STUDYING COGNITION VIA INFORMATION SEARCH IN TWO-PERSON GUESSING GAME EXPERIMENTS

Miguel Costa-Gomes, University of York, and
Vincent Crawford, University of California, San Diego

Berkeley Psychology and Economics Seminar, 6 March 2007

Barcelona JOCS Seminar, 26 March 2007

Context and Motivation

The paper, which is still in progress, analyzes the information search data from the experiments whose decisions are reported in:

Costa-Gomes and Crawford, "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review* 2006 ("CGC"),

(CGC also use the search data in some of their econometric estimates)

A literature review and discussion of modeling issues are in:

Crawford, "Look-ups as the Windows of the Strategic Soul: Studying Cognition via Information Search in Game Experiments," to appear in Andrew Caplin and Andrew Schotter, editors, *Perspectives on the Future of Economics: Positive and Normative Foundations*, Oxford University Press, 2008

These papers are part of a larger experimental project studying strategic thinking—how people model other people's decisions in games

First goal is a structural alternative to equilibrium as a model of initial responses

Such a model should explain why equilibrium often predicts initial responses well in simple games, establishing the robustness of some equilibrium conclusions

It should also predict systematic deviations from equilibrium in complex games

Needed for comparative statics and mechanism design, which involve "new" games; and for applications that now rely on equilibrium in complex games without clear precedents

Second goal is better models of learning, distinguishing reinforcement from beliefs-based and more sophisticated learning rules and giving insight into the kinds of imperfect analogies people can learn from, and how

Needed for better predictions regarding selection among multiple equilibria

Experiments that study strategic thinking

CGC joins two strands of experimental papers that study strategic thinking

Those I will focus on randomly and anonymously paired subjects to play series of different but related games, with no feedback

Suppressing learning from experience and repeated-game effects allows the designs to elicit subjects' initial responses, game by game

This allows subjects to focus on predicting others' responses, "uncontaminated" by adaptive learning (which can make even amoebas converge to equilibrium)

("Eureka!" learning remains possible, but can be tested for and seems rare)

The first strand includes papers that elicit decisions in games alone, as in:

Stahl and Wilson, "Experimental Evidence on Players' Models of Other Players" *Journal of Economic Behavior and Organization* 1994 ("SW")

Stahl and Wilson, "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior* 1995

on matrix games, and

Nagel, "Unraveling in Guessing Games: An Experimental Study," *American Economic Review* 1995

Ho, Camerer, and Weigelt, "Iterated Dominance and Iterated Best Response in Experimental 'p-Beauty Contests'," *American Economic Review* 1998 ("HCW")

on *n*-person guessing games

The second strand includes papers that monitor both decisions in games and subjects' searches for hidden but freely accessible information about payoffs

(Monitoring originated in MouseLab studies of individual decisions as in Payne, Bettman, and Johnson, *The Adaptive Decision Maker*, 1993)

Camerer, Johnson, Rymon, and Sen, "Cognition and Framing in Sequential Bargaining for Gains and Losses," in Kenneth Binmore, Alan Kirman, and Piero Tani, editors, *Frontiers of Game Theory*, 1993 ("CJ")

Johnson, Camerer, Sen, and Rymon, "Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining," *Journal of Economic Theory* 2002 ("CJ")

(Subjects played two-person, three-round alternating-offers bargaining games, in which the "pie" varies across rounds to simulate discounting at a common rate; each game was presented to subjects as a series of searchable pies; goal is to test backward-induction and social-preferences explanations of behavior)

Camerer and Johnson, “Thinking about Attention in Games: Backward and Forward Induction,” in Isabel Brocas and Juan Carrillo (editors), *The Psychology of Economic Decisions, Volume Two: Reasons and Choices*, Oxford, 2004

(Subjects played simple extensive-form games with independently searchable individual payoffs; the goal is to test forward induction (also exposits CJ papers))

Costa-Gomes, Crawford, and Broseta, “Cognition and Behavior in Normal-Form Games: An Experimental Study,” *Econometrica* 2001 (“CGCB”)

(Subjects played series of two-person matrix games with independently searchable individual payoffs; games had various patterns of iterated dominance or unique pure-strategy equilibria without dominance; goal was to test equilibrium and iterated dominance against alternatives such as SW’s “level-k” types)

Most of these experiments give support to a particular structural non-equilibrium model of initial responses based on level-*k* types, described below

Level-*k* models have also been used to resolve empirical puzzles

Applications of level-*k* Models

Crawford, “Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions,” *American Economic Review* 2003 (D-day as data…)

Camerer, Ho, and Chong, “A Cognitive Hierarchy Model of Games,” *Quarterly Journal of Economics* 2004 (see sections on entry games and zero-sum betting)

Crawford and Iriberri, “Fatal Attraction: Salience, Naivete, and Sophistication in Experimental ‘Hide-and-Seek’ Games,” *American Economic Review* in press (explains why “Any government wanting to kill an opponent…would not try it at a meeting with government officials” is a fallacy only if you’re a game theorist)

Crawford and Iriberri, “Level-*k* Auctions: Can a Non-Equilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?,” manuscript 2006

Crawford, Gneezy, and Rottenstreich, “The Power of Focal Points is Limited: Even Minute Payoff Asymmetry May Yield Large Coordination Failures,” 2007

Outline

CGC's experimental design

Level-*k* model

Types

CGC's results for decisions

Puzzles left unresolved by CGC's analysis of decisions

CGC's analysis of information search

Search data for representative R/TS and Baseline subjects

How does cognition show up in search?

Types as models of cognition and search

Derivation of types' search implications

CGC's econometric analysis of guesses and search

Possible answers to the unresolved puzzles

CGC’s experimental design

CGC’s experiments elicit subjects’ initial responses to a series of 16 dominance-solvable two-person guessing games, relatives of Nagel’s and HCW’s

In each game, each player has his own lower and upper limit, both strictly positive; but players are not required to guess between their limits

Guesses outside the limits are automatically adjusted up to the lower or down to the upper limit as necessary (a trick to enhance separation of types via search).

Each player also has his own target, and his payoff increases with the closeness of his adjusted guess to his target times the other’s adjusted guess

The targets and limits vary independently across players and 16 games, with the targets either both less than one, both greater than one, or mixed

(In previous guessing experiments, the targets and limits were always the same for both players, and varied either only across treatments or not at all)

The 16 games subjects played are finitely dominance-solvable in 3-52 rounds, with essentially (because the only thing about a guess that matters is its adjusted guess) unique equilibria determined by the targets and limits in a simple way

TABLE 3—STRATEGIC STRUCTURES OF THE GAMES

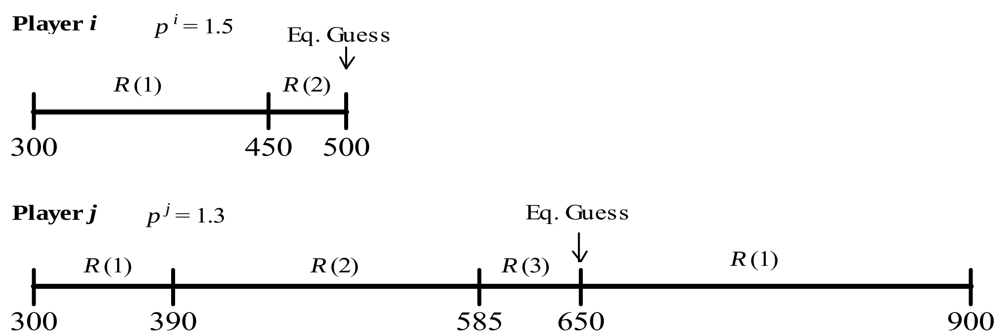
Game ij	Order played	Targets	Equilibrium	Rounds of dominance	Pattern of dominance	Dominance at both ends
1. $\alpha 2\beta 1$	6	Low	Low	4	A	No
2. $\beta 1\alpha 2$	15	Low	Low	3	A	No
3. $\beta 1\gamma 2$	14	Low	Low	3	A	Yes
4. $\gamma 2\beta 1$	10	Low	Low	2	A	No
5. $\gamma 4\delta 3$	9	High	High	2	S	No
6. $\delta 3\gamma 4$	2	High	High	3	S	Yes
7. $\delta 3\delta 3$	12	High	High	5	S	No
8. $\delta 3\delta 3$	3	High	High	5	S	No
9. $\beta 1\alpha 4$	16	Mixed	Low	9	S/A	No
10. $\alpha 4\beta 1$	11	Mixed	Low	10	S/A	No
11. $\delta 2\beta 3$	4	Mixed	Low	17	S/A	No
12. $\beta 3\delta 2$	13	Mixed	Low	18	S/A	No
13. $\gamma 2\beta 4$	8	Mixed	High	22	A	No
14. $\beta 4\gamma 2$	1	Mixed	High	23	A	Yes
15. $\alpha 2\alpha 4$	7	Mixed	High	52	S/A	No
16. $\alpha 4\alpha 2$	5	Mixed	High	51	S/A	No

Notes: Game identifiers: limits α for 100 and 500, β for 100 and 900, γ for 300 and 500, or δ for 300 and 900; targets 1 for 0.5, 2 for 0.7, 3 for 1.3, 4 for 1.5. Low targets are <1; high targets are >1; mixed targets are one <1, one >1. High equilibrium is determined by players' upper limits; low equilibrium is determined by players' lower limits. Rounds of dominance refers to the number player i needs to identify his equilibrium guess. Alternating dominance (A) occurs first for one player, then the other, then the first, etc.; simultaneous dominance (S) occurs for both players at once; and simultaneous then alternating dominance (S/A) is simultaneous in the first round and then alternating. Dominance at both ends refers to whether guesses are eliminated near both of a player's limits.

For example, in game γ4δ3 (#5 in Table 3), player *i*'s limits and target are [300, 500] and 1.5 and player *j*'s are [300, 900] and 1.3

The product of targets 1.5 × 1.3 > 1, so players' equilibrium adjusted guesses are determined (at least indirectly) by their upper limits; *i*'s equilibrium adjusted guess equals his upper limit of 500, but *j*'s is below his upper limit at 650

The way in which equilibrium is determined here is general in CGC's games



(Guesses in *R*(*k*) are eliminated in round *k* of iterated dominance.)

CGC's design exploits the discontinuity of the equilibrium correspondence by including some games that differ mainly in whether the product is slightly greater, or slightly less, than 1; equilibrium responds strongly to such differences, but empirically plausible non-equilibrium decision rules are largely unmoved by them

The way in which equilibrium is jointly determined by both players' parameters also helps to separate the search implications of equilibrium and other rules

Level-*k* model

CGC's analysis of decisions follows SW's, Nagel's, and CGCB's in using a structural non-equilibrium model of initial responses in which each subject's decisions are determined by one of several decision rules or *types*

Types will play a central role in linking decisions and search: CGC's analysis, like CGCB's, takes a procedural view of decision-making in which a subject's type determines his search and his type and search then determine his decision

CGC's types assume risk-neutrality, with no social preferences; plausible, given the design, and mostly justified by data

The empirical strategy was to throw in any type from previous work that had a chance of being important, then to test carefully for omitted types and overfitting

Types

L0 is an "anchoring type," here uniform random over the feasible decisions

(*L0* usually has 0 estimated frequency (or is confounded with the error structure), but its role is crucial, representing *L1*'s beliefs, *L2*'s beliefs about *L1*'s beliefs)

L1 best responds to *L0*

(*L1* understands the rules, but makes no attempt to model others' decisions)

L2 (*L3*) best responds to *L1* (*L2*), and so on

(*Lk* for $k > 0$ is rational, and *k*-level rationalizable and so coincides with equilibrium in games that are *k*-dominance solvable; but uses a simplified model of others that may make it deviate from equilibrium)

(Previous analyses have considered alternative definitions: e.g. SW's *L2* best responds to a noisy *L1*; and CHC's *L2* best responds to an estimated mixture of *L1* and *L0*; CCG discuss the evidence)

Equilibrium makes its equilibrium decision

D1 (*D2*) does one round (two rounds) of deletion of dominated decisions and then best responds to a uniform prior over the other's remaining decisions

(By a quirk of our notation, *L2* (not *L1*) is *D1*'s cousin, and *L3* is *D2*'s; theorists tend instinctively to identify *Lk* with *Dk*-1, etc., but they are cognitively very different: *Lk* starts with a naïve prior over the other's decisions and iterates the best-response mapping, while *Dk*-1 starts with reasoning based on iterated knowledge of rationality and closes the process with a naïve prior at the end)

(*Lk* and *Dk*-1 decisions are perfectly confounded in Nagel's games and weakly separated in other previous experiments, but CGC strongly separate them and show that empirically, *Lk* is common (for *k* = 1, 2, maybe 3), *Dk*-1 doesn't exist)

Sophisticated best responds to the probabilities of other's decisions, as estimated in CGC from subjects' observed choice frequencies; included to test for subjects whose understanding goes beyond mechanical rules

(CGCB and CGC show that *Sophisticated* doesn't exist (in the lab))

CGC's results for decisions

The large strategy spaces and independent variation of targets and limits in CGC's design enhance separation of types' implications for decisions, to the point where many subjects' types can be precisely identified from guesses alone

Game	*ai*	*bi*	*pi*	*aj*	*bj*	*pj*	*L1*	*L2*	*L3*	*D1*	*D2*	*E*	*S*
1	100	900	1.5	300	500	0.7	600	525	630	600	611.25	750	630
2	300	900	1.3	300	500	1.5	520	650	650	617.5	650	650	650
3	300	900	1.3	300	900	1.3	780	900	900	838.5	900	900	900
4	300	900	0.7	100	900	1.3	350	546	318.5	451.5	423.15	300	420
5	100	500	1.5	100	500	0.7	450	315	472.5	337.5	341.25	500	375
6	100	500	0.7	100	900	0.5	350	105	122.5	122.5	122.5	100	122
7	100	500	0.7	100	500	1.5	210	315	220.5	227.5	227.5	350	262
8	300	500	0.7	100	900	1.5	350	420	367.5	420	420	500	420
9	300	500	1.5	300	900	1.3	500	500	500	500	500	500	500
10	300	500	0.7	100	900	0.5	350	300	300	300	300	300	300
11	100	500	1.5	100	900	0.5	500	225	375	262.5	262.5	150	300
12	300	900	1.3	300	900	1.3	780	900	900	838.5	900	900	900
13	100	900	1.3	300	900	0.7	780	455	709.8	604.5	604.5	390	695
14	100	900	0.5	300	500	0.7	200	175	150	200	150	150	162
15	100	900	0.5	100	500	0.7	150	175	100	150	100	100	132
16	100	900	0.5	100	500	1.5	150	250	112.5	162.5	131.25	100	187

Of 88 subjects, 43 made guesses that complied *exactly* (within 0.5) with one type's guesses in 7-16 of the games (20 *L1*, 12 *L2*, 3 *L3*, and 8 *Equilibrium*)

E.g. CGC's Figure 2 shows the "fingerprints" of the 12 subjects whose apparent types were *L2*. Of their 192 guesses, 138 (72%) were exact

Given how strongly CGC's design separates types' guesses, and that guesses could take 200-800 different rounded values, these subjects' exact compliance rates are far higher than could occur by chance

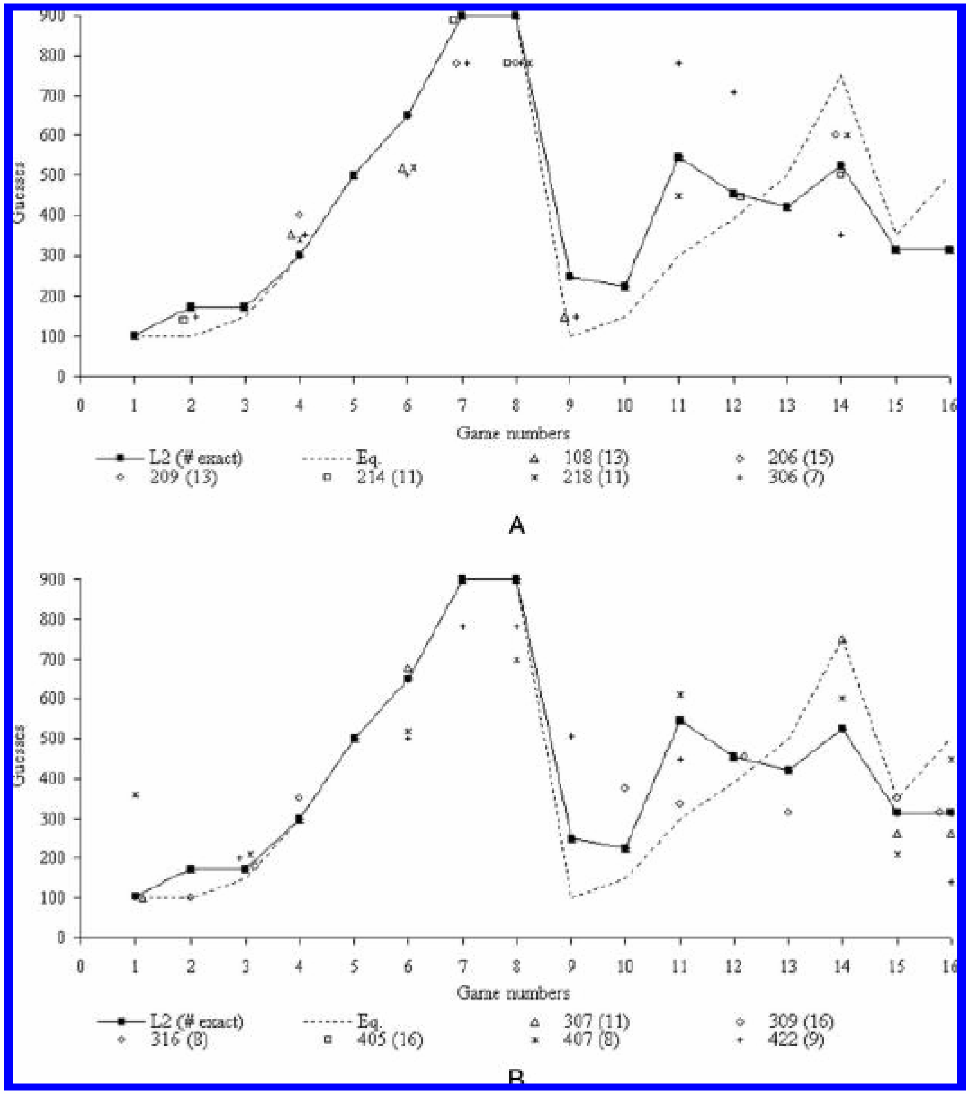


FIGURE 2. "FINGERPRINTS" OF 12 APPARENT *L2* SUBJECTS

Notes: Only deviations from *L2*'s guesses are shown. Of these subjects' 192 guesses, 138 (72 percent) were exact *L2* guesses.

Further, because the types specify precise, well-separated guess sequences in a very large space of possibilities, their high exact compliance rules out (intuitively or econometrically) alternative interpretations of their behavior

In particular, because the types build in risk-neutral, self-interested rationality and perfect models of the game, the deviations from equilibrium of the 35 subjects whose apparent types are *L1*, *L2*, or *L3* can be attributed to non-equilibrium beliefs rather than irrationality, risk aversion, altruism, spite, or confusion

(In SW's or CGCB's matrix games, even a perfect fit does not distinguish a subject's best-fitting type from nearby omitted types; and in Nagel's and HCW's guessing games, with each subject playing one game, the ambiguity is worse)

CGC's other 45 subjects' types are less apparent from their guesses; but econometric estimates and specification analysis still turns up only *L1*, *L2*, *L3*, and *Equilibrium* in significant numbers

Dk and other types, including *Sophisticated*, which is clearly separated from *Equilbrium* here, don't exist, at least in these games

Puzzles left unresolved by CGC's analysis of decisions

To go further, it is necessary to distinguish CGC's three treatments

In the Baseline treatment, subjects played the games with other subjects, looking up the targets and limits through a MouseLab interface as explained below

The Open Boxes ("OB") treatment was identical to the Baseline, except subjects did not need to look up their targets and limits; they were continually displayed

The analysis of decisions described so far pools the decision data from the Baseline and OB treatments, which did not differ significantly

CGC also ran six different Robot/Trained Subjects ("R/TS") treatments, one each for types *L1*, *L2*, *L3*, *D1*, *D2*, and *Equilibrium*

R/TS treatments were identical to the Baseline, except subjects played against a "robot" (described as "the computer") and the computer played according to a pre-specified, announced type; subjects were trained to identify the computer's type's guesses; and subjects were paid for their payoffs against the computer

Puzzle A. What are Those Baseline "*Equilibrium*" Subjects Really Doing?

Consider the eight Baseline subjects with near-*Equilibrium* fingerprints

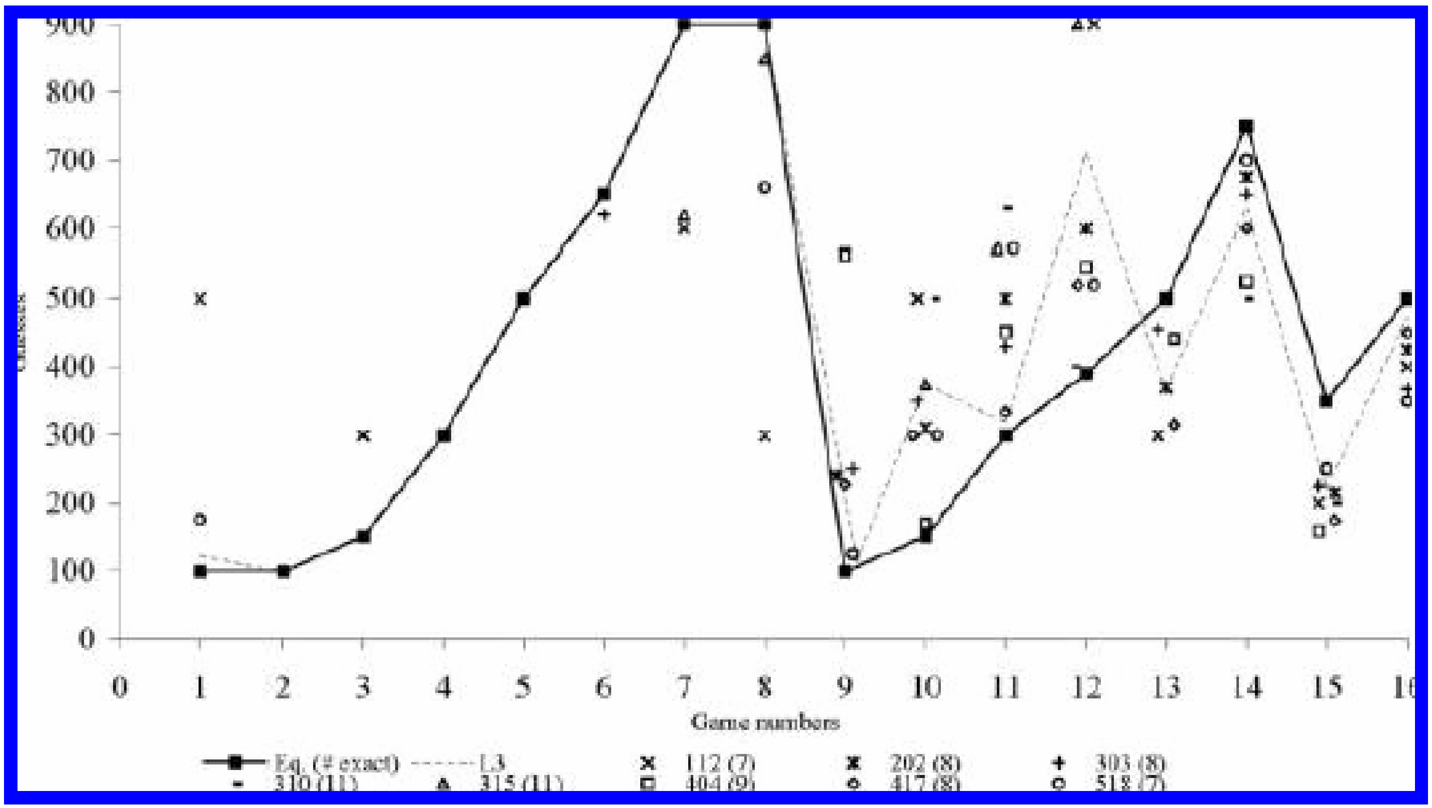


FIGURE 4. "FINGERPRINTS" OF EIGHT APPARENT *EQUILIBRIUM* SUBJECTS

Notes: Only deviations from *Equilibrium*'s guesses are shown. Of these subjects' 128 guesses, 69 (54 percent) were exact *Equilibrium* guesses.

Ordering the games by strategic structure as in Figure 4, with the 8 games with mixed targets (Table 3) on the right, shows that these subjects' deviations from equilibrium almost always occur with mixed targets

Thus these subjects, whose exact compliance with *Equilibrium* guesses is off the scale by any normal standard, are actually following a rule that only mimics *Equilibrium*, and that only in games without mixed targets

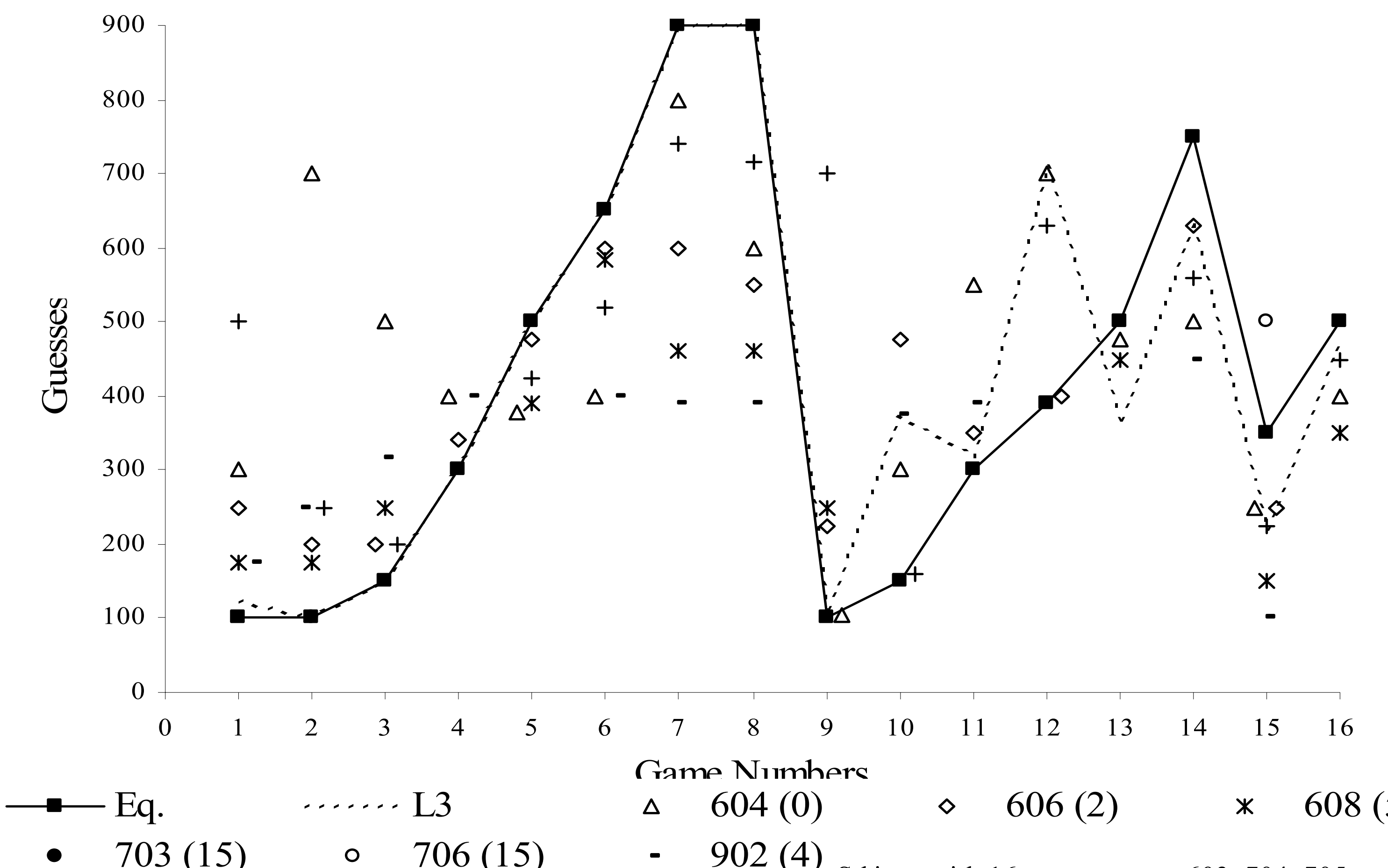
Yet all the ways we teach people to identify equilibria (best-response dynamics, equilibrium checking, iterated dominance) work just as well with mixed targets

Whatever these subjects are doing, it's something we haven't thought of yet

And whatever it is, it has a structure: *All* 44 of these subjects' deviations from *Equilibrium* (solid line) when it is separated from *L3* (dotted line) are in the direction of (and sometimes beyond) *L3* guesses; but this could reflect no more than the fact that *Equilibrium* guesses are more extreme than other types'

Equilibrium R/TS subjects' compliance is equally high with and without mixed targets, so training eliminates whatever the Baseline subjects are doing

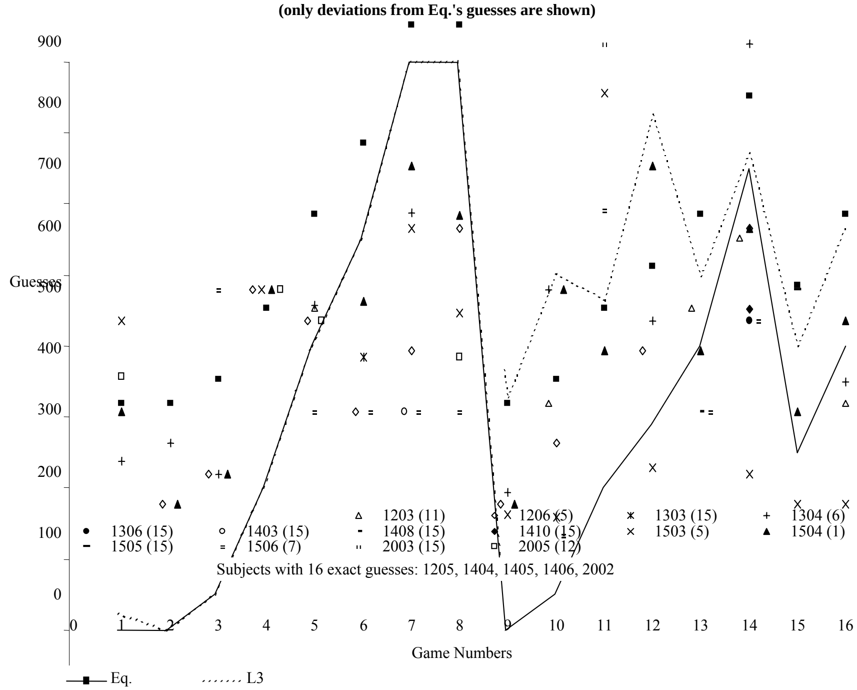
Fingerprints of 10 UCSD Equilibrium R/TS Subjects
(only deviations from Eq.'s guesses are shown)



Subjects with 16 exact guesses: 603, 704, 705

Fingerprints of 18 York Equilibrium R/TS Subjects

(only deviations from Eq.'s guesses are shown)



Puzzle B. Why are *Lk* the only non-*Equilibrium* types that exist?

A careful analysis of CGC's decision data, including specification tests not described here, reveals many subjects of types *L1*, *L2*, *Equilibrium*, or hybrids of *L3* and/or *Equilibrium*, but no other types that do better than a completely random model of guesses for more than one of their 88 Baseline and OB subjects

Why, out of the enormous number of possibilities do these rules predominate?

Why, for instance, don't we get *Dk* rules, which are closer to what we teach?

Answering this question may be the key to a deeper theory of bounded rationality

I suggest possible ways to resolve both puzzles after discussing search analysis

CGC's analysis of information search

In CGC's design, within a publicly announced structure, each game was presented via MouseLab, which normally concealed the targets and limits but allowed subjects to look them up as often as desired, one at a time

(Click option, versus rollover option used by CJ; opening and closing boxes conscious decisions, though they are quickly subordinated to their purposes)

With search costs as low as subjects' searches make them seem, free access made the entire structure effectively public knowledge, so the results can be used to test theories of behavior in complete-information versions of the games

These designs also maintain tight control over subjects' motives for search by making information from previous plays completely irrelevant to current payoffs

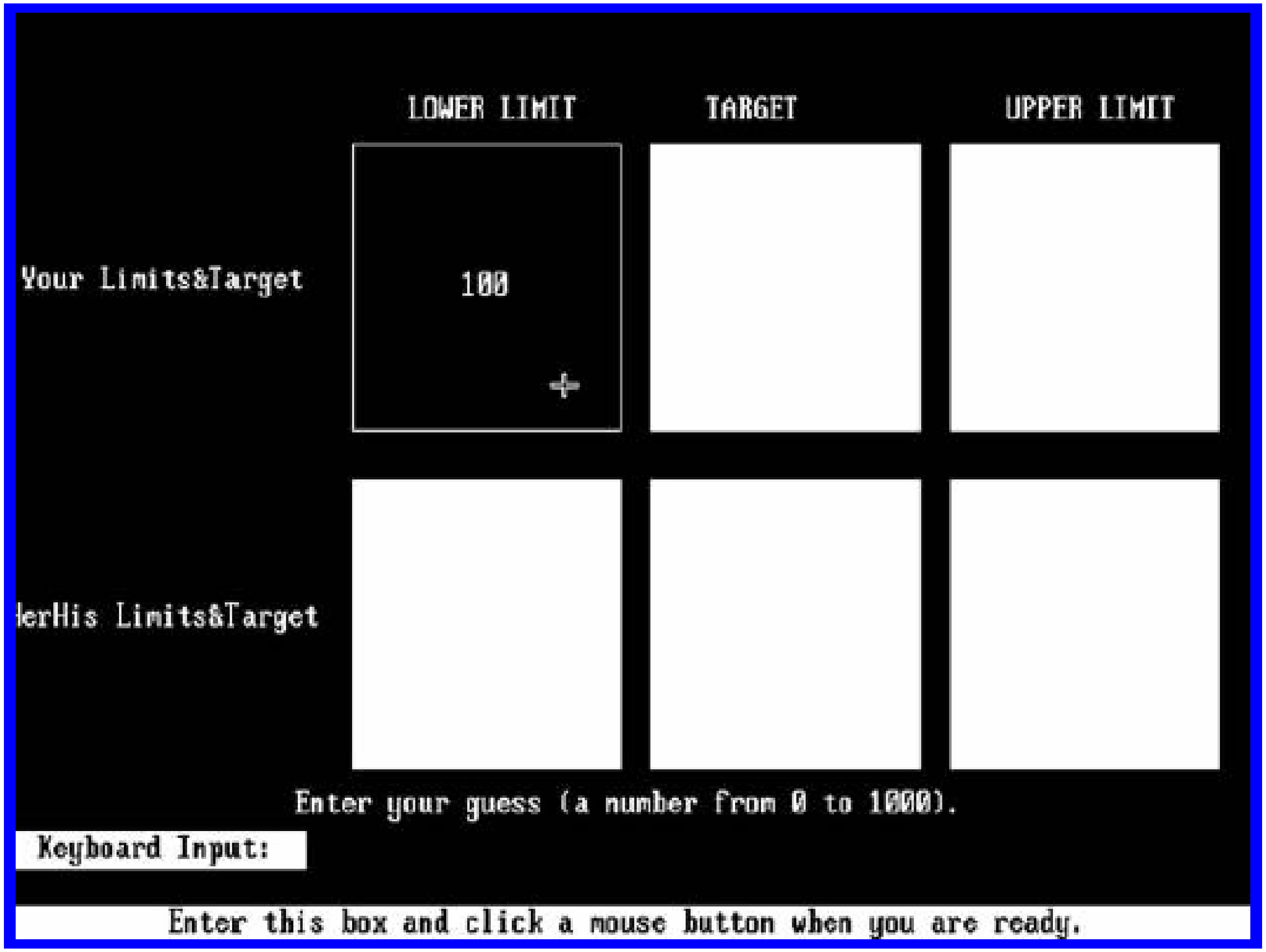


FIGURE 6. SCREEN SHOT OF THE MOUSELAB DISPLAY

Design desiderata for studying cognition via search

CGC's design combines the strengths of CJ's and CGCB's designs

CJ's design allows subjects to search for a small number of hidden parameters (pies) within a simple, publicly announced structure, but makes their search patterns essentially one-dimensional, and so less informative than they could be

CGCB's design makes search roughly three-dimensional (up-down in own payoffs, left-right in other's payoffs, transitions from own to other's payoffs) and independently separates the implications of leading types for search and decisions

But search is complex (8-16 payoffs in games with no common structure beyond being matrix games)

CGC's design has a simple parametric structure like CJ's, but makes subjects' search patterns high-dimensional, with leading types' search implications (almost) independent of the game

Search data for representative R/TS and Baseline subjects

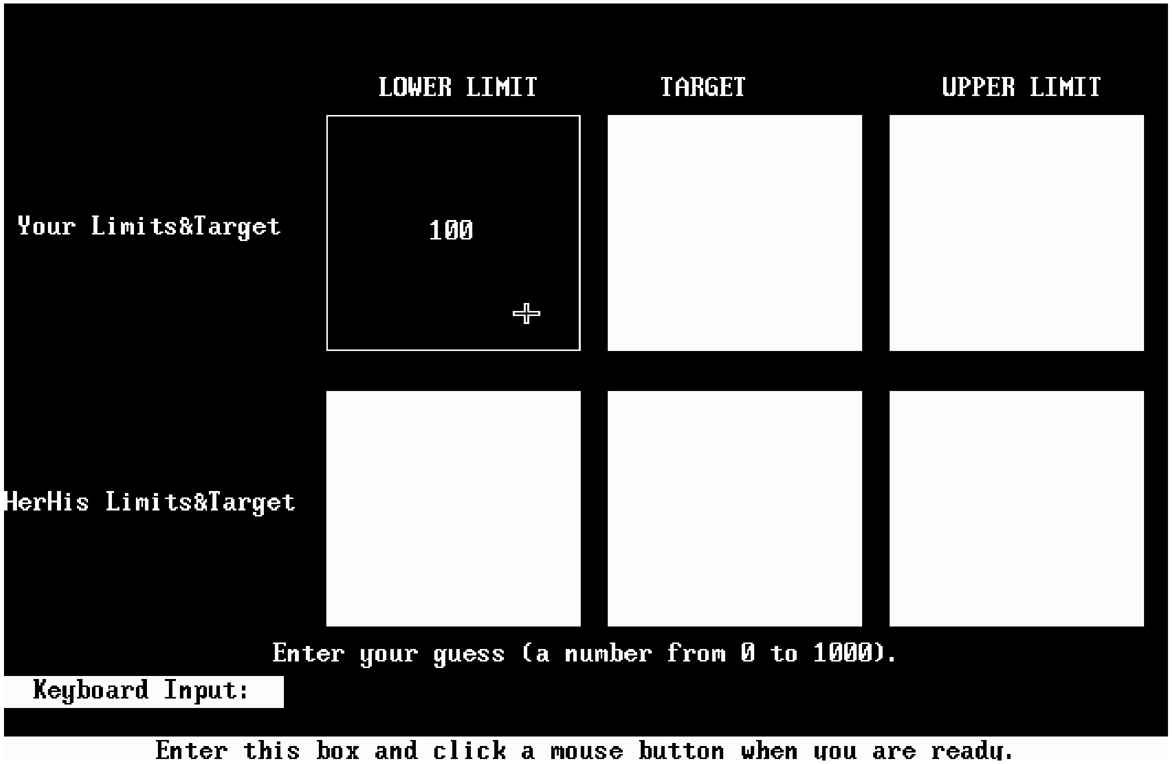
Start by comparing the search data for representative R/TS and Baseline subjects whose guesses conform closely to their assigned or estimated type with the implications of CGC's theory of cognition and search

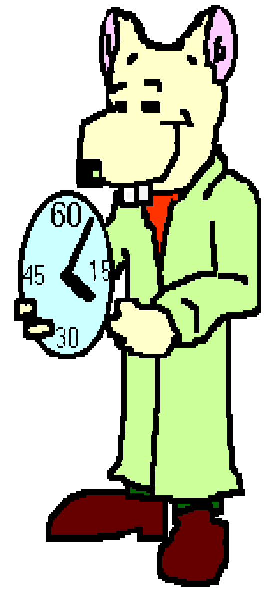
(CGC's theory is close to CGCB's, and was therefore almost completely specified before these data were generated)

Eyeballing compliance with types' search implications will suggest that there some usable structure in the data, and then we can figure out how to model it

But first….

SPEAK RODENT LIKE A NATIVE IN ONE EASY LESSON!





	a	***p***	***B***
You (*i*)	1	2	3
S/he (*j*)	4	5	6

MouseLab Box Numbers

Selected R/TS Subjects' Information Searches and Assigned Types' Search Implications

MouseLab box numbers

	a	*p*	*b*
You (*i*)	1	2	3
S/he (*j*)	4	5	6

Types' Search Implications

L1	{[4,6],2}
L2	{([1,3],5),4,6,2}
L3	{([4,6],2),1,3,5}
D1	{(4,[5,1], (6,[5,3]),2}
D2	{(1,[2,4]),(3,[2,6]),(4,[5,1],(6,[5,3]),5,2}
Eq	{[2,5],4} if pr. tar.<1,{[2,5],6} if > 1

Subject	**904**	**1716**	**1807**	**1607**	**1811**	**2008**	**1001**	**1412**	**805**	**1601**	**804**	**1110**	**1202**	**704**	**1205**	**1408**	**2002**
Type(#rt.)	L1 (16)	L1 (16)	L1 (16)	L2 (16)	L2 (16)	L2 (16)	L3 (16)	L3 (16)	D1 (16)	D1 (16)	D1 (3)	D2 (14)	D2 (15)	Eq (16)	Eq (16)	Eq (15)	Eq (16)
Alt.(#rt.)											L2 (16)						
Est. style	late	often	early	often	early				early								
Game																	
1	123456 4623	146462 134646 23	462513	135462 1313	134446 5213*4 6	111313 131313 5423	462135 21364* 246231 52	146231 564623 1	154356 423213 2642	254514 36231	154346 5213	135464 2646*1 313	246466 135464 641321 342462 422646 124625 5*1224 654646	123456 363256 565365 626365 652651 452262 6526	123456 424652 562525 6352*4 65	123123 456445 632132 11	142536 125365 253616 361454 613451 213452 63
2	123456 4231	462462 13	462132 25	135461 354621 3	134653 125642 313562 52	131313 566622 333	462135 642562 223146 2562*6 2	462462 546231 546231	514535 615364 23	514653 6213	515135 365462 3	135134 642163 451463 211136 414262 135362 *14654 6	123645 132462 426262 241356 462*13 524242 466135 6462	123456 525123 652625 635256 262365 456	123456 244565 565263 212554 146662 654251 44526* 31	123456 456123 643524 1	143625 361425 142523 625656 3

The subjects' frequencies of making their assigned types' (and when relevant, alternate types') exact guesses are in parentheses after the assigned type.
ct's look-up sequence means that the subject entered a guess there without immediately confirming it.

Selected Baseline Subjects' Information Searches and Estimated Types' Search Implications

MouseLab box numbers

	a	*p*	*b*
You (*i*)	1	2	3
S/he (*j*)	4	5	6

Types' Search Implications

L1	{[4,6],2}
L2	{([1,3],5),4,6,2}
L3	{([4,6],2),1,3,5}
D1	{(4,[5,1], (6,[5,3]),2}
D2	{(1,[2,4]),(3,[2,6]),(4,[5,1],(6,[5,3]),5,2}
Eq	{[2,5],4} if pr. tar.<1,{[2,5],6} if > 1

Subject	**101**	**118**	**413**	**108**	**206**	**309**	**405**	**210**	**302**	**318**	**417**	**404**	**202**	**310**	**315**
Type(#rt.)	L1 (15)	L1 (15)	L1 (14)	L2 (13)	L2 (15)	L2 (16)	L2 (16)	L3 (9)	L3 (7)	L1 (7)	Eq (8)	Eq (9)	Eq (8)	Eq (11)	Eq (11)
Alt.(#rt.)								Eq (9)	Eq (7)	D1 (5)	L3 (7)	L2 (6)	D2 (7)		
Alt.(#rt.)								D2 (8)			L2 (5)		L3 (7)		
Est. style	early/late	early	late	early	early	late	early	early	early	early	early	early	early	early/late	early
Game															
1	146246 213	246134 626241 32*135	123456 545612 3463*	135642	533146 213	1352	144652 313312 546232 12512	123456 123456 213456 254213 654	221135 465645 213213 45456* 541	132456 465252 13242* 1462	252531 464656 446531 641252 462121 3	462135 464655 645515 21354* 135462 426256 356234 131354 645	123456 254613 621342 *525	123126 544121 565421 254362 *21545 4*	213465 624163 564121 325466
2	46213	246262 2131	123564 62213*	135642 3	531462 31	135263 1526*2 *3	132456 253156 456545 463123 156562 62	123456 465562 231654 456*2	213546 566213 545463 21*266 54123	132465 132*46 2	255236 62*365 243563	462461 352524 261315 463562	123456 445613 255462 513565 23	123546 216326 231456 *62	134652 124653 656121 3
3	462*46	246242 466413 *426	264231	135642 53	535164 2231	135263	312456 5231*1 236545 5233** 513	123455 645612 3	265413 232145 563214 563214 523*65 4123	134652 1323*4	521363 641526 5263*6 52	462135 215634 *52	123456 123562 3	123655 463213	132465 544163 *3625

These data suggest the following conclusions:

(i) Search is so heterogeneous and noisy that should study it at individual level

(ii) There is little difference between the look-up sequences of R/TS and Baseline subjects of a given type (assigned for R/TS, apparent for Baseline); perhaps unsurprising, because R/TS subjects were not trained in search strategies

(iii) A subject's type's predicted sequence is unusually dense in his searches, at least for *L1* and *L2*, and can quickly learn to read the algorithms many subjects are using in the data (CGC's econometric analysis measures search compliance as the density of a type's relevant sequences in the subject's sequence; Table 7 shows that many subjects' types can be reliably identified from search alone)

(iv) For some subjects search is an important check on decisions; e.g. Baseline subject 309, with 16 exact *L2* guesses, misses some of *L2*'s relevant look-ups, avoiding deviations from *L2* only by luck (s/he has a Eureka! moment between games 5 and 6, and from then on complies perfectly); reminiscent of CJ's finding that in their alternating-offers bargaining games, 10% of the subjects *never* looked at the last-round pie and 19% never looked at the second-round pie

How does cognition show up in search?

CJ gave roughly equal weight to look-up durations and to the numbers of look-ups of each pie (“acquisitions”) and the transitions between pies

Rubinstein, “Instinctive and Cognitive Reasoning: A Study of Response Times,” *EJ* 2007, which considers some matrix games, considered only durations

Gabaix, Laibson, Moloche, and Weinberg, “Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model,” *American Economic Review* 2006, focused on acquisitions and considered some aspects of order too

The analyses were at a high level of aggregation, across subjects and over time

CGCB and CGC presume which look-ups subjects make, in which order, reveals as much information about cognition as durations or transition frequencies

(Simple theories of cognition more readily suggest roles for which look-ups subjects make, and in which orders, than for durations; no claim is intended that durations are irrelevant—CGCB (Table IV) present some results on durations—just that they don’t deserve the priority they have been given)

Types as models of cognition and search

CGC's and CGCB's models of cognition, search, and decisions are based on an individual-level, procedural view of decision-making, in which a subject's type determines his search, and his type and search then determine his decision

Each type is naturally associated with algorithms that process payoff information into decisions

The analysis uses the algorithms as models of cognition, deriving a type's search implications under simple assumptions about how cognition determines search

(Because a type's search implications depend not only on what decisions it specifies, but why, something like a types-based model seems necessary here)

With their derived search implications, the types then provide a kind of basis for the enormous space of possible decision and search sequences, imposing enough structure to describe subjects' behavior in a comprehensible way, and to make it meaningful to ask how subjects' decisions and searches are related

Without further assumptions, nothing precludes a subject's scanning and memorizing the information and "going into his brain" to figure out what to do, in which case his searches will reveal nothing about cognition

But subjects' actual searches appear to contain a lot of information

We need enough additional assumptions to extract the signal from the noise in subjects' searches, but not so many that they distort the meaning of the signal

CGC's (like CGCB's) assumptions are conservative in that they rest on types' minimal search implications and add as little structure to these as possible

Types' minimal search implications in CGC's games can be derived from their *ideal guesses*, those they would make if they had no limits; with automatic adjustment of guesses and quasiconcave payoffs, a subject's ideal guess is all he needs to know to ensure his adjusted guess is optimal

The left side of Table 4 lists formulas for types' ideal guesses in CGC's games.

The right side lists types' search implications in our model, first in terms of our notation, then in terms of the box numbers in which MouseLab records the data

TABLE 4—TYPES' IDEAL GUESSES AND RELEVANT LOOK-UPS

Type	Ideal guess	Relevant look-ups
L1	$p^i[a^j + b^j]/2$	$\{[a^j, b^j], p^i\} \equiv \{[4, 6], 2\}$
L2	$p^iR(a^j, b^j; p^j[a^i + b^i]/2)$	$\{([a^i, b^i], p^j), a^j, b^j, p^i\} \equiv \{([1, 3], 5), 4, 6, 2\}$
L3	$p^iR(a^j, b^j; p^jR(a^i, b^i; p^i[a^j + b^j]/2))$	$\{([a^j, b^j], p^i), a^i, b^i, p^j\} \equiv \{([4, 6], 2), 1, 3, 5\}$
D1	$p^i(\max\{a^j, p^ja^i\} + \min\{p^jb^i, b^j\})/2$	$\{(a^j, [p^j, a^i]), (b^j, [p^j, b^i]), p^i\} \equiv \{(4, [5, 1]), (6, [5, 3]), 2\}$
D2	$p^i[\max\{\max\{a^j, p^ja^i\}, p^j\max\{a^i, p^ia^j\}\} + \min\{p^j\min\{p^ib^j, b^i\}, \min\{p^jb^i, b^j\}\}]/2$	$\{(a^i, [p^i, a^j]), (b^i, [p^i, b^j]), (a^j, [p^j, a^i]), (b^j, [p^j, b^i]), p^j, p^i\} \equiv \{(1, [2, 4]), (3, [2, 6]), (4, [5, 1]), (6, [5, 3]), 5, 2\}$
Eq.	$\{a^i$ if $p^ia^j \leq a^i$ or $\min\{p^ia^j, b^i\}$ if $p^ia^j > a^i\}$ if $p^ip^j < 1$ or $\{b^i$ if $p^ib^j \geq b^i$ or $\max\{a^i, p^ib^j\}$ if $p^ib^j < b^i\}$ if $p^ip^j > 1$	$\{[p^i, p^j], a^j\} \equiv \{[2, 5], 4\}$ if $p^ip^j < 1$ or $\{[p^i, p^j], b^j\} \equiv \{[2, 5], 6\}$ if $p^ip^j > 1$
Soph.	[no closed-form expression; *Sophisticated*'s search implications are the same as *D2*'s]	$\{(a^i, [p^i, a^j]), (b^i, [p^i, b^j]), (a^j, [p^j, a^i]), (b^j, [p^j, b^i]), p^j, p^i\} \equiv \{(1, [2, 4]), (3, [2, 6]), (4, [5, 1]), (6, [5, 3]), 5, 2\}$

Notes: The most basic operations are represented by the innermost look-ups, grouped within square brackets; these can appear in any order, but *may not* be separated by other look-ups. Other operations are represented by look-ups grouped within parentheses or curly brackets; these can appear in any order, and *may* be separated by other look-ups. *Equilibrium*'s minimal search implications are derived not directly from *Equilibrium*'s ideal guesses, but from p^ia^j when $p^ip^j < 1$ and p^ib^j when $p^ip^j > 1$ via Observation 1 (see on-line Appendix H).

p is a target; *a* (*b*) is a lower (upper) limit; *i* and *j* are the player and his partner; and *R( )* is the automatic adjustment function.

Derivation of types' search implications

Evaluating a formula for a type's ideal guess requires a series of *operations*, some of which are *basic* in that they logically precede any other operation

E.g. [$a^j + b^j$] is the only basic operation for *L1*'s ideal guess, $p^i\,[a^j + b^j]/2$

The search implications in Table 4 assume that subjects perform basic operations one at a time via adjacent look-ups, remember their results, and otherwise rely on repeated look-ups rather than memory

Basic operations will then be represented by adjacent look-up pairs that can appear in either order, but cannot be separated by other look-ups

Such pairs are grouped within square brackets, as in {[a^j, b^j], p^i} for *L1*

Other operations can appear in any order and their look-ups can be separated

They are represented by look-ups grouped within curly brackets or parentheses

L1, *L2*, *L3*, *D1*, *D2* search implications are easy to derive directly from Table 4's formulas

Equilibrium is more interesting

It can use any workable method to find its ideal guess; we allow any method, and seek the one with minimal search requirements

Equilibrium-checking (conjecturing guesses and checking them for consistency with equilibrium) is less demanding than other methods, but requires more luck than almost all of our subjects appeared to have

Accordingly, we allow *Equilibrium* to use both targets to determine whether the equilibrium is High or Low, and then to enter its own target times its partner's lower (upper) limit when the product of targets is < (>) 1, which CGC's Observation 1 shows ensures its adjusted guess is in equilibrium

This has the same search requirements as equilibrium-checking except that it requires the targets to be adjacent; and thereby avoids the need for luck

(Unlike in CGCB's and CJ's designs, *Equilibrium*'s search implications are just as simple as *L1*'s, and simpler than other boundedly rational types'!)

Aside: Background evidence on cognition and search

(i) CJ's Robot/Trained Subjects' searches, which led them to characterize subgame-perfect equilibrium via backward induction search in terms of transitions between the second- and third-round pies

(ii) CGCB's Trained Subjects' searches, which suggest a similar view of *Equilibrium* search in matrix games

(iii) CGC's R/TS and Baseline subjects with high compliance with their types' guesses, whose searches suggest a similar view of *L1* and *L2* search (CGC's specification analysis turned up only one clear violation of the proposed characterization of types' search implications: Baseline subject 415, whose apparent type was *L1* with 9 exact guesses, had 0 *L1* search compliance in 9 of the 16 games because s/he had no adjacent [a^j,b^j] pairs as we required for *L1*. Her/his look-up sequences were unusually rich in (a^j,p^i,b^j) and (b^j,p^i,a^j) triples, in those orders. Because the sequences were *not* rich in such triples with other superscripts, this is clear evidence that 415 was an *L1* who happened to be more comfortable with several numbers in working memory than our characterization of search assumes, or than our other subjects were. But because this violated our assumptions on search, this subject was "officially" estimated to be *D1*.)

CGC's econometric analysis of guesses and search

In the econometric analysis we summarize a subject's compliance with a type's search implications in a game by the density of the type's look-up sequence in the subject's observed look-up sequence, with some refinements

We extend CGCB's and CGC's maximum likelihood error-rate models of decisions to explain search compliance as well as decisions

Most guesses-and-search type estimates reaffirm guesses-only estimates

For some subjects the guesses-and-search estimate resolves a tension between guesses-only and search-only estimates in favor of a type other than the guesses-only estimate

In more extreme cases, a subject's guesses-only type estimate is excluded because it has 0 search compliance in 8 or more games, like subject 415

Search refines and sharpens conclusions and confirms the absence of significant numbers of types other than *L1*, *L2*, *Equilibrium*, or hybrids of *L3* or *Equilibrium*

Table 1—Summary of Baseline and OB Subjects' Estimated Type Distributions

Type	Apparent from guesses	Econometric from guesses	Econometric from guesses, excluding random	Econometric from guesses, with specification test	Econometric from guesses and search, with specification test
L1	20	43	37	27	29
L2	12	20	20	17	14
L3	3	3	3	1	1
D1	0	5	3	1	0
D2	0	0	0	0	0
Eq.	8	14	13	11	10
Soph.	0	3	2	1	1
Unclassified	45	0	10	30	33

Note: The far-right-hand column includes 17 OB subjects classified by their econometric-from-guesses type estimates.

Aside on Searchmetrics:

Our econometric analysis focuses on the order of look-ups and ignores duration, following CJ and CGCB

We view search for hidden payoff information as just another kind of decision—not the kind conventionally studied, but potentially also useful in helping to identify the decision rules that best describe subjects' behavior

Combine guessmetrics with a maximum-likelihood error-rate model of search as in CGCB (but subject-by-subject, not mixture model)

The main econometric problem is extracting signals from subjects' highly idiosyncratic, noisy look-up sequences, without a well-tested model that implies strong restrictions on how cognition drives search

Subjects vary in the location of look-ups relevant to their types in their sequences. Filter this out via subject-specific nuisance parameter called style ("early" or "late"), assumed constant across games for each subject. (58 of 71 Baseline subjects' estimated styles are "early," 10 are "late," and 3 are tied.)

Quantify compliance with a type's search implications as the density of the type's relevant look-up sequence in the subject's look-up sequence. If style is early, start at the beginning of the sequence and continue until the type's relevant sequence is first completed. Compliance is the length of the relevant sequence divided by the length of the sequence that first completes it. This definition makes compliance meaningfully comparable across games, styles.

We assume that a subject's type and style determine his search and guess in a given game, each with error; and we further assume that, given type and style, errors in search and guesses are independent of each other and across games. This strong but useful simplifying assumption makes the log-likelihood separable across guesses and search, avoiding some complications in CGCB.

To avoid stronger distributional assumptions, we discretize compliance into three categories: $C_H \equiv [0.67, 1.00]$, $C_M \equiv [0.33, 0.67]$, and $C_L \equiv [0, 0.33]$.

Subject *i*'s guesses-and-search log-likelihood is:

$$\sum_c \left[m_c^{isk} \ln(\zeta_c) + (m_c^{isk} - n_c^{isk}) \ln(1-\varepsilon) + n^{isk} \ln(\varepsilon) + \sum_{g \in N_c^{isk}} \ln d_g^k (R_g^i(x_g^i), \lambda) \right] \equiv$$

$$(G-n^{ik})\ln(G-n^{ik}) + n^{ik}\ln(n^{ik}) + \sum_{g \in N^{ik}} \ln d_g^k(R_g^i(x_g^i), \lambda) - G\ln G + \sum_c \left[m_c^{isk} \ln m_c^{isk} \right] - 2G\ln G,$$

where m_c^{isk} is the number of games for which subject *i* has type-*k* style-*s* compliance *c*. (The search term is convex in the m_c^{isk}, and therefore favors types for which compliance varies less across games, because such types "explain" search behavior better. See CGCB, Section 4.D.)

The maximum-likelihood estimates of ε and ζ_c, given *k* and *s*, are n^{ik}/G and m_c^{isk}/G, the sample frequencies with which subject *i*'s adjusted guesses are non-exact for that *k* and *i* has compliance *c* for that *k* and *s*. The maximum likelihood estimate of *λ* is the standard logit precision.

The maximum likelihood estimate of subject *i*'s type *k* maximizes the above log-likelihood over *k* and *s*, given the estimated ε and *λ*.

The model favors such types without regard to whether compliance is high or low. This seems appropriate because compliance is neither meaningfully comparable across types nor guaranteed to be high for the "true" type (which could be cognitively very difficult). But it means we need to rule out estimates where a type wins simply because its compliance is very low in all games.

A few subjects' type estimates change (Table 1) when search is included:

For some subjects a tension between guesses-only and search-only type estimates is resolved in favor of the search estimates. (The search part of the likelihood has weight only about 1/6 of the guesses part, because our theory of search makes much less precise predictions than our theory of guesses—a necessary evil, given the noisiness and idiosyncrasy of search behavior.)

For other subjects the guesses-only type estimate has 0 search compliance in 8 or more games, and so we rule it out a priori.

Possible answers to unresolved puzzle A. What are Those Baseline "*Equilibrium*" Subjects Really Doing?

(i) Can we tell how Baseline *Equilibrium* subjects find equilibrium in games without mixed targets: best-response dynamics, equilibrium checking, iterated dominance, or something else that doesn't "work" with mixed targets? Can check by refining characterization of *Equilibrium* search and redoing searchmetrics

Aside: Refined characterization of *Equilibrium* search
Equilibrium's ideal guess can be identified by (1) evaluating a formula, (2) equilibrium-checking, (3) iterated dominance, or (4) best-response dynamics

(1) Two ways to evaluate a formula: using *Equilibrium*'s ideal guess, or using Observation 1's proxy for *Equilibrium*'s ideal guess

Because they are logically related, our theory cannot distinguish them

The latter is less stringent, and yields requirements

(1) $\{[p^i,p^j],a^j\} \equiv \{[2, 5], 4\}$ if $p^ip^j < 1$ or $\{[p^i,p^j],b^j\} \equiv \{[2, 5], 6\}$ if $p^ip^j > 1$

(2) Equilibrium-checking's requirements are almost the same, usually requiring both of the partner's limits but excluding one in some cases, depending on luck

I omit the requirements here, noting only that this method also requires $[p^i, p^j]$

(3) Iterated dominance we assume requires one or more complete rounds, stopping when there is a clear up-or-down direction in which dominance eliminates guesses, enough to guess whether the equilibrium is High or Low

Once the required rounds are completed, the player can use Observation 1's proxy for *Equilibrium*'s ideal guess; this adds a p^i times either a^j (Low equilibrium) or b^j (High) to his sequence

As it happens, the search requirements for *k* rounds are independent of *k*; thus, the search requirements for iterated dominance are like CGC's characterization for *D2* (*D2*, not *D1*, because unlike *D1*, a *k*-round iterated-dominance player must delete *k* rounds of dominated guesses for himself too)

(3) $\{(a^i, [p^i, a^j]), (b^i, [p^i, b^j]), (a^j, [p^j, a^i]), (b^j, [p^j, b^i]), p^j, p^i\}$
$\equiv \{(1, [2, 4]), (3, [2, 6]), (4, [5, 1]), (6, [5, 3]), 5, 2\}$

(4) For best-response dynamics we assume the subject does only one complete round: that is, starting with a trial guess for one player, best-responding for the other, and then best-responding back for the first player

We also assume the subject can infer from whether the iterated best response goes up or down (if it changes) whether equilibrium is High or Low

(4) {([a^i,p^j] or [b^i,p^j] or [(a^j,p^i] or [(b^j,p^i]), p^i,p^j,(all but at most one of a^i,b^i,a^j, and b^j)}

The main difference among *Equilibrium* methods is that methods 1 and 2 have a [p^i, p^j] requirement and methods 3 and 4 do not

We know from the absence of Baseline *Dk* subjects in CGC's guesses-and-search estimates that method 3's requirements don't fit the data well

Its also seems, from the data, that [p^i, p^j] are comparatively rare for Baseline apparent *Equilibrium* subjects, and even for R/TS *Equilibrium* subjects

Thus searchmetrics may favor best-response dynamics, truncated 1-2 rounds

(ii) Is there any difference in Baseline *Equilibrium* subjects' search patterns in games with and without mixed targets? If so, how does the difference compare to the differences for *L1*, *L2*, or *L3* subjects?

(Our 20 Baseline *L1* subjects' compliance with *L1* guesses is almost the same with and without mixed targets (Figure 1); unsurprising, distinction is irrelevant to *L1*. But 12 *L2* and 3 *L3* (Figures 2-3) subjects' compliance with types' guesses is lower with mixed targets. This is curious, because for *L2* and *L3*, unlike for *Equilibrium*, games with mixed targets require no deeper understanding.)

(iii) Can we tell how R/TS *Equilibrium* subjects with high compliance manage to find their *Equilibrium* guesses even with mixed targets? How does their search in those games differ from Baseline *Equilibrium* subjects' search?

(CGC strove to make the R/TS *Equilibrium* training as neutral as possible, but something must come first. R/TS subjects were taught equilibrium-checking first, then best-response dynamics, then iterated dominance. To the extent that they used one of those methods, it explains why they have equal compliance with and without mixed targets. If they used something else that deviates from equilibrium with mixed targets, it might be a clue to what Baseline *Equilibrium* subjects did.)

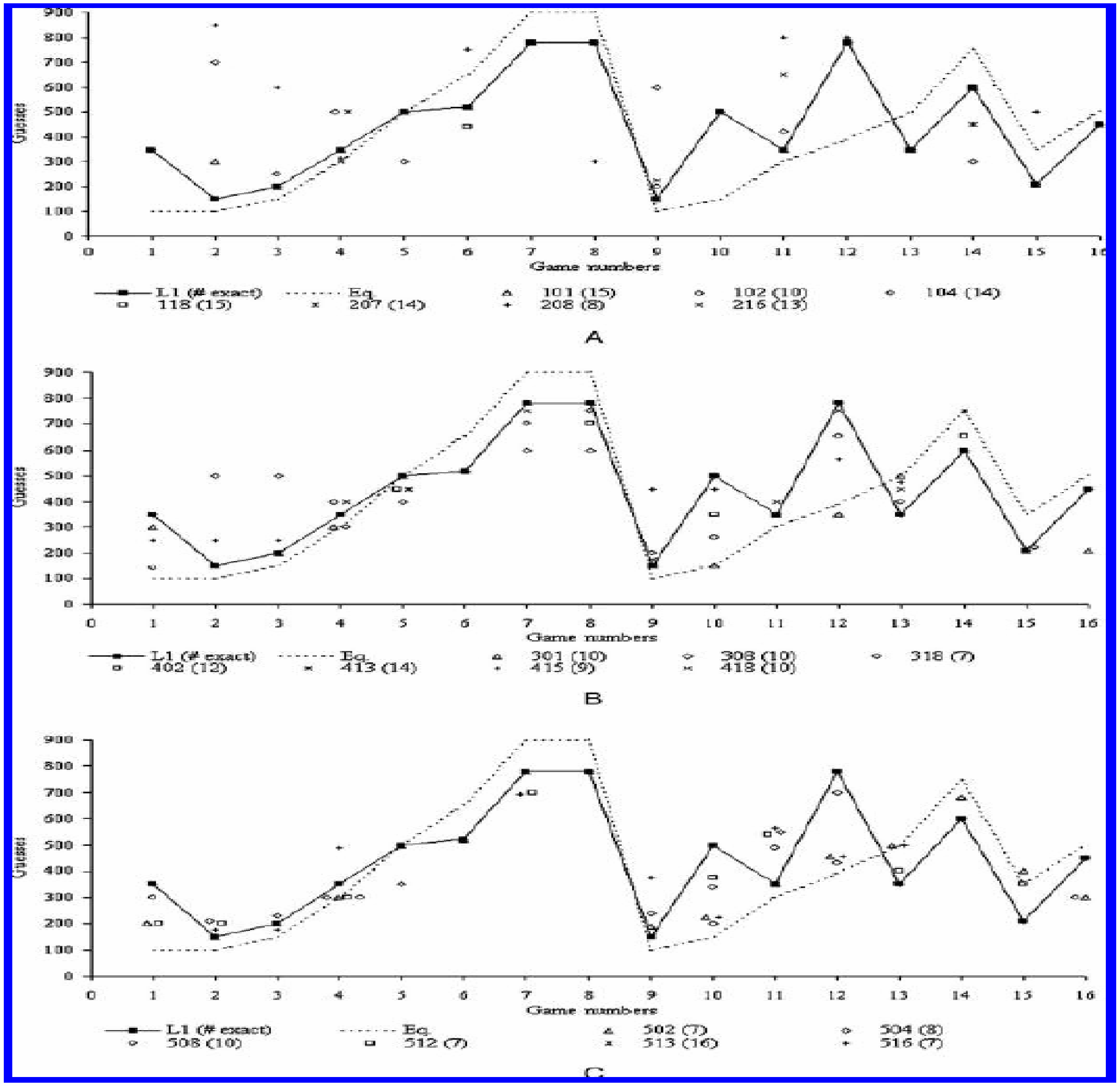


FIGURE 1. "FINGERPRINTS" OF 20 APPARENT *L1* SUBJECTS

Notes: Only deviations from *L1*'s guesses are shown. Of these subjects' 320 guesses, 216 (68 percent) were exact *L1* guesses.

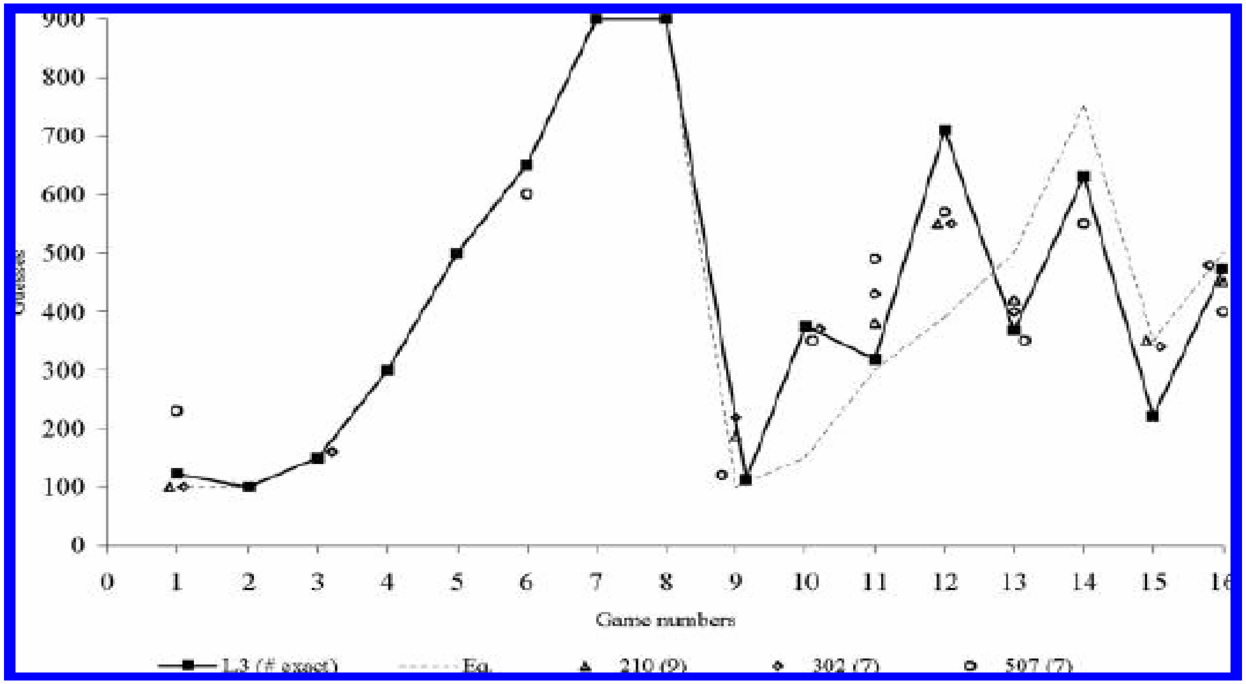


FIGURE 3. "FINGERPRINTS" OF THREE APPARENT *L3* SUBJECTS

Notes: Only deviations from *L3*'s guesses are shown. Of these subjects' 48 guesses, 23 (48 percent) were exact *L3* guesses.

(Note that CGC's Baseline subjects with high compliance for some type are like robot *un*trained subjects. These don't usually exist because you can't tell robot subjects how they will be paid without teaching them how the robot works. Thus CGC's design provides an unusual opportunity to separate the effects of training and strategic uncertainty, by comparing Baseline and R/TS subjects: Either *Equilibrium* is natural with mixed targets, but subjects don't see it without training; or *Equilibrium* is unnatural, and/or subjects don't believe that others, even with training, will make *Equilibrium* guesses with mixed targets.)

Possible answers to unresolved puzzle B. Why Are *Lk* the Only Types other than *Equilibrium* with Nonnegligible Frequencies?

(i) Most R/TS subjects could reliably identify their type's guesses, even *Equilibrium* or *D2*. (These average rates are for exact compliance, and so are quite high. Individual subjects' compliance was usually bimodal within type, on very high and very low.)

R/TS Subjects' Exact Compliance with Assigned Type's Guesses

	L1	***L2***	***L3***	***D1***	***D2***	***Eq.***
Number of subjects	25	27	18	30	19	29
% Compliance|Passed UT2	80.0	91.0	84.7	62.1	56.6	70.3
% Failed UT2	0.0	0.0	0.0	3.2	5.0	19.4

(ii) But there are noticeable signs of differences in difficulty across types:

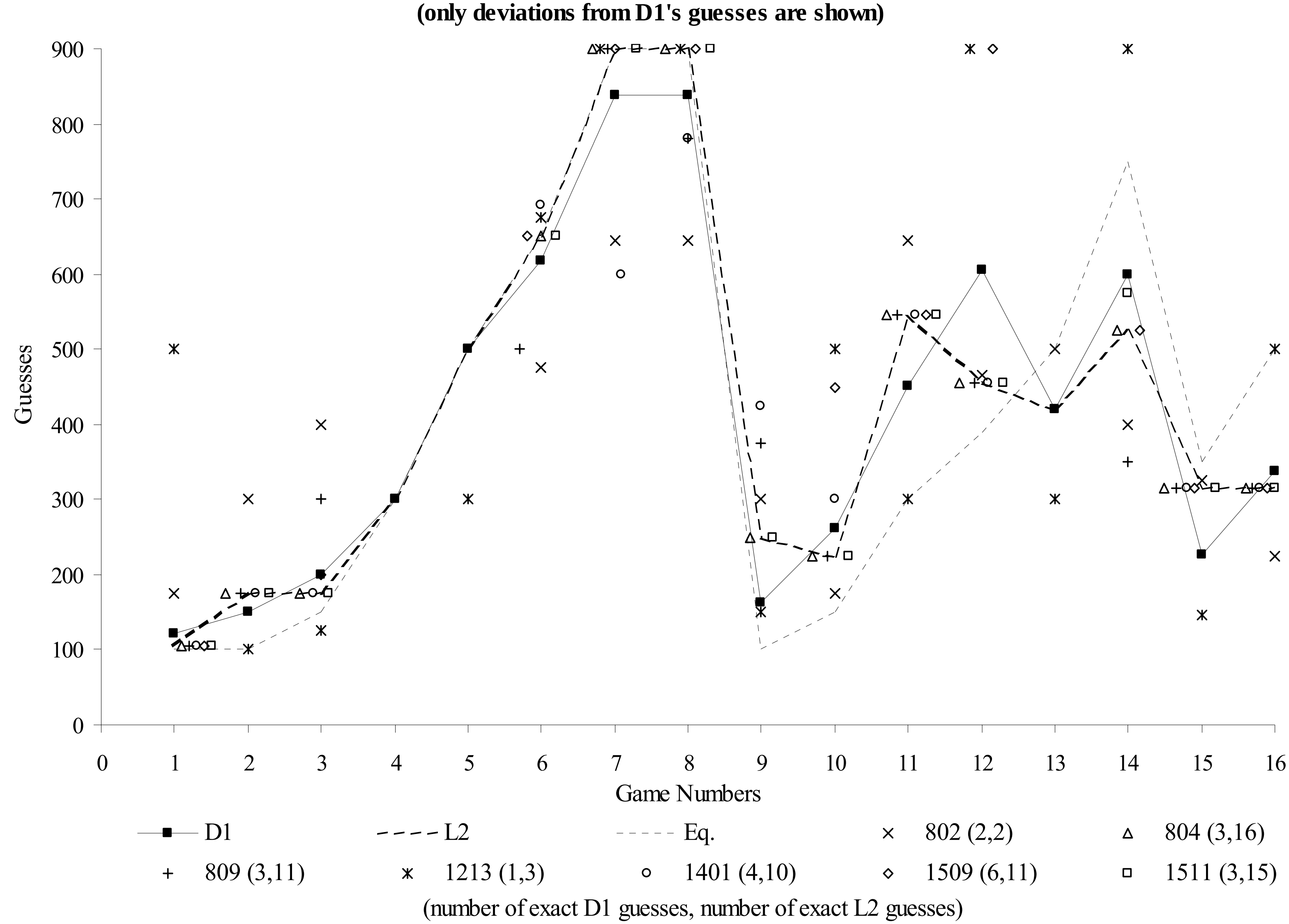
(a) No one ever failed an *Lk* Understanding Test, while some failed the *Dk* and many failed the *Equilibrium* Understanding Test.

(b) For those who passed, compliance was highest for *Lk* types, then *Equilibrium*, then *Dk* types. This suggests that *Dk* is even harder than *Equilibrium*, but could just be an artifact of the more stringent screening of the *Equilibrium* Test.

(c) Among *Lk* and *Dk* types, compliance was higher for lower *k* as expected, except *L1* was lower than *L2* or *L3* compliance. (We suspect that this is because *L1* best responds to a random *L0* robot, which some subjects think they can outguess; while *L2* and *L3* best respond to a deterministic *L1* or *L2* robot.)

(d) Remarkably, 7 of 19 R/TS *D1* subjects passed the *D1* Understanding Test, in which *L2* answers are wrong; and then "morphed" into *L2*s when making their guesses, significantly reducing their earnings. E.g. R/TS *D1* subject 804 made 16 exact *L2* (and so only 3 exact *D1*) guesses. (Recall that it is *L2* that is *D1*'s cousin.) This kind of morphing, in this direction, is the only kind that occurred. We view this as pretty compelling evidence that *Dk* types are unnatural.

Fingerprints of 7 R/TS Subjects who morphed from D1 to L2



Appendix 1: A “Theory” of Optimal Search for Hidden Payoff Information

I now sketch a simple model of optimal search with costs of storing numbers in working memory, which rationalizes the stylized facts of search behavior in these experiments.

The model views search for hidden payoff information as just another decision, and takes the formula that relates a type's desired decisions to the hidden parameters as given. Thus a subject's type determines both an optimal search pattern and an optimal decision.

Occurrence

The usual rationality assumption implies that a player will look up all costlessly available information that might affect his beliefs and best respond to his beliefs.

When, as here, observing a parameter will normally cause a nonnegligible change in beliefs and the optimal decision, this conclusion extends to all relevant information that is available at a sufficiently small but non-0 cost. (There is a lot of evidence that subjects perceive the cost of a look-up as close to negligible.)

Thus, if a type’s decision depends on a hidden parameter, then that parameter must appear in the type’s look-up sequence. But so far any order will do.

Adjacency

Assume that there is a cost of keeping numbers in working memory, which starts out small, but is much larger, even for one number, than the cost of a look-up; and that this cost increases with the number of stored numbers and is proportional to storage time. (Thus a player's "lifetime" total memory cost is the time integral of an increasing function of the number of stored numbers.) (There is some evidence for these assumptions too.)

Given these assumptions, a player minimizes his total memory plus look-up cost for evaluating an expression like *L1*'s ideal guess, $p^i\,[a^j+b^j]/2$, containing a basic operation like $[a^j+b^j]$, by processing $[a^j+b^j]$ separately, storing the result (in the meanwhile "forgetting" a^j and b^j), and combining the result with p^i.

(The alternative, processing $p^i\,a^j$ separately, storing the result, then processing $p^i\,b^j$ separately and combining it with $p^i\,a^j$, requires leaving more numbers in working memory longer: The sequence of numbers in memory for the method in the previous paragraph is 1, 2, 1, 2, 1; the sequence for the method in this paragraph is 1, 2, 1, 2, 3, 2, 1. The previous method also saves by eliminating the repeated look-up of p^i, but this is of second-order importance.)

I have only illustrated the cost savings from giving priority to basic operations, but I conjecture that the argument is general. If so, it justifies assuming that subjects perform basic operations one at a time via adjacent look-ups, remembering the results, and otherwise relying on repeated look-ups rather than memory.

The argument also seems likely to extend to CJ's extensive-form games, justifying their focus on transitions between pies from adjacent rounds.

The theory implies more than this, regarding both the order of operations (basic ones should come first) and how non-basic operations are executed. I defer such implications in favor of mentioning an issue regarding CGC's search data:

Many Baseline subjects usually look first at their apparent type's relevant sequence and then make irrelevant look-ups or stop (e.g. 108, 118, and 206, labeled "early" in the above look-up data). Others make irrelevant look-ups first, and look at the relevant sequence only near the end (e.g. 413, labeled "late"). Others repeat the relevant sequence many times (e.g. 101, labeled "early/late"). The theory is actually consistent with this kind of heterogeneity when look-up costs are negligible (but storage costs are not). Because MouseLab allows a subject to enter a tentative guess without confirming it (the *s in the look-up data), this kind of storage has zero cost in CGC's and CGCB's designs; and so subjects can satisfy their curiosity (early or late) without running up storage costs.

Appendix 2: Costa-Gomes, Crawford, and Broseta's Matrix-Game Experiments

CGCB adapted CJ's methods to study cognition via search for hidden payoffs in matrix games, eliciting initial responses to 18 games with various patterns of iterated dominance or unique pure-strategy equilibria without dominance (CGCB, Figure 2).

CGCB's design strongly separates leading types' implications for decisions.

Previous experiments (e.g. SW) found systematic deviations from the equilibrium decisions when players have pecuniary preferences (in games that probably disable social preferences).

CGCB's results for decisions replicated most patterns in previous experiments, with high equilibrium compliance with in games solvable by one or two rounds of iterated dominance but lower compliance in games solvable by three rounds or by the circular logic of equilibrium without dominance (CGCB, Table II).

CGCB's design replicated previous results in a way that allowed a more precise assessment of subjects' cognition, which confirms the view of subjects' behavior suggested by analyses of decisions alone, with some differences.

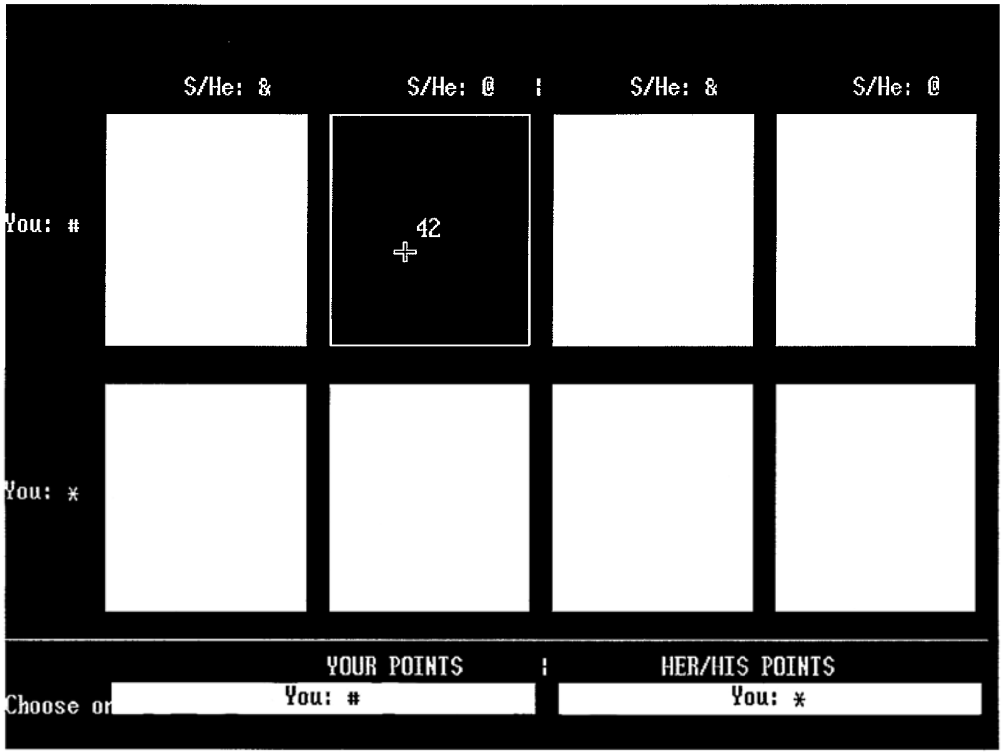
Monitoring Search via MouseLab in Matrix Games

Within a publicly announced structure, CGCB presented each game to subjects as a matrix via MouseLab, which normally concealed payoffs but allowed subjects to look up their own and their partner's payoffs for each decision combination as often as desired, one at a time (click option in MouseLab).

Row and Column players' payoffs were spatially separated to ease cognition and make search more informative.

(Subjects were always framed as Row players, although each played each of our games once as Row and once as Column player, in a sequence that disguised those relationships and randomized away effects of patterns in their structures.)

(Subjects were not allowed to write down the payoffs, and the frequencies with which they looked them up made clear that they did not memorize them.)



CGCB's Figure 1. MouseLab Screen Display (for a 2×2 game)

Separation of Types' Implications for Decisions in CGCB's Design: Figure 2

2A (1,2)	*A,P,N*	*D12,L2,E,S*
A	72,93	31,46
D	84,52	55,79

2B (1,2)	*A,P,N*	*D12,L2,E,S*
D	94,23	38,57
A	45,89	14,18

3A (2,1)	*D*	*A*
D12,L2,E,S	75,51	42,27
A,P,N	48,80	89,68

3B (2,1)	*D*	*A*
A,P,N	21,92	87,43
D12,L2,E,S	55,36	16,12

4A (2,1)		*D*	*A*
A,P,N	59,58	46,83	85,61
D12,L2,E,S	38,29	70,52	37,23

4B (2,1)	*A*	*D*
D12,L2,E,S	31,32	68,46
P	72,43	47,61
A,N	91,65	43,84

4C (1,2)	*D12,L2,E,S*	*A,P,N*
	28,37	57,58
A	22,36	60,84
D	51,69	82,45

4D (1,2)	*D12,L2,E,S*	*P*	*A,N*
D	42,64	57,43	80,39
A	28,27	39,68	61,87

5A (3,2)	*A,P,N*	*D12,L2,E,S*
A	53,86	24,19
P,N,D1,L2,S	79,57	42,73
D2,E	28,23	71,50

5B (3,2)	*A,P,N*	*D12,L2,E,S*
A	76,93	25,12
D2,E	43,40	74,62
P,N,D1,L2,S	94,16	59,37

6A (2,3)	*A*	*D2,E,S*	*P,N,D1,L2*
D12,L2,E,S	21,26	52,73	75,44
A,P,N	88,55	25,30	59,81

6B (2,3)	*D2,E*	*A*	*P,N,D1,L2,S*
A,P,N	42,45	95,78	18,96
D12,L2,E,S	64,76	14,27	39,61

7A (∞,∞)	*N,D12,L2,S*	*A,P*	*E*
L2,E,S	87,32	18,37	63,76
A,P,N,D12	65,89	96,63	24,30

7B (∞,∞)	*N,D12,L2,S*	*A,P*	*E*
A,P,N,D12	67,91	95,64	31,35
L2,E,S	89,49	23,53	56,78

8A (∞,∞)	*L2,E,S*	*A,P,N,D12*
E	72,59	26,20
A,P	33,14	59,92
N,D12,L2,S	28,83	85,61

8B (∞,∞)	*L2,E,S*	*A,P,N,D12*
A,P	46,16	57,88
E	71,49	28,24
N,D12,L2,S	42,82	84,60

9A (1,2)	*D12,L2,E,S*	*A,P,N*
	22,14	57,55
	30,42	28,37
A	15,60	61,88
D	45,66	82,31

9B (2,1)			*A*	*D*
A,P,N	56,58	38,29	89,62	32,86
D12,L2,E,S	15,23	43,31	61,16	67,46

(*A* = *Altruistic*, *P* = *Pessimistic* (minimax), *N* = *Naïve* (CGCB's name for *L1*) and *Optimistic* (maximax, decisions not separated from *Naïve*'s), *E* = *Equilibrium*, *S* = *Sophisticated*, *D12* = *D1* and *D2*, *D* = dominant decision = all types but *A*.)

Separation of Types' Implications for Search in CGCB's Design

CGCB's design also makes it possible to test and compare types via search

They make two assumptions about how cognition affects search, *Occurrence* and *Adjacency* that are close to CGC's characterization of cognition and search

In CGCB's display, a subject's searches can vary in three main dimensions:

(i) the extent to which his transitions are up-down in his own payoffs, which under *Occurrence* and *Adjacency* is (for a Row player) naturally associated with rationality in the decision-theoretic sense

(ii) the extent to which his transitions are left-right in other's payoffs, which under *Occurrence* and *Adjacency* is associated with thinking about other's incentives

(iii) the extent to which he makes transitions from own to other's payoffs and back for the same decision combination, which under *Occurrence* and *Adjacency* is associated with interpersonal fairness or competitiveness comparisons

This variation allows strong separation of types' implications for search

The independent separation of types' implications for decisions and search is an important strength of the design: Searches and decisions together, and their relationships, yield a much clearer view of a subject's type than decisions alone

Some types' implications under Occurrence and Adjacency in game 3A (Column has dominant decision, "nonstrategic" Rows pick B and "strategic" Rows pick T)

	S/He: L	S/He: R	S/He: L	S/He: R
You: T	75	42	51	27
You: B	48	89	80	68
	Your	Points	Her/His	Points
	You: T		You: B	

Naïve (*L1*) compares expected payoffs of own decisions given a uniform prior over other's, via either up-down or left-right own payoff comparisons. Occurrence requires look-ups 75, 48, 42, and 89. Adjacency requires either the set of comparisons {(75,42), (48,89)} *or* the set of comparisons {(75,48), (42,89)}.

	S/He: L	S/He: R	S/He: L	S/He: R
You: T	75	42	51	27
You: B	48	89	80	68
	Your	Points	Her/His	Points
	You: T		You: B	

L2 needs to identify other's *Naïve* decision and *L2*'s best response to it; Occurrence requires all other's look-ups plus 75 and 48, the own look-ups for other's *Naïve* decision. Adjacency requires either the set of comparisons {(51,27), (80,68)} or the set of comparisons {(51,80), (27,68)} to identify other's Naïve decision, plus the comparison (75,48) to identify *L2*'s best response.

If *Equilibrium* has a dominant decision it needs only to identify it. If not, it can use iterated dominance or equilibrium-checking, decision combination by combination or via "best-response dynamics." Occurrence requires look-ups 51, 27, 80, 68, 75, and 48. Adjacency requires comparisons (51,27), (80,68), and (75,48).

CGCB's Results

The most frequent estimated types are *Naïve* (*L1*) and L2, each nearly half of the population

Incorporating search compliance into the econometric analysis shifts the estimated type distribution toward *Naïve*, at the expense of *Optimistic* and *D1*

Part of this shift occurs because *Naïve*'s search implications explain more of the variation in subjects' searches and decisions than *Optimistic*'s, which are too unrestrictive to be useful in the sample; another part occurs because *Naïve*'s search implications explain more of the variation in subjects' searches and decisions than *D1*'s, which are more restrictive, but too weakly correlated with subjects' decisions

D1 also loses some frequency to *L2*, even though their decisions are weakly separated in CGCB's design, because *L2*'s search implications explain much more of the variation in subjects' searches and decisions

Overall, CGCB's analysis of decisions and search yields a significantly different interpretation of behavior than their analysis of decisions alone. The analysis suggests a strikingly simple view of behavior, with *Naïve* and *L2* 65-90% of the population and *D1* 0-20%, depending on confidence in their model of search