for other cognitive decision modeling applications, perhaps, for example, to describe changes in attribute and cue importance weights over the course of a decision process. By the same token, a variant of our model that dropped the assumptions of stochastic sampling and finite-horizon dynamic accumulation but still relied on the recurrent bidirectional neural network architecture could describe an alternative hypothesis about strategic deliberation, albeit without accounting for altogether realistic behavioral patterns. Indeed, the dual accumulator model proposed here grew out of an earlier attempt to apply a simplified version of Kosko's (1988) recurrent bidirectional neural network to game theoretic decision making (Bhatia & Golman, 2019). In this work we showed that that a variant of Kosko's (1988) model solves two-player games *rationally* through the process of constraint satisfaction; that is, that this model only selects rationalizable strategies, that it finds a pure-strategy Nash equilibrium if it is able to converge on a single strategy profile, and that any pure-strategy Nash equilibrium is a stable stationary state of the network (see Bhatia & Golman, 2019 for details).

Extensions

We have made various modeling assumptions for parsimony:

1. As in Roe et al. (2001), Usher and McClelland (2004), Bhatia (2013), and related papers, we assume that the decision terminates when a time limit is reached. Other models assume that decisions are made when an activation level reaches a threshold. Using an activation threshold might provide a better account of reaction time (RT) data, but it would complicate application of the model across games with varying payoff levels.

2. We do not assume any additive random error in activation as an additional source of stochasticity, nor do we assume any primacy or recency effects in accumulation, although such elements are commonly assumed in accumulator models (see, e.g., Busemeyer & Townsend, 1993). An error term could improve fits to data and accommodate rare choices of strictly dominated strategies.

3. We assume that the decision maker alternates deterministically between the two layers of the network, sampling one of the opponent's strategies to accumulate activation for his own strategies and then sampling one of his own

strategies to accumulate activation for the opponent's. A more intuitively reasonable specification might allow the decision maker to switch between the two layers stochastically, such that the decision maker is more likely to accumulate activation in his preference nodes than in his belief nodes. This assumption would introduce an additional degree of freedom into the model, perhaps providing too much flexibility.

4. We assume salience is entirely exogenous. An alternative specification might allow payoffs to contribute to salience (see Wright & Leyton-Brown, 2016). In the 11–20 game, for example, it has been suggested that the strategy 20 is particularly salient because it offers the highest sure payoff (as well as because it is the only round number). Additionally, in coordination games with a unique Pareto optimal equilibrium (such as the minimum-effort coordination game, the stag hunt, or the Kreps game), players may find the payoff-dominant equilibrium strategy to be more salient because of its potential high payoffs for all players. Indeed, while risk dominance tends to predict better than payoff dominance, payoff dominance still has some pull, and our model appears to slightly underestimate the probabilities with which players choose payoff dominant equilibrium strategies in coordination games.

Each of these modeling choices might be reconsidered in future work.

We have limited the application of our model to two-player games. There are two ways in which it can be extended to multiplayer games. The first, applicable only in symmetric games, is to assume aggregate mental representations for opponent actions. Decision makers would still have two layers of nodes in such a model: one for their own preferences and one for beliefs regarding the opponents. The latter nodes would, however, represent group-level action propensities (rather than action propensities for an individual opponent). This type of representation and deliberation process would be tractable both for the decision maker and the theorist wishing to model the decision. The second way to incorporate multiplayer strategic reasoning is to assume a separate representation (in the form of a separate layer) for each opponent, as well as the sequential sampling and updating of nodes across different layers. Thus, decision makers would sequentially consider one of their opponent's actions, and update their own preferences as well beliefs about other opponents' action tendencies, in response to the sampled action. This would continue with all the players in the game (including the decision maker) being sampled one after the other. Although such a model is capable of making predictions in multiplayer games, the deliberation process it assumes may be too difficult for most decision makers. Developing, refining, and testing models for multiplayer games is an important topic for future work.

We have also focused exclusively on one-shot, simultaneous-move games with complete information. Sequential-move games, repeated games (with the same players), and games with asymmetric information require players to engage in hypothetical, contingent reasoning, which is notoriously difficult for people (see, e.g., Evans, 2007). Backward induction, the perfectly rational method for analyzing sequential games (when they terminate after

a finite number of moves), is clearly too cognitively demanding in many situations, as evidence from the centipede game (as well as many other games) indicates (McKelvey & Palfrey, 1992). We might speculate that an accumulator model could work well to describe individual moves in a larger game, but a complete model would require some specification of a heuristic process that players could use to estimate the continuation value of intermediate game states (see Hotaling & Busemeyer, 2012 for relevant work on this topic).

In general, extensions of the dual accumulator model to more complex strategic settings necessitate additional assumptions regarding the underlying representation and deliberation processes. In the most general case, the model would possess multiple types of representations (for multiple players, multiple game stages, or multiple exogenous variables relevant to the decision) and would sample and update these representations as it deliberates. By formulating a cognitive model that makes explicit the representations and computations that may be involved in simple two-player games, our article takes a first step toward the creation of cognitive models applicable to more complex types of strategic decision making.

We have also not addressed learning in this article, but we must acknowledge that behavior will change with experience. Many important questions remain open about how people learn in strategic games. How does an individual encode payoffs when they are not explicitly given as part of the choice task? How does an individual react after observing the results of a round of play? Payoff encoding could be accomplished with a simple form of associative learning, according to which a decision maker changes the connection strength between his own chosen strategy and the strategy chosen by the opponent, based on their realized payoffs. With enough experience, an associative learning model would be able to represent the entire game as a bipartite network, allowing for deliberation in the sophisticated manner proposed in this article. Additionally, information about a representative opponent's choice tendencies observed in previous rounds of a particular game could be incorporated into the model using its salience mechanism: Strategies frequently chosen by the opponent would be more salient and would thus be more likely to be considered by the decision maker during the choice process. In a sense, observing choices in previous rounds of play would prime the decision maker to consider these choices and lead to more sampling of them. This approach would not distinguish observation of previous rounds of play from other forms of priming, and the question remains open whether people do respond similarly to these two kinds of experiences. A saliency learning mechanism could be compared against and informed by other learning dynamics, such as instance-based learning, rule learning, reinforcement learning, belief-based learning, and experience-weighted attraction (Camerer & Ho, 1999; Erev & Roth, 1998; Frey & Goldstone, 2013; Fudenberg & Levine, 1998; Gonzalez, Ben-Asher, Martin, & Dutt, 2015; Marchiori & Warglien, 2008; Sgroi & Zizzo, 2009). We hope that future work will explore the interaction between learning and deliberation in games.

Future work should also incorporate social preferences, such as altruism, reciprocity, inequality aversion, and fairness concerns. Considering utility to be equivalent to monetary payoff is clearly wrong. Many different forms of social utility functions have been proposed and would be compatible with the dual accumulator model (see, e.g., Bolton & Ockenfels, 2000; Charness & Rabin, 2002; Fehr & Schmidt, 1999; Levine, 1998; Van Lange, 1999). By embedding social preferences within a dual accumulator process, we could model both how people value payoffs to others and how they deliberate over strategies to select mutually beneficial social outcomes.

We have applied our model to make predictions across games with consistent parameter values across all of these games, but we have considered only a set of games with comparable payoff magnitudes. Because activation strength grows according to payoffs, comparisons across games require consistent units of payoff. In fitting the model to data from games that vary in payoff magnitude, some form of payoff normalization may be necessary (McKelvey, Palfrey, & Weber, 2000; Rangel & Clithero, 2012). The precise normalization procedure that should be used is a topic for future research.

Future work might also fit our model along with related accumulator models to data spanning strategic and nonstrategic choices. We would expect an individual's response times and the noisiness of sampling to be correlated on the individual level across choice tasks (see, e.g., Frydman & Nave, 2016 for a related approach). We believe that describing strategic decision making using the same theoretical constructs used to describe nonstrategic and nonpreferential choice tasks leads to a more cohesive understanding of the cognitive mechanisms at play in choice.

Conclusion

This article has proposed a model of strategic deliberation and decision making, inspired by existing cognitive models of nonstrategic choice. Our proposed model assumes that decision makers represent the strategies available to themselves and the strategies available to their opponents in a network with two separate sets of nodes. They sample strategies from each set of nodes sequentially and update the activation of nodes corresponding to their preferences for their own strategies and to their beliefs about the opponent's preferences. We show that our dual accumulator model is able to account for a number of behavioral patterns, including limited iterated reasoning, payoff sensitivity, risk-reward tradeoffs in coordination games, and effects of salient labeling. It also offers predictions about attention during deliberation, provisional choices during deliberation, and correlations between choices and response times. And it generates a good quantitative fit and more accurate out-of-sample predictions than other popular behavioral game theoretic models. In summary, the dual accumulator model provides a cognitively and behaviorally realistic account of strategic deliberation and decision making and offers coherence in our understanding of strategic and nonstrategic judgment and decision making.

References

Acevedo, M., & Krueger, J. I. (2005). Evidential reasoning in the prisoner's dilemma. *The American Journal of Psychology, 118,* 431–457.

Agranov, M., Caplin, A., & Tergiman, C. (2015). Naive play and the process of choice in guessing games. *Journal of the Economic Science Association, 1,* 146–157. http://dx.doi.org/10.1007/s40881-015-0003-5

Agranov, M., Potamites, E., Schotter, A., & Tergiman, C. (2012). Beliefs and endogenous cognitive levels: An experimental study. *Games and Economic Behavior, 75,* 449–463. http://dx.doi.org/10.1016/j.geb.2012.02.002

Anderson, S. P., Goeree, J. K., & Holt, C. A. (1998). Rent seeking with bounded rationality: An analysis of the all-pay auction. *Journal of Political Economy, 106,* 828–853. http://dx.doi.org/10.1086/250031

Anderson, S. P., Goeree, J. K., & Holt, C. A. (2001). Minimum-effort coordination games: Stochastic potential and logit equilibrium. *Games and Economic Behavior, 34,* 177–199. http://dx.doi.org/10.1006/game .2000.0800

Arad, A., & Rubinstein, A. (2012). The 11–20 money request game: A Level-*k* reasoning study. *The American Economic Review, 102,* 3561–3573. http://dx.doi.org/10.1257/aer.102.7.3561

Bacharach, M., Gold, N., & Sugden, R. (2006). *Beyond individual choice: Teams and frames in game theory.* Princeton, NJ: Princeton University Press. http://dx.doi.org/10.1515/9780691186313

Bar-Hillel, M. (2015). Position effects in choice from simultaneous displays: A conundrum solved. *Perspectives on Psychological Science, 10,* 419–433. http://dx.doi.org/10.1177/1745691615588092

Bartels, D. M., & Johnson, E. J. (2015). Connecting cognition and consumer choice. *Cognition, 135,* 47–51.

Basu, K. (1994). The traveler's dilemma: Paradoxes of rationality in game theory. *The American Economic Review, 84,* 391–395.

Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological Review, 120,* 522–543. http://dx.doi.org/10.1037/a0032457

Bhatia, S. (2014). Sequential sampling and paradoxes of risky choice. *Psychonomic Bulletin & Review, 21,* 1095–1111. http://dx.doi.org/10 .3758/s13423-014-0650-1

Bhatia, S. (2016). The dynamics of bidirectional thought. *Thinking & Reasoning, 22,* 1–46. http://dx.doi.org/10.1080/13546783.2016 .1187205

Bhatia, S. (2017). Choice rules and accumulator networks. *Decision, 4,* 146–170. http://dx.doi.org/10.1037/dec0000038

Bhatia, S., & Golman, R. (2019). Bidirectional constraint satisfaction in rational strategic decision making. *Journal of Mathematical Psychology, 88,* 48–57. http://dx.doi.org/10.1016/j.jmp.2018.11.004

Bhatia, S., & Mullett, T. L. (2016). The dynamics of deferred decision. *Cognitive Psychology, 86,* 112–151. http://dx.doi.org/10.1016/j .cogpsych.2016.02.002

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *The American Economic Review, 90,* 166–193. http://dx.doi.org/10.1257/aer.90.1.166

Bosch-Doménech, A., Montalvo, J. G., Nagel, R., & Satorra, A. (2002). One, two, (three), infinity, . . .: Newspaper and lab beauty-contest experiments. *The American Economic Review, 92,* 1687–1701. http://dx .doi.org/10.1257/000282802762024737

Branas-Garza, P., Espinosa, M. P., & Rey-Biel, P. (2011). Travelers' types. *Journal of Economic Behavior & Organization, 78,* 25–36. http://dx.doi .org/10.1016/j.jebo.2010.12.005

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review, 113,* 409–432. http://dx.doi.org/10.1037/0033-295X.113.2.409

Breitmoser, Y. (2012). Strategic reasoning in p-beauty contests. *Games and Economic Behavior, 75,* 555–569. http://dx.doi.org/10.1016/j.geb .2012.02.010

Brocas, I., Carrillo, J. D., Wang, S. W., & Camerer, C. F. (2014). Imperfect choice or imperfect attention? Understanding strategic thinking in private information games. *The Review of Economic Studies, 81,* 944–970. http://dx.doi.org/10.1093/restud/rdu001

Brown, G. W. (1951). Iterative solutions of games by fictitious play. In T. C. Koopmans (Ed.), *Activity analysis of production and allocation* (pp. 374–376). New York, NY: Wiley.

Burchardi, K. B., & Penczynski, S. P. (2014). Out of your mind: Eliciting individual reasoning in one shot games. *Games and Economic Behavior, 84,* 39–57. http://dx.doi.org/10.1016/j.geb.2013.12.005

Burnham, T. C., Cesarini, D., Johannesson, M., Lichtenstein, P., & Wallace, B. (2009). Higher cognitive ability is associated with lower entries

in a p-beauty contest. *Journal of Economic Behavior & Organization, 72,* 171–175. http://dx.doi.org/10.1016/j.jebo.2009.05.015

Busemeyer, J. R., & Johnson, J. G. (2004). Computational models of decision making. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 133–154). Malden, MA: Blackwell Publishing. http://dx.doi.org/10.1002/9780470752937 .ch7

Busemeyer, J. R., & Rieskamp, J. (2014). Psychological research and theories on preferential choice. *Handbook of choice modeling* (pp. 49–72). Cheltenham, UK: Edward Elgar Publishing.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432–459. http://dx.doi.org/10 .1037/0033-295X.100.3.432

Camerer, C. F. (2003a). *Behavioral game theory: Experiments in strategic interaction.* Princeton, NJ: Princeton University Press.

Camerer, C. F. (2003b). Behavioural studies of strategic thinking in games. *Trends in Cognitive Sciences, 7,* 225–231. http://dx.doi.org/10.1016/ S1364-6613(03)00094-9

Camerer, C. F. (2008). Behavioral game theory and the neural basis of strategic choice. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics, 2nd ed.: Decision making and the brain* (pp. 193–206). Waltham, MA: Academic Press.

Camerer, C., & Ho, T. (1999). Experience-weighted attraction learning in normal form games. *Econometrica, 76,* 827–874. http://dx.doi.org/10 .1111/1468-0262.00054

Camerer, C. F., Ho, T., & Chong, J. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics, 119,* 861–898. http:// dx.doi.org/10.1162/0033553041502225

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

Camerer, C. F., & Smith, A. (2012). Cognitive hierarchies and emotions in behavioral game theory. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 346–363). New York, NY: Oxford University Press.

Capra, C. M., Goeree, J. K., Gomez, R., & Holt, C. A. (1999). Anomalous behavior in a traveler's dilemma? *The American Economic Review, 89,* 678–690. http://dx.doi.org/10.1257/aer.89.3.678

Carpenter, J., Graham, M., & Wolf, J. (2013). Cognitive ability and strategic sophistication. *Games and Economic Behavior, 80,* 115–130. http://dx.doi.org/10.1016/j.geb.2013.02.012

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics, 117,* 817–869. http:// dx.doi.org/10.1162/003355302760193904

Clithero, J. A. (2018). Improving out-of-sample predictions using response times and a model of the decision process. *Journal of Economic Behavior & Organization, 148,* 344–375. http://dx.doi.org/10.1016/j.jebo .2018.02.007

Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences, 26,* 139–153. http://dx.doi.org/10.1017/S0140525X03000050

Colman, A. M., & Bacharach, M. (1997). Payoff dominance and the Stackelberg heuristic. *Theory and Decision, 43,* 1–19. http://dx.doi.org/ 10.1023/A:1004911723951

Colman, A. M., Pulford, B. D., & Lawrence, C. L. (2014). Explaining strategic coordination: Cognitive hierarchy theory, strong Stackelberg reasoning, and team reasoning. *Decision, 1,* 35–58. http://dx.doi.org/10 .1037/dec0000001

Cooper, D. J., Fatas, E., Morales, A. J., & Qi, S. (2016). *Consistent depth of reasoning in Level-k models* (Working paper). Retrieved from https://pdfs .semanticscholar.org/a3b8/c4c08cf41e5a3338be43bf15b06c6309a018.pdf

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1989). Communication in the battle of the sexes game: Some experimental results. *The RAND Journal of Economics, 20,* 568–587.

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1993). Forward induction in the battle-of-the-sexes games. *The American Economic Review, 83,* 1303–1316.

Costa-Gomes, M. A., & Crawford, V. P. (2006). Cognition and behavior in two-person guessing games: An experimental study. *The American Economic Review, 95,* 1737–1768. http://dx.doi.org/10.1257/aer.96.5.1737

Costa-Gomes, M. A., Crawford, V. P., & Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica, 69,* 1193–1235. http://dx.doi.org/10.1111/1468-0262.00239

Costa-Gomes, M. A., Crawford, V. P., & Iriberri, N. (2009). Comparing models of strategic thinking in Van Huyck, Battalio, and Beil's coordinating games. *Journal of the European Economic Association, 7,* 365–376. http://dx.doi.org/10.1162/JEEA.2009.7.2-3.365

Crawford, V. P., Costa-Gomes, M. A., & Iriberri, N. (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature, 51,* 5–62. http://dx.doi.org/10.1257/jel.51.1.5

Crawford, V. P., Gneezy, U., & Rottenstreich, Y. (2008). The power of focal points is limited: Even minute payoff asymmetry may yield large coordination failures. *The American Economic Review, 98,* 1443–1458. http://dx.doi.org/10.1257/aer.98.4.1443

Crawford, V. P., & Iriberri, N. (2007). Fatal attraction: Naïveté, and sophistication in experimental "hid-and-seek" games. *The American Economic Review, 97,* 1731–1750. http://dx.doi.org/10.1257/aer.97.5.1731

Dai, J., & Busemeyer, J. R. (2014). A probabilistic, dynamic, and attribute-wise model of intertemporal choice. *Journal of Experimental Psychology: General, 143,* 1489–1514. http://dx.doi.org/10.1037/a0035976

Devetag, G., Di Guida, S., & Polonio, L. (2016). An eye-tracking study of feature-based choice in one-shot games. *Experimental Economics, 19,* 177–201. http://dx.doi.org/10.1007/s10683-015-9432-5

Diederich, A. (1997). Dynamic stochastic models for decision making under time constraints. *Journal of Mathematical Psychology, 41,* 260–274. http://dx.doi.org/10.1006/jmps.1997.1167

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review, 88,* 848–881.

Evans, J. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement.* New York, NY: Psychology Press. http://dx.doi.org/10.4324/9780203947487

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics, 114,* 817–868. http://dx.doi.org/10.1162/003355399556151

Fragiadakis, D. E., Knoepfle, D. T., & Niederle, M. (2016). *Who is strategic?* Unpublished manuscript, Stanford University, Stanford, CA.

Frey, S., & Goldstone, R. L. (2013). Cyclic game dynamics driven by iterated reasoning. *PLoS ONE, 8,* e56416. http://dx.doi.org/10.1371/journal.pone.0056416

Frith, C. D., & Frith, U. (1999). Interacting minds—A biological basis. *Science, 286,* 1692–1695. http://dx.doi.org/10.1126/science.286.5445.1692

Frydman, C., & Nave, G. (2016). Extrapolative beliefs in perceptual and economic decisions: Evidence of a common mechanism. *Management Science, 63,* 2340–2352. http://dx.doi.org/10.1287/mnsc.2016.2453

Fudenberg, D., & Kreps, D. M. (1993). Learning mixed equilibria. *Games and Economic Behavior, 5,* 320–367. http://dx.doi.org/10.1006/game.1993.1021

Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games* (Vol. 2). Cambridge, MA: MIT press.

Fudenberg, D., Strack, P., & Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *The American Economic Review, 108,* 3651–3684. http://dx.doi.org/10.1257/aer.20150742

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2,* 493–501. http://dx.doi.org/10.1016/S1364-6613(98)01262-5

Georganas, S., Healy, P. J., & Weber, R. A. (2015). On the persistence of strategic sophistication. *Journal of Economic Theory, 159,* 369–400. http://dx.doi.org/10.1016/j.jet.2015.07.012

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103,* 650–669. http://dx.doi.org/10.1037/0033-295X.103.4.650

Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer & P. M. Todd (Eds.), *Evolution and cognition. Simple heuristics that make us smart* (pp. 3–34). New York, NY: Oxford University Press.

Gill, D., & Prowse, V. (2016). Cognitive ability, character skills, and learning to play equilibrium: A Level-*k* analysis. *Journal of Political Economy, 124,* 1619–1676. http://dx.doi.org/10.1086/688849

Glöckner, A., Hilbig, B. E., & Jekel, M. (2014). What is adaptive about adaptive decision making? A parallel constraint satisfaction account. *Cognition, 133,* 641–666. http://dx.doi.org/10.1016/j.cognition.2014.08.017

Goeree, J. K., & Holt, C. A. (1999). Stochastic game theory: For playing games, not just for doing theory. *Proceedings of the National Academy of Sciences of the United States of America, 96,* 10564–10567. http://dx.doi.org/10.1073/pnas.96.19.10564

Goeree, J. K., & Holt, C. A. (2001). Ten little treasures of game theory and ten intuitive contradictions. *The American Economic Review, 91,* 1402–1422. http://dx.doi.org/10.1257/aer.91.5.1402

Goeree, J. K., & Holt, C. A. (2004). A model of noisy introspection. *Games and Economic Behavior, 46,* 365–382. http://dx.doi.org/10.1016/S0899-8256(03)00145-3

Goeree, J. K., & Holt, C. A. (2005). An experimental study of costly coordination. *Games and Economic Behavior, 51,* 349–364. http://dx.doi.org/10.1016/j.geb.2004.08.006

Goeree, J. K., Holt, C. A., & Palfrey, T. R. (2005). Regular quantal response equilibrium. *Experimental Economics, 8,* 347–367. http://dx.doi.org/10.1007/s10683-005-5374-7

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30,* 535–574. http://dx.doi.org/10.1146/annurev.neuro.29.051605.113038

Golman, R. (2011). Why learning doesn't add up: Equilibrium selection with a composition of learning rules. *International Journal of Game Theory, 40,* 719–733. http://dx.doi.org/10.1007/s00182-010-0265-3

Golman, R. (2012). Homogeneity bias in models of discrete choice with bounded rationality. *Journal of Economic Behavior & Organization, 82,* 1–11. http://dx.doi.org/10.1016/j.jebo.2011.12.011

Golman, R., Hagmann, D., & Miller, J. H. (2015). Polya's bees: A model of decentralized decision-making. *Science Advances, 1,* e1500253. http://dx.doi.org/10.1126/sciadv.1500253

Gonzalez, C., Ben-Asher, N., Martin, J. M., & Dutt, V. (2015). A cognitive model of dynamic cooperation with varied interdependency information. *Cognitive Science, 39,* 457–495. http://dx.doi.org/10.1111/cogs.12170

Goodie, A. S., Doshi, P., & Young, D. L. (2012). Levels of theory-of-mind reasoning in competitive games. *Journal of Behavioral Decision Making, 25,* 95–108. http://dx.doi.org/10.1002/bdm.717

Grosskopf, B., & Nagel, R. (2008). The two-person beauty contest. *Games and Economic Behavior, 62,* 93–99. http://dx.doi.org/10.1016/j.geb.2007.03.004

Hargreaves Heap, S., Rojo Arjona, D., & Sugden, R. (2014). How portable is Level-0 behavior? A test of Level-k theory in games with non-neutral frames. *Econometrica, 82,* 1133–1151. http://dx.doi.org/10.3982/ECTA11132

Harsanyi, J. C., & Selten, R. (1988). *A general theory of equilibrium selection in games*. Cambridge, MA: MIT Press.

Hart, S. (1992). Games in extensive and strategic forms. *Handbook of Game Theory with Economic Applications, 1,* 19–40.

Haruvy, E., & Stahl, D. O. (2007). Equilibrium selection and bounded rationality in symmetric normal-form games. *Journal of Economic Behavior & Organization, 62,* 98–119. http://dx.doi.org/10.1016/j.jebo .2005.05.002

Hedden, T., & Zhang, J. (2002). What do you think I think you think?: Strategic reasoning in matrix games. *Cognition, 85,* 1–36. http://dx.doi .org/10.1016/S0010-0277(02)00054-9

Ho, T., Camerer, C., & Weigelt, K. (1998). Iterated dominance and iterated best response in experimental "p-beauty contests." *The American Economic Review, 88,* 947–969.

Hofbauer, J., & Sigmund, K. (2003). Evolutionary game dynamics. *Bulletin of the American Mathematical Society, 40,* 479–519. http://dx.doi .org/10.1090/S0273-0979-03-00988-1

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General, 128,* 3–31. http://dx.doi.org/10.1037/0096-3445.128.1.3

Hotaling, J. M., & Busemeyer, J. R. (2012). DFT-D: A cognitive-dynamical model of dynamic decision making. *Synthese, 189,* 67–80. http://dx.doi.org/10.1007/s11229-012-0157-0

Hyndman, K. B., Terracol, A., & Vaksmann, J. (2015). *Beliefs and (in) stability in normal-form games* (Working paper). Retrieved from https:// econpapers.repec.org/scripts/redir.pf?u=https%3A%2F%2Fmpra.ub .uni-muenchen.de%2F47221%2F1%2FMPRA_paper_47221.pdf; h=repec:pra:mprapa:47221

Johnson, E. J., Camerer, C. F., Sen, S., & Rymon, T. (2002). Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining. *Journal of Economic Theory, 104,* 16–47. http:// dx.doi.org/10.1006/jeth.2001.2850

Kagel, J. H., Harstad, R. M., & Levin, D. (1987). Information impact and allocation rules in auctions with affiliated private values: A laboratory study. *Econometrica, 55,* 1275–1304. http://dx.doi.org/10.2307/1913557

Kocher, M. G., & Sutter, M. (2006). Time is money—Time pressure, incentives, and the quality of decision-making. *Journal of Economic Behavior & Organization, 61,* 375–392. http://dx.doi.org/10.1016/j.jebo .2004.11.013

Kosko, B. (1988). Bidirectional associative memories. *IEEE Transactions on Systems, Man, and Cybernetics, 18,* 49–60. http://dx.doi.org/10 .1109/21.87054

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience, 13,* 1292–1298. http://dx.doi.org/10.1038/nn.2635

Kreps, D. M. (1995). Nash equilibrium. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new Palgrave: Game theory* (pp. 176–177). New York, NY: Norton.

Krueger, J. I. (2007). From social projection to social behavior. *European Review of Social Psychology, 18,* 1–35. http://dx.doi.org/10.1080/ 10463280701284645

Krueger, J. I., DiDonato, T. E., & Freestone, D. (2012). Social projection can solve social dilemmas. *Psychological Inquiry, 23,* 1–27. http://dx .doi.org/10.1080/1047840X.2012.641167

Levine, D. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics, 1,* 593–622. http://dx.doi.org/10.1006/ redy.1998.0023

Li, X., & Camerer, C. F. (2019). *Using visual salience in empirical game theory*. Retrieved from https://ssrn.com/abstract=3308886 or http://dx .doi.org/10.2139/ssrn.3308886

Lindner, F. (2014). Decision time and steps of reasoning in a competitive market entry game. *Economics Letters, 122,* 7–11. http://dx.doi.org/10 .1016/j.econlet.2013.10.019

Lindner, F., & Sutter, M. (2013). Level-k reasoning and time pressure in the 11–20 money request game. *Economics Letters, 120,* 542–545. http://dx.doi.org/10.1016/j.econlet.2013.06.005

Loomes, G. (2015). Variability, noise, and error in decision making under risk. In G. Keren & G. Wu (Eds.), *The Wiley Blackwell handbook of judgment and decision making* (pp. 658–695). Chichester, UK: John Wiley & Sons.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions*. New York, NY: Wiley Sons.

Marchiori, D., & Warglien, M. (2008). Predicting human interactive learning by regret-driven neural networks. *Science, 319,* 1111–1113. http:// dx.doi.org/10.1126/science.1151185

McKelvey, R. D., & Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica, 60,* 803–836. http://dx.doi.org/10.2307/ 2951567

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior, 10,* 6–38. http:// dx.doi.org/10.1006/game.1995.1023

McKelvey, R. D., Palfrey, T. R., & Weber, R. A. (2000). The effects of payoff magnitude and heterogeneity on behavior in 2 × 2 games with unique mixed strategy equilibria. *Journal of Economic Behavior & Organization, 42,* 523–548. http://dx.doi.org/10.1016/S0167- 2681(00)00102-5

Mehta, J., Starmer, C., & Sugden, R. (1994). The nature of salience: An experimental investigation of pure coordination games. *The American Economic Review, 84,* 658–673.

Misyak, J. B., & Chater, N. (2014). Virtual bargaining: A theory of social decision-making. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*. Advance online publication. http://dx.doi.org/10.1098/rstb.2013.0487

Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: Virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences, 18,* 512–519. http://dx.doi.org/10 .1016/j.tics.2014.05.010

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review, 85,* 1313–1326.

Newell, B., & Bröder, A. (2008). Cognitive processes, models and metaphors in decision research. *Judgment and Decision Making, 3,* 195–204.

Noguchi, T., & Stewart, N. (2018). Multialternative decision by sampling: A model of decision making constrained by process data. *Psychological Review, 125,* 512–544. http://dx.doi.org/10.1037/rev0000102

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104,* 266–300. http://dx.doi.org/10.1037/0033-295X.104.2.266

Oppenheimer, D. M., & Kelso, E. (2015). Information processing as a paradigm for decision making. *Annual Review of Psychology, 66,* 277– 294. http://dx.doi.org/10.1146/annurev-psych-010814-015148

Osborne, M. J., & Rubinstein, A. (1998). Games with procedurally rational players. *The American Economic Review, 88,* 834–847.

Polonio, L., Di Guida, S., & Coricelli, G. (2015). Strategic sophistication and attention in games: An eye-tracking study. *Games and Economic Behavior, 94,* 80–96. http://dx.doi.org/10.1016/j.geb.2015.09.003

Rangel, A., & Clithero, J. A. (2012). Value normalization in decision making: Theory and evidence. *Current Opinion in Neurobiology, 22,* 970–981. http://dx.doi.org/10.1016/j.conb.2012.07.011

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108. http://dx.doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111,* 159–182. http:// dx.doi.org/10.1037/0033-295X.111.1.159

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873– 922.

Rieskamp, J. (2006). Perspectives of probabilistic inferences: Reinforcement learning and an adaptive network compared. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1355–1370. http://dx.doi.org/10.1037/0278-7393.32.6.1355

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review, 108,* 370–392. http://dx.doi.org/10.1037/0033-295X.108.2.370

Rogers, B. W., Palfrey, T. R., & Camerer, C. F. (2009). Heterogeneous quantal response equilibrium and cognitive hierarchies. *Journal of Economic Theory, 144,* 1440–1467. http://dx.doi.org/10.1016/j.jet.2008.11.010

Rubinstein, A. (2007). Instinctive and cognitive reasoning: A study of response times. *Economic Journal, 117,* 1243–1259. http://dx.doi.org/10.1111/j.1468-0297.2007.02081.x

Schelling, T. (1960). *The strategy of conflict.* Cambridge, MA: Harvard University Press.

Rubinstein, A., Tversky, A., & Heller, D. (1997). Naïve strategies in competitive games. In W. Albers, W. Guth, P. Hammerstein, B. Moldovanu, & E. van Damme (Eds.), *Understanding strategic interaction: Essays in honor of Reinhard Selten* (pp. 394–402). Berlin, Germany: Springer-Verlag. http://dx.doi.org/10.1007/978-3-642-60495-9_30

Schmidt, D., Shupp, R., Walker, J. M., & Ostrom, E. (2003). Playing safe in coordination games: The roles of risk dominance, payoff dominance, and history of play. *Games and Economic Behavior, 42,* 281–299. http://dx.doi.org/10.1016/S0899-8256(02)00552-3

Selten, R., & Chmura, T. (2008). Stationary concepts for experimental 2x2-games. *The American Economic Review, 98,* 938–966. http://dx.doi.org/10.1257/aer.98.3.938

Sgroi, D., & Zizzo, D. J. (2009). Learning to play 3 × 3 games: Neural networks as bounded-rational players. *Journal of Economic Behavior & Organization, 69,* 27–38. http://dx.doi.org/10.1016/j.jebo.2008.09.008

Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of preferences by constraint satisfaction. *Psychological Science, 15,* 331–336. http://dx.doi.org/10.1111/j.0956-7976.2004.00678.x

Stahl, D. O., & Haruvy, E. (2008). Level-n bounded rationality and dominated strategies in normal-form games. *Journal of Economic Behavior & Organization, 66,* 226–232. http://dx.doi.org/10.1016/j.jebo.2006.04.002

Stahl, D. O., II, & Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization, 25,* 309–327. http://dx.doi.org/10.1016/0167-2681(94)90103-1

Stahl, D. O., & Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior, 10,* 218–254. http://dx.doi.org/10.1006/game.1995.1031

Stewart, N., Gächter, S., Noguchi, T., & Mullett, T. L. (2016). Eye movements in strategic choice. *Journal of Behavioral Decision Making, 29,* 137–156. http://dx.doi.org/10.1002/bdm.1901

Straub, P. (1995). Risk dominance and coordination failures in static games. *The Quarterly Review of Economics and Finance, 35,* 339–363. http://dx.doi.org/10.1016/1062-9769(95)90048-9

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review, 121,* 179–205. http://dx.doi.org/10.1037/a0036137

Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. *Proceedings of the National Academy of Sciences of the United States of America, 109,* 9659–9664. http://dx.doi.org/10.1073/pnas.1119569109

Turner, B. M., Schley, D. R., Muller, C., & Tsetsos, K. (2018). Competing theories of multialternative, multiattribute preferential choice. *Psychological Review, 125,* 329–362. http://dx.doi.org/10.1037/rev0000089

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108,* 550–592. http://dx.doi.org/10.1037/0033-295X.108.3.550

Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychological Review, 111,* 757–769. http://dx.doi.org/10.1037/0033-295X.111.3.757

Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review, 80,* 234–248.

Van Lange, P. A. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology, 77,* 337–349. http://dx.doi.org/10.1037/0022-3514.77.2.337

von Neumann, J. V., & Morgenstern, O. (1944). *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

Wang, J. T., Spezio, M., & Camerer, C. F. (2010). Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth-telling and deception in games. *The American Economic Review, 100,* 984–1007. http://dx.doi.org/10.1257/aer.100.3.984

Webb, R. (2018). The (neural) dynamics of stochastic choice. *Management Science, 65,* 230–255. http://dx.doi.org/10.1287/mnsc.2017.2931

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13,* 103–128. http://dx.doi.org/10.1016/0010-0277(83)90004-5

Wright, J. R., & Leyton-Brown, K. (2016). *Models of Level-0 behavior for predicting human behavior in games.* Advance online publication. Retrieved from https://arxiv.org/abs/1609.08923v1

Zhang, J., Hedden, T., & Chia, A. (2012). Perspective-taking and depth of theory-of-mind reasoning in sequential-move games. *Cognitive Science, 36,* 560–573.

(Appendices follow)

Appendix A

Model Simulation

Here we present pseudocode for simulating the dual accumulator model to generate predictions for the games studied in this article:

1. Select model parameters λ and T

2. Set all initial node activations: $A_{1i}(0) = 0$ and $A_{2j}(0) = 0$

3. Sample one of the opponent's strategies with probability
$$p_j = \frac{e^{\lambda(A_{2j}(0) + \sigma_{2j})}}{\sum_{k=1}^{M} e^{\lambda(A_{2k}(0) + \sigma_{2k})}} = \frac{e^{\lambda\sigma_{2j}}}{\sum_{k=1}^{M} e^{\lambda\sigma_{2k}}}$$

4. Update activations in preference node so that $A_{1i}(1) = A_{1i}(0) + u_1(s_{1i}; s_{2j}) = u_1(s_{1i}; s_{2j})$, where j is the strategy sampled in Step 3.

5. Sample one of the decision maker's own strategies with probability $q_i = \frac{e^{\lambda(A_{1i}(1) + \sigma_{1i})}}{\sum_{k=1}^{N} e^{\lambda(A_{1i}(1) + \sigma_{1k})}} = \frac{e^{\lambda(u_1(s_{1i}; s_{2j}) + \sigma_{1i})}}{\sum_{k=1}^{N} e^{\lambda(u_1(s_{1k}; s_{2j}) + \sigma_{1k})}}$, where j is the strategy sampled in Step 3.

6. Update activations in belief node so that $A_{2j}(1) = A_{2j}(0) + u_2(s_{1i}; s_{2j})$, where i is the strategy sampled in Step 5.

7. Repeat Steps 3–6 until $t = T$, with strategy sampling probabilities $p_j = \frac{e^{\lambda(A_{2j}(t-1) + \sigma_{2j})}}{\sum_{k=1}^{M} e^{\lambda(A_{2k}(t-1) + \sigma_{2k})}}$ and $q_i = \frac{e^{\lambda(A_{1i}(t) + \sigma_{1i})}}{\sum_{k=1}^{N} e^{\lambda(A_{1k}(t) + \sigma_{1k})}}$ and node activation updating rules $A_{1i}(t) = A_{1i}(t-1) + u_1(s_{1i}; s_{2j})$ and $A_{2j}(t) = A_{2j}(t-1) + u_2(s_{2j}; s_{1i})$.

8. Choose decision maker's strategy node i with the highest activation (if multiple nodes have equal activation choose randomly between them).

9. Repeat Steps 2–8 1,000 times to get the average choice probabilities for each of the decision maker's strategies, for the selected parameter values.

Appendix B

Model Fitting Details

For each model except for the two specifications of cognitive hierarchy theory, we exhaustively searched over a uniform grid of parameter values to minimize the *MSE* from each subject's observed choices over all games. For the cognitive hierarchy models, we iteratively estimated the individual levels of reasoning and the distribution of levels of reasoning in the population until the estimate of the individual levels converged. That is, we estimated each subject's level of reasoning given the previously estimated distribution of levels in the population (by minimizing *MSE* from each subject's observed choices over all games) and then estimated the distribution of levels of reasoning in the population given the previous estimate of each individual's level, and repeated. All our fits were implemented in MATLAB. We have made our code and the parameter estimates from our model fits available in the online supplemental materials. We provide a summary of our approach below.

Minimizing *MSE*

1. For each value of the parameters under consideration, generate the predicted choice probabilities for each game. These choice probabilities can be represented with a 3×12 matrix $\hat{y}$ (three choices in each of 12 games in the Stahl-Wilson data set).

2. For each individual, find the parameters that produce the choice probabilities with the lowest *MSE* across all games. The *MSE* for each subject is $\frac{1}{g}\sum_{j=1}^{g} \frac{1}{c}\sum_{i=1}^{c} (y_{i,j} - \hat{y}_{i,j})^2$ where $g = 12$ is the number of games, $c = 3$ is the number of choices per game, $y_{i,j}$ is the an indicator variable for whether the subject took the ith choice in the jth game, and $\hat{y}_{i,j}$ is the predicted probability of choosing the ith choice in the jth game).

Predicted Choice Probabilities for Each Model

Dual accumulator model. For every pair of parameter values $\lambda \in \{0, .01, .02, \ldots, 0.1, 0.2, \ldots, 10\}$ and $T \in \{1, 2, \ldots, 50\}$ we computed the predicted choice distribution in each of our games by running 10,000 simulations.[14] See the pseudocode in Appendix A.

Level-*k* reasoning. We estimated each individual subject's level by first identifying the choice that a given level $k > 0$ would make in each of our games (given that Level-0 types choose uniformly), and then assigning each subject the level that minimized the *MSE* from that subject's observed choices. We consider only $k < 7$. The pseudo code for calculating the choice probabilities for the levels for a particular game is as follows:

---

[14] When $T = 1$, the value of λ makes no difference (as we assume no salience effects in these games), so we fix a single value of λ while $T = 1$.

(Appendices continue)

1. Set the choice probabilities for a Level-0 type to be $\frac{1}{c}$ for each strategy, where $c = 3$ is the number of strategies.

2. Calculate the choice probabilities for Level-1 by finding the best response(s) to a Level-0 player.

3. Calculate the choice probability for Level-2 by finding the best response(s) to a Level-1 player.

4. Continue up through Level 6.

Level-k reasoning with noise. For each level k of reasoning and noise component $\varepsilon \in \{0, .01, \ldots, 1\}$, we calculated the predicted choice distribution in each of our games by computing a weighted average of the choice probabilities generated by a particular level of reasoning (see above) and the uniform distribution using weights $1 - \varepsilon$ and ε.

Logit quantal response equilibrium. We found a logit equilibrium for each $\eta \in \{0.1, 0.2, \ldots, 10\}$ by repeatedly iterating logit responses to a uniform distribution until the mixed strategy distribution converged. (In games with multiple logit equilibria, we thus selected this particular equilibrium as a unique prediction and dismissed any other logit equilibria.) Our convergence criterion is when the average absolute value of the difference in the probabilities from one iteration to the next is lower than a threshold, which we set at 0.01.

Noisy introspection. For each pair of parameter values $\mu \in \{.001, .002, \ldots, .01, .02, \ldots, 0.1, 0.2, \ldots, 10\}$ and $\tau \in \{1, 1.1, \ldots, 10, 11, \ldots, 100, 110, \ldots, 1000\}$, we found the noisy introspection prediction for each game in the data set. The pseudocode for calculating these predictions is as follows:

1. Calculate the logit response with parameter $1/\mu$ to a uniformly mixed strategy.

2. Repeatedly calculate the next telescopic iterated logit response, where the parameter each inner logit response is divided by an additional factor of τ.

3. Terminate Step 2 when the average absolute value of the difference in the probabilities from one iteration to the next is lower than a threshold, which we set at 0.01.

Cognitive hierarchy model. Unlike the other models where parameters could be assessed for each individual independently, the cognitive hierarchy model has the feature that each individual is best responding to a perceived distribution of other individuals. For example, a Level-2 player could behave just like a Level-1 player if the proportion of Level-0 players in the population is large relative to the proportion of Level-1 players, but would instead best respond to Level-1 players if the proportion of Level-1 players in the population is very large relative to the proportion of Level-0 players. Thus, estimating the type of an individual de-

pends on knowing the distribution they are responding to. But estimating this distribution requires knowing the types of each individual. We used an iterative fitting algorithm (expectation maximization, implemented in MATLAB), where we alternated making estimates of the individual levels and the distribution of the levels until they converge. We used two versions of this model, each embodying different assumptions about the distribution of levels. For the Poisson cognitive hierarchy specification, we maintained Camerer et al.'s (2004) assumption that the levels are distributed according to a Poisson distribution. For the empirical cognitive hierarchy specification, we simply used the empirical distribution of the levels. For both specifications we considered only up to Level-6. The pseudo code for estimating the levels of each subject for the Poisson model is as follows:

1. Make an initial guess about the distribution of levels. For a Poisson distribution this means guessing the mean. Our initial guess was 1.

2. For each game:

   a. For a Level-0, set the probability of choosing each strategy to $\frac{1}{c}$ where $c = 3$ is the number of strategies.

   b. Calculate the probabilities for Level-1 by finding the best response to a Level-0 player. That is find the strategy with the highest payoff against a Level-0 player.

   c. For each level higher than 1, calculate the choice probabilities for that level by finding the best response to the mixture of all lower levels, with mixing weights proportional to the probability of that level in the assumed Poisson distribution.

3. Estimate the level of each subject by finding the level with predicted choice probabilities that minimize the *MSE* of that subject's observed choices.

4. Estimate a new distribution of levels. for the Poisson specification, this amounts to finding the mean estimated level.

5. Repeats Steps 2–4 with the newly estimated Poisson distribution. Continue until the estimated mean of the distribution changes by less than some threshold from one iteration to the next (our threshold was 0.01). It is possible that this threshold of convergence will never be reached, because the algorithm could cycle between two estimates of tau, where the mean of the estimated level for the first value of tau is the second value of tau and vice versa. We did not observe any such nonconvergence in fitting this data set.

(*Appendices continue*)

The pseudo code for the empirical distribution model is as follows:

1. Make an initial guess about the distribution of levels. For the empirical distribution this means guessing the proportion of subjects at each level. Our initial guess was a uniform distribution over Levels 0–6.

2. For each game:

   a. For a Level-0, set the probability of choosing each strategy to $\frac{1}{c}$ where $c = 3$ is the number of strategies.

   b. Calculate the probabilities for Level-1 by finding the best response to a Level-0 player. That is find the strategy with the highest payoff against a Level-0 player.

   c. For each level higher than 1, calculate the choice probabilities for that level by finding the best response to the mixture of all lower levels, with mixing weights proportional to the probability of that level in the assumed distribution.

3. Estimate the level of each subject by finding the level with predicted choice probabilities that minimize the *MSE* of that subject's observed choices.

4. Estimate a new distribution of levels. For the empirical specification, this amounts to finding the proportion of subjects estimated to be each level.

5. Repeats Steps 2–4 with the newly estimated empirical distribution. Continue until the mean absolute difference in proportions between iterations is less than some threshold (our threshold was 0.01). As in the Poisson version convergence to this threshold is not guaranteed and cycling between empirical distributions is possible, but we did not observe such nonconvergence in fitting this data set.

Hold-One-Out Analysis

For our hold-one-out analysis, we fit the models to each subject's observed choices in 11 of the games, and then calculate the *MSE*s of the models' predictions for the held-out 12th game. We repeat this process, holding out each game once. We then average the out-of-sample *MSE*s across all games and all subjects.