Beyond fictitious play beliefs:
Incorporating pattern recognition and similarity matching

Leonidas Spiliopoulos

University of New South Wales

Abstract

Belief models capable of detecting 2- to 5-period patterns in repeated games by matching the current historical context to similar realizations of past play are presented. The models are implemented in a cognitive framework, ACT-R, and vary in how they implement similarity-based categorization—using either an exemplar or a prototype approach. Empirical estimation is performed on the elicited-belief data from two experiments (Nyarko and Schotter, 2002; Rutström and Wilcox, 2009) using repeated games with a unique, albeit significantly different, stage mixed-strategy Nash equilibrium. Model comparisons are performed by cross-validation both within and between these two datasets, and using data from completely unrelated non-strategic tasks. Subjects' beliefs are best described by 2-period pattern detection. Parameter estimates exhibited considerable instability across the two belief-elicitation datasets, and surprisingly, using median values from a wide variety of unrelated studies led to better predictions.

1. Introduction

A significant portion of the experimental game theory literature is devoted to games with a unique mixed-strategy Nash equilibrium (MSNE), which prescribes the empirical marginal distribution over actions to be played at equilibrium. Furthermore, in repeated games play must be independently distributed over rounds; players have no vested interest in searching for historical patterns if an opponent is randomizing effectively. Empirical investigations using both non-strategic and strategic tasks cast significant doubt on the validity of this assumption: people are generally not effective at randomizing[1] and actively search for, and recognize, patterns in their environment.[2] The two most relevant papers incorporating pattern recognition directly into beliefs are Spiliopoulos (2012) and Rutström and Wilcox (2009). I will advance this literature by proposing variants of pattern detection models based on cognitive processes, selecting amongst these variants using novel model evaluation methods, and investigating parameter stability and predictive performance across experimental tasks.

Research on pattern recognition in strategic games is surprisingly scarce. Theoretical convergence results are difficult to attain and require relatively strict assumptions about the belief updating process (Aoyagi, 1996; Sonsino, 1997; Fudenberg and Levine, 1998). Only a handful of empirical papers have examined pattern detection in games. In the context of a repeated Battle-of-the-Sexes game, Sonsino and Sirota (2003) empirically investigate the resulting patterns of alternating equilibrium outcomes. West and Lebiere (2001) examine agents modeled as simple neural networks that condition their action choices on the recent history of play. Lebiere and West (1999) and Lebiere et al. (2000) examine the behavior of agents modeled on the ACT-R framework (Anderson, 2007) and find qualitative similarities with summary statistics from experimental data such as aggregate payoff performance and empirical marginal probabilities of action choices.

[1] Studies in non-strategic tasks find a deficit in the ability of humans to randomize (Bar-Hillel and Wagenaar, 1991; Rapoport and Budescu, 1997). In strategic games, results are mixed but the weight of evidence points to a predominance of inefficient randomization when subjects are not experts in the task that they face.

Professionals in their field of specialization/experience are efficient randomizers in soccer penalty kicks (Palacios-Huerta, 2003; Chiappori et al., 2002) but not in tennis serves (Walker and Wooders, 2001)—although even in the latter the empirical marginal probabilities are in line with the MSNE prescription. Professionals' ability to randomize efficiently in their domain of specialization is found not to transfer to other settings in the lab (Chiappori et al., 2002)—professionals did not randomize well even when playing against computerized Nash opponents. Although Palacios-Huerta and Volij (2008) find that professionals in the lab randomize more efficiently than non-professional subjects and generally conform more closely to the MSNE, Wooders (2010) finds that other implications of the MSNE are not consistent with their conclusion. Further research seems needed to clarify the conditions under which professionals transfer specialized knowledge to other domains.

The evidence regarding inexperienced subjects is more homogeneous in finding inefficient randomization, starting from the reexamination of O'Neill (1987) in Brown and Rosenthal (1990), and the aforementioned studies pitting professional versus inexperienced subjects (Palacios-Huerta and Volij, 2008; Levitt et al., 2010), although Wooders (2010) rejects the minimax hypothesis only for one of four actions in inexperienced subjects. Notably, Shachat (2002) does not find evidence of subjects playing the MSNE even when an external randomization device is made available. Finally, Scroggin (2007) finds significant evidence of exploitable patterns that can be explained by a local representativeness bias in two classic experiments (Mookherjee and Sopher, 1994; Rapoport and Budescu, 1992).

[2] In non-strategic tasks, the sequence-learning literature concludes that humans are capable of detecting patterns of up to 5 temporal lags (Clegg et al., 1998; Gomez, 1997; Nissen and Bullemer, 1987; Remillard and Clark, 2001). Furthermore, searching for patterns appears to be a widespread human trait, as subjects search and believe they have found patterns even in random sequences (Wolford et al., 2004; Gaissmaier and Schooler, 2008).

Spiliopoulos (2013) presents evidence of 2-period pattern detection influencing marginal and conditional distributions over action choices, thereby explaining the strategic adaptation of human subjects to different computer opponents in a repeated 2×2 game with a unique MSNE.

The following two studies specifically examine pattern recognition in belief formation. Spiliopoulos (2012) investigates a repeated constant-sum 2×2 game with fixed partner matching using belief elicitation, i.e. subjects reported the probability they assigned to an opponent's action. Subjects were modeled as being one of two types. The first (non-pattern detecting) type formed beliefs according to standard weighted fictitious play (Cheung and Friedman, 1997) and the second type according to a 2-period pattern detecting variant of weighted fictitious play. A large proportion of subjects were identified as the second type, and the probability of a subject belonging to the second type increased with the degree of exploitability–in terms of deviation from independent behavior across rounds–that an opponent exhibited. Importantly, evidence of pattern recognition was found both in the elicited beliefs and the realized action choices of subjects. Rutström and Wilcox (2009) examine the behavioral effects of two belief-elicitation procedures in a repeated 2×2 game with fixed matching. Although not specifically investigating pattern recognition, they modeled belief formation as a convex combination of standard weighted fictitious play and a 2-period weighted fictitious play model that conditioned on both players' history of play—they estimated the weight attributed to the pattern detecting model to lie between 0.214 and 0.352, depending on the belief elicitation treatment.

This paper advances the literature in the following directions. Firstly, instead of arbitrarily assuming an upper bound of 2-period pattern recognition, the depth employed by subjects will be empirically determined from a permissible range of 2- to 5-period patterns.[3]

Secondly, the process for matching historical patterns with the current history will allow for two different types of similarity-based matching (exemplar and prototype)—similar, but not necessarily identical, histories will also influence the belief formation process. The exemplar approach to categorization, pioneered by Medin and Schaffer (1978), Nosofsky (1984, 1986) and extended by Nosofsky and Palmeri (1997), performs similarity comparisons individually with respect to each observation. In contrast, the prototype approach advocated by Posner and Keele (1968) and Reed (1972) computes similarity comparisons with prototypes of each category that are created by the averaging of observations. I directly compare the exemplar and prototype models by embedding them in the ACT-R (Anderson, 1996; Anderson and Lebiere, 1998; Anderson, 2007) cognitive framework that has performed successfully in a wide range of domains of human cognition.[4] Close attention will be paid to realistic cognitive constraints and how they relate to the belief formation process.

[3] This is important, as not accounting for pattern recognition can lead to biased parameter estimates due to model misspecification (Spiliopoulos, 2012).

[4] Examples include problem solving (Gunzelmann and Anderson, 2001), decision making (Gonzalez et al., 2003; Gonzalez and Dutt, 2011; Marewski and Melhorn, 2011), visual search and attention (Anderson et al., 1997), cognitive arithmetic (Lebiere, 1999), spatial reasoning (Boeddinghaus et al., 2006) and language processing (Budiu, 2004)—the reader is referred to http://act-r.psy.cmu.edu/publications/index.php for an extensive list of ACT-R related papers.

Specifically, I will examine the implications of working memory constraints of 4 ± 1 items (Cowan, 2001),[5] as well as the constraints on the long-term declarative-memory module with respect to forgetting, noisy retrieval, and context effects.

Finally, model selection is performed by examining the generalizability of models across two different datasets/experiments, the elicited-belief experiments of Nyarko and Schotter (2002), or NS for short, and Rutström and Wilcox (2009), or RW. Previous studies have advocated the use of three distinct variants of cross-validation (CV) performance that I taxonomize below. The *concurrent* implementation of all three CV techniques is, to the best of my knowledge, novel and will provide insights into the stability of parameter estimates[6] and the robustness of the predictive performance of learning models.

Within-dataset-CV (wCV) For a given dataset, model comparisons will be made on the basis of out-of-sample data from the same study to overcome problems of over-fitting. This type of CV corresponds to the procedures advocated by Gigerenzer and Todd (1999), but is not as widespread in economics as in other disciplines—notable exceptions include Camerer and Ho (1999), Stahl and Wilson (1995), and Wilcox (2011).

Between-dataset-CV (bCV) The same models will be estimated on data from one study (NS or RW), and their predictive cross-validation performance will be evaluated on the other dataset. The datasets differ in their game payoffs, subjects and general characteristics, but are generated from a very similar task, i.e. a repeated game with a unique-stage MSNE. This type of CV examination (commonly referred to as the generalization criterion) has been strongly advocated by Busemeyer and Wang (2000), Yechiam and Busemeyer (2008), and Ahn et al. (2008).

Generalized-between-dataset-CV (gCV) Models will not be estimated using the two datasets of interest, but instead parameter values will be fixed at values estimated from other unrelated experimental tasks. Cross-validation performance can then be calculated for both NS and RW datasets. This type of CV is essentially a stricter case of bCV, as the data/experimental space over which a model is expected to perform well is significantly larger and more diverse/less similar.

The key findings of this paper are: a) subjects recognize 2-period patterns by conditioning on the immediately prior lag, b) parameter estimates of models show some instability across the within- and between-dataset procedures, and c) models calibrated from unrelated tasks (gCV) performed better than those estimated from each of the two related tasks (bCV). The implications of these results are important for the learning literature and are discussed in the text.

[5]The classic paper on working memory span by Miller (1956) concludes that the constraint is 7 ± 2 items. However, Cowan (2001) surveys the recent literature and argues strongly in favor of revising this number to 4 ± 1 items.

[6]The stability of parameter estimates is a useful additional criterion in model selection—for example, see recent studies regarding the stability of parameters in Prospect Theory models (Zeisberger et al., 2010; Glöckner and Pachur, 2012).

Table 1: Comparison of the two datasets

	Nyarko and Schotter (2002)				Rutström and Wilcox (2009)			
			Column				Column	
			Green	Red			Left	Right
	Row	Green	6, 2	3, 5	Row	Up	19, 0	0, 1
		Red	3, 5	5, 3		Down	0, 1	1, 0
# subjects	28				92			
# rounds	60				36			
MSNE marginal pr.	$red_r = 0.6, red_c = 0.6$				$up = 0.5, left = 0.05$			
Empirical marginal pr. (last 10 rounds)	$\widehat{red_r} = 0.5, \widehat{red_c} = 0.56$				$\widehat{up} = 0.64, \widehat{left} = 0.3$			
Elicited beliefs, b (last 10 rounds)	$b\left(red_r\right) = 0.56, b\left(red_c\right) = 0.6$				$b\left(up\right) = 0.65, b\left(left\right) = 0.33$			
MSNE π per round (exper. units)	Row	4.2	Col.	3.8	Row	0.95	Col.	0.5
MSNE π per round (dollars)		\$0.21		\$0.19		\$0.19		\$0.10

2. Experimental datasets

The two datasets examined[7] are Experiment 1 in Nyarko and Schotter (2002) and the scoring rule (SR) treatment in Rutström and Wilcox (2009)—Table 1 compares their main characteristics. Both experiments presented subjects with the complete payoff matrix and full feedback regarding opponents' choices and realized payoffs, allowing them to calculate foregone payoffs. Convergence of action choices to the MSNE is significantly weaker in RW than in NS due to the large distance of the MSNE from equiprobable play, which is exacerbated by the limited number of rounds to learn. Elicited beliefs are well-calibrated to the empirical marginal probabilities of action choices for the last ten rounds of each experiment.

The RW experiment is less conducive to the emergence of pattern recognition for the following reasons:

1. The stark asymmetry in payoffs induces MSNE marginal probabilities for the Column player that are extreme, therefore Row's learning should, at least initially, be focused on the marginal probabilities of Column's play, rather than on Column's conditional probabilities, i.e. n-period patterns.

[7] These datasets were chosen on the following grounds:

1. The experimental procedure includes a belief-elicitation stage before subjects indicate their action choices. This allows direct model estimation on the stated-belief series, instead of latent estimation of the belief model from action choice data. The latter has two undesirable properties. Firstly, it requires the researcher to postulate a decision rule to map beliefs to a distribution over action choices; therefore, tests of any hypotheses regarding the belief model will essentially be a joint test of these hypotheses together with the assumptions regarding the decision rule. This is the Duhem-Quine problem, see Smith (2002) for a discussion pertinent to experimental economics. Direct estimation on stated beliefs minimizes the number of other assumptions imposed that are not central to the questions being asked. Secondly, the latent estimation of beliefs leads to significantly less power in hypothesis testing.
2. No evidence that subjects participated in hedging of payoffs between the belief-elicitation and action choice tasks.
3. A large number of rounds with fixed matching between subjects to ensure that a sufficiently large history of play exists to permit pattern detection.
4. The datasets should have very different characteristics to ensure that model comparisons are challenging, e.g. different degrees of payoff asymmetry, MSNE probabilities, number of rounds, and incentives.

2. The MSNE payoffs (π) per round in dollars (subjects' real financial incentives) are roughly equivalent with the exception of the Column player in RW, who has a weaker incentive to randomize efficiently. In praxis, the empirical marginal probabilities for the Column player are far from the MSNE prescription, therefore the Row player should be able to exploit this effectively even with a non-pattern detecting algorithm.
3. Finally, RW has significantly fewer rounds than NS, 36 compared to 60, making inference about patterns more difficult.

3. Models

Two different approaches to pattern recognition using exemplar and prototype models are presented in Sections 3.1 and 3.2. In the categorization literature, the exemplar and prototype models (Nosofsky and Zaki, 2002) were not originally implemented in the ACT-R framework. Anderson and Betz (2001) subsequently implemented an exemplar approach in ACT-R; also, Instance-based Learning Theory (Gonzalez et al., 2003; Gonzalez and Dutt, 2011) can be viewed as an exemplar-based approach to modeling non-strategic decisions.[8] To the best of my knowledge the prototype model has not been implemented in ACT-R. This paper will embed both exemplar and prototype models in a single cognitive framework (ACT-R) presented in Section 3.4, as this is conducive to the model comparison process[9]—see Marewski and Olsson (2009) for an excellent exposition regarding model comparisons.

The declarative memory module in ACT-R is a symbolic system represented by chunks, each comprised of a number of slots capable of storing a single symbol or piece of information. Each memory chunk encodes the relevant information, in this case an opponent's action choice, and the historical context in which it was observed, which is a function of the history of play. At any time t, the memory module will hold the set of chunks created in all the previous rounds. Associated with each chunk is an activation level, which is determined by the frequency and timing of a chunk's use. The difference between the exemplar and prototype models is how the history of play is represented and encoded in memory chunks.

Let a_t be the action chosen by a player at time t from the action space $A = \{0, 1\}$ and a'_t be the action chosen by that player's opponent from $A' = \{0, 1\}$. The depth of pattern recognition is denoted by n and implies conditioning on the $(n-1)$ prior lags of play, e.g. $n = 2$ denotes patterns of two consecutive time

[8] Although similar, these models predict either categories or actual choices/decisions, not beliefs about a strategic opponent.

[9] I abide by the following two principles in defining the competing models within the common cognitive framework:

1. *Maximization of common auxiliary assumptions:* In competitive model comparisons, it is important to maximize the number of auxiliary assumptions that are common between the models, thereby facilitating the attribution of significant results to the main assumptions of interest.
2. *Minimization of auxiliary assumptions:* Whilst maximizing the number of common assumptions between models, the overall number of auxiliary assumptions should be minimized. This principle specifically attenuates the Duhem-Quine problem (Smith, 2002) that any hypothesis test is really a joint test of the hypothesis and all other assumptions made.

periods, or equivalently conditioning on the single prior lag. Define the context $\boldsymbol{\omega}_t$ at time $t \geq n$[10] (prior to making and observing action choices of round t) as the $2 \cdot (n-1)$-tuple $(a_{t-1}, \ldots, a_{t-n+1}, a'_{t-1}, \ldots, a'_{t-n+1})$ and ω_t^m as the mth element of $\boldsymbol{\omega}_t$. The depth of pattern recognition, and by extension the context, is constrained by the size of working memory—in order to encode a chunk the current context must be available in working memory. Therefore, assuming that working memory is exhausted, chunks storing n-period patterns require a working memory capacity of $2 \cdot (n-1) + 1$.[11]

3.1. Exemplar approach

After every round, an exemplar is constructed from the concatenation of a'_t (an opponent's action at time t) and $\boldsymbol{\omega_t}$ (the current context). If the current exemplar does not already exist in memory as a chunk then it is created, otherwise the existing chunk's activation will be updated (the updating procedure is described later). The stored context in a chunk j is denoted as $\boldsymbol{\omega_j}$, determined by the context at time t_j when the chunk was first created. The opponent's action choice given the stored context is referred to as the value of a chunk $V_j = a'_{t_j}$, which denotes the category the context is assigned to. See below for a tabular representation of an exemplar stored as a memory chunk—note, the information stored in each chunk does not change over time in the exemplar model. The maximum number of chunks in the exemplar model is $|A|^{n-1}|A'|^n$, however some of the possible contexts may not be observed during game play due to the limited number of rounds.

Slot type	Category	$\boldsymbol{\omega_j}$					
Slot ID	V_j	1	$\cdots$	$n-1$	n	$\cdots$	$M = 2(n-1)$
Chunk c_j	a'_{t_j}	a_{t_j-1}	$\cdots$	a_{t_j-n+1}	a'_{t_j-1}	$\cdots$	a'_{t_j-n+1}

3.2. Prototype approach

This model does not store individual exemplars, but rather stores and updates representative prototypes of categories by averaging over all the realizations of historical contexts for $\forall s < t$ for each category a'—the stored prototypical context is now time-dependent and represented by $\bar{\boldsymbol{\omega}}_{\boldsymbol{j}}(t)$. The maximum number of chunks stored is equal to the size of the opponent's action-space, therefore in this case there are only two chunks representing two categories or possible action choices. At the end of each round t, only chunk j's prototypical context $\bar{\boldsymbol{\omega}}_{\boldsymbol{j}}(t)$ is updated, where $V_j = a'_t$. Let $R = \{r : a'_r = 1, \forall n \leq r < t\}$; then the state of the chunks representing the prototypes at time t (at the time of decision making, but before updating the prototypes) is represented in tabular form below, where E is the expectation operator. In contrast to the

[10]This constraint ensures that sufficient time has lapsed for the observation of at least one n-period context.

[11]For example, 2-period patterns require the following items in working memory: own and opponent lagged actions a'_{t-1}, a_{t-1} and the subsequently observed opponent's current action choice a'_t. In general, n-period patterns require conditioning on $n-1$ prior lags for each subject's historical action profile (of which there are two), plus the current action played by one's opponent.

exemplar model, the information stored in each chunk changes over time as the prototype contexts $\bar{\boldsymbol{\omega}}_{\boldsymbol{j}}$ are updated. Note that the prototype context $\bar{\boldsymbol{\omega}}_{\boldsymbol{j}}$ could also be directly defined as a function of lagged actions a and a' since historical contexts are defined as $\boldsymbol{\omega_t} = (a_{t-1}, \ldots, a_{t-n+1}, a'_{t-1}, \ldots, a'_{t-n+1})$.

Slot type	Category	$\bar{\boldsymbol{\omega}}_{\boldsymbol{j}}\ (t)$					
Slot ID	V_j	1	$\cdots$	$n-1$	n	$\cdots$	$M = 2(n-1)$
Chunk c_1	$a' = 1$	$E_{s\in R}(\omega^1_{s-1})$	$\cdots$	$E_{s\in R}(\omega^{n-1}_{s-n+1})$	$E_{s\in R}(\omega^n_{s-1})$	$\cdots$	$E_{s\in R}(\omega^M_{s-n+1})$
Chunk c_2	$a' = 0$	$E_{s\notin R}(\omega^1_{s-1})$	$\cdots$	$E_{s\notin R}(\omega^{n-1}_{s-n+1})$	$E_{s\notin R}(\omega^n_{s-1})$	$\cdots$	$E_{s\notin R}(\omega^n_{s-n+1})$

3.3. Comparative differences between the exemplar and prototype approaches

There are important differences between the exemplar and prototype approaches to categorization (Ashby and Maddox, 1993); however, the exemplar model can emulate many characteristics of the prototype model, although the reverse is not true (Love and Tomlinson, 2010). This section highlights the differences for this particular application, clarifying under what circumstances the two models might diverge in their behavior and their relative performance.

The similarity comparison of the current context with the stored individual exemplars or archetypal prototypes leads to a categorization boundary that partitions the context space into areas where the current context is more likely to be assigned to either category.[12] The prototype approach presupposes that the categorization problem is linearly separable since the categorization boundary is an $(|A'| - 1)$-dimensional hyperplane (Love and Tomlinson, 2010). In contrast, the exemplar approach has a non-linear categorization boundary capable of capturing more complex categorization problems. The two approaches exist on the opposite ends of the bias-variance tradeoff (Briscoe and Feldman, 2011)—the prototype approach has higher bias but lower variance than the exemplar approach. Consequently, the exemplar approach is more prone to over-fitting to noise and less robust to outliers than the prototype approach. This is more important for strategic games than standard categorization problems as action choices (the category in this case) are stochastic with respect to the context, i.e. for the same context a subject may choose different actions at two different points in time.[13]

The exemplar and prototype approaches can reach different conclusions depending on the distance of a new observation from the prototype and previously observed exemplars (Homa et al., 1981). For example, consider a candidate set of new (unique) observations that are all equidistant from the prototype, but differ in their distance from prior exemplars. The prototype model would classify each of these observations as

[12] Any contextual points exactly on the categorization boundary would have an equal probability of being assigned to each category.

[13] In standard categorization tasks this relationship is usually deterministic—for example, a shape with four sides of equal length at right angles *always* belongs to the category of squares.

equally similar, but the exemplar approach would classify the observations that are closer to an existing exemplar (or cluster of exemplars) as more similar.

Furthermore, the prototype model may be problematic if there exist significant correlations between the elements of the context, as the averaging of contexts to create a prototype leads to the loss of this information (Medin et al., 1982; Nosofsky, 1986; Love and Tomlinson, 2010). For example, assume that a particular category is associated with the 2-tuple contexts $\{(1,1),(0,0)\}$ and the other category with the contexts $\{(1,0),(0,1)\}$. The prototypical context for both categories will be $(0.5, 0.5)$—it is impossible to distinguish between the two, despite the clear relationship between contexts and categories (contexts whose elements are not equal belong to one category, whereas contexts with matching elements belong to the other category).

Summarizing, the two approaches would make divergent predictions under the following circumstances:

1. if the categorization problem is not linearly separable (Love and Tomlinson, 2010).
2. in the presence of outliers, or atypical observations (Briscoe and Feldman, 2011).
3. if observations are noisy (Briscoe and Feldman, 2011).
4. if a new observation is identical (or close) to an old exemplar, for a given distance from the prototype (Homa et al., 1981).
5. if the elements of the observed contexts exhibit significant correlation (Medin et al., 1982; Nosofsky, 1986; Love and Tomlinson, 2010).

In terms of practical implementation, there is a significantly greater computational cost for the exemplar approach as each exemplar must be stored individually, rather than be averaged into a prototype, and the similarity process involves a large number of calculations for each exemplar. Therefore, the cost-benefit ratio of these two approaches will depend on the number of rounds of a repeated game (and by extension the number of exemplars observed) and the size of the action-spaces of the players. Finally, if the computational and storage constraints are binding for this particular problem, the increased demands of the exemplar approach may necessarily limit its application to lower pattern detection depths than the prototype approach.

3.4. ACT-R cognitive framework

The underlying cognitive architecture describing memory is based on the Rational Analysis of Memory model (Schooler and Anderson, 1997) incorporated in the general ACT-R framework (Anderson, 1996; Anderson and Lebiere, 1998; Anderson, 2007). Each memory chunk has an activation level that controls the likelihood and speed with which this chunk can be retrieved from the declarative memory module. Whenever a chunk is observed (exemplar model) or modified (prototype model) its activation level is increased, whereas an inactive chunk loses activation over time. The total activation of chunk j, A_j, is comprised of three different types of activation. Let B_j represent the base-level activation, t_q represent the time elapsed since this chunk was observed in the history of play for each of q times in the past, and γ be the rate of activation decay:

$$B_j = \ln \sum_q t_q^{-\gamma} \tag{1}$$

The context-dependent activation of a chunk depends on the degree of matching between the current context and the stored context of each memory chunk. To allow for non-exact matching, i.e. retrieval of chunks that may have a similar but not exact same historical context, a similarity function must be defined. Let the attention weights w_m denote the attention a subject pays to each element in the context, where $w_m \geq 0$ and $\sum w_m = 1$. The dissimilarity function, $\varDelta_j$ between the current context $\boldsymbol{\omega_t}$ and the context encoded in a memory chunk j, $\boldsymbol{\omega_j}$ (or $\bar{\boldsymbol{\omega}}_{\boldsymbol{j}}(t)$ for the prototype model), is given by a distance function employing the city-block metric as recommended by Gärdenfors (2004) for separable dimensions, see Eq. 2 below. Note that the distance is equal to zero if $\boldsymbol{\omega_j} = \boldsymbol{\omega_t}$, i.e. if the current context is identical to the chunk's stored context.

$$\varDelta_j = \sum_m w_m |\omega_t^m - \omega_j^m| \tag{2}$$

The total activation of a chunk A_j is equal to the base-level activation minus the product of the dissimilarity function and the matching penalty parameter μ determining the importance of exact matching:

$$A_j = B_j - \mu \varDelta_j \tag{3}$$

This setup bears some similarities to (and deviations from) a normative Bayesian model of information integration as discussed in Massaro and Friedman (1990).[14] The ACT-R cognitive architecture was inspired by the principle that the statistical properties of the environment should be reflected in cognitive mechanisms, such as memory. The activation level A_j can be viewed as capturing the posterior log-odds of the relevance of chunk j conditional on the current context (or alternatively, the probability of a specific n-period pattern occurring). Specifically, B_j is an estimate of the log-prior odds of needing the chunk (or observing a pattern) and $\mu\varDelta_j$ adjusts the log prior odds given the current context. In this sense, $\mu\varDelta_j$ corresponds to an estimate of the log likelihood given the current context, although technically this adjustment cannot be termed Bayesian, as its units are in distance (determined by the dissimilarity function) instead of probability. The integration of the contextual values for each lagged action belongs to the family of multidimensional scaling models discussed in Massaro and Friedman (1990).

[14] Note that Massaro and Friedman (1990) present the normative Bayesian model where the relationship between categories and evidence is known, and does not need to be learned. In our case, the relationship between categories and the context must also be learned in real time, requiring an extra dimension of Bayesian updating after each round of play; also, the prior unconditional probabilities of each category are unknown and must be updated.

Let τ be a fixed threshold level that fixes the minimum activation level necessary for a chunk to be accessible for retrieval. Accessing chunks is prone to an error $\epsilon_j \sim N(0, \sigma^2)$, which will be approximated by the logistic function for tractability. The standard ACT-R probability of retrieving chunk j, r_j, is given by Eq. 4, where $s = \sqrt{3}\sigma/\pi$. This is a sigmoid function with an inflection point at $\tau = A_j$ corresponding to a 50% probability of chunk retrieval. If $A_j < \tau \Rightarrow r_j < 0.5$ (similarly, $A_j > \tau \Rightarrow r_j > 0.5$)—a chunk with activation below the threshold may still be retrieved due to the activation noise s.

$$r_j = \left[1 + e^{\frac{\tau - A_j}{s}}\right]^{-1} \tag{4}$$

The value V_j of each memory chunk can be viewed as a prediction, or the vote a memory chunk casts in favor of each possible action that may be played. Beliefs are constructed by the process of blended retrieval (Gonzalez and Lebiere, 2005); each memory chunk's vote is weighted by the probability of retrieving that particular chunk given the current context. This technique belongs to the group of ensemble learning models since each chunk acts as a single model supporting a particular hypothesis. Ensemble models can exhibit better and more robust predictive performance than any of the individual constituent models (Dietterich, 2000).[15]

The beliefs assigned to category $a'_t = 1$ given the current context $\boldsymbol{\omega_t}$ and the strength of an equiprobable prior δ are given by:

$$p(a'_t = 1|\boldsymbol{\omega_t}) = \frac{\delta + \sum_{\{j:V_j=1\}} r_j}{2\delta + \sum_j r_j} \tag{5}$$

The effect of τ on belief formation is important because by directly setting the minimum level of activation necessary for non-zero probability of retrieval, this parameter skews the information included in belief formation. If τ is high, proportionately more weight is given to chunks with high activation. The level of τ captures the tradeoff between the cost of using more information, i.e. memory chunks that are harder to retrieve and therefore take longer, versus the benefit of having this additional information available to incorporate into beliefs.

3.5. Baseline models

Three baseline models are included in the analysis: a theoretical baseline model (the MSNE beliefs, *msne*), a statistical baseline model (the empirical marginal probabilities, *emp*) and a baseline learning model (the weighted fictitious play model, *wfp*).[16] Since *wfp* is the standard belief formation model used

[15] Furthermore, the blended retrieval process can be viewed as a Bayes classifier using approximately optimal weights (Gonzalez and Lebiere, 2005). Recall the discussion earlier linking activation levels A_j to Bayesian calculations—since r_j is a function of A_j these weights can be considered approximately optimal in the Bayesian sense, albeit tempered by biological cognitive constraints such as forgetting, noise, and threshold effects.

[16] Note, the baseline models *msne* and *emp* have the same within- and between-dataset CV as they have no estimated parameters, in contrast to *wfp*.

in the literature, and outperforms the other two baseline models, this is the most competitive baseline to compare the effectiveness of pattern-detecting models.

The MSNE beliefs are given by the unique mixed-strategy Nash solution of a game (note, this model does not require any empirical data to make predictions):

$$msne_{NS}(a'_t = red) = 0.6 \text{ for both column and role players}$$

$$msne_{RW}(a'_t = up) = 0.5, msne_{RW}(a'_t = left) = 0.05$$

The empirical marginal probability model is not a real-time learning model that can be employed by subjects, but serves as a statistical baseline to determine whether the ACT-R model can make more accurate predictions on new data when previous experimental data exists. If $I(a'_t)$ is an indicator function that is equal to one if $a'_t = 1$ and zero otherwise, the empirical marginal probabilities over an estimation dataset T_{est} are given by:

$$emp(a'_t = 1) = \mid T_{est} \mid^{-1} \sum_{t \in T_{est}} I(a'_t)$$

Weighted fictitious play (Cheung and Friedman, 1997) is presented in Eq. 6, where $I\left(a'_{t-1}\right)$ and $I\left(a'_{t-u-1}\right)$ are indicator functions whose values are equal to one if $a'_{t-1} = 1$ and $a'_{t-u-1} = 1$ respectively, and zero otherwise. The forgetting parameter is denoted as $\tilde{\gamma}$ and plays a similar role to its counterpart in the ACT-R model; however, in *wfp* the forgetting function is an exponential function instead of the inverse power function in ACT-R.[17]

$$wfp(a'_t = 1) = \frac{I(a'_{t-1}) + \sum_{u=1}^{t-2} \tilde{\gamma}^u \cdot I(a'_{t-u-1})}{1 + \sum_{u=1}^{t-2} \tilde{\gamma}^u}. \tag{6}$$

4. Estimation and results

The prototype and exemplar models will be estimated with different degrees of freedom to ascertain the robustness of cross-validation performance and to establish which parameters truly require estimation. For ease of exposition, let a model m be denoted as $ex(\theta; n)$ and $pr(\theta; n)$ for exemplar and prototype models respectively—θ is the set of free (estimated) parameters and n is the depth of pattern recognition. For example, $pr(\tau, \delta; 3)$ is a prototype model of depth $n = 3$ where τ and δ are the only estimated parameters. Fixed parameters in each model will be set to the following values. All w_m will be equal, implying the same

[17] An assumption of perfect memory with no forgetting is captured by $\gamma = 0$ in ACT-R, in contrast to $\tilde{\gamma} = 1$ in *wfp*.

degree of attention or importance of similarity matching for every lagged action choice. The δ parameter will be set to zero, implying no prior beliefs about an opponent's behavior. The remaining parameters $\tau, \gamma, \sigma, \mu$ do not have any standard values that can be justified using the principle of insufficient reason. However, an online database exists of the estimated parameter values over a wide range of applications for different ACT-R studies (Wong et al., 2010) irrelevant to strategic decision making. These parameters will be set to the median values across all the studies reported in this database[18]—these values are $\tau = -0.3, \gamma = 0.4, \sigma = 0.3, \mu = 1.5$.

The following restriction is imposed on the estimated attention vector $\boldsymbol{w}$ to reduce the number of free parameters. Assume that subjects direct a fixed proportion of attention to their own and an opponent's lagged choices, and fixed proportions of attention to each particular lag. Let w_o be the attention directed to a subject's own lagged choices and $(1 - w_o)$ be the attention directed to an opponent's lagged choices. For pattern detection of depth n, let $w_1, ..., w_{n-1}$ be the attentional weights directed towards lagged choices (both their own and an opponent's), the sum of which is constrained to one. For example, the attentional weight on the second own lag would be $w_o \times w_2$, the weight on the opponent's first lag $(1 - w_o) \times w_1$. This reduces the number of free parameters of $\boldsymbol{w}$ from $2n - 3$ to $n - 1$.

A brief discussion of the estimation procedure follows (for a more detailed exposition the reader is referred to Appendix B). In NS the first ten rounds are not included in any measure of fit or cross-validation performance, and in RW the first eleven rounds are not included.[19] Let $sb_{i,t}$ denote player i's stated belief at time t and $\hat{b}_{i,t}(\theta_i)$ denote a model's empirically estimated beliefs as a function of the individual parameter set θ_i—models are always estimated per individual to account for the possibility of parameter heterogeneity. The objective function of an estimation procedure seeks to minimize the root mean squared deviation (RMSD) between stated beliefs and model-predicted beliefs over the set of time periods included in an estimation dataset T_{est}: $\min_{\theta_i = \theta_i^\star} \sqrt{\mid T_{est} \mid^{-1} \sum_{t \in T_{est}} \left(sb_{i,t} - \hat{b}_{i,t}(\theta_i) \right)^2}$. Similarly, the cross-validation performance is again measured as the RMSD between stated beliefs and model-predicted beliefs over the observations in the cross-validation dataset (for the optimal parameter set derived from the estimation procedure). The cross-validation performance is averaged over all observations and individuals, therefore the reported RMSD is the error of a subject's stated belief for a single round of play.

Within-dataset-CV is performed according to the k-fold cross-validation procedure (see Shiffrin et al. (2008) for a discussion of this and other model evaluation techniques). The two datasets are divided into folds consisting of five temporally sequential rounds, yielding ten folds for NS composed of rounds $\{(11, 12, 13, 14, 15), \ldots, (56, 57, 58, 59, 60)\}$, and five folds for RW $\{(12, 13, 14, 15, 16), \ldots, (32, 33, 34, 35, 36)\}$.

[18]The database can be found at http://www-abc.mpib-berlin.mpg.de/actrdb/ and the reported median parameter values are those reported in the database on 1st October, 2011. The fact that researchers using ACT-R have recourse to such a database is important, as it facilitates the examination of cognitive models' ability to generalize across different data and even domains. If these general parameters are found to predict well in the investigated tasks, it will further support the appropriateness of modeling strategic behavior with ACT-R.

[19]This is done for two reasons: first, to allow for beliefs to settle as they are often very noisy in the first few rounds and, second, to allow for a large enough history of play to be available to subjects to permit pattern search. Note that information from these prior periods is still used to form beliefs.

The model is estimated by withholding one fold from the chosen dataset (either NS or RW) as the cross-validation set T_{cv}, and estimating the model on the remaining $k-1$ folds, which form T_{est}—this estimation procedure is repeated until all k folds have served once as the cross-validation set. In this manner, out-of-sample predictions are collected for all folds, and by extension for all the observations. The within-dataset-CV performance per model is the RMSD calculated over all cross-validation folds and subjects—it is denoted by $wCV_y[m]$ where y denotes the dataset under examination.

For the bCV procedure, parameters are estimated using all the observations from one of the two datasets, NS or RW—the parameters are then fixed and used to predict behavior in the other dataset. Hence, the objective function minimizes the RMSD calculated over all the rounds in the estimation dataset, whereas the between-dataset-CV measure is derived from the other dataset's predictions. Let ${}_xbCV_y[m]$ denote the RMSD for between-dataset-CV, where x is the estimation dataset and y is the cross-validation dataset. In calculating this, allowance is made for the individual parameter heterogeneity resulting from the estimation procedure—each player in the CV dataset is assumed to have an equal probability of sharing the same parameter values with each of the subjects in the estimation dataset (see Appendix B.2 for calculation details).

Finally, the gCV procedure requires no parameter estimation, only calculation of the cross-validation performance for the set of ACT-R derived parameters—$gCV_y[m]$ denotes the generalized-between-dataset RMSD for prediction of dataset y.

Due to the absence of general theoretical results regarding the distribution of CV statistics, the standard procedure used when comparing models based on cross-validation is simply to accept the model with the best CV performance. I propose the following non-parametric test procedure to allow for an examination of statistical significance. A paired sign-test or Wilcoxon signed-rank test can be executed on the mean CV per subject for each model. The null hypothesis of no difference implies that each model should outperform the other in approximately 50% of subjects. These non-parametric tests are less powerful than their parametric counterparts,[20] and therefore make it harder to reject the null hypothesis of no pattern recognition.

The results from within-dataset-CV evaluation are reported in Table 2, whilst the between-dataset-CV evaluation is presented in Table 5. The mean RMSD for cross-validation procedures on both datasets are denoted as $\overline{wCV}[m]$, $\overline{bCV}[m]$ and $\overline{gCV}[m]$. Note, the empirical marginal probability and the weighted fictitious play baseline models are estimated and cross-validated in an identical manner as the ACT-R models, i.e. on exactly the same partitions of the datasets.[21]

[20]The normality assumptions inherent in standard parametric tests were violated by the data and therefore are not reported. In the interests of transparency, results will be presented for both of these tests, however the signed-rank test is preferable as it incorporates more information and has greater power.

[21]The MSNE belief model does not require any empirical data as its prediction is calculated from the game payoff matrix.

Table 2: Within-dataset and generalized-between-dataset cross-validation performance per dataset (RMSD)

		NS				RW				Average of NS & RW			
	msne	0.306				0.412				0.359			
	emp	0.302				0.323				0.312			
	wfp	0.300				0.300				0.300			
		n				n				n			
		2	3	4	5	2	3	4	5	2	3	4	5
gCV	pr	0.292	0.291	0.292	0.296	0.315	0.317	0.317	0.318	0.303	0.304	0.304	0.307
	ex	0.308	0.317	0.332	0.330	0.315	0.331	0.334	0.336	0.311	0.324	0.333	0.333
wCV	$pr(\tau)$	0.275	0.275	0.274	0.275	0.288	0.289	0.291	0.289	0.282	0.282	0.282	0.282
	$ex(\tau)$	0.271	0.274	0.277	0.282	0.286	0.290	0.296	0.303	0.279	0.282	0.286	0.293
	$pr(\tau, \delta)$	0.266	0.263	0.265	0.264	0.274	$\underline{0.272}$	0.275	0.276	0.270	0.268	0.270	0.270
	$ex(\tau, \delta)$	0.266	0.261	0.260	0.262	0.277	0.277	0.279	0.279	0.271	0.269	0.269	0.270
	$pr(\tau, \delta, \boldsymbol{w})$	0.267	0.263	0.265	0.265	0.275	0.277	0.280	0.280	0.271	0.270	0.273	0.273
	$ex(\tau, \delta, \boldsymbol{w})$	0.259	0.264	0.263	0.265	0.275	0.279	0.283	0.284	$\underline{0.267}$	0.271	0.273	0.274
	$pr(\tau, \delta, \gamma, \boldsymbol{w})$	0.261	$\underline{0.257}$	0.258	0.259	0.283	0.284	0.288	0.286	0.272	0.270	0.273	0.273
	$ex(\tau, \delta, \gamma, \boldsymbol{w})$	0.258	0.261	0.262	0.265	0.284	0.286	0.288	0.287	0.271	0.274	0.275	0.276

Underscored values represent the models with the lowest RMSD within each set of data.

4.1. Within-dataset prediction results

The best average within-dataset performance for both datasets is $\overline{wCV}\,[ex(\tau, \delta, \boldsymbol{w}; 2)] = 0.267$, which is insignificantly different from the next best performance $\overline{wCV}\,[pr(\tau, \delta; 3)] = 0.268$. The latter is preferable, on the grounds of employing 120 degrees of freedom less,[22] and because prototype models are the best performers in each individual dataset. For these reasons $pr(\tau, \delta; 3)$ shall be referred to as the best performing model in this instance. Note that the gCV performance of models, not estimated on the NS or RW datasets, is clearly trumped by models with free parameters fitted to these datasets—this could be due to either or both of the following reasons. Firstly, there may exist significant subject heterogeneity that is not addressed by the gCV procedure (as the same parameter estimates are used for each subject in this case). Secondly, parameters of the model are affected by characteristics of specific tasks and are not stable across different domains.

4.1.1. Subjects' beliefs incorporate pattern detection not modeled in weighted fictitious play

The performance of the *wfp* model is $wCV_{NS}[wfp] = 0.3$, $wCV_{RW}\,[wfp] = 0.3$ and $\overline{wCV}\,[wfp] = 0.3$. In contrast, the best predictive performance of the pattern detecting models is $wCV_{NS}\,[pr(\tau, \delta, \gamma, w; 3)] = 0.257$, $wCV_{RW}\,[pr(\tau, \delta; 3)] = 0.272$ and $\overline{wCV}\,[pr(\tau, \delta; 3)] = 0.268$.[23] These are economically significant differences, in particular, for the NS dataset using *wfp* instead of *msne* leads to a reduction in RMSD of 0.006, whereas

[22] One additional parameter is estimated for the exemplar model, w_0, for each of 28 subjects in NS and 92 subjects in RW—note that $\boldsymbol{w} = w_0$ when $n = 2$.

[23] The larger error for the RW dataset compared to NS verifies the hypothesis that pattern recognition may play a less important role in RW due to the smaller number of rounds and high degree of asymmetry.

Table 3: Within-dataset evidence of pattern recognition with respect to *wfp*

	NS $pr(\tau, \delta, \gamma, \boldsymbol{w}; 3)$	RW $pr(\tau, \delta; 3)$	Average $pr(\tau, \delta; 3)$
Sign test	$p = 0.0037$	$p < 0.0001$	$p < 0.0001$
Signed-rank test	$p = 0.0005\,(z = -3.46)$	$p < 0.0001\,(z = 5.179)$	$p < 0.0001\,(z = 6.17)$

the improvement from using $pr(\tau, \delta, \gamma, \boldsymbol{w}; 3)$ compared to *msne* is 0.049. In the RW dataset, where the empirical marginal probabilities are still very far from the MSNE, the reduction in RMSD from using *wfp* instead of *msne* is 0.112, whereas the improvement from using $pr(\tau, \delta; 3)$ instead of *wfp* is 0.028. The improvements in RMSD for both datasets testify to the added value of extending belief models to include pattern recognition.

I examine the statistical significance of these results by testing the null hypothesis that there is no difference between the predictive performance of the best pattern detecting model for each dataset and *wfp*—results are reported in Table 3. The hypothesis of no difference is clearly rejected at the 1% level of significance in each dataset by both tests (sign and signed-rank), as is the case comparing the average performance on both datasets. The Bonferroni correction for multiple comparisons supports this conclusion—since 40 models are being compared for each dataset, the corresponding Bonferroni corrected significance level is $0.05/40 = 0.00125$.

In sum, pattern detection is both economically and statistically significant in these datasets, therefore misspecifying the belief formation model using *wfp* may lead to biased parameter estimates (Spiliopoulos, 2012). It is advisable for future studies to first test a hypothesis of pattern recognition in subjects' behavior, and only if this is ruled out empirically proceed to use non-pattern detecting models of learning.

4.1.2. Subjects predominantly use pattern recognition of depth $n = 2$

The depth of the best performing within-dataset model for both datasets is $n = 3$, but this performance is very close to the best performing model with $n = 2$. Table 4 presents formal tests of the null hypothesis of no difference between the within-dataset-CV for $n = 2$ and $n = 3$ pattern detecting models. The null hypothesis cannot be rejected at any reasonable level of significance for either dataset or the average of both datasets (note the high p-values for the latter, $p = 1$ for a sign test and $p = 0.778$ for a signed-rank test).

In short, the economically significant improvement in belief modeling is captured primarily by 2-period pattern detection. Although results are statistically indistinguishable from models where $n = 3$, Occam's razor dictates that 2-period pattern detection is preferable; the latter incorporates less historical information, requires smaller working memory capacity, and has fewer cognitive computational costs.[24]

[24] The reader is reminded that, had the comparison been performed on in-sample fits, the conclusion would have been that the best performing depth of pattern recognition is $n = 5$, as models of less depth are nested within higher depth models.

Table 4: Depth of pattern recognition

Models	NS $ex(\tau, \delta, \gamma, \boldsymbol{w}; 2)/pr(\tau, \delta, \gamma, \boldsymbol{w}; 3)$	RW $pr(\tau, \delta; 2)/pr(\tau, \delta; 3)$	Average $ex(\tau, \delta, \boldsymbol{w}; 2)/pr(\tau, \delta; 3)$
Sign test	$p = 0.185$	$p = 0.594$	$p = 1$
Signed-rank test	$p = 0.716\,(z = -0.364)$	$p = 0.683\,(z = 0.409)$	$p = 0.778\,(z = 0.282)$

4.1.3. Comparison of exemplar and prototype models

The cross-validation performance of both exemplar and prototype models is often very similar; specifically, as the number of free parameters increases the models' performance tends to converge.[25] The most striking difference in performance is obtained when comparing the exemplar and prototype models with fixed parameters. The difference is more pronounced in the NS dataset than in RW; however, the prototype model performs better in both instances. This is preliminary evidence that the prototype model may be more appropriate than the exemplar model, but further investigations are required to claim this with a sufficient degree of confidence. A prototype model is preferred on the basis of its simplicity (reduced computational and memory costs) as it requires the storage and processing of only $\mid A' \mid$ chunks compared to a maximum of $\mid A \mid^{n-1} \cdot \mid A' \mid^{n}$ chunks for the exemplar model.[26]

4.2. Between-dataset prediction results

Predicting behavior in completely different datasets (and games) is more demanding than within-dataset-CV. Two important observations arise from a comparison of the bCV with the wCV procedure (see Tables 2 and 5). Firstly, as expected, the predictive performance of models has degraded in the bCV procedure compared to wCV—this points to parameter instability across the two datasets. Secondly, note that the best performing models in Table 5 are gCV_{NS} $[pr(3)]$, ${}_{NS}bCV_{RW}$ $[ex(\tau; 4)]$, $\overline{gCV}$ $[pr(2)]$. These models have fewer degrees of freedom than the best performing wCV models. Surprisingly, the latter model that uses fixed parameters estimated from other unrelated ACT-R studies outperforms models using parameters estimated from the NS and RW data.[27] This is counter-intuitive as we might expect parameters estimated in similar tasks to be more appropriate than those estimated on very different tasks. This is evidence that a theory of parameter variation linked to task characteristics is an important future endeavor. At the same time, the good performance of the ACT-R model using parameter estimates from a wide range of unrelated studies

[25]This is a common problem as different non-linear models may exhibit a large degree of mimicking for certain parameter values. This is exacerbated by the number of degrees of freedom—the more free parameters, the greater the likelihood that parameter values will exist for which the models behave similarly.

[26]The exact number of chunks needed in an exemplar model may be less than this value if not all possible combinations of action histories are realized. Note that this is increasing in n for the exemplar model but independent of n for the prototype model; therefore, the latter may be a more efficient categorization algorithm for pattern detection of large depth despite the information loss associated with storing prototypes.

[27]Note that NS and RW parameter estimates are a distribution of values for the subject pool and therefore are more detailed than the homogeneous estimates from the online ACT-R database.

Table 5: Between-dataset and generalized-between-dataset cross-validation performance per dataset (RMSD)

		NS				RW				Average			
	msne	0.306				0.412				0.359			
	wfp	0.318				0.307				0.313			
		n				n				n			
		2	3	4	5	2	3	4	5	2	3	4	5
gCV	*pr*	0.292	$\underline{0.291}$	0.292	0.296	0.315	0.317	0.317	0.318	$\underline{0.303}$	0.304	0.304	0.307
	ex	0.308	0.317	0.332	0.330	0.315	0.331	0.334	0.336	0.311	0.324	0.333	0.333
bCV	$pr(\tau)$	0.310	0.306	0.308	0.312	0.340	0.339	0.342	0.345	0.325	0.323	0.325	0.329
	$ex(\tau)$	0.315	0.310	0.312	0.311	0.320	0.310	$\underline{0.308}$	0.313	0.318	0.310	0.310	0.312
	$pr(\tau, \delta)$	0.301	0.297	0.299	0.302	0.328	0.326	0.328	0.332	0.314	0.311	0.314	0.317
	$ex(\tau, \delta)$	0.303	0.296	0.299	0.297	0.329	0.324	0.328	0.329	0.316	0.310	0.314	0.313
	$pr(\tau, \delta, \boldsymbol{w})$	0.305	0.301	0.305	0.307	0.327	0.327	0.326	0.326	0.316	0.314	0.316	0.317
	$ex(\tau, \delta, \boldsymbol{w})$	0.306	0.302	0.308	0.305	0.329	0.321	0.319	0.322	0.318	0.312	0.314	0.313
	$pr(\tau, \delta, \gamma, \boldsymbol{w})$	0.313	0.310	0.313	0.313	0.331	0.333	0.332	0.332	0.322	0.321	0.323	0.323
	$ex(\tau, \delta, \gamma, \boldsymbol{w})$	0.319	0.314	0.317	0.314	0.335	0.330	0.328	0.331	0.327	0.322	0.323	0.322

Underscored values represent the models with the lowest RMSD within each set of data.

Table 6: Between-dataset evidence of pattern recognition with respect to *wfp*

	NS	RW	Average
	$pr(3)$	$ex(\tau; 4)$	$pr(2)$
Sign test	$p = 0.002$	$p = 0.0929$	$p = 0.027$
Signed-rank test	$p = 0.003\,(z = 3.598)$	$p = 0.141\,(z = -1.474)$	$p = 0.228\,(z = 1.206)$

is encouraging evidence that this model can serve as a general cognitive platform for human cognition, including strategic interactions and decision making.

Statistical tests comparing the best performing models to *wfp* are shown in Table 6. There is evidence from the *bCV* procedure confirming that pattern detection is an important determinant of beliefs but it is weaker than that found in the *wCV* procedure—this is to be expected if parameters are not stable across datasets. In the *NS* dataset, the best pattern detection model is statistically significant at the 5% level for both tests, in contrast to the RW dataset where we find no significant difference at the 5% level. Similarly, comparing the best model for both datasets, $pr(2)$ with *wfp*, the former is found to be significantly different from the latter at the 5% level by a sign test but not by a signed-rank test. Bonferroni multiple comparison corrections would reject the null hypothesis for the NS dataset at approximately the 10% level, but would clearly not reject the null hypothesis for the RW dataset at any meaningful level.[28]

Furthermore, comparison of the prototype and exemplar models using the *gCV* procedure with weighted fictitious play puts the former at a disadvantage, as the latter is estimated on data from one of the two

[28] However, the Bonferonni correction provides a lower bound that is conservative in rejecting the null hypothesis if individual tests are positively correlated—this is particularly true in this case as many models are nested.

experiments under investigation (NS and RW). The prototype and exemplar models in this case are calibrated on completely unrelated studies, thus they are capable of making predictions without observing any empirical data on strategic games. Another interesting comparison is with the MSNE, which is also capable of making predictions in the absence of any empirical data. The prototype model makes more accurate predictions than the MSNE, as $gCV\,[pr(2)] = 0.303$ compared to 0.359 for the MSNE—beliefs are more accurate by a magnitude of 0.056.

4.2.1. Pattern recognition depth and prototype versus exemplar models

The conclusions reached from the wCV procedure carry over to the bCV procedure. The depth of pattern recognition is confirmed as $n = 2$ by the best performing model across both datasets $\overline{gCV}\,[pr(2)]$ on the basis of Occam's razor. Note that the RMSD is virtually the same for $n = 3, 4$ but the difference between $\overline{gCV}\,[pr(2)]$ and $\overline{gCV}\,[pr(5)]$, 0.004, is found to be statistically significant according to both a sign test and Wilcoxon signed-rank test, $z = -2.2, p = 0.028$ and $p = 0.009$ respectively. Also, the prototype models moderately outperform the exemplar models, especially when all parameters are fixed rather than estimated.

4.2.2. Stability of parameter estimates across datasets

The benefit of performing both within- and between-dataset comparisons is that the within-dataset analysis has clarified which parameters are required to effectively predict behavior in these games—namely the parameters (τ, δ) from the best performing $pr(\tau, \delta; 3)$ model. Table 7 compares various percentiles of the parameter distributions (arising from the individual heterogeneity in these parameters) using NS and RW as the estimation datasets. The last column presents exact Kolmogorov-Smirnov tests of the null hypothesis that both distributions are identical. This assumption cannot be rejected for the threshold parameter τ at any reasonable level of significance, but can be rejected for δ at the 10% level. The robustness of τ across these two games is an encouraging sign for the ACT-R model; however, it should be noted that the median values, 1.533 and 1.339 for NS and RW respectively, are quite far from the standard ACT-R parameter value of -0.3.[29] This is in line with the findings of the ACT-R community identifying systematic variation in the threshold parameter τ across studies (Anderson and Lebiere, 1998). It is therefore recommended that τ be estimated for each task or experimental dataset.

A comparison with the between-dataset analysis highlights whether these parameters are stable or whether a meta-theory is required to explain how these parameters vary. Given the observed inadequacy of parameter estimates from NS to perform well on RW data, is it possible to link parameter values to game characteristics without observing actual behavior? This cannot be accomplished in this study as it would

[29] These results are statistically significant at the 1% level using Wilcoxon signed-rank tests of the null hypothesis that the medians are equal to -0.3, $p < 0.0001$ and $p = 0.01$ respectively.

Table 7: Distributions of parameter estimates for $pr(\tau, \delta; 3)$ using the bCV procedure

		Percentile (%)					
Estimation dataset	Parameter	10	25	50	75	90	K-S p-value
NS	τ	-3.692	-0.097	1.533	2.249	2.926	0.569
RW	τ	-4.557	0.301	1.339	2.587	3.067	
NS	δ	0.000	0.010	0.083	0.609	2.809	0.082
RW	δ	0.000	0.001	0.011	0.263	2.863	

require a far larger number of games and datasets to robustly make such associations. The discussion below outlines a plan of how this could be achieved in future work.

The prior strength parameter δ can be linked to the payoff structure of a game, that is, it may be associated with behavior in the one-shot heuristic literature. Also, non-uniform priors biased in favor of specific actions could be incorporated and linked to the underlying payoff structure. Note that the results presented above are consistent with this interpretation as δ is higher for the NS than for the RW dataset, i.e. the strength of an equiprobable prior is larger in the NS game, whose MSNE is closer to equiprobable play than the RW game.

The threshold parameter τ captures the cost-benefit tradeoff from retrieval of memory chunks. The higher the retrieval cost of a memory chunk, the higher the threshold should be, ceteris paribus, as this limits the retrieval of memories. A cost-benefit analysis of retrieving additional information depends upon two characteristics of a game: the financial or monetary incentives associated with the game, and whether the game exhibits focal points which may simplify the game, thereby not requiring much cognitive effort to perform well in the game.

In general, findings of parameter instability may have different implications, especially when considered concurrently with the degree of generalizability of a model (Yechiam and Busemeyer, 2008). A model may be correctly specified yet show parameter instability and low generalizability if certain parameter values are task specific. In this case, the model may be viewed as incomplete since it does not include a theory of parameter variation; however, for each task the combination of model structure and estimated parameters is still correctly specified.

Model misspecification in light of parameter instability can be ascertained by considering whether the unstable parameters should reflect a stable, latent variable or whether there is a plausible reason for parameter variation, *given the task variations.* This qualification is important, as it is possible to justify the variation of almost any parameter on some basis. For example, the relative weights attached to gains and losses in risky choice in Yechiam and Busemeyer (2008) were found to be stable, and arguably represent stable internal traits given the task variations in their experimental design; however, these relative weights could likely still be manipulated by the prior context (such as reference points). In our discussion above, τ

and δ arguably could be linked to variations in the experimental game characteristics, rather than be solely attributed to stable, internal latent variables.

Alternatively, parameter instability may arise from over-specification or undue model complexity—in non-linear models, parameters are often poorly identified, and therefore estimates may vary considerably without a significant loss of generalizability. This can be ruled out in two different ways. First, by rigorously estimating and cross-validating nested subsets of the full model (by imposing parameter constraints)—this approach has been used extensively in this paper. Second, poor identification can be revealed by the covariance of parameter estimates over cross-validation folds.

Concluding, it is imperative that future research involving model comparisons systematically address these possibilities. The most common progression of a field of study is for models to become ever more complex as new results are obtained that cannot be explained by earlier models. The progression to more complex models is not necessarily wrong, but greater caution must be exercised and robust evidence presented to justify such a move.

5. Discussion

The finding that subjects detect patterns of depth 2 (and possibly 3), requiring 3 (and 5) items in working memory respectively, is in accord with the 4 ± 1 item constraint found in a wide variety of tasks (Cowan, 2001). Importantly, this finding is based on the within-dataset, between-dataset and generalized-between-dataset cross-validation procedures, testifying to the robustness of this result. Whether this result is a general one, and extends to other strategic games, remains an important question. While a definitive answer requires further empirical research, there are various reasons why this may be the case.

The first is the constraint on working memory, which would limit pattern detection to 2-3 periods in any application. The second is that conditioning beliefs only on the single most recent lagged information may be a fast and frugal heuristic, along the lines of Gigerenzer and Selten (2002). Due to the recursive and coupled nature of players' choices in repeated games, information lagged by one period contains some information from all the prior lags as well, albeit with some loss of information. However, this loss of information may actually turn out to be an advantage as it prevents subjects from over-fitting their beliefs to noise in action profiles. Thirdly, this argument is further strengthened if players do not exhibit significant patterns of higher depth. If players' decision rules include a component that actively, but inefficiently, attempts to randomize, then this should also be restricted by working memory capacity. Therefore, patterns arising from inefficient randomization are likely restricted to depths of two or three, making this level of pattern detection appropriate. Finally, the applicability of pattern detection falls as n increases because a longer history is required before it can be employed. These arguments suggest that robust pattern detection in general human interactions may be achieved with 2-period patterns.

The robustness of pattern detection in between-dataset comparisons, in contrast to the instability of parameter estimates of the models, has important implications for the learning literature. While the majority of studies allow heterogeneous parameter estimates of learning models, they rarely allow for pattern detection. Hypothesis tests of the increasingly complex decision rules used in the literature–for example, incorporating forward-looking behavior–are contingent on the assumptions of the underlying belief model; therefore it is important not to misspecify the belief models. Incorporating pattern detection in beliefs will likely lead to even more interesting findings on how decision rules incorporate this information.

Although this paper focuses on repeated games with a unique mixed-strategy Nash equilibrium, there is nothing precluding the use of pattern detecting beliefs in other games. For example, in repeated prisoner's dilemma games, often-found strategies such as tit-for-tat are essentially pattern detecting strategies using information from prior lags. There is scope for unifying behavior in these two games by modeling pattern detection using the cognitive memory framework discussed herein. This would also permit a thorough investigation as to how cognitive restrictions in working and long-term memory may affect specific strategy use, along the lines of Stevens et al. (2011).

Pattern detecting models can be extended to permit forward-looking behavior, providing beliefs about an opponent's behavior further into the future than just the current round. Experiments using stated beliefs have so far only elicited beliefs about the current round; collecting beliefs further into the future would allow the empirical estimation of forward-looking beliefs.

Finally, using cognitive-architecture models that specify computational processes, it is possible to derive response time predictions. The collection of such data would permit stronger evaluation of competing models, a particularly important consideration in the game theory learning literature, which often suffers from low econometric power. In the cognitive psychology literature, comparisons of empirical response time distributions have contributed significantly to the evaluation of competing models—see Luce (1996) for a general discussion. Some specific examples include distinguishing between serial and parallel memory search processes (Donkin and Nosofsky, 2012), instinctive and cognitive reasoning in games (Rubinstein, 2007), and compensatory and non-compensatory processes in decision making from memory (Marewski and Melhorn, 2011).

Acknowledgements

The author would like to thank Andreas Ortmann for encouraging this investigation and providing advice throughout its implementation, Lael Schooler, Elisabet Rutström, Ben Newell for constructive feedback and discussions, and Yaw Nyarko, Elisabet Rutström, Andrew Schotter and Nathaniel Wilcox for kindly providing experimental data. This work has benefited from feedback by seminar participants at the Center for Adaptive Behaviour and Cognition (Max-Planck-Institute for Human Development), the Sydney Experimental Seminar series, and the ESA European Conference (Luxembourg 2011). The research was funded by a UNSW Vice Chancellor Postdoctoral Fellowship and benefited from hardware donated under the Nvidia Professor Partnership Program. Any errors are the responsibility of the author.

References

Ahn, W.-Y., J. Busemeyer, E.-J. Wagenmakers, and J. Stout (2008, December). Comparison of Decision Learning Models Using the Generalization Criterion Method. *Cognitive Science 32*(8), 1376–1402.

Anderson, J. (2007). *How can the human mind occur in the physical universe?*, Volume 3. Oxford University Press, USA.

Anderson, J. and J. Betz (2001). A Hybrid Model of Categorization. *Psychonomic Bulletin & Review 8*(4), 629.

Anderson, J., M. Matessa, and C. Lebiere (1997). ACT-R: A Theory of Higher Level Cognition and Its Relation to Visual Attention. *Human-Computer Interaction 12*, 439–462.

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist 51*(4), 355–365.

Anderson, J. R. and C. Lebiere (1998). *The atomic components of thought.* Routledge.

Aoyagi, M. (1996). Evolution of Beliefs and the Nash Equilibrium of Normal Form Games. *Journal of Economic Theory 70*(2), 444–469.

Ashby, F. G. and W. T. Maddox (1993). Relations between Prototype, Exemplar, and Decision Bound Models of Categorization. *Journal of Mathematical Psychology 37*, 372–400.

Bar-Hillel, M. and W. Wagenaar (1991). The perception of randomness. *Advances in Applied Mathematics 12*(4), 428–454.

Boeddinghaus, J., M. Ragni, M. Knauff, and B. Nebel (2006). Simulating Spatial Reasoning Using ACT-R. In *Proceedings of the Seventh International Conference on Cognitive Modeling*, pp. 62–67.

Briscoe, E. and J. Feldman (2011). Conceptual complexity and the bias/variance tradeoff. *Cognition 118*(1), 2–16.

Brown, J. and R. Rosenthal (1990). Testing the Minimax Hypothesis: A Re-Examination of O' Neill's Game Experiment. *Econometrica 58*(5), 1065–1081.

Budiu, R. (2004). Interpretation-based processing: a unified theory of semantic sentence comprehension. *Cognitive Science 28*(1), 1–44.

Busemeyer, J. and Y. Wang (2000). Model Comparisons and Model Selections Based on Generalization Criterion Methodology. *Journal of Mathematical Psychology 44*(1), 171–189.

Camerer, C. and T.-H. Ho (1999). Experience-weighted attraction learning in normal form games. *Econometrica 67*(4), 827–874.

Cheung, Y. and D. Friedman (1997). Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior 19*(1), 46–76.

Chiappori, P. A., S. Levitt, and T. Groseclose (2002). Testing mixed strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review 92*(4), 1138–1151.

Clegg, B. A., G. J. DiGirolamo, and S. W. Keele (1998). Sequence learning. *Trends in Cognitive Sciences 2*(8), 275–281.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences 24*(1), 87–114.

Dietterich, T. G. (2000). Ensemble Methods in Machine Learning. In *Lecture Notes in Computer Science: Multiple Classifier Systems Vol. 1857*, pp. 1–15. Berlin: Springer.

Donkin, C. and R. M. Nosofsky (2012). The structure of short-term memory scanning: an investigation using response time distribution models. *Psychonomic Bulletin & Review 19*(3), 363–394.

Fudenberg, D. and D. Levine (1998). Learning in games. *European Economic Review 42*, 631–639.

Gaissmaier, W. and L. J. Schooler (2008). The smart potential behind probability matching. *Cognition 109*(3), 416–22.

Gärdenfors, P. (2004). Conceptual spaces.

Gigerenzer, G. and R. Selten (2002). *Bounded Rationality: The Adaptive Toolbox*, Volume 1. The MIT Press.

Gigerenzer, G. and P. Todd (1999). *Simple heuristics that make us smart.* Oxford University Press, USA.

Glöckner, A. and T. Pachur (2012). Cognitive models of risky choice: parameter stability and predictive accuracy of prospect theory. *Cognition 123*(1), 21–32.

Gomez, R. L. (1997). Transfer and complexity in artificial grammar learning. *Cognitive Psychology 33*(2), 154–207.

Gonzalez, C. and V. Dutt (2011). Instance-based learning: integrating sampling and repeated decisions from experience. *Psychological Review 118*(4), 523–551.

Gonzalez, C. and C. Lebiere (2005). Instance-based cognitive models of decision making. *Transfer of knowledge in economic decision making. New York: Palgrave McMillan*, 1–28.

Gonzalez, C., J. Lerch, and C. Lebiere (2003). Instance-based learning in dynamic decision making. *Cognitive Science 27*(4), 591–635.

Gunzelmann, G. and J. R. Anderson (2001). An ACT-R Model of the Evolution of Strategy Use and Problem Difficulty. In E. M. Altman, A. Cleeremans, C. D. Schunn, and W. D. Gray (Eds.), *Proceedings of the Fourth International Conference on Cognitive Modeling*, pp. 109–114. Mahwah, NJ: Lawrence Erlbaum Associates.

Homa, D., S. Sterling, and L. Trepel (1981). Limitations of Exemplar-Based Generalization and the Abstraction of Categorical Information. *Journal of Experimental Psychology: Human Learning and Memory 7*(6), 418–439.

Lebiere, C. (1999). The dynamics of cognition : An ACT-R model of cognitive arithmetic. *Kognitionswissenschaft*, 5–19.

Lebiere, C., D. Wallach, and R. L. West (2000). A memory-based account of the prisoner's dilemma and other 2x2 games. *Proceedings of International Conference on Cognitive Modeling*, 185–193.

Lebiere, C. and R. L. West (1999). A dynamic ACT-R model of simple games. In *Proceedings of the Twenty-first Conference of the Cognitive Science Society*, pp. 296–301.

Levitt, S. D., J. A. List, and D. H. Reiley (2010). What Happens in the Field Stays in the Field: Exploring Whether Professionals Play Minimax in Laboratory Experiments. *Econometrica 78*(4), 1413–1434.

Love, B. C. and M. Tomlinson (2010). Mechanistic models of associative and rule-based category learning. *The Making of Human Concepts*, 53–74.

Luce, R. D. (1996). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

Marewski, J. N. and K. Melhorn (2011). Using the ACT-R architecture to specify 39 quantitative process models of decision making. *Journal of Decision Making 6*(6), 439–519.

Marewski, J. N. and H. Olsson (2009). Beyond the Null Ritual. *Zeitschrift für Psychologie 217*(1), 49–60.

Massaro, D. W. and D. Friedman (1990). Models of integration given multiple sources of information. *Psychological Review 97*(2), 225–252.

Medin, D. L., M. W. Altom, and S. M. Edelson (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition 8*(1), 37–50.

Medin, D. L. and M. M. Schaffer (1978). Context theory of classification learning. *Psychological Review 85*(3), 207–238.

Miller, G. (1956). The magical number seven, plus or minus two. *Psychological Review 63*, 81–97.

Mookherjee, D. and B. Sopher (1994). Learning Behavior in an Experimental Matching Pennies Game. *Games and Economic Behavior 7*(1), 62–91.

Nissen, M. J. and P. Bullemer (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology 19*(1), 1–32.

Nosofsky, R. and S. Zaki (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition 28*(5), 924.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition 10*(1), 104–114.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General 115*(1), 39–61.

Nosofsky, R. M. and T. J. Palmeri (1997). An exemplar-based random walk model of speeded classification. *Psychological Review 104*(2), 266–300.

Nyarko, Y. and A. Schotter (2002). An experimental study of belief learning using elicited beliefs. *Econometrica 70*(3), 971–1005.

O'Neill, B. (1987). Nonmetric test of the minimax theory of two-person zerosum games. In *Proceedings of the National Academy of Sciences*, Volume 84, pp. 2106–9.

Palacios-Huerta, I. (2003). Professionals play minimax. *Review of Economic Studies 70*, 395–415.

Palacios-Huerta, I. and O. Volij (2008). Experientia docet: Professionals play minimax in laboratory experiments. *Econometrica 76*(1), 71–115.

Posner, M. and S. Keele (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology 77*(3), 241–248.

Rapoport, A. and D. Budescu (1992). Generation of random series in two-person strictly competitive games. *Journal of Experimental Psychology: General 121*(3), 352.

Rapoport, A. and D. Budescu (1997). Randomization in individual choice behavior. *Psychological Review 104*(603-617).

Reed, S. (1972). Pattern recognition and categorization. *Cognitive Psychology 3*(3), 382–407.

Remillard, G. and J. M. Clark (2001). Implicit learning of first-, second-, and third-order transition probabilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition 27*(2), 483–498.

Rubinstein, A. (2007). Instinctive and Cognitive Reasoning: A Study of Response Times. *The Economic Journal 117*(523), 1243–1259.

Rutström, E. E. and N. T. Wilcox (2009). Stated beliefs versus inferred beliefs: A methodological inquiry and experimental test. *Games and Economic Behavior 67*(2), 616–632.

Schooler, L. and J. Anderson (1997). The role of process in the rational analysis of memory. *Cognitive Psychology 32*, 219–250.

Scroggin, S. (2007). Exploitable actions of believers in the law of small numbers in repeated constant-sum games. *Journal of Economic Theory 133*(1), 219–235.

Shachat, J. (2002). Mixed Strategy Play and the Minimax Hypothesis. *Journal of Economic Theory 104*(1), 189–226.

Shiffrin, R. M., M. D. Lee, W. Kim, and E.-J. Wagenmakers (2008). A Survey of Model Evaluation Approaches With a Tutorial on Hierarchical Bayesian Methods. *Cognitive Science 32*(8), 1248–1284.

Smith, V. (2002). Method in experiment: rhetoric and reality. *Experimental Economics 5*, 91–110.

Sonsino, D. (1997). Learning to Learn, Pattern Recognition, and Nash Equilibrium. *Games and Economic Behavior 18*(2), 286–331.

Sonsino, D. and J. Sirota (2003). Strategic pattern recognition—experimental evidence. *Games and Economic Behavior 44*(2), 390–411.

Spiliopoulos, L. (2012). Pattern recognition and subjective belief learning in a repeated constant-sum game. *Games and Economic Behavior 75*(2), 921–935.

Spiliopoulos, L. (2013). Strategic adaptation of humans playing computer algorithms in a repeated constant-sum game. *Journal of Autonomous Agents and Multi-Agent Systems 27*(1), 131–160.

Stahl, D. O. and P. W. Wilson (1995). On Players' Models of Other Players: Theory and Experimental Evidence. *Games and Economic Behavior 10*(1), 218–254.

Stevens, J. R., J. Volstorf, L. J. Schooler, and J. Rieskamp (2011). Forgetting Constrains the Emergence of Cooperative Decision Strategies. *Frontiers in Psychology 1*, 1–12.

Walker, M. and J. Wooders (2001). Minimax play at Wimbledon. *The American Economic Review 91*(5), 1521–1538.

West, R. and C. Lebiere (2001). Simple games as dynamic, coupled systems: randomness and other emergent properties. *Journal of Cognitive Systems Research 1*(4), 221–239.
Wilcox, N. T. (2011). Stochastically more risk averse: A contextual theory of stochastic discrete choice under risk. *Journal of Econometrics 162*(1), 89–104.
Wolford, G., S. Newman, and M. Miller (2004). Searching for patterns in random sequences. *Canadian Journal of Experimental Psychology 58*(4).
Wong, T., E. Cokely, and L. J. Schooler (2010). An Online Database of ACT-R Parameters : Towards a Transparent Community-based Approach to Model Development. In *Proceedings of ICCM*, Volume 1, pp. 282–286.
Wooders, J. (2010). Does Experience Teach? Professionals and Minimax Play in the Lab. *Econometrica 78*(3), 1143–1154.
Yechiam, E. and J. R. Busemeyer (2008). Evaluating generalizability and parameter consistency in learning models. *Games and Economic Behavior 63*, 370–394.
Zeisberger, S., D. Vrecko, and T. Langer (2010). Measuring the time stability of Prospect Theory preferences. *Theory and Decision 72*(3), 359–386.

Appendix A. A numerical example of the prototype approach

This section provides a numerical example of the prototype updating procedure for $n = 3$. Table A.8 presents a realization of the history of play over time (or rounds) t, and the subsequent evolution of the prototypes for two chunks, corresponding to each of the two possible action choices.

Firstly, note that time starts at $t = 4$, since for $n = 3$ this is the smallest t that will allow a context to be fully observed (since for $n = 3$ it is necessary to observe two prior lags). After the observation of the outcomes at $t = 4$, a new chunk is created for c_1 since $a'_4 = 0$. The prototype of this chunk, c_2, is equal to the observed history of play at this point in time.

At $t = 5$ a new chunk c_1 is created for $V_j = 1$ since this is the first time that $a'_t = 1$ has been observed—again the prototype is simply equal to the history of play at that point. Note, that the prototype of chunk c_2 does not change with respect to $t = 4$ since $a'_5 = 1$, i.e. the subject has not observed any new information regarding $a'_t = 0$. The maximum number of chunks–equal to two since this is the size of the action-space–has now been achieved, therefore from now on chunks will simply be updated.

At $t = 6$, chunk c_2 will be updated since $a'_6 = 0$ and chunk c_1 will remain unchanged with respect to $t = 5$. The prototypical context of c_2 is updated in the following fashion by averaging over the historical contexts in the rounds when $a'_t = 0$—in this case at $t = 4$ and $t = 6$. The first element of the context $\bar{\boldsymbol{\omega}}_{\mathbf{2}}\,(t)$ is the average of the values of a_{t-1} for times $t = 4, 6$, i.e. is equal to ${}^{(0+1)}/_{2} = 0.5$. Likewise, for the second, third and fourth elements, the respective calculations are: ${}^{(1+0)}/_{2} = 0.5$, ${}^{(1+1)}/_{2} = 1$, ${}^{(0+0)}/_{2} = 0$.

At time $t = 7$, chunk c_1 is updated in a similar function using the observations at $t = 5, 7$ where $a'_t = 1$.

At time $t = 8$, chunk c_1 will again be updated as $a'_8 = 1$ and chunk c_2 will remain unchanged. The calculations for the prototypical context are now the averaged historical contexts at times $t = 5, 7, 8$. The elements of chunk c_1's context are calculated in order as follows: ${}^{(0+0+1)}/_{3} = 0.33$, ${}^{(0+1+0)}/_{3} = 0.33$, ${}^{(0+0+1)}/_{3} = 0.33$, ${}^{(1+1+0)}/_{3} = 0.66$.

Table A.8: Example of prototype creation and updating

Time/Round	History of play					Chunk $c_1, V_j = 1$				Chunk $c_2, V_j = 0$			
	a'_t	a_{t-1}	a_{t-2}	a'_{t-1}	a'_{t-2}	$\bar{\boldsymbol{\omega}}_{\mathbf{1}}(t)$				$\bar{\boldsymbol{\omega}}_{\mathbf{2}}(t)$			
4	0	0	1	1	0	-	-	-	-	0	1	1	0
5	1	0	0	0	1	0	0	0	1	0	1	1	0
6	0	1	0	1	0	0	0	0	1	0.5	0.5	1	0
7	1	0	1	0	1	0	0.5	0	1	0.5	0.5	1	0
8	1	1	0	1	0	0.33	0.33	0.33	0.66	0.5	0.5	1	0

From the state of the prototypes at $t = 8$ it is beginning to become clear how the contexts of the two prototypes are beginning to diverge. For example, if $a'_{t-2} = 1$ this is stronger evidence of an observation belonging to chunk c_1 than c_2 as the corresponding element in $\bar{\boldsymbol{\omega}}_{\mathbf{1}}(t)$ is 0.66 and in $\bar{\boldsymbol{\omega}}_{\mathbf{2}}(t)$ is 0—this would increase the predicted likelihood of $a'_t = 1$ since this is the value V_1 of chunk c_1.

Appendix B. Econometric details

Appendix B.1. Within-dataset CV

Each dataset is divided into folds consisting of five temporally sequential rounds: ten folds for NS composed of rounds $\{(11, 12, 13, 14, 15), \ldots, (56, 57, 58, 59, 60)\}$, and five folds for RW composed of rounds $\{(12, 13, 14, 15, 16), \ldots, (32, 33, 34, 35, 36)\}$. The model is estimated by withholding one fold from a single dataset as the cross-validation set T_{cv}, and estimating the model on the remaining $k-1$ folds, which form T_{est}—this estimation procedure is repeated until all k folds have served once as the cross-validation set.

Let the set of estimation and cross-validation sets for a dataset y that can be created from all possible combinations of folds be denoted by $\left(T^k_{est}, T^k_{cv}\right) \in \mathbb{K} : k = 1, ..., K_y$. Let $sb_{i,t}$ denote player i's stated belief at time t and $\hat{b}_{i,t}(\theta_{i,k})$ denote a model's empirically estimated beliefs as a function of the individual parameter set $\theta_{i,k}$ per individual i and estimation dataset T^k_{est}.

Choose a dataset y, NS or RW, and perform the following procedures:

1. Repeat the following procedure for each subject, i, in the chosen dataset y.
   (a) Repeat the following minimization for each estimation set T^k_{est} where $k = 1, ..., K_y$:

$$\min_{\theta_{i,k}=\theta^\star_{i,k}} \sqrt{\mid T^k_{est} \mid^{-1} \sum_{t \in T^k_{est}} \left(sb_{i,t} - \hat{b}_{i,t}(\theta_{i,k})\right)^2}$$

   (b) Calculate the cross-validation performance for each T^k_{cv} of a subject i using the optimal parameters $\theta^\star_{i,k}$ estimated from the corresponding estimation set T^k_{est}:

$$wCV^k_{i,y}[m] = \mid T^k_{cv} \mid^{-1} \sum_{t \in T^k_{cv}} \left(sb_{i,t} - \hat{b}_{i,t}\left(\theta^\star_{i,k}\right)\right)^2$$

2. Finally, calculate the cross-validation criterion $wCV_y[m]$ for dataset y by summing over all K_y cross-validation sets and individuals i:

$$wCV_y[m] = \sqrt{\frac{1}{I_y}\sum_i \frac{1}{K_y}\sum_k wCV_{i,y}^k[m]}$$

Appendix B.2. Between-dataset CV

Let $sb_{i,t}$ denote player i's stated belief at time t and $\hat{b}_{i,t}(\theta_i)$ denote a model's empirically estimated beliefs as a function of the individual parameter set θ_i. Furthermore, let the subjects in the estimation dataset x be indexed $i = 1, ..., I_x$ and for the cross-validation dataset y $i' = 1, ..., I_y$. The estimation procedure is performed per subject i in the dataset x, whilst the cross-validation procedure is performed on the dataset y.

Choose an estimation dataset x and cross-validation dataset y and perform the following procedures.

1. Repeat the following estimation procedure for each subject i in dataset T_x:

$$\min_{\theta_i=\theta_i^\star} \sqrt{|\, T_x \,|^{-1} \sum_{t\in T_x} \left(sb_{i,t} - \hat{b}_{i,t}(\theta_i)\right)^2}$$

2. Perform the following procedure for each subject i' in the cross-validation dataset T_y:
   (a) Calculate the errors for each subject i' for each set of optimal parameters $\theta_i^\star$ where $i = 1, ..., I_x$

$$|\, T_y \,|^{-1} \sum_{t\in T_y} \left(sb_{i',t} - \hat{b}_{i',t}(\theta_i^\star)\right)^2$$

3. Calculate the cross-validation criterion by averaging the errors calculated in the previous step over all i' and over each i—this incorporates uncertainty over the type/parameter heterogeneity of a player i' by giving equal weighting to each of the parameter sets estimated for subjects i in the estimation dataset.

$${}_xbCV_y[m] = \sqrt{\frac{1}{I_y}\sum_{i'} \frac{1}{I_x}\sum_i |\, T_y \,|^{-1} \sum_{t\in T_y} \left(sb_{i',t} - \hat{b}_{i',t}(\theta_i^\star)\right)^2}$$

Appendix B.3. Generalized-between-dataset CV

Let $\bar{\theta}$ be the set of parameter values derived from the ACT-R database, and therefore not estimated using any data from NS and RW. As this is a single set of parameters without any individual heterogeneity, the cross-validation criterion for a dataset y is:

$$gCV_y[m] = \sqrt{\frac{1}{I_y}\sum_{i'} |\, T_y \,|^{-1} \sum_{t\in T_y} \left(sb_{i',t} - \hat{b}_{i',t}(\bar{\theta})\right)^2}$$