Cognitive load and mixed strategies: On brains and minimax

Duffy, Sean and Naddeo, JJ and Owens, David and Smith, John

Rutgers University-Camden, Haverford College

8 June 2016

Cognitive load and mixed strategies: On brains and minimax*

Sean Duffy† J. J. Naddeo‡ David Owens§ John Smith¶

October 3, 2018

Abstract

It is well-known that laboratory subjects often do not play mixed strategy equilibria games according to the theoretical predictions. However, little is known about the role of cognition in these strategic settings. We conduct an experiment where subjects play a repeated hide and seek game against a computer opponent. Subjects play with either fewer available cognitive resources (high cognitive load treatment) or with more available cognitive resources (low cognitive load treatment). Surprisingly, we find some evidence that subjects in the high load treatment earn more than subjects in the low treatment. However, we also find that subjects in the low treatment exhibit a greater rate of increase in earnings across rounds, thus suggesting more learning. Our evidence is consistent with subjects in the low load treatment over-experimenting. Further, while we observe that subjects do not mix in the predicted proportions and that their actions exhibit serial correlation, we do not find strong evidence these are related to their available cognitive resources. This suggests that the standard laboratory deviations from the theoretical predictions are not associated with the availability of cognitive resources. Our results shed light on the extent to which cognitive resources affect (and do not affect) behavior in games with mixed strategy equilibria.

Keywords: bounded rationality, experimental economics, working memory load, cognition, learning, over-experimentation

JEL: C72, C91

*We thank Carlos Alós-Ferrer, Roberto Barbera, Carlos Cueva, Guillaume Fréchette, Johannes Hoelzemann, Hannu Kivimaki, Andreu Mas-Colell, Ken Morris, William Neilson, Tibor Neugebauer, Sebastian Olschewski, Lisa Saal, Adam Sanjurjo, and Rei Sayag. John Smith thanks Biblioteca de Catalunya. This research was supported by Rutgers Research Council Grants #202071, #202167, #202256, and #18-AA-00143.

†Rutgers University-Camden, Department of Psychology

‡Georgetown University, Department of Economics

§Haverford College, Department of Economics

¶Corresponding Author; Rutgers University-Camden, Department of Economics, 311 North 5th Street, Camden, New Jersey, USA 08102; Email: smithj@camden.rutgers.edu; Phone: +1-856-225-6319; Fax: +1-856-225-6602

1 Introduction

It is common to use the Nash equilibrium, or the mutual best response, as a prediction of behavior in games. In games of conflict, another prediction of behavior is minimax, whereby the player can employ a strategy that guarantees a minimum payoff. In zero-sum games, these predictions coincide, thereby increasing the plausibility of these theoretical predictions.

In experimental settings, subjects often do not mix according to these theoretical predictions. They frequently deviate from the predicted mixture proportions and their actions exhibit serial correlation.[1] A criticism of this literature is that subjects are often inexperienced in settings where strategic mixing is required. Prompted by this criticism, many studies have examined mixing behavior in settings where decision makers have ample experience: the field.[2] Although some deviations are still detected, this literature mostly finds that the mixing in field settings is closer to the theoretical predictions than in the laboratory.

In order to better understand the robustness of these deviations from the theoretical predictions, researchers have examined whether experience in mixing in a field setting translates to successfully mixing in a novel experimental setting.[3] Our paper is complementary in that we seek to better understand the role of cognition in games with mixed strategy equilibria. Also similar to this literature, we are specifically interested in the earnings of the subjects and their conformity to the theoretical predictions.

To our knowledge, Geng, Peng, Shachat, and Zhong (2015) is the only other study that investigates the relationship between cognition and mixing behavior.[4] The authors do not find evidence that higher measures of cognitive ability[5] are related to behavior consistent with the

[1] See O'Neill (1987), Brown and Rosenthal (1990), Batzilis et al. (2017), Binmore, Swierzbinski, and Proulx (2001), Geng, Peng, Shachat, and Zhong (2014), Mookherjee and Sopher (1994, 1997), O'Neill (1991), Ochs (1995), Palacios-Huerta and Volij (2008), Rapoport and Amaldoss (2000, 2004), Rapoport and Boebel (1992), Rosenthal, Shachat, and Walker (2003), Shachat (2002), Van Essen and Wooders (2015). In fact, Martin et al. (2014) find evidence that chimpanzee subjects are closer to the theoretical predictions than human subjects.

[2] See Azar and Bar-Eli (2011), Bailey and McGarrity (2012), Bar-Eli et al. (2007), Buzzacchi and Pedrini (2014), Chiappori, Levitt, and Groseclose (2002), Coloma (2007), Emara, Owens, Smith, and Wilmer (2017), Hsu, Huang, and Tang (2007), Kovash and Levitt (2009), McGarrity and Linnen (2010), Palacios-Huerta (2003a), Reed, Critchfield, and Martens (2006), Walker and Wooders (2001).

[3] See Palacios-Huerta and Volij (2008), Levitt, List, and Reiley (2010), and Van Essen and Wooders (2015).

[4] See Palacios-Huerta et al. (2014) for a study of brain activity during a game with a mixed strategy equilibrium.

[5] Raven's standard progressive matrices test (Raven and De Lemos, 1990) and a score on a math test.

theoretical predictions: proportions of the mixture or serial correlation. Further, Geng et al. do not find a relationship between measures of cognitive ability and earnings in these games. However, one potential drawback of employing measures of cognitive ability is that these measures are possibly also correlated with other (observable or unobservable) characteristics of the subjects, for instance educational opportunities. Thus, these correlations can make inferences problematic.

We take a complementary approach as we seek to better understand behavior in games with mixed strategy equilibria by experimentally manipulating the available cognitive resources available to subjects. Our study follows other cognitive load experiments that observe behavior or judgments while the subject has some information committed to memory. This manipulation allows a within-subject design, in the sense that our subjects are placed into different cognitive load treatments, and such is not possible with measures of cognitive ability.[6] In our experiment, subjects are directed to either remember a large number (high cognitive load treatment) or a small number (low cognitive load treatment).

In our design, subjects play against two distinct computer opponents[7] in an experimental session and are told of this fact. Each computer opponent is programmed to play either one of two *Exploitative* strategies (designed to exploit suboptimal mixing by the subjects) or one of two *Naive* strategies (designed to allow subjects the possibility of exploiting the computer). Therefore, subjects face opponents who are playing only one of a few well-defined strategies, and this facilitates the analysis of the extent to which the strategic behavior is optimal given the strategy of the opponent. Further, using computer opponents in an experiment that manipulates cognitive load avoids concerns regarding subjects' beliefs about their opponents' cognitive load.

The cognitive load manipulation is designed to diminish the working memory capacity

[6] We note that Carpenter, Graham, and Wolf (2013) find that the cognitive load manipulation is more effective on subjects with a higher measure of cognitive ability. However, Allred, Duffy, and Smith (2016) do not find such a relationship. Here we find some evidence of such a relationship. In particular, we find that higher measures of cognitive ability are associated with better strategic outcomes but that cognitive load can mitigate this relationship.

[7] Also see Messick (1967), Fox (1972), Shachat and Swarthout (2004, 2012), Coricelli (2005), Levitt, List, and Reiley (2010), Spiliopoulos (2012, 2013), Samson and Kostyszyn (2015), Shachat, Swarthout, and Wei (2015), and Bayer and Renou (2016a).

of subjects and to produce a diminished ability to make computations. Since both of these abilities are important in learning, we aim to determine whether subjects in the high load treatment have less success detecting and exploiting Naive computer strategies, and have less success against Exploitative computer strategies. We therefore compare the payoffs earned by subjects in the cognitive load treatments. In addition, researchers have found that subjects have difficulty detecting and producing random sequences.[8] The ability to mix in a manner consistent with the theoretical predictions would seem to be dependent on the computational ability of the subject because it is a difficult and subtle cognitive task. Therefore, we seek to determine whether actions of subjects in the high load treatment are farther from the theoretical predictions.

To our surprise, we find some evidence that subjects in the high cognitive load treatment earn more than subjects in the low load treatment. On the other hand, we find that subjects in the low load treatment exhibit an increase in earnings across rounds, whereas we do not find such a relationship for subjects in the high load treatment. In addition, we find that the response times of subjects in the low load treatment decrease at a faster rate than the response times of subjects in the high load treatment. We interpret these results as suggesting that subjects in the low cognitive load treatment exhibit a significantly faster rate of learning than do subjects in the high load treatment.

Further, consistent with the previous literature, the behavior in our experiment exhibits mixture proportions and serial correlation that are inconsistent with the theoretical predictions. However, we do not find strong evidence that cognitive load is related to either the mixing proportions or the serial correlation.

The contributions of this paper are as follows. We are the first to attempt to better understand mixing behavior by using the cognitive load manipulation. Our analysis also shows that subjects with fewer cognitive resources do not necessarily exhibit worse performance, particularly in the early rounds, than subjects with more cognitive resources. On the other hand, early round experimentation, which would facilitate learning, can lead to lower payoffs in

[8] For instance, see Wagenaar (1972), Bar-Hillel, and Wagenaar (1991), Rapoport and Budescu (1992), Budescu and Rapoport (1994), Rabin (2002), and Oskarsson et al. (2009).

these rounds. Our analysis suggests that subjects with more cognitive resources exhibit more learning than subjects with fewer cognitive resources. This is consistent with the contention that subjects in the low load treatment have the available cognitive resources to sufficiently remember and analyze previous outcomes. However, the diminished payoffs from the early round experimentation are not compensated by a corresponding increase in payoffs in the latter rounds. We therefore refer to this as *over-experimentation*. Further, we do not find strong evidence that the standard experimental results on mixing (suboptimal mixture proportions and serial correlation) are related to the available cognitive resources of the subject. Our results shed light on the extent to which cognitive resources affect (and do not affect) behavior in games with mixed strategy equilibria.

2 Related literature

There is a large and growing experimental literature that examines the relationship between measures of cognitive ability and strategic behavior.[9] We take a complementary approach in that, rather than measure cognitive ability, we manipulate the subjects' available cognitive resources.

The cognitive load manipulation is well-studied in nonstrategic settings. Cognitive load has been found to make subjects more impulsive and less analytical (Hinson, Jameson, and Whitney, 2003), more risk averse (Whitney, Rinehart, and Hinson, 2008; Benjamin, Brown, and Shapiro, 2013; Gerhardt, Biele, Heekeren, and Uhlig, 2016), more impatient (Benjamin, Brown, and Shapiro, 2013), make more mistakes (Rydval, 2011),[10] exhibit less self control over their actions (Shiv and Fedorikhin, 1999; Ward and Mann, 2000, Mann and Ward, 2007),

[9] See Al-Ubaydli, Jones, and Weel (2016), Ballinger et al. (2011), Baghestanian and Frey (2016), Bayer and Renou (2016a,2016b), Benito-Ostolaza, Hernández, and Sanchis-Llopis (2016), Brañas-Garza, Espinosa, and Rey-Biel (2011), Brañas-Garza, Garcia-Muñoz, and Hernan Gonzalez (2012), Brañas-Garza and Smith (2016), Burks et al. (2009), Burnham et al. (2009), Carpenter, Graham, and Wolf (2013), Chen, Huang, and Wang (2018), Corgnet et al. (2016), Coricelli and Nagel (2009), Devetag and Warglien (2003), Fehr and Huck (2016), Georganas, Healy, and Weber (2015), Gill and Prowse (2016), Grimm and Mengel (2012), Jones (2014), Jones (2008), Kiss, Rodriguez-Lara, and Rosa-García (2016), Lohse (2016), Palacios-Huerta (2003b), Proto, Rustichini, and Sofianos (2018), Putterman, Tyran, and Kamei (2011), Rydval (2011), Rydval and Ortmann (2004), and Schnusenberg and Gallo (2011).

[10] Drichoutis and Nayga (2018) find that high cognitive load does not increase internal inconsistency on a GARP budget allocation task. Lee, Amir, and Ariely (2009) find that subjects under a high load make fewer intransitive choices.

fail to process available and relevant information (Gilbert, Pelham, and Krull, 1988; Swann et al., 1990), more susceptible to anchoring effects (Epley and Gilovich, 2006), perform worse on gambling tasks (Hinson, Jameson, and Whitney, 2002), perform worse on visual judgment tasks (Morey and Cowan, 2004; Allen, Baddeley, and Hitch, 2006; Morey and Bieler, 2013; Zokaei, Heider, and Husain, 2014; Allred, Crawford, Duffy, and Smith, 2016), offer different allocation decisions (Cornelissen, Dewitte, and Warlop, 2011; Schulz et al., 2014),[11] give different evaluations of the fairness of outcomes (van den Bos et al., 2006), less dishonest (van't Veer, Stel, and van Beest, 2014), and more influenced by visual salience (Milosavljevic, Navalpakkam, Koch, and Rangel, 2012).[12]

While many cognitive load studies occur in individual decision settings, only a few involve strategic settings, and none entail the study of mixing behavior. To our knowledge, studies of cognitive load in strategic settings only include Milinski and Wedekind (1998), Roch et al. (2000), Cappelletti, Güth, and Ploner (2011), Carpenter, Graham, and Wolf (2013), Duffy and Smith (2014), Samson and Kostyszyn (2015), Allred, Duffy, and Smith (2016), and Buckert, Oechssler, and Schwieren (2017).

We note that Milinski and Wedekind (1998), Duffy and Smith (2014), and Buckert, Oechssler, and Schwieren (2017) study the effect of cognitive load on behavior in repeated game settings. The authors find that the cognitive load affects the ability of subjects in repeated games to employ information from previous repetitions.

With the exception of Carpenter et al. (2013) and Samson and Kostyszyn (2015), the previous literature on cognitive load in strategic settings describe experiments where the subjects are placed under a cognitive load and play against a human opponent, who is either under a cognitive load or not. One of the drawbacks of conducting a cognitive load experiment in a strategic setting with a human opponent is that the subjects' beliefs about the distribution of the cognitive load of the opponents and their beliefs about the effect of the cognitive load on their opponents are not well specified and are difficult to measure. A design such as ours, which employs a computer opponent, can address this critique. Further, it allows us to ob-

[11] Although Hauge et al. (2016) does not find an effect.

[12] Deck and Jahedi (2015) study several effects at a time and find that subjects under a cognitive load are less patient, more risk averse, perform worse on arithmetic tasks, and are more prone to anchoring effects.

serve the effect of cognitive load on subjects playing against a small set of distinct varieties of opponent strategies.

3 Experimental design

3.1 Hide and seek game

Subjects play a repeated, deterministic version of the zero-sum, "hide and seek" game (Rosenthal, Shachat, and Walker, 2003) against a computer opponent while under an experimentally manipulated cognitive load.

Subjects select either "Up" or "Down" as the "Evader" and the computer selects either Up or Down as the "Pursuer." Subjects always play as the row player.[13] If the computer correctly guesses the subject's choice then the subject earns 0. The payoff to the two outcomes characterized as successful evasion ({Up,Down} and {Down,Up}) are unequal, with one yielding a payoff of 2 points to the Evader, and the other 1 point. To mitigate concerns about order or presentation effects, sessions were conducted in which both successful evasion outcomes yield the higher payoff. Both of the corresponding payoff matrices are presented in Figure 1. Roughly half of the subjects played the version on the left.

<table>
<tr><td></td><td></td><td colspan="2">Pursuer</td></tr>
<tr><td></td><td></td><td>Up</td><td>Down</td></tr>
<tr><td>Evader</td><td>Up</td><td>0</td><td>1</td></tr>
<tr><td></td><td>Down</td><td>2</td><td>0</td></tr>
</table>

<table>
<tr><td></td><td></td><td colspan="2">Pursuer</td></tr>
<tr><td></td><td></td><td>Up</td><td>Down</td></tr>
<tr><td>Evader</td><td>Up</td><td>0</td><td>2</td></tr>
<tr><td></td><td>Down</td><td>1</td><td>0</td></tr>
</table>

Figure 1: Both versions of the hide and seek game, where Evader payoffs are provided

Each point corresponded to $1.50. The computer's actions were presented in red, and the subject's actions and payoffs were presented in blue. We provide a screenshot in Figure 2.

<<Figure 2 here>>

[13] In Rosenthal, Shachat, and Walker (2003), actions were labeled "Left" and "Right" and the roles as both row and column were played by subjects.

To simplify the analysis that follows, we recode the data from both treatments to correspond to the game on the left, where successfully evading with Down earns 2. Subjects play 100 repetitions of the same version of the game. Following each game stage, subjects receive feedback for that period, including their action, the action of the computer opponent, and the amount of points earned in that period.

3.2 Theoretical predictions

Subject could guarantee a minimum expected payoff of $\frac{2}{3}$ by playing the minimax strategy: randomly and independently playing Up with probability $\frac{2}{3}$ and Down with probability $\frac{1}{3}$. In the zero-sum version of the game, the Pursuer would seek to keep the subject at the minimum expected payoff, and this would be accomplished by playing Up with probability $\frac{1}{3}$ and Down with probability $\frac{2}{3}$. We refer to this strategy of the Pursuer as the *minimax strategy*.[14]

3.3 Computer opponent strategies

Subjects are randomly allocated into computer opponent strategy treatments. There are two Naive computer strategies. These are strategies that, once detected, can be exploited by subjects in a straightforward manner. One of these Naive computer strategies mixes between Up and Down with equal probability.[15] We refer to this as the *Naive* $50-50$ strategy. The best response to this strategy is to play Down in each period.[16]

The other Naive computer strategy mixes with the overall frequency corresponding to the theoretical predictions, but in the deterministic pattern of Up-Down-Down-(repeat). We refer to this as the *Naive Pattern* strategy. The best response to this strategy is to play Down-Up-Up-(repeat).

There are two Exploitative computer strategies. One plays the first 5 periods according to the minimax strategy: Up with probability $\frac{1}{3}$ and Down with probability $\frac{2}{3}$. Then in the

[14] These strategies are identical to the equilibrium of the zero-sum version of the game.

[15] This computer opponent strategy also appears in Shachat and Swarthout (2004) and Levitt, List, and Reiley (2010).

[16] In Shachat and Swarthout (2004), subjects played against computer opponents that selected their actions according to a fixed probability distribution. The 50-50 mixture is near the threshold where their subjects notice that their opponent is not playing optimally. We also note that their experiment did not contain cognitive load treatments and their subjects played against this strategy for 200 rounds.

remaining 45 repetitions, the computer plays the minimax strategy with probability 0.5 and with probability 0.5 selects the action that would have minimized the subject's payoffs given the proportion of the subject's previous 4 decisions.[17] We refer to this as the *Exploitative Mix* strategy.

The other Exploitative strategy also begins playing the minimax strategy for the first 5 periods, then in the remaining 45 repetitions plays the minimax strategy with probability 0.5 and seeks to exploit the Win-Stay-Lose-Shift tendency[18] with probability 0.5. In particular, if a subject displays behavior consistent with the Win-Stay-Lose-Shift strategy in 2 or 3 of the previous 3 decisions then the computer selects the action that minimizes the subject's payoffs anticipating the Win-Stay-Lose-Shift strategy. On the other hand, if a subject exhibits behavior consistent with the Win-Stay-Lose-Shift strategy in 0 or 1 of the previous 3 decisions then the computer selects the action that minimizes the subject's payoffs anticipating the Win-Shift-Lose-Stay strategy. We refer to this as the *Exploitative WSLS* strategy.

Each subject plays 50 consecutive rounds against a Naive computer strategy and 50 consecutive rounds against an Exploitative computer strategy. With probability 0.5 subjects first play against a Naive computer opponent. Subjects are not informed about whether their opponent strategy is Naive or Exploitative.

In order to strike a balance between revealing too little to the subjects and too much to the subjects, we told them the following about the computer strategies: Before the first period, subjects were told, "How does the computer decide what to play? A number of possible strategies have been programmed. Some computer strategies can be exploited by you. Some computer strategies are designed to exploit you. One of these possible strategies has been selected for the first 50 periods." After the first 50 periods, subjects were told, "The computer strategy from the first 50 periods is definitely not the same as that in the second 50 periods."

[17] A similar non-stationary computer opponent strategy appears in Levitt, List, and Reiley (2010).

[18] See Imhof, Fudenberg, and Nowak (2007), Spiliopoulos (2013), Wang and Xu (2014), and Wang, Xu, and Zhou (2014).

3.4 Cognitive load treatments

Before each repetition of the game, a cognitive load is imposed on subjects by directing them to remember a number. Subjects in the *low* cognitive load treatment are required to remember a one-digit number that ranges from 1 to 9. Subjects in the *high* cognitive load treatment are required to remember a six-digit number that ranges from 100000 to 999999. Each number is independently drawn with replacement from a uniform distribution on the specified range.

In both treatments, a new number is given for each period. After playing an iteration of the game and receiving feedback, subjects are asked for the number. Subjects play 50 consecutive repetitions in the high load treatment and 50 consecutive repetitions in the low load treatment. In this sense, cognitive load is a within-subject manipulation. With probability 0.5 subjects play first in the high load treatment. Subjects are not given feedback about their performance on the memorization tasks.

3.5 Incentivization scheme

Each subject earns a $5 show-up fee. Additional payments are designed to decouple the material incentives from the game in any period with material incentives from the memorization task in that period. Subjects complete 100 repetitions of the game and 100 memorization tasks. Those who correctly complete all 100 memorization tasks are paid for 30 randomly selected game outcomes, those who correctly complete 99 are paid for 29, those who correctly complete 98 are paid for 28, and so on, until subjects who correctly complete 70 or fewer memorization tasks are not paid for any of the game outcomes.

3.6 Experimental procedure

At the start of every period, subjects were given 15 seconds to commit a number to memory[19] then proceeded to the game.[20] After receiving feedback on the game[21] they were asked for the

[19] Subjects could click to proceed to the next stage but after 15 seconds would proceed automatically.

[20] Subjects were given 20 seconds to reflect on their action in the game. They could click to proceed to the next stage but there was no penalty for not responding before the 20 seconds elapsed.

[21] Subjects were given 20 seconds to reflect on the game feedback. They could click to proceed to the next stage but after 20 seconds would proceed automatically.

memorization number. Finally, subjects were informed of the number of periods completed (out of 50) under the computer opponent strategy and cognitive load treatments.[22]

Prior to the incentivized games and memorization tasks, subjects were given an unincentivized test of their understanding of the hide and seek game. Specifically, they were asked to report the number of points that they would earn for all 4 combinations of own actions and computer actions. They received feedback on these responses. In addition, they were given an unincentivized opportunity to memorize a six-digit number and an unincentivized opportunity to memorize a one-digit number. Unlike the incentivized portion of the experiment, subjects were given feedback about their performance on these memorization tasks.

After completing the incentivized portion of the experiment, subjects reported their gender, whether they were an economics major, whether they have taken a game theory course, an optional estimate of their grade point average[23] (GPA), and a rating of the difficulty in recalling the large and the small memorization numbers. These difficulty ratings were elicited on a scale of 1 ("Very Difficult") to 7 ("Not Very Difficult"). After these questions were completed, subjects learned their earnings. Subsequently, the experimenter took an image of the right hand of the subjects with a digital scanner[24] and then they were paid in cash.[25]

A total of 130 subjects participated in the experiment. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).[26] Of the 130 subjects, 78 were students at Rutgers University-Camden and 52 were students at Haverford College.[27] There were 13 sessions conducted at Camden and 3 at Haverford. None of the 16 sessions lasted longer than 60 minutes and the subjects earned an average of $33. Table 1 lists the distribution of subjects within the cognitive load and computer opponent treatments.[28]

[22] Subjects could click to proceed to the next period but would automatically do so after 10 seconds.

[23] Grade point average ranges from 0.0 to 4.0, and is increasing in performance.

[24] We employed a Canon CanoScan 4507B002 LiDE110 Color Image Scanner. We report the analysis of this data in Duffy et al. (2018).

[25] Screenshots of the entire set of instructions are available at https://osf.io/bha7c/.

[26] The z-Tree code is available at https://osf.io/bha7c/.

[27] The Haverford subjects were recruited via ORSEE (Greiner, 2015).

[28] See Table $A1$, in the Supplemental Online Appendix, for more on the distribution in the first and second blocks of 50 rounds.

Table 1: Distribution of subjects within treatments

	High load	Low load	Total
Naive 50-50	32	37	69
Naive Pattern	22	39	61
Expl. WSLS	37	27	64
Expl. Mix	39	27	66
Total	130	130	260

We list the distribution of subjects across cognitive load and computer opponent treatments. There are a total of 260 observations because each of the 130 subjects played against both a Naive and an Exploitative computer opponent strategy and played a block in the high cognitive load treatment and a block in the low cognitive load treatment.

3.7 Discussion of the experimental design

We employ only 4 strategies for the computer opponent. As a result, we can compare strategic behavior against a few well-defined opponent strategies. In the case of human opponents, the number of these strategies would either be unknown or only vaguely defined. Further, we allude to the Naive and Exploitative strategies in the description to the subjects because these communicate the range of possible computer strategies. Moreover, this description increases the plausibility of the claim that the computer strategy is not identical in both blocks of 50 rounds.

In the Exploitative computer strategies, the computer attempts exploitation with probability 0.5, rather than with certainty. A deterministic strategy could be detected by some subjects, rendering this strategy easier to exploit.

Given the Exploitative strategies, the reader might ask why the subject should mix. The answer, applicable to all strategies of the opponent, is that mixing according to minimax is the only way to avoid being exploited and attaining an expected payoff of less than $\frac{2}{3}$. Further, any concerns about incentives to mix against our Exploitative strategies also apply to settings with human opponents.

Recall that subjects memorize a different number every period, keeping cognitive load relatively constant across periods. On the other hand, memorizing only a single number across several periods could produce a non-constant load, as the number could be rehearsed across periods. It is not known whether this taxing of cognitive resources would increase or decrease across periods, however it would likely not be constant. We also note that we did not exclude repeated numbers or numbers that we considered to be easier to remember than other numbers.

Computer strategies, both Naive and Exploitative, are only vaguely described to subjects as "...can be exploited by you..." and "...designed to exploit you..." We acknowledge that the meaning of these sentences may be ambiguous to some subjects. However, it is important to the design of the experiment so that subjects attempt to discern a strategy on the part of the computer, and direct cognitive resources towards this goal. Our language clarifies that a motivation can be discerned. Language that more precisely specifies the four strategies has two potential drawbacks. First, anything specific about the strategies themselves would render all four of them too easy to exploit. Second, wording that pins down the computer's game-theoretic motivation[29] could exacerbate differences between subjects who are familiar with game theory or economics, and those who are not. Our description was selected because it clarifies that subjects should try to learn something about the computer opponent without being specific about its precise nature.

The goals of our incentive scheme are as follows: strongly incentivize the memorization task, keep incentives for memorization in each period independent from incentives for the game decision in that particular period, and maintain identical game decision incentives for high and low load memorization periods. Our solution to this is to not provide feedback on the memorization task and to pay a number of randomly selected game outcomes that is decreasing in the number of incorrect memorization tasks. Only 1 subject out of 130 failed to correctly perform at least 70 memorization tasks, suggesting that the incentive scheme was properly calibrated. In addition, as feedback was not given on the memorization task, it is not clear whether subjects realized that they were near or below 70 correct. Finally, while incorrectly answering a memorization task decreases incentives, this affects high and low load trials equally and we are primarily interested in the difference between these treatments.

[29] For example, the instructions could have explained "Some computer opponents are programmed to detect and best-respond to your strategy."

The reader might worry about the design feature that we allowed subjects to affect the timing of when they proceed to the next stage. In particular, the reader might argue that it would have been preferable to force the subjects to remain in the game stage for a fixed period of time.

First, it is not clear what the ideal compelled response time would be. Out of 13,000 game decisions, only 612 (4.7%) took longer than 3 seconds, only 360 (2.8%) took longer than 4 seconds, only 220 (1.7%) took longer than 5 seconds, and only 42 (0.3%) took longer than 10 seconds. Either the compelled response time would be fast and would greatly affect behavior or it would not be fast and would only affect a small number of decisions.

Further, if a response time is compelled, it is not clear how the game should proceed if the subject does not offer a response within the window provided. Does the program randomly select an action for the subject? Does the computer opponent regard this selection as the same as if it was selected by the subject? Would the subject treat this round differently than a round where the action was selected by the subject? In a sequence of one-shot games, these matters are not important. However, in a game such as ours, with a repeated nature, these matters are important. In the end, it is our opinion that allowing subjects to proceed at their own pace is the best design.

Finally, we load subjects during the feedback stage because we want to leave subjects with less unloaded time for deliberation. For instance, during the feedback stage, subjects could simply decide on the action for the subsequent period. This would circumvent the load treatment during the decision stage.

4 Results

4.1 Summary statistics

We begin with the summary statistics of the main variables of interest.[30] Correct is a dummy variable indicating that the memorization task is correctly completed, Down is a dummy indicating that the Down action is selected, and Earnings is the amount earned in a particular

[30] The dataset is available at https://osf.io/bha7c/.

game outcome: 0, 1, or 2 points. Female, Economics, and Game Theory are dummies indicating gender, that the subject was an economics major, and that the subject reported having taken a game theory course. GPA refers to the subject's self-reported grade point average. Table 2 lists the means of these variables and Table 3 lists the mean Earnings and Down within the cognitive load and computer opponent treatments.

Table 2: Summary statistics

	Pooled	High load	Low load
Correct	0.929	0.880	0.979
Down	0.555	0.551	0.558
Earnings	0.733	0.737	0.730
Female	0.531	-	-
Economics	0.169	-	-
Game Theory	0.177	-	-
GPA (optional)	3.365	-	-

The Pooled means for Correct, Down, and Earnings have 13,000 observations. Female, Economics, and Game Theory have 130 observations. GPA has 103 observations. Both the High load and the Low load means have 6500 observations.

Table 3: Earnings and Down within treatments

	High load	Low load	Total
Naive 50-50	0.779***	0.794***	0.787***
	(0.615)	(0.585)	(0.599)
Naive Pattern	0.855***	0.753***	0.790***
	(0.494)	(0.524)	(0.513)
Expl. WSLS	0.707*	0.735**	0.719***
	(0.559)	(0.568)	(0.563)
Expl. Mix	0.664	0.601**	0.638*
	(0.523)	(0.561)	(0.538)

We provide mean Earnings and Down (in parentheses) by treatment. The number of observations within each cell is 50 for every subject in the treatment, as indicated in Table 1. We perform a one-sample t-test about whether Earnings are significantly different than the theoretical prediction of 0.6667. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

We observe that only the Exploitative Mix, high load treatment is not significantly different than the theoretical prediction. The Naive-Exploitative strategy treatments are successful in that Earnings against Naive opponents is larger than that against Exploitative opponents, according to a Mann-Whitney test $Z = -7.56$, $p < 0.001$. Table 4 reports the Spearman correlation coefficients.

Table 4: Spearman non-parametric correlation coefficients

	1	2	3	4	5	6
1 Correct	1.00					
2 Down	0.011	1.00				
3 Earnings	0.008	0.034***	1.00			
4 Female	−0.014	−0.004	−0.023**	1.00		
5 Economics	−0.008	−0.005	0.005	−0.316***	1.00	
6 Game Theory	−0.007	−0.018*	0.012	−0.210*	0.436***	1.00
7 GPA (optional)	0.077***	0.007	0.025**	0.076	−0.234*	−0.065

Each correlation between variables 1, 2, or 3, and variables 4, 5, or 6 has 13,000 observations. Each correlation between variables 1, 2, or 3, and variable 7 has 10,3000 observations. Each correlation between variables 4, 5, or 6, and variable 7 has 103 observations. Each correlation among variables 1, 2, and 3 has $13,000$ observations. Each correlation among variables 4, 5, and 6 has 130 observations. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

We observe that higher GPA subjects tend to earn more and are more likely to correctly perform the memorization task. While we do not observe a relationship between Earnings and either Economics or Game Theory, we observe a negative relationship between Earnings and Female. Finally, we do not observe a correlation between Earnings and either Economics or Game Theory. This suggests that our description of the computer strategies did not advantage those with exposure to the study of games.[31]

We define Round to be the number of periods under a particular computer opponent treatment and cognitive load treatment. Therefore, Round ranges from 1 to 50. Figures 3 and 4 demonstrate Earnings and Down across Rounds.

<<Figures 3 and 4 Here>>

The high load memorization tasks are correct (5718 of 6500) with a significantly lower frequency than the low load memorization tasks (6362 of 6500), according to a Mann-Whitney

[31] Apparently, being an economics major is not good for one's GPA: we note a negative relationship between GPA and Economics.

test $Z = -22.03$, $p < 0.001$. This suggests that our cognitive load manipulation is successful[32] although it should be noted that the success rate is relatively high in the high load treatment. As each of the 130 subjects attempt 50 high load memorization tasks and 50 low load memorization tasks, Table 5 presents a characterization of the subject-level distribution of the number of correct memorization tasks by cognitive load treatment and the number pooled across treatments.

Table 5: Distribution of subjects by number of correct memorization tasks

	Within blocks of 50							
	$46-50$	$41-45$	$36-40$	$31-35$	$26-30$	$23-25$	< 23	Total
High load	72	30	13	10	2	3	0	130
Low load	125	5	0	0	0	0	0	130

	Across both blocks of 50							
	$96-100$	$91-95$	$86-90$	$81-85$	$76-80$	$71-75$	< 71	Total
Pooled	64	31	16	11	5	2	1	130

The upper panel characterizes the subject-level distribution of the number of correct memorization tasks by cognitive load treatment. The lower panel characterizes the subject-level distribution of the correct memorization tasks across both cognitive load treatments.

Table 5 shows that 111 of the 130 subjects successfully completed more than 85% of their memorization tasks correctly. This suggests that the incentives were sufficient to elicit cognitive effort on these tasks.

Recall that the incentive scheme is designed to keep incentives for memorization in each period independent from incentives for the game decision in that period. Therefore, even if the subject is confident that the memorization task would not be correctly completed in that period, the subject should exert effort in the game stage. Even if the reader acknowledges that the design implements this goal, it is possible that the subjects did not act accordingly.

In order to investigate this possibility, we conduct an analysis of earnings while restricting attention to a specific cognitive load treatment. We use the Correct variable as the independent variable. In some of the regressions below, we include dummy variables identifying the

[32] As a robustness check, we run a repeated measures logistic regression with Correct as the dependent variable and High load as the independent variable. The High load estimate is negative and significant, $t = -13.56$, $p < 0.001$.

computer opponent treatment. In addition, we consider specifications that account for the repeated nature of the observations. In these repeated measures regressions, we estimate an exchangeable covariance matrix, clustered by subject. In other words, we assume a unique correlation between any two observations involving a particular subject. However, we assume that observations involving two different subjects are statistically independent. We also consider specifications that control for Female, Economics, and Game Theory. We refer to this collection of variables as *Demographics*. We also account for self-reported GPA. Recall that a response to GPA was optional and only 103 of 130 subjects provided a response. In the upper panel we report the results restricted to the high load treatment and in the lower panel we report the results from the low load treatment. This analysis is summarized in Table 6.

Table 6: Earnings restricted to cognitive load treatment

	High load				
	(1)	(2)	(3)	(4)	(5)
Correct	0.018	−0.001	−0.007	−0.005	−0.006
	(0.031)	(0.031)	(0.031)	(0.031)	(0.035)
Strategy dummies	*No*	*Yes*	*Yes*	*Yes*	*Yes*
Repeated measures	*No*	*No*	*Yes*	*Yes*	*Yes*
Demographics	*No*	*No*	*No*	*Yes*	*Yes*
GPA	*No*	*No*	*No*	*No*	*Yes*
AIC	15660	15630	15607	15617	12373
Observations	6500	6500	6500	6500	5150
	Low load				
	(6)	(7)	(8)	(9)	(10)
Correct	0.079	0.087	0.072	0.073	0.053
	(0.069)	(0.069)	(0.069)	(0.069)	(0.075)
Strategy dummies	*No*	*Yes*	*Yes*	*Yes*	*Yes*
Repeated measures	*No*	*No*	*Yes*	*Yes*	*Yes*
Demographics	*No*	*No*	*No*	*Yes*	*Yes*
GPA	*No*	*No*	*No*	*No*	*Yes*
AIC	15632	15601	15545	15554	12358
Observations	6500	6500	6500	6500	5150

In the upper panel we report the results restricted to the high load treatment and in the lower panel we report the results from the low load treatment. The repeated measures regressions estimate an exchangeable covariance matrix, clustered by subject. We do not provide the estimates of the intercepts, the individual demographics variables, the covariance estimates, or the strategy dummies. AIC refers to the Akaike information criterion (Akaike, 1974). † denotes $p < 0.1$.

In none of the specifications is there a relationship between Correct and Earnings. This suggests that, even if subjects incorrectly perform the memorization task, there are no significant differences in their earnings.[33]

4.2 Earnings differences

Now we examine differences in earnings by cognitive load treatment. We conduct regressions similar to that in Table 6, but we also include the interactions between High load and the strategy dummies. Including these interactions is attractive in that it does not assume that the cognitive load treatment will have the same effect in each computer opponent treatment. However, these interactions make the interpretation of the High load coefficient more difficult. We therefore include differences in Least Square Means (LSM) estimates[34] of the earnings in the high and low load treatments. This analysis is summarized in Table 7.

Table 7: Regressions of Earnings

	(1)	(2)	(3)	(4)	(5)	(6)
High load	0.007	0.063*	0.069*	0.079*	0.097**	0.422**
	(0.014)	(0.028)	(0.033)	(0.033)	(0.037)	(0.124)
GPA	–	–	–	–	0.046*	0.093**
					(0.023)	(0.029)
GPA*High load	–	–	–	–	–	−0.097**
						(0.035)
Strategy dummies	*No*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Repeated measures	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*
Demographics	*No*	*No*	*No*	*Yes*	*Yes*	*Yes*
AIC	31283.0	31221.7	31199.0	31212.3	24767.3	24764.7
Observations	13,000	13,000	13,000	13,000	10,300	10,300
LSM Difference:						
High load-Low load	0.007	0.030*	0.030*	0.029*	0.047**	0.048**
	(0.014)	(0.014)	(0.014)	(0.014)	(0.016)	(0.016)

The repeated measures regressions estimate an exchangeable covariance matrix, clustered by subject. We do not provide the estimates of the intercepts, the individual demographics variables, the covariance estimates, the strategy dummies, or the High load-strategy dummies interactions. AIC refers to the Akaike

[33] In order to address the concern that the Correct variable does not sufficiently capture whether the subject knows that the memorization task will be incorrect, we offer a different specification in Table $A2$ in the Supplementary Online Appendix. Our results are robust to this specification.

[34] The earnings estimates at the population means.

information criterion (Akaike, 1974). LSM refers to the Least Square Means. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

In each specification that includes strategy dummies, we find that subjects in the high cognitive load treatment earn a significantly larger amount than subjects in the low load treatment.[35] We also find a positive relationship between GPA and Earnings. However, the estimate of the GPA-High load interaction suggests the positive relationship between GPA and Earnings is driven by subjects in the low load treatment.[36,37]

4.3 Earnings differences across rounds

To better understand the evidence for higher earnings in the high cognitive load treatment, we now consider the trajectory of earnings across rounds. We define Second half to be a dummy variable indicating whether the round was in the second half of the block of 50. In other words, for the first 25 rounds within the block of 50, the Second half variable is 0, and 1 otherwise. Other than the inclusion of the Second half variable and the interaction of Second half and High load, the analysis is identical to that summarized in Table 7. As we are interested in the differences across rounds, we provide the LSM estimates of the differences in the estimates in the second half and the first half of rounds, for both cognitive load treatments. We also provide the LSM estimates of the differences in the high and the low cognitive load treatments for both the first half and the second half of rounds. This analysis is summarized in Table 8.

[35] Table $A3$ in the Supplimentary Online Appendix performs regression (3) but restricted by computer opponent treatment.

[36] The reader might worry about whether high cognitive load in the first block of 50 affects behavior in the second block. We therefore supplemented the regressions in Table 7 with a First block dummy variable and its interaction with High load. We find that First block dummy is positive and significant at 0.05, and we find that the interaction is not significant. This perhaps indicates that subjects experienced fatigue in the second block, regardless of the load in the first block. This analysis is available from the corresponding author upon request.

[37] In regressions $(4) - (6)$ we note that neither the Economics nor the Game Theory variables are significant. This suggests that our description of the computer strategies did not advantage those with an exposure to the study of games.

Table 8: Regressions of Earnings across rounds

	(1)	(2)	(3)	(4)	(5)	(6)
High load	0.043*	0.099**	0.105**	0.115**	0.130**	0.455***
	(0.01998)	(0.032)	(0.036)	(0.036)	(0.040)	(0.125)
Second half	0.054**	0.054**	0.054**	0.054**	0.050*	0.050*
	(0.020)	(0.020)	(0.020)	(0.020)	(0.022)	(0.022)
Second half*High load	−0.072*	−0.072*	−0.072*	−0.072*	−0.066*	−0.066*
	(0.028)	(0.028)	(0.028)	(0.028)	(0.032)	(0.032)
GPA	–	–	–	–	0.046*	0.093**
					(0.023)	(0.029)
GPA*High load	–	–	–	–	–	−0.097**
						(0.035)
Strategy dummies	*No*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Repeated measures	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*
Demographics	*No*	*No*	*No*	*Yes*	*Yes*	*Yes*
AIC	31287.0	31225.6	31202.8	31216.2	24773.4	24770.7
Observations	13,000	13,000	13,000	13,000	10,300	10,300
LSM Differences:						
High load-Low load	0.043*	0.066**	0.066**	0.065**	0.080***	0.081***
for First half	(0.020)	(0.020)	(0.020)	(0.020)	(0.023)	(0.023)
High load-Low load	−0.029	−0.006	−0.006	−0.007	0.014	0.015
for Second half	(0.020)	(0.020)	(0.020)	(0.020)	(0.023)	(0.023)
Second half-First half	−0.018	−0.018	−0.018	−0.018	−0.017	−0.017
for High load	(0.020)	(0.020)	(0.020)	(0.020)	(0.022)	(0.022)
Second half-First half	0.054**	0.054**	0.054**	0.054**	0.050*	0.050*
for Low load	(0.020)	(0.020)	(0.020)	(0.020)	(0.022)	(0.022)

The repeated measures regressions estimate an exchangeable covariance matrix, clustered by subject. We do not provide the estimates of the intercepts, the individual demographics variables, the covariance estimates, the strategy dummies, or the High load-strategy dummies interactions. AIC refers to the Akaike information criterion (Akaike, 1974). LSM refers to the Least Square Means. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

Within the first half of rounds, Earnings are higher in the high cognitive load treatment than in the low cognitive load treatment. However, there is no significant difference between these in the second half of rounds. We also observe that the subjects in the low load treatment exhibit higher Earnings in the second half than in the first half of rounds. On the other hand, in the high cognitive load treatment, Earnings in the second half are lower than those in the first half, although this difference is not significant.[38,39]

These results are consistent with the claim that subjects in the low treatment are experimenting in early rounds and benefiting from this experimentation in the latter rounds. By contrast, Earnings in the high cognitive load treatment do not indicate that the subjects experimented. However, it seems as if the experimentation in the low load treatment is excessive in that the early round diminished earnings associated with the experimentation are not sufficiently compensated by improved earnings in the latter rounds. We consider this to be evidence of over-experimentation.

4.4 Differences in response time

Research finds a positive relationship between the time spent deciding on a choice and the difficulty of the choice.[40] In other words, decisions where one option is clearly better than the others tends to take less time than decisions where this is not the case. We define the Response time to be the time spent deliberating on the decision. Therefore, a higher value indicates more time spent on the game decision.

Somewhat surprisingly, in the Naive opponent treatments, subjects in the high load treatment (mean = 0.704, SD = 0.470) have longer Response times than subjects in the low load treatment (mean = 0.627, SD = 0.496), according to a Mann-Whitney test, $Z = 6.206$, $p < 0.001$. However, in the Exploitative opponent treatments, subjects in the high load treatment (mean = 0.589, SD = 0.476) have shorter Response times than subjects in the low load treatment (mean = 0.728, SD = 0.508), according to a Mann-Whitney test, $Z = 11.288$, $p < 0.001$.

In order to better understand the analysis of Earnings across rounds, here we study Response time across rounds. Figure 5 illustrates Response time across rounds.

<<Figure 5 here>>

[38] Table $A4$ in the Supplimentary Online Appendix performs regression (3) but restricted by computer opponent treatment.

[39] See Tables $A5$ and $A6$ in the Supplimentary Online Appendix for a similar analysis but with Round, rather than Second half, as an independent variable.

[40] See Wilcox (1993), Moffatt (2005), Rubinstein (2007), Alós-Ferrer, Granić, Shi, and Wagner (2012), Chen and Fischbacher (2015), and Alós-Ferrer, Granić, Kern, and Wagner (2016).

We run the analogous regressions as summarized in Table 8 but we employ Response time as the dependent variable. Because Response time is constrained to be nonnegative, we perform the analysis by taking the natural log of Response time. We note that Response time ranges from 0 to 29 seconds. However, our z-Tree output lists response times in integers. Presumably, responses in less than 0.5 seconds are reported as 0.[41] In order to avoid taking the log of 0, we add 1 to all values. This analysis is summarized in Table 9.

Table 9: Regressions of the natural log of Response time across rounds

	(1)	(2)	(3)	(4)	(5)	(6)
High load	-0.053^{***}	-0.191^{***}	-0.236^{***}	-0.212^{***}	-0.214^{***}	-0.333^{***}
	(0.012)	(0.019)	(0.047)	(0.048)	(0.056)	(0.084)
Second half	-0.189^{***}	-0.189^{***}	-0.189^{***}	-0.189^{***}	-0.169^{***}	-0.169^{***}
	(0.012)	(0.012)	(0.010)	(0.010)	(0.012)	(0.012)
Second half*High load	0.042^{*}	0.042^{*}	0.042^{**}	0.042^{**}	0.017	0.017
	(0.017)	(0.017)	(0.015)	(0.015)	(0.016)	(0.016)
GPA	–	–	–	–	0.020	0.003
					(0.055)	(0.056)
GPA*High load	–	–	–	–	–	$0.036^{\dagger}$
						(0.019)
Strategy dummies	*No*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Repeated measures	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*
Demographics	*No*	*No*	*No*	*Yes*	*Yes*	*Yes*
AIC	18004.5	17789.0	14504.4	14510.8	11456.0	11458.5
Observations	13,000	13,000	13,000	13,000	10,300	10,300
LSM Differences:						
High load-Low load	-0.053^{***}	-0.049^{***}	-0.053^{***}	-0.053^{***}	-0.027^{*}	-0.027^{*}
for First half	(0.012)	(0.012)	(0.010)	(0.010)	(0.012)	(0.012)
High load-Low load	-0.011	-0.007	-0.011	-0.011	-0.010	-0.010
for Second half	(0.012)	(0.012)	(0.010)	(0.010)	(0.012)	(0.012)
Second half-First half	-0.147^{***}	-0.147^{***}	-0.147^{***}	-0.147^{***}	-0.152^{***}	-0.152^{***}
for High load	(0.012)	(0.012)	(0.010)	(0.010)	(0.012)	(0.012)
Second half-First half	-0.189^{***}	-0.189^{***}	-0.189^{***}	-0.189^{***}	-0.169^{***}	-0.169^{***}
for Low load	(0.012)	(0.012)	(0.010)	(0.010)	(0.012)	(0.012)

The repeated measures regressions estimate an exchangeable covariance matrix, clustered by subject. We do not provide the estimates of the intercepts, the individual demographics variables, the covariance estimates, the strategy dummies, or the High load-strategy dummies interactions. AIC refers to the Akaike information criterion (Akaike, 1974). LSM refers to the Least Square Means. *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

[41] There were 3452 observations with a response time of 0 seconds.

We observe that Response time decreases across rounds for subjects in both the high and low load treatments. Further, subjects in the low cognitive load treatment have a longer Response time in the first half than subjects in the high load treatment. However, there are no significant differences in the second half of rounds. These results are consistent with the interpretation of the analysis summarized in Table 8 that subjects in the low load treatment exhibit a greater amount of learning than subjects in the high load treatment.[42]

Interestingly, we do not find a relationship between Response time and GPA, analogous to that found in Table 8. On the other hand, we find some evidence that higher GPA subjects in the high load treatment exhibit a differentially larger Response time.

Overall the analysis summarized in Table 9 is consistent with the contention that subjects in the low load treatment are learning the strategy of the opponent better than subjects in the high load treatment.

4.5 Mixture proportions

We now test whether the subjects mixed in the proportions as consistent with the theoretical predictions: Up with probability $\frac{2}{3}$ and Down with probability $\frac{1}{3}$. Here we restrict attention to observations against Exploitative computer strategies because there are difficulties interpreting the mixing proportions against the Naive computer strategies.

Similar to Palacios-Huerta and Volij (2008), and Levitt, List, and Reiley (2010), we conduct a binomial χ^2 test on each subject.[43,44] Performing a joint test on the 76 subjects in the high load treatment by summing their test statistics, we reject the hypothesis that, on aggregate, they mix with these proportions, $\chi^2(76, 1) = 1026.22$, $p < 0.001$. We also conduct a joint binomial χ^2 test on 53 the subjects in the low load treatment by summing their test statistics, and again we reject the hypothesis that they mix in proportions as consistent with the theoretical predictions, $\chi^2(53, 1) = 774.08$, $p < 0.001$.

[42] See Table $A7$ in the Supplimentary Online Appendix for a similar analysis but with Round, rather than Second half, as an independent variable.

[43] See the Supplemental Online Appendix for the subject-level data. Note that one subject selected Down in every period and therefore we cannot perform a binomial χ^2 test on this subject.

[44] We note that there does not exist a significant Spearman correlation between the χ^2 statistic and the Female, Game Theory, Economics, and GPA variables.

Next we test the hypothesis that the subjects in the high and low load treatments have identical distributions. We conduct a two-sample Kolmogorov-Smirnov test[45] on the distribution of the individual χ^2 test statistics and we cannot reject the hypothesis that they are identically distributed, $K = 0.164$, $p = 0.37$.[46]

We also test for differences between the treatments using a Mann-Whitney test on the percentage of Down actions against an Exploitative computer opponent. We find that the subjects in the high load treatment (54.05%) had a significantly different mixture than subjects in the low load treatment (56.44%), $Z = 1.910$, $p = 0.056$. However, the difference between high load (53.42%) and low load (56.30%) subjects is not significant when we restrict attention to the final 25 periods of the 50 period block, $Z = 1.622$, $p = 0.105$.

To further explore this, we conduct regressions with Down as the dependent variable. We conduct logistic regressions due to the discrete nature of the variable. We estimate an exchangeable log odds ratio, clustered by subject. In other words, we assume a unique relationship between any two observations involving a particular subject. However, we assume that observations involving two different subjects are statistically independent. The regressions are estimated using Generalized Estimating Equations (GEE). Since GEE is not a likelihood-based method, Akaike's Information Criterion is not available. Therefore, we provide the Quasilikelihood information criterion (QIC), Pan (2001). We run specifications restricted to the Exploitative WSLS treatment, to the Exploitative Mix treatment, and to the Pooled Exploitative treatments. This analysis is summarized in Table 10.

[45] See Gibbons and Chakraborti (1992).

[46] This qualitative result is not changed when we restrict attention to the last 25 rounds of each 50 period block, $K = 0.129$, $p = 0.70$. We also cannot reject that they come from identical distributions when we restrict attention to the Exploitative WSLS treatment ($K = 0.183$, $p = 0.67$) or to the Exploitative Mix treatment ($K = 0.141$, $p = 0.91$).

Table 10: Logistic regressions of Down

	Expl. WSLS		Expl. Mix		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
High load	-0.038	-0.022	-0.152	$-0.253^{\dagger}$	-0.152	-0.174
	(0.150)	(0.149)	(0.123)	(0.150)	(0.123)	(0.134)
Strategy dummies	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Repeated measures	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Demographics	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*
QIC	4403.1	4422.2	4561.6	4545.6	8964.7	8975.4
Observations	3200	3200	3300	3300	6500	6500
LSM Difference:						
High load-Low load	-0.038	-0.022	-0.152	$-0.253^{\dagger}$	-0.095	-0.117
	(0.150)	(0.149)	(0.123)	(0.150)	(0.097)	(0.102)

The repeated measures regressions estimate an exchangeable log odd ratio, clustered by subject. We do not provide the estimates of the intercepts, the individual demographics variables, the log odds estimates, the strategy dummies, or the High load-strategy dummies interactions. QIC refers to the Quasi-likelihood information criterion (Pan, 2001). LSM refers to the Least Square Means. * denotes $p < 0.05$, and † denotes $p < 0.1$.

Here in the Exploitative Mix treatment, we find some weak evidence of differences in Down by cognitive load treatment. However, in the other specifications, we do not find significant differences between the cognitive load treatments.[47]

Therefore, consistent with the previous literature, we find that the subjects do not mix in the proportions as consistent with the theoretical predictions. However, we do not find strong evidence of a significant difference between the mixture proportions of subjects in the high and low load treatments. This suggests that the availability of cognitive resources is not related to the observed deviations from the theoretical predictions of mixture proportions in this game.

4.6 Serial correlation

Next we investigate whether the actions in our data exhibit serial correlation. As in the previous subsection, we restrict attention to observations against an Exploitative opponent

[47] As expected, in the specifications that are restricted to a single opponent treatment, the LSM differences are equal to the High load coefficient estimates. We also note that the High load coefficient estimate of (5) is identical to that in (3) because of the designation as the reference strategy treatment.

because the best response to a Naive opponent is serially correlated.[48]

In order to detect serial correlation, we perform tests of runs, as described in Gibbons and Chakraborti (1992). A run (r) is defined to be a sequence of one or more identical actions followed by a different action or no action at all. Given the number of Up actions (n_U) and the number of Down actions (n_D) selected by a subject, we are able to calculate the probability of observing any feasible number of runs.[49] For every subject, given n_D and n_U we calculate the probability density function of the number of runs:

$$f(r|n_U, n_D) = \begin{cases} \frac{2\binom{n_U-1}{(\frac{r}{2})-1}\binom{n_D-1}{(\frac{r}{2})-1}}{\binom{n_D+n_U}{n_U}} \text{ if } r \text{ is even} \\ \\ \frac{\binom{n_U-1}{\frac{r-1}{2}}\binom{n_D-1}{\frac{r-3}{2}}+\binom{n_U-1}{\frac{r-3}{2}}\binom{n_D-1}{\frac{r-1}{2}}}{\binom{n_D+n_U}{n_U}} \text{ if } r \text{ is odd.} \end{cases}$$

From the density function, we can calculate the cumulative distribution function:

$$F(r|n_U, n_D) = \sum_{k=1}^{r} f(k|n_U, n_D),$$

which is the probability of observing r or fewer runs. Similar to Walker and Wooders (2001), Palacios-Huerta and Volij (2008), and Levitt, List, and Reiley (2010), we calculate two statistics, $F(r-1|n_U, n_D)$ and $F(r|n_U, n_D)$, for each subject.[50] At a 5% level of significance, we would reject the null hypothesis of independence, if either $F(r|n_U, n_D) < 0.025$ or if $1 - F(r-1|n_U, n_D) < 0.025$. Because we plan to run one-sample Kolmogorov-Smirnov tests on these probabilities, as Walker and Wooders (2001), for each subject we take a draw from the uniform distribution with $F(r|n_U, n_D)$ as the upper bound and $F(r-1|n_U, n_D)$ as the lower bound. This leaves us with a single probability estimate for each subject.[51] If the actions are selected independently then these probabilities would be distributed as a uniform between

[48] Optimal behavior against Naive strategies would either have the largest run possible (Naive 50-50) or have runs of 2 followed by runs of 1 (Naive pattern).

[49] Given $n_U > 0$ and $n_D > 0$, there must be at least 2 runs and the maximum possible number of runs is equal to $2 * \min(n_U, n_D) + 1$.

[50] See the Supplemental Online Appendix for the subject-level data.

[51] We note that there does not exist a significant Spearman correlation between these probability estimates, and the Female, Game Theory, Economics, and GPA variables.

0 and 1.

We perform a one-sample Kolmogorov-Smirnov test that the 53 probabilities associated with subjects in the low load treatment are uniformly distributed between 0 and 1. We reject the hypothesis that the probabilities are distributed as a uniform, $K = 0.174$, $p = 0.071$. Figure 6 illustrates the test on subjects in the low load treatment. We also perform a one-sample Kolmogorov-Smirnov test that the 76 probabilities associated with subjects in the high load treatment are uniformly distributed between 0 and 1. Again, we reject the hypothesis, $K = 0.246$, $p < 0.001$. Figure 7 illustrates the test on the subjects in the high load treatment.

<<Figures 6 and 7 about here>>

While subjects in neither cognitive load treatments appear to be mixing in an independent fashion, it remains to be seen whether the distribution associated with subjects in the high load treatment is different from the distribution associated with subjects in the low load treatment. We perform a two-sample Kolmogorov-Smirnov test and we cannot reject the hypothesis that the distributions are identical, $K = 0.158$, $p = 0.42$.[52]

To further investigate the question of serial correlation, we define the Switch from previous variable, which assumes a 1 if the choice of action is identical to that in the previous period, and a 0 otherwise. We conduct regressions with this as the dependent variable. We note that there are only 49 observations per subject in the regressions that follow because the first round of the block does not have a previous response. The regressions are otherwise identical to those summarized in Table 10. This analysis is summarized in Table 11.

[52] This qualitative result remains unchanged when we restrict the analysis to the final 25 rounds in the 50 period block, $K = 0.091$, $p = 0.97$. We also cannot reject that they come from the same distribution when we restrict attention to the Exploitative WSLS treatment ($K = 0.188$, $p = 0.64$). However, we can reject the hypothesis that they are the same in the Exploitative Mix treatment ($K = 0.346$, $p = 0.047$).

Table 11: Logistic regressions of Switch from previous

	Expl. WSLS		Expl. Mix		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
High load	-0.186	-0.202	0.304*	0.394*	0.304*	0.326*
	(0.158)	(0.154)	(0.146)	(0.179)	(0.146)	(0.159)
Strategy dummies	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Repeated measures	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Demographics	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*
QIC	4355.3	4373.5	4481.2	4485.2	8836.6	8854.4
Observations	3136	3136	3234	3234	6370	6370
LSM Difference:						
High load-Low load	-0.186	-0.202	0.304*	0.394*	0.059	0.076
	(0.158)	(0.154)	(0.146)	(0.179)	(0.107)	(0.112)

The repeated measures regressions estimate an exchangeable log odd ratio, clustered by subject. We do not provide the estimates of the intercepts, the individual demographics variables, the log odds estimates, the strategy dummies, or the High load-strategy dummies interactions. QIC refers to the Quasi-likelihood information criterion (Pan, 2001). LSM refers to the Least Square Means. ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

We see that with an Exploitative Mix opponent, subjects in the high load treatment switch their actions with a larger probability than the subjects in the low load treatment. However, this difference is not significant in the Exploitative WSLS treatment or the Pooled analysis. We interpret this as providing weak evidence of differences in serial correlation by cognitive load treatment.

5 Discussion

The experimental literature largely finds that subjects do not mix in the proportions consistent with the theoretical predictions and that actions exhibit serial correlation. We find these features in our data, however we find only weak evidence that they are related to cognitive load. Therefore, we do not find evidence that these standard experimental results on mixing are associated with the available cognitive resources of the subject.

These results are reminiscent of the findings reported in Geng et al. (2015). These authors do not find a relationship between measures of cognitive ability and either the mixture proportions or serial correlation. Although the design of Geng et al. (adolescent subjects, human opponents, measures of cognitive ability) exhibit notable differences from our design (college student subjects, computer opponents, cognitive load manipulation), neither study finds evidence of a relationship between cognition and either mixing proportions or serially correlated actions.

We also find surprising evidence that subjects in the high cognitive load treatment earn more than subjects in the low cognitive load treatment, particularly in the early rounds. This is consistent with the explanation that subjects in the high load treatment employ a simple, stable strategy and subjects in the low load treatment engage in experimentation during those early rounds.

In addition, we find that subjects in the low load treatment experience increased earnings across rounds, while those in the high load treatment do not.[53] An interpretation of this result is that the subjects with greater available cognitive resources exhibit more learning than subjects with less.[54] Our analysis of response time is also consistent with this interpretation. This result has an intuitive appeal because remembering and analyzing previous outcomes would seem to require available cognitive resources. For instance, Hinson, Jameson, and Whitney (2003) find that subjects who are under a cognitive load are more impulsive and less analytical. Our results are also consistent with Milinski and Wedekind (1998), Duffy and Smith (2014), and Buckert, Oechssler, and Schwieren (2017), which report that cognitive load affects the ability of subjects in repeated games to employ information from previous repetitions.

We refer to our results as over-experimentation, since the early round diminished earnings associated with the experimentation are not sufficiently compensated by the improved earnings due to the learning. It is possible that subjects in the low load treatment overestimated the benefit from experimenting or underestimated the cost to early round experimentation.

Whereas our results shed light on the role of cognition affecting (and not affecting) behavior in games with mixed strategy equilibria, we acknowledge that there is much work to be done

[53] Geng et al. (2015) did not study the trajectory of earnings across rounds.

[54] Gill and Prowse (2016) find a similar result, albeit in a different setting. These authors observe that subjects with higher measured cognitive ability exhibit a faster convergence to the equilibrium prediction in a repeated beauty contest game.

on the topic. We leave it to future research to determine whether there is a sufficient number of rounds where subjects under a low load would earn more than subjects under a high load. Also, it is possible that subjects play a computer opponent differently than a human opponent because the computer might be expected to employ a more stable strategy. Further, our design allows time between feedback and the choice in the game where subjects are not under a cognitive load. We are interested to know the effects of imposing a cognitive load during this time. These and other interesting questions are a matter for future research.

References

Akaike, Hirotugu (1974): "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, 19(6), 716–723.

Allen, Richard J., Baddeley, Alan D., and Hitch, Graham J. (2006): "Is the binding of visual features in working memory resource-demanding?" *Journal of Experimental Psychology: General*, 135(2), 298–313.

Allred, Sarah, Crawford, L. Elizabeth, Duffy, Sean, and Smith, John (2016): "Working memory and spatial judgments: Cognitive load increases the central tendency bias," *Psychonomic Bulletin and Review*, 23(6), 1825–1831.

Allred, Sarah, Duffy, Sean, and Smith, John (2016): "Cognitive load and strategic sophistication," *Journal of Economic Behavior and Organization*, 125, 162–178.

Alós-Ferrer, Carlos, Granić, Đura-Georg, Kern, Johannes, and Wagner, Alexander K. (2016): "Preference Reversals: Time and Again," *Journal of Risk and Uncertainty*, 52(1), 65–97.

Alós-Ferrer, Carlos, Granić, Đura-Georg, Shi, Fei, and Wagner, Alexander K. (2012): "Choices and preferences: Evidence from implicit choices and response times," *Journal of Experimental Social Psychology*, 48(6), 1336–1342.

Al-Ubaydli, Omar, Jones, Garett, and Weel, Jaap (2016): "Average player traits as predictors of cooperation in a repeated prisoner's dilemma," *Journal of Behavioral and Experimental Economics*, 64, 50–60.

Azar, Ofer H. and Bar-Eli, Michael (2011): "Do soccer players play the mixed-strategy Nash equilibrium?" *Applied Economics*, 43(25), 3591–3601.

Baghestanian, Sascha and Frey, Seth (2016): "GO Figure: Analytic and Strategic Skills are Separable," *Journal of Behavioral and Experimental Economics*, 64, 71–80.

Bailey, Brett James and McGarrity, Joseph P. (2012): "The Effect of Pressure on Mixed-Strategy Play in Tennis: The Effect of Court Surface on Service Decisions," *International Journal of Business and Social Science*, 3(20), 11–18.

Ballinger, T. Parker, Hudson, Eric, Karkoviata, Leonie, and Wilcox, Nathaniel T. (2011): "Saving behavior and cognitive abilities," *Experimental Economics*, 14, 349–374.

Bar-Eli, Michael, Azar, Ofer H., Ritov, Ilana, Keidar-Levin, Yael, and Schein, Galit (2007): "Action bias among elite soccer goalkeepers: The case of penalty kicks," *Journal of Economic Psychology*, 28(5), 606–621.

Bar-Hillel, Maya, and Wagenaar, Willem A. (1991): "The perception of randomness," *Advances in Applied Mathematics*, 12(4), 428–454.

Batzilis, Dimitris, Jaffe, Sonia, Levitt, Steven, List, John A., and Picel, Jeffrey (2017): "Behavior in Strategic Settings: Evidence from a Million Rock-Paper-Scissors Games," working paper, Harvard University.

Bayer, Ralph and Renou, Ludovic (2016a): "Logical omniscience at the laboratory," *Journal of Behavioral and Experimental Economics*, 64, 41–49.

Bayer, Ralph and Renou, Ludovic (2016b): "Logical abilities and behavior in strategic-form games," *Journal of Economic Psychology*, 56, 39–59.

Benito-Ostolaza, Juan M., Hernández, Penélope, and Sanchis-Llopis, Juan A. (2016): "Are individuals with higher cognitive ability expected to play more strategically?" *Journal of Behavioral and Experimental Economics*, 64, 5–11.

Binmore, Ken, Swierzbinski, Joe, and Proulx, Chris (2001): "Does minimax work? An experimental study," *Economic Journal*, 111(473), 445–464.

Brañas-Garza, Pablo, Espinosa, Maria Paz, and Rey-Biel, Pedro (2011): "Travelers' types," *Journal of Economic Behavior and Organization*, 78, 25–36.

Brañas-Garza, Pablo, Garcia-Muñoz, Teresa and Hernan Gonzalez, Roberto (2012): "Cognitive effort in the Beauty Contest Game," *Journal of Economic Behavior and Organization*, 83(2), 254–260.

Brañas-Garza, Pablo and Smith, John (2016): "Cognitive abilities and economic behavior," *Journal of Behavioral and Experimental Economics*, 64, 1–4.

Brown, James N. and Rosenthal, Robert W. (1990): "Testing the minimax hypothesis: a re-examination of O'Neill's game experiment," *Econometrica*, 58(5), 1065–1081.

Buckert, Magdalena, Oechssler, Jörg, and Schwieren, Christiane (2017): "Imitation under stress," *Journal of Economic Behavior and Organization*, 139, 252–266.

Budescu, David V. and Rapoport, Amnon (1994): "Subjective randomization in one-and two-person games," *Journal of Behavioral Decision Making*, 7(4), 261–278.

Burks, Stephen V., Carpenter, Jeffrey P., Götte, Lorenz and Rustichini, Aldo (2009): "Cognitive Skills Explain Economic Preferences, Strategic Behavior, and Job Attachment," *Proceedings of the National Academy of Sciences*, 106(19), 7745–7750.

Burnham, Terence C., Cesarini, David, Johannesson, Magnus, Lichtenstein, Paul and Wallace, Björn (2009): "Higher cognitive ability is associated with lower entries in a p-beauty contest," *Journal of Economic Behavior and Organization*, 72, 171–175.

Buzzacchi, Luigi and Pedrini, Stefano (2014): "Does player specialization predict player actions? Evidence from penalty kicks at FIFA World Cup and UEFA Euro Cup," *Applied Economics*, 46(10), 1067–1080.

Cappelletti, Dominique, Güth, Werner, and Ploner, Matteo (2011): "Being of two minds: Ultimatum offers under cognitive constraints," *Journal of Economic Psychology*, 32(6), 940–950.

Carpenter, Jeffrey, Graham, Michael and Wolf, Jesse (2013): "Cognitive Ability and Strategic Sophistication," *Games and Economic Behavior*, 80(1), 115–130.

Chen, Chun-Ting, Huang, Chen-Ying and Wang, Joseph Tao-yi (2018): "A Window of Cognition: Eyetracking the Reasoning Process in Spatial Beauty Contest Games," *Games and Economic Behavior*, 111, 143–158.

Chen, Fadong and Fischbacher, Urs (2015): "Cognitive Processes of Distributional Preferences: A Response Time Study," Working paper, University of Konstanz.

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

Coloma, Germán (2007): "Penalty Kicks in Soccer An Alternative Methodology for Testing Mixed-Strategy Equilibria," *Journal of Sports Economics*, 8(5), 530–545.

Corgnet, Brice, Espín, Antonio M., Hernán-González, Roberto, Kujal, Praveen, and Rassenti, Stephen (2016): "To trust, or not to trust: Cognitive reflection in trust games," *Journal of Behavioral and Experimental Economics*, 64, 20–27.

Coricelli, Giorgio (2005): "Strategic interaction in iterated zero-sum games," working paper, University of Southern California.

Coricelli, Giorgio and Nagel, Rosemarie (2009): "Neural correlates of depth of strategic reasoning in medial prefrontal cortex," *Proceeings of the National Academy of Science*, 106 (23), 9163–9168.

Cornelissen, Gert, Dewitte, Siegfried, and Warlop, Luk (2011): "Are Social Value Orientations expressed automatically? Decision making in the dictator game," *Personality and Social Psychology Bulletin*, 37(8), 1080–1090.

Deck, Cary and Jahedi, Salar (2015): "The Effect of Cognitive Load on Economic Decision Making: A Survey and New Experiments," *European Economic Review*, 78, 97–119.

Devetag, Giovanna and Warglien, Massimo (2003): "Games and phone numbers: Do short-term memory bounds affect strategic behavior?" *Journal of Economic Psychology*, 24, 189–202.

Drichoutis, Andreas C. and Nayga, Rodolfo (2018): "Economic rationality under cognitive load," working paper, Agricultural University of Athens.

Duffy, Sean, Miller, Dillon, Naddeo, Joseph, Owens, David, and Smith, John (2018): "Are digit ratios (2D:4D and rel2) related to strategic behavior, academic performance, or the efficacy of the cognitive load manipulation?" working paper, Rutgers University-Camden.

Duffy, Sean and Smith, John (2014): "Cognitive Load in the Multi-player Prisoner's Dilemma Game: Are There Brains in Games?" *Journal of Behavioral and Experimental Economics*, 51, 47–56.

Emara, Noha, Owens, David, Smith, John, and Wilmer, Lisa (2017): "Serial correlation and its effects on outcomes in the National Football League," *Journal of Behavioral and Experimental Economics*, 69, 125–132.

Epley, Nicholas and Gilovich, Thomas (2006): "The anchoring-and-adjustment heuristic: Why the adjustments are insufficient," *Psychological Science*, 17(4), 311–318.

Fehr, Dietmar and Huck, Steffen (2016): "Who knows it is a game? On strategic awareness and cognitive ability," *Experimental Economics*, 19(4), 713–726.

Fischbacher, Urs (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10(2), 171–178.

Fox, John (1972): "The learning of strategies in a simple, two-person zero-sum game without saddlepoint," *Behavioral Science*, 17(3), 300–308.

Gerhardt, Holger, Biele, Guido P., Heekeren, Hauke R., and Uhlig, Harald (2016): "Cognitive Load Increases Risk Aversion," working paper, Humboldt-Universität zu Berlin.

Geng, Sen, Peng, Yujia, Shachat, Jason, and Zhong, Huizhen (2015): "Adolescents, Cognitive Ability, and Minimax Play," *Economics Letters*, 128, 54–58.

Georganas, Sotiris, Healy, Paul J., and Weber, Roberto A. (2015): "On the Persistence of Strategic Sophistication," *Journal of Economic Theory*, 159(A), 369–400.

Gibbons, Jean Dickinson and Chakraborti, Subhabrata (1992): *Nonparametric Statistical Inference*, Marcel Dekker, New York.

Gilbert, Daniel T., Pelham, Brett W., and Krull, Douglas S. (1988): "On Cognitive Busyness: When Person Perceivers Meet Persons Perceived," *Journal of Personality and Social Psychology*, 54(5), 733–740.

Gill, David and Prowse, Victoria (2016): "Cognitive Ability, Character Skills, and Learning to Play Equilibrium: A Level-k Analysis," *Journal of Political Economy*, 126(4), 1619–1676.

Greiner, Ben (2015): "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE," *Journal of the Economic Science Association*, 1(1), 114–125.

Grimm, Veronika, and Mengel, Friederike (2012): "An experiment on learning in a multiple games environment," *Journal of Economic Theory*, 147(6), 2220–2259.

Hauge, Karen Evelyn, Brekke, Kjell Arne, Johansson, Lars-Olof, Johansson-Stenman, Olof, and Svedsäter, Henrik (2016): "Keeping others in our mind or in our heart? Distribution games under cognitive load," *Experimental Economics*, 19(3), 562–576.

Hinson, John M., Jameson, Tina L., and Whitney, Paul (2002): "Somatic markers, working memory, and decision making," *Cognitive, Affective, and Behavioral Neuroscience*, 2 (4), 341–353.

Hinson, John M., Jameson, Tina L., and Whitney, Paul (2003): "Impulsive Decision Making and Working Memory," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(2), 298–306.

Hsu, Shih-Hsun, Huang, Chen-Ying, and Tang, Cheng-Tao (2007): "Minimax play at Wimbledon: Comment," *American Economic Review*, 97(1), 517–523.

Imhof, Lorens A., Fudenberg, Drew and Nowak, Martin A. (2007): "Tit-for-tat or win-stay, lose-shift?" *Journal of Theoretical Biology*, 247(3), 574–580.

Jones, Garett (2008): "Are smarter groups more cooperative? Evidence from prisoner's dilemma experiments, 1959-2003," *Journal of Economic Behavior and Organization*, 68, 489–497.

Jones, Matthew T. (2014): "Strategic Complexity and Cooperation: An Experimental Study," *Journal of Economic Behavior and Organization*, 106, 352–366.

Kiss, H. J., Rodriguez-Lara, I., and Rosa-García, A. (2016): "Think Twice Before Running! Bank Runs and Cognitive Abilities," *Journal of Behavioral and Experimental Economics*, 64, 12–19.

Kovash, Kenneth and Levitt, Steven D. (2009): "Professionals do not play minimax: evidence from major League Baseball and the National Football League," working paper, National Bureau of Economic Research.

Lee, Leonard, Amir, On, and Ariely, Dan (2009): "In search of homo economicus: Cognitive noise and the role of emotion in preference consistency," *Journal of Consumer Research*, 36(2), 173–187.

Levitt, Steven D., List, John A., and Reiley, David H. (2010): "What happens in the field stays in the field: Exploring whether professionals play minimax in laboratory experiments," *Econometrica*, 78(4), 1413–1434.

Lohse, Johannes (2016): "Smart or Selfish-When Smart Guys Finish Nice," *Journal of Behavioral and Experimental Economics*, 64, 28–40.

Mann, Traci and Ward, Andrew (2007): "Attention, Self-Control, and Health Behaviors," *Current Directions in Psychological Science*, 16(5), 280–283.

Martin, Christopher Flynn, Bhui, Rahul, Bossaerts, Peter, Matsuzawa, Tetsuro, and Camerer, Colin (2014): "Chimpanzee choice rates in competitive games match equilibrium game theory predictions," *Scientific Reports*, 4, 5182.

McGarrity, Joseph P. and Linnen, Brian (2010): "Pass or Run: An Empirical Test of the Matching Pennies Game Using Data from the National Football League," *Southern Economic Journal*, 3, 791–810.

Messick, David M. (1967): "Interdependent decision strategies in zero-sum games: A computer-controlled study," *Behavioral Science*, 12(1), 33–48.

Milinski, Manfred, and Wedekind, Claus (1998): "Working memory constrains human cooperation in the Prisoner's Dilemma," *Proceedings of the National Academy of Sciences* 95(23), 13755–13758.

Milosavljevic, Milica, Navalpakkam, Vidhya, Koch, Christof, and Rangel, Antonio (2012): "Relative visual saliency differences induce sizable bias in consumer choice," *Journal of Consumer Psychology*, 22(1), 67–74.

Moffatt, Peter G. (2005): "Stochastic choice and the allocation of cognitive effort," *Experimental Economics*, 8(4), 369–388.

Mookherjee, Dilip and Sopher, Barry (1997): "Learning and decision costs in experimental constant sum games," *Games and Economic Behavior*, 19(1), 97–132.

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

Morey, Candice C. and Bieler, Malte (2013): "Visual short-term memory always requires general attention," *Psychonomic Bulletin and Review*, 20(1), 163–170.

Morey, Candice C. and Cowan, Nelson (2004): "When visual and verbal memories compete: Evidence of cross-domain limits in working memory," *Psychonomic Bulletin and Review*, 11(2), 296–301.

Ochs, Jack (1995): "Games with unique, mixed strategy equilibria: An experimental study," *Games and Economic Behavior*, 10(1), 202–217.

O'Neill, Barry (1987): "Nonmetric test of the minimax theory of two-person zerosum games," *Proceedings of the National Academy of Sciences*, 84(7), 2106–2109.

O'Neill, Barry (1991): "Comments on Brown and Rosenthal's reexamination," *Econometrica*, 59(2), 503–507.

Oskarsson, An T., Van Boven, Leaf, McClelland, Gary H., and Hastie, Reid (2009): "What's next? Judging sequences of binary events," *Psychological Bulletin*, 135(2), 262–285.

Palacios-Huerta, Ignacio (2003a): "Professionals play minimax," *Review of Economic Studies*, 70(2), 395–415.

Palacios-Huerta, Ignacio (2003b) "Learning to Open Monty Hall's Doors," *Experimental Economics*, 6(3), 235–251.

Palacios-Huerta, Ignacio, Olivero, Antonio, Bestmann, Sven, Vila, Jose Florensa, and Apesteguia, Jose (2014): "Mapping Minimax in the Brain," *Beautiful Game Theory: How Soccer Can Help Economics*, Palacios-Huerta, Ignacio, Princeton University Press, Princeton, New Jersey, 58–67.

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

Pan, Wei (2001): "Akaike's information criterion in generalized estimating equations," *Biometrics*, 57(1), 120-125.

Proto, Eugenio, Rustichini, Aldo, and Sofianos, Andis (2018): "Intelligence Personality and Gains from Cooperation in Repeated Interactions," *Journal of Political Economy*, forthcoming.

Putterman, Louis, Tyran, Jean-Robert and Kamei, Kenju (2011): "Public goods and voting on formal sanction schemes," *Journal of Public Economics*, 95, 1213–1222.

Rabin, Matthew (2002): "Inference by Believers in the Law of Small Numbers," *Quarterly Journal of Economics*, 117(3), 775–816.

Rapoport, Amnon and Amaldoss, Wilfred (2000): "Mixed strategies and iterative elimination of strongly dominated strategies: an experimental investigation of states of knowledge," *Journal of Economic Behavior and Organization*, 42(4), 483–521.

Rapoport, Amnon and Amaldoss, Wilfred (2004): "Mixed-strategy play in single-stage first-price all-pay auctions with symmetric players," *Journal of Economic Behavior and Organization*, 54(4), 585–607.

Rapoport, Amnon and Boebel, Richard B. (1992): "Mixed strategies in strictly competitive games: a further test of the minimax hypothesis," *Games and Economic Behavior*, 4(2), 261–283.

Rapoport, Amnon and Budescu, David V. (1992): "Generation of random series in two-person strictly competitive games," *Journal of Experimental Psychology: General*, 121(3), 352–363.

Raven, John C., De Lemos, Marion M., 1990. Standard Progressive Matrices. Oxford, Psychologists Press, Oxford, United Kingdom.

Reed, Derek D., Critchfield, Thomas S., and Martens, Brian K. (2006): "The generalized matching law in elite sport competition: Football play calling as operant choice," *Journal of Applied Behavior Analysis*, 39(3), 281–297.

Roch, Sylvia G., Lane, John A. S., Samuelson, Charles D., Allison, Scott T. and Dent, Jennifer L. (2000): "Cognitive Load and the Equality Heuristic: A Two-Stage Model of Resource Overconsumption in Small Groups," *Organizational Behavior and Human Decision Processes*, 83(2), 185–212.

Rosenthal, Robert W., Shachat, Jason and Walker, Mark (2003): "Hide and seek in Arizona," *International Journal of Game Theory*, 32(2), 273–293.

Rubinstein, Ariel (2007): "Instinctive and cognitive reasoning: A study of response times," *Economic Journal*, 117(523), 1243–1259.

Rydval, Ondrej (2011): "The Causal Effect of Cognitive Abilities on Economic Behavior: Evidence from a Forecasting Task with Varying Cognitive Load," working paper, CERGE-EI.

Rydval, Ondrej and Ortmann, Andreas (2004): "How financial incentives and cognitive abilities affect task performance in laboratory settings: An illustration," *Economics Letters*, 85, 315–320.

Samson, Katarzyna and Kostyszyn, Patrycjusz (2015): "Effects of Cognitive Load on Trusting Behavior – An Experiment Using the Trust Game," *PLoS ONE*, 10(5), 0127680.

Schulz, Jonathan F., Fischbacher, Urs, Thöni, Christian, Utikal, Verena (2014): "Affect and fairness: Dictator games under cognitive load," *Journal of Economic Psychology*, 41, 77–87.

Schnusenberg, Oliver and Gallo, Andrés (2011): "On cognitive ability and learning in a beauty contest," *Journal for Economic Educators*, 11(1), 13–24.

Shachat, Jason M. (2002): "Mixed strategy play and the minimax hypothesis," *Journal of Economic Theory*, 104(1), 189–226.

Shachat, Jason and Swarthout, J. Todd (2004): "Do we detect and exploit mixed strategy play by opponents?" *Mathematical Methods of Operations Research*, 59(3), 359–373.

Shachat, Jason and Swarthout, J. Todd (2012): "Learning about learning in games through experimental control of strategic interdependence," *Journal of Economic Dynamics and Control*, 36(3), 383–402.

Shachat, Jason, Swarthout, J. Todd, and Wei, Lijia (2015): "A hidden Markov model for the detection of pure and mixed strategy play in games." *Econometric Theory*, 31(4), 729–752.

Shiv, Baba and Fedorikhin, Alexander (1999): "Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making," *Journal of Consumer Research*, 26(3), 278–292.

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

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

Swann, William B., Hixon, Gregory, Stein-Seroussi, Alan, and Gilbert, Daniel T. (1990): "The Fleeting Gleam of Praise: Cognitive Processes Underlying Behavioral Reactions to Self-Relevant Feedback," *Journal of Personality and Social Psychology*, 59 (1), 17–26.

Van den Bos, Kees, Peters, Susanne L., Bobocel, D. Ramona, and Ybema, Jan Fekke (2006): "On preferences and doing the right thing: Satisfaction with advantageous inequity when cognitive processing is limited," *Journal of Experimental Social Psychology*, 42, 273–289.

van't Veer, Anna E., Stel, Mariëlle, and van Beest, Ilja (2014): "Limited capacity to lie: Cognitive load interferes with being dishonest," *Judgment and Decision Making*, 9(3), 199–206.

Van Essen, Matt, and Wooders, John (2015): "Blind Stealing: Experience and Expertise in a Mixed-Strategy Poker Experiment," *Games and Economic Behavior*, 91, 186–206.

Wagenaar, Willem A. (1972): "Generation of random sequences by human subjects: A critical survey of literature," *Psychological Bulletin*, 77(1), 65–72.

Walker, Mark and Wooders, John (2001): "Minimax play at Wimbledon," *American Economic Review*, 91(5), 1521–1538.

Wang, Zhijian, and Xu, Bin (2014): "Incentive and stability in the Rock-Paper-Scissors game: an experimental investigation," working paper, arXiv preprint arXiv:1407.1170.

Wang, Zhijian, Xu, Bin, and Zhou, Hai-Jun (2014): "Social cycling and conditional responses in the Rock-Paper-Scissors game," *Scientific Reports*, 4, 5830.

Ward, Andrew and Mann, Traci (2000): "Don't Mind If I Do: Disinhibited Eating Under Cognitive Load," *Journal of Personality and Social Psychology*, 78 (4), 753–763.

Whitney, Paul, Rinehart, Christa A., and Hinson, John M. (2008): "Framing effects under cognitive load: The role of working memory in risky decisions," *Psychonomic Bulletin and Review*, 15(6), 1179–1184.

Wilcox, Nathaniel T. (1993): "Lottery choice: Incentives, complexity and decision time," *Economic Journal*, 103(421), 1397–1417.

Wooders, John (2010): "Does experience teach? Professionals and minimax play in the lab," *Econometrica*, 78(3), 1143–1154.

Zokaei, Nahid, Heider, Maike, and Husain, Masud (2014): "Attention is required for maintenance of feature binding in visual working memory," *Quarterly Journal of Experimental Psychology*, 67(6), 1191–1213.

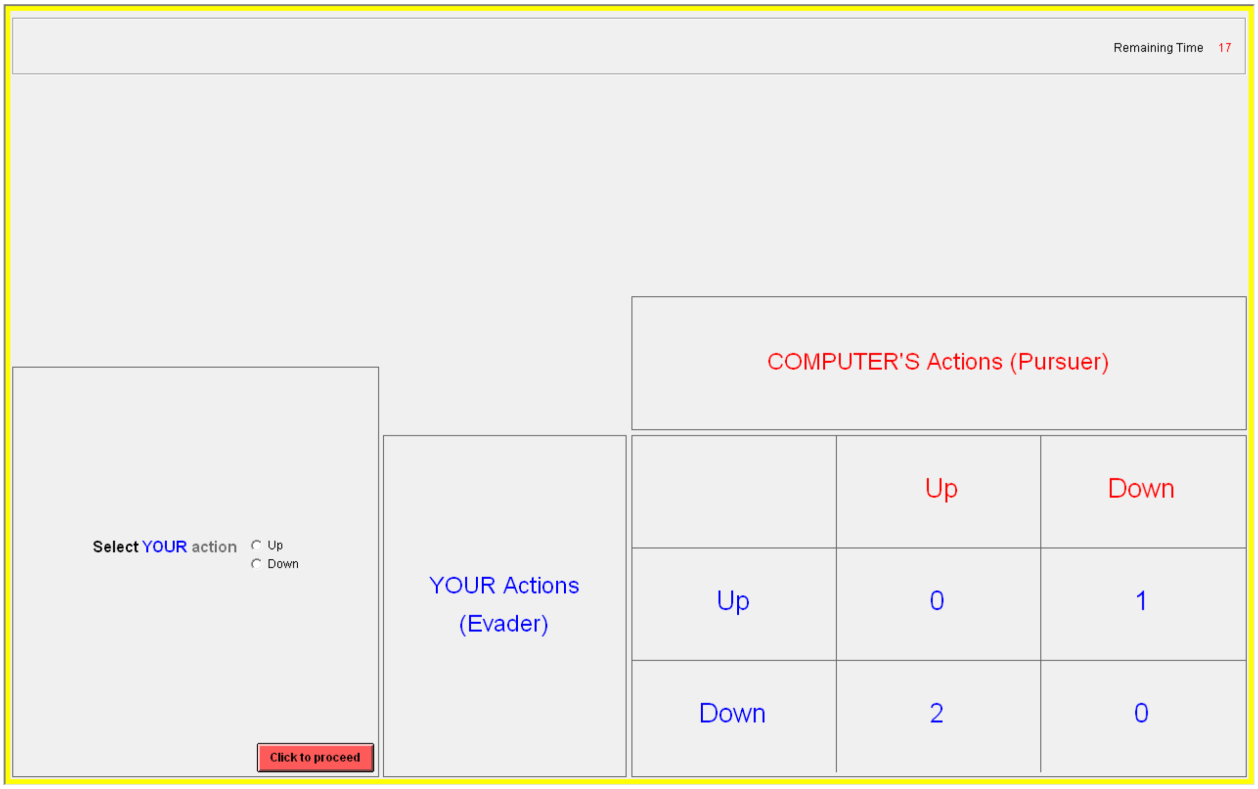

Figure 2: Screenshot of the game stage

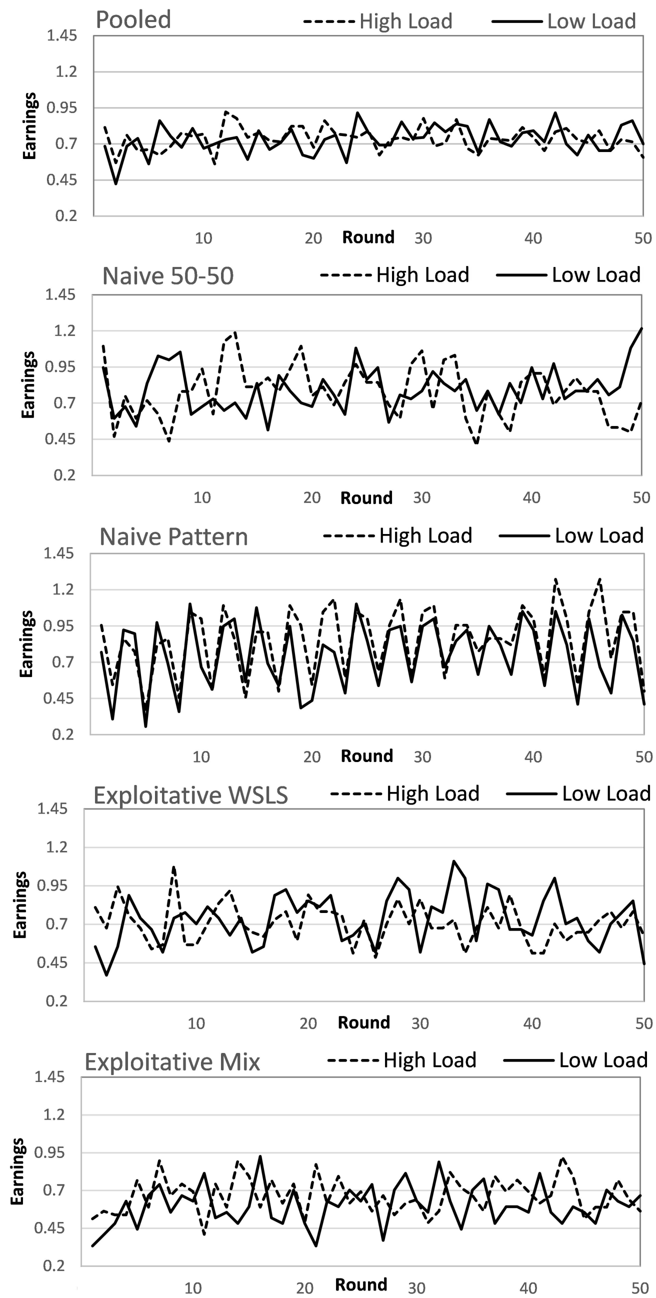

Figure 3: Earnings across rounds

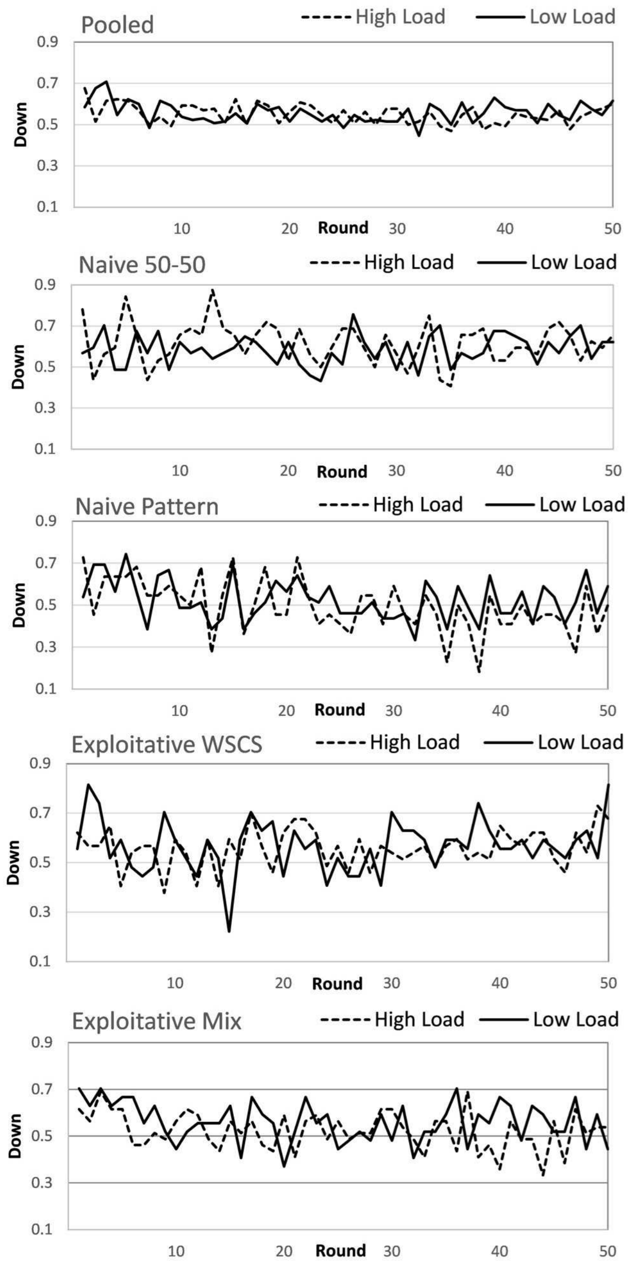

Figure 4: Down across rounds

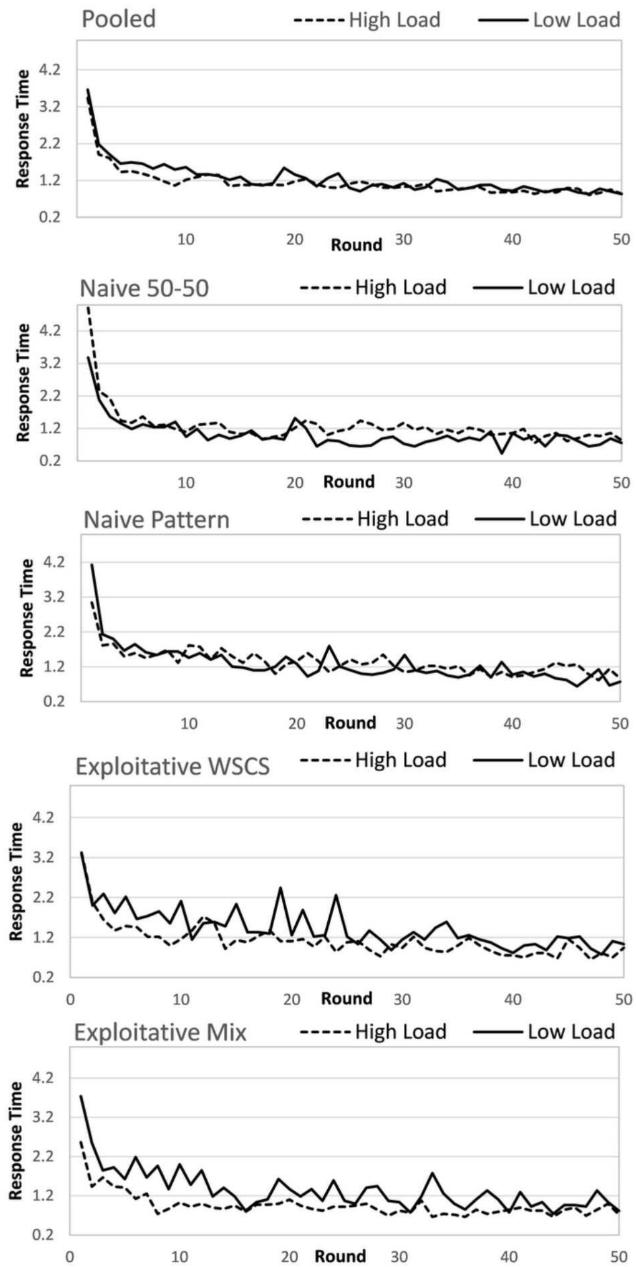

Figure 5: Response times across rounds

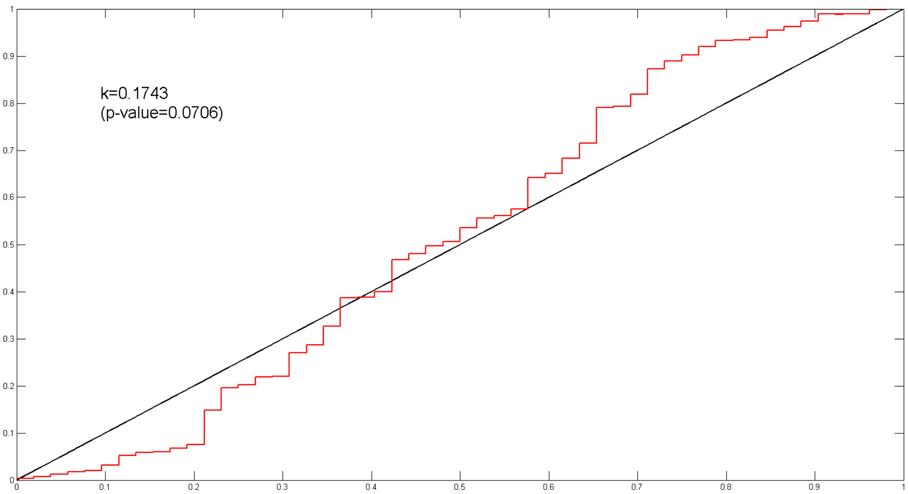

Figure 6: Kolmogorov-Smirnov Test of Runs for subjects in the low load treatment

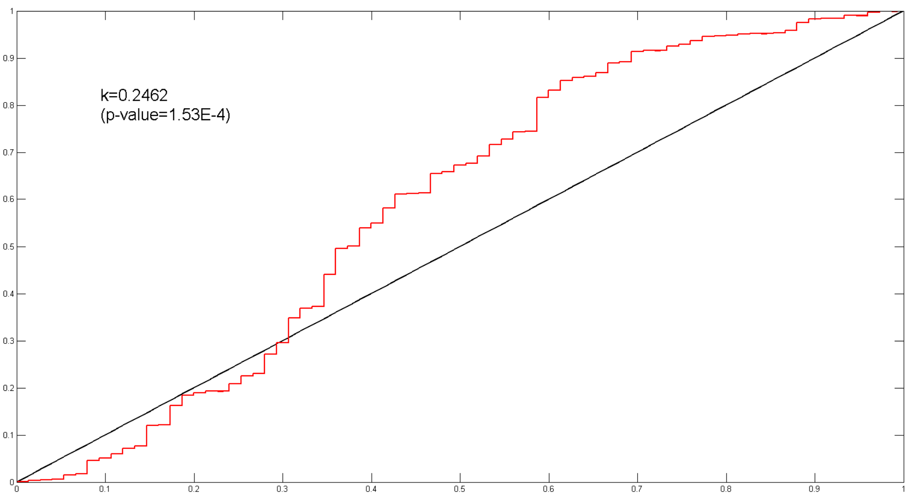

Figure 7: Kolmogorov-Smirnov Test of Runs for subjects in the high load treatment

Supplemental Online Appendix

Table 1 lists the distribution of subjects within the treatments. Table $A1$ also includes the distribution within the first and second blocks of 50 rounds.

Table $A1$: Distribution of subjects within treatments

	High load	Low load	Total
Naive 50-50	32	37	69
	$(17, 15)$	$(19, 18)$	$(36, 33)$
Naive Pattern	22	39	61
	$(10, 12)$	$(24, 15)$	$(34, 27)$
Expl. WSLS	37	27	64
	$(17, 20)$	$(12, 15)$	$(29, 35)$
Expl. Mix	39	27	66
	$(16, 23)$	$(15, 12)$	$(31, 35)$
Total	130	130	260

We list the distribution of subjects across cognitive load and computer opponent treatments. In parenthesis we list the distribution in the first and second blocks. There are a total of 260 observations because each of the 130 subjects played against both a Naive and an Exploitative computer opponent strategy and played a block in the high cognitive load treatment and a block in the low cognitive load treatment.

The analysis summarized in Table 6 examines the relationship between Correct and Earnings. One might be concerned that Correct does not distinguish an incorrect response that is off by one digit and an incorrect response that is off by all 6 digits. Here we offer a different specification of correct: Longest Consecutive Subsequence (LCS). The LCS is the longest correct consecutive subsequence of the response. For the high load treatment responses, the raw value ranges between 0 and 6. However, below we normalize the values so that they range between 0 and 1. The mean of the LCS within the high load treatment is 0.934. In the low load treatment, Correct is identical to LCS. Therefore, we only present the analysis for the high load treatment. We summarize this analysis in Table $A2$ below.

Table $A2$: Earnings restricted to cognitive load treatment

	High load				
	(1)	(2)	(3)	(4)	(5)
LCS	0.018	-0.007	-0.020	-0.017	-0.038
	(0.051)	(0.051)	(0.052)	(0.052)	(0.058)
Strategy dummies	No	Yes	Yes	Yes	Yes
Repeated measures	No	No	Yes	Yes	Yes
Demographics	No	No	No	Yes	Yes
GPA	No	No	No	No	Yes
AIC	15658.9	15629.1	15605.8	15615.9	12371.6
Observations	6500	6500	6500	6500	5150

The repeated measures regressions estimate an exchangeable covariance matrix, clustered by subject. We do not provide the estimates of the intercepts, the individual demographics variables, the covariance estimates, or the strategy dummies. AIC refers to the Akaike information criterion (Akaike, 1974). † denotes $p < 0.1$.

Again, in none of the specifications is there a relationship between LCS and Earnings in that period. We interpret this as providing additional evidence that subjects correctly expended effort on the game stage even when the memorization task was not correct in that particular round.

We employ only 4 strategies for the computer opponent. As a result, we can compare strategic behavior against a few well-defined opponent strategies. Here we summarize regression (3) of Table 7 but restricted by computer opponent treatment. The specification includes repeated measures and is summarized in Table $A3$.

Table $A3$: Regressions of Earnings across rounds

	Naive 50-50	Naive Pattern	Expl. WSLS	Expl. Mix
High load	-0.015	0.101	-0.028	$0.063^\dagger$
	(0.032)	(0.061)	(0.035)	(0.032)
AIC	8832.8	6907.8	7805.5	7459.7
Observations	3450	3050	3200	3300

The regressions estimate an exchangeable covariance matrix, clustered by subject. We do not provide the estimates of the intercepts or the covariance estimates. AIC refers to the Akaike information criterion (Akaike, 1974). † denotes $p < 0.1$.

Here we summarize regression (3) of Table 8 but restricted by computer opponent treatment. The specification includes repeated measures and is summarized in Table $A4$.

Table $A4$: Regressions of Earnings across rounds

	Naive 50-50	Naive Pattern	Expl. WSLS	Expl. Mix
High load	0.044	0.090*	0.024	0.089*
	(0.044)	(0.041)	(0.045)	(0.042)
Second half	0.049	$0.061^{\dagger}$	0.062	0.041
	(0.040)	(0.035)	(0.044)	(0.041)
Second half*High load	−0.119	0.023	$-0.103^{\dagger}$	−0.053
	(0.059)	(0.058)	(0.058)	(0.053)
AIC	8837.8	7035.1	7811.5	7468.1
Observations	3450	3050	3200	3300
LSM Differences:				
High load-Low load	0.044	0.090	0.024	0.089*
for First half	(0.043)	(0.067)	(0.045)	(0.042)
High load-Low load	$-0.075^{\dagger}$	$0.113^{\dagger}$	$-0.079^{\dagger}$	0.036
for Second half	(0.043)	(0.067)	(0.045)	(0.042)
Second half-First half	−0.070	$0.084^{\dagger}$	−0.041	−0.011
for High load	(0.043)	(0.044)	(0.038)	(0.034)
Second half-First half	0.049	$0.061^{\dagger}$	0.062	0.041
for Low load	(0.040)	(0.033)	(0.044)	(0.041)

The regressions estimate an exchangeable covariance matrix, clustered by subject. We do not provide the estimates of the intercepts or the covariance estimates. AIC refers to the Akaike information criterion (Akaike, 1974). LSM refers to the Least Square Means. ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

Recall that we defined Round to be the number of periods under a particular computer opponent treatment and cognitive load treatment. Therefore, Round ranges from 1 to 50. Whereas Table 8 summarized the analysis that employed the Second half variable, here we employ the Round variable and the interaction of Round and cognitive load treatment. This analysis is summarized in Table $A5$.

Table $A5$: Regressions of Earnings across rounds

	(1)	(2)	(3)	(4)	(5)	(6)
High load	0.058*	0.114**	0.120**	0.130**	0.146**	0.471***
	(0.029)	(0.038)	(0.041)	(0.042)	(0.046)	(0.127)
Round	0.0019**	0.0019**	0.0019**	0.0019**	0.0020*	0.0020*
	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0008)	(0.0008)
Round*High load	−0.0020*	−0.0020*	−0.0020*	−0.0020*	$-0.0019^{\dagger}$	$-0.0019^{\dagger}$
	(0.00098)	(0.00098)	(0.00097)	(0.00097)	(0.0011)	(0.0011)
GPA	–	–	–	–	0.046*	0.093**
					(0.023)	(0.029)
GPA*High load	–	–	–	–	–	−0.097**
						(0.035)
Strategy dummies	*No*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Repeated measures	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*
Demographics	*No*	*No*	*No*	*Yes*	*Yes*	*Yes*
AIC	31300.9	31239.6	31216.8	31230.1	24785.8	24783.1
Observations	13,000	13,000	13,000	13,000	10,300	10,300

The repeated measures regressions estimate an exchangeable covariance matrix, clustered by subject. We do not provide the estimates of the intercepts, the individual demographics variables, the covariance estimates, the strategy dummies, or the High load-strategy dummies interactions. AIC refers to the Akaike information criterion (Akaike, 1974). *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

The positive and significant high load coefficient suggests that subjects in the high load treatment earn more in the early rounds. However, we observe a positive and significant Round coefficient in addition to a negative and significant Round-High load interaction. This indicates that the subjects in the low load treatment exhibit improved earnings across rounds, however, the subjects in the high load treatment do not exhibit such an improvement. This finding is robust to the specification of the analysis and is consistent with the results of Table 8.

Here we summarize regression (3) of Table $A5$ but restricted by computer opponent treatment. The specification includes repeated measures and is summarized in Table $A6$.

Table $A6$: Regressions of Earnings across rounds

	Naive 50-50	Naive Pattern	Expl. WSLS	Expl. Mix
High load	0.099	0.0464	0.0577	0.0848
	(0.061)	(0.0786)	(0.0622)	(0.0566)
Round	$0.0025^{\dagger}$	0.0017	0.0018	0.0014
	(0.0014)	(0.0012)	(0.0015)	(0.0014)
Round*High load	-0.0045^{*}	0.0022	$-0.0034^{\dagger}$	-0.0009
	(0.0021)	(0.0019)	(0.0020)	(0.0018)
AIC	8850.3	6922.2	7825.3	7481.5
Observations	3450	3050	3200	3300

The regressions estimate an exchangeable covariance matrix, clustered by subject. We do not provide the estimates of the intercepts or the covariance estimates. AIC refers to the Akaike information criterion (Akaike, 1974). ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

We run the analogous regressions as summarized in Table 9 but we employ the Round variable rather than the Second half variable. This analysis is summarized in Table $A7$.

Table $A7$: Regressions of the natural log of Response time across rounds

	(1)	(2)	(3)	(4)	(5)	(6)
High load	-0.076^{***}	-0.214^{***}	-0.258^{***}	-0.235^{***}	-0.223^{***}	-0.342^{***}
	(0.017)	(0.022)	(0.048)	(0.050)	(0.057)	(0.084)
Round	-0.0084^{***}	-0.0084^{***}	-0.0084^{***}	-0.0084^{***}	-0.0075^{***}	-0.0075^{***}
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Round*High load	0.0017^{**}	0.0017^{**}	0.0017^{***}	0.0017^{***}	0.0007	0.0007
	(0.0006)	(0.0006)	(0.0005)	(0.0005)	(0.0006)	(0.0006)
GPA	–	–	–	–	0.0200	0.0032
					(0.0549)	(0.0557)
GPA*High load	–	–	–	–	–	$0.0355^{\dagger}$
						(0.0183)
Strategy dummies	*No*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Repeated measures	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*
Demographics	*No*	*No*	*No*	*Yes*	*Yes*	*Yes*
AIC	17751.6	17530.9	14154.4	14160.8	11212.3	11214.7
Observations	13,000	13,000	13,000	13,000	10,300	10,300

The repeated measures regressions estimate an exchangeable covariance matrix, clustered by subject. We do not provide the estimates of the intercepts, the individual demographics variables, the covariance estimates, the strategy dummies, or the High load-strategy dummies interactions. AIC refers to the Akaike information criterion (Akaike, 1974). *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$, and † denotes $p < 0.1$.

Table $A7$ provides evidence consistent with that in Table 9. Subjects in the high load treatment take less time to reach their decisions in the early rounds. This is possibly done in an effort to quickly proceed to the memorization task. We also observe that Response time decreases across rounds for subjects in the low load treatment. In addition, the Round-High load interaction suggests that subjects in the low load treatment exhibit a greater increase in the decision speed across rounds than subjects in the high load treatment. This is consistent with the interpretation that subjects in the low load treatment exhibit a greater amount of learning than subjects in the high load treatment.

Low Load against Exploitative opponent

Subject	#Down	χ^2	p-value	Runs	$F(r)$	$F(r-1)$	U draw
2	22	2.56	0.110	25	0.483	0.371	0.387
3	22	2.56	0.110	31	0.956	0.921	0.939
4	25	6.25	0.012	21	0.099	0.058	0.058
6	23	3.61	0.057	19	0.033	0.017	0.031
9	28	11.56	< 0.001	32	0.977	0.956	0.962
13	27	9.61	0.002	26	0.576	0.460	0.506
16	28	11.56	< 0.001	20	0.068	0.037	0.060
20	22	2.56	0.110	42	0.999	0.999	0.999
22	49	94.09	< 0.001	3	1.000	0.040	0.202
27	24	4.84	0.028	37	0.999	0.999	0.999
30	22	2.56	0.110	34	0.995	0.989	0.990
31	29	13.69	< 0.001	33	0.992	0.983	0.989
33	25	6.25	0.012	32	0.969	0.942	0.955
34	18	0.16	0.689	29	0.959	0.916	0.934
43	31	18.49	< 0.001	26	0.718	0.612	0.650
47	33	24.01	< 0.001	16	0.014	0.006	0.012
49	27	9.61	0.002	34	0.994	0.987	0.990
52	22	2.56	0.110	28	0.796	0.704	0.714
53	27	9.61	0.002	24	0.351	0.250	0.286
62	25	6.25	0.0124	30	0.902	0.841	0.872
64	17	0.01	0.9203	21	0.270	0.172	0.218
65	32	21.16	< 0.001	24	0.550	0.434	0.497
67	50	–	–	1	0	0	0
68	29	13.69	< 0.001	26	0.629	0.516	0.556
70	24	4.84	0.028	19	0.0314	0.016	0.020
71	28	11.56	< 0.001	30	0.921	0.869	0.902
77	27	9.61	0.002	31	0.949	0.911	0.920
78	33	24.01	< 0.001	15	0.0061	0.002	0.003
102	30	16.00	< 0.001	25	0.559	0.438	0.535

Camden subjects are labeled $1-78$. Haverford subjects are labeled $101-152$. Among the 50 decisions against an Exploitative opponent, we report the number of down actions, the χ^2 statistic and the corresponding p-value as discussed in subsection 4.5. We report the number of runs, the two CDF statistics, and the draw of the uniform between these, as discussed in subsection 4.6.

Low Load against Exploitative opponent

Subject	#Down	χ^2	p-value	Runs	$F(r)$	$F(r-1)$	U draw
103	34	27.04	< 0.001	21	0.341	0.223	0.326
105	48	88.36	< 0.001	4	0.118	0.041	0.075
110	25	6.25	0.012	21	0.098	0.058	0.067
112	35	30.25	< 0.001	26	0.936	0.892	0.932
119	28	11.56	< 0.001	26	0.598	0.483	0.575
120	25	6.25	0.012	26	0.558	0.442	0.467
121	34	27.04	< 0.001	16	0.021	0.009	0.017
125	27	9.61	0.002	36	0.999	0.998	0.998
128	36	33.64	< 0.001	23	0.805	0.663	0.793
130	30	16.00	< 0.001	25	0.559	0.438	0.480
131	25	6.25	0.012	25	0.442	0.335	0.388
134	22	2.56	0.110	28	0.796	0.704	0.791
136	25	6.25	0.012	22	0.159	0.098	0.149
137	27	9.61	0.002	17	0.008	0.003	0.007
140	41	53.29	< 0.001	14	0.237	0.151	0.195
141	24	4.84	0.028	25	0.447	0.339	0.399
142	25	6.25	0.012	28	0.763	0.665	0.683
144	23	3.61	0.057	20	0.062	0.033	0.052
145	28	11.56	< 0.001	26	0.598	0.483	0.561
146	25	6.25	0.012	29	0.841	0.763	0.818
147	24	4.84	0.028	24	0.339	0.240	0.270
149	36	33.64	< 0.001	19	0.282	0.167	0.219
150	25	6.25	0.012	27	0.665	0.558	0.641
151	30	16.00	< 0.001	29	0.912	0.850	0.889
152	23	3.61	0.057	33	0.987	0.973	0.973

Camden subjects are labeled $1-78$. Haverford subjects are labeled $101-152$. Among the 50 decisions against an Exploitative opponent, we report the number of down actions, the χ^2 statistic and the corresponding p-value as discussed in subsection 4.5. We report the number of runs, the two CDF statistics, and the draw of the uniform between these, as discussed in subsection 4.6.

High Load against Exploitative opponent

Subject	#Down	χ^2	p-value	Runs	$F(r)$	$F(r-1)$	U draw
1	36	33.64	< 0.001	19	0.282	0.167	0.189
5	24	4.84	0.028	28	0.767	0.669	0.716
7	26	7.84	0.005	28	0.767	0.669	0.743
8	28	11.56	< 0.001	26	0.598	0.483	0.501
10	9	5.29	0.021	17	0.829	0.576	0.613
11	31	18.49	< 0.001	19	0.063	0.033	0.059
12	31	18.49	< 0.001	21	0.177	0.109	0.122
14	22	2.56	0.110	32	0.977	0.956	0.974
15	30	16.00	< 0.001	19	0.051	0.026	0.045
17	23	3.61	0.057	31	0.949	0.911	0.925
18	22	2.56	0.110	16	0.004	0.001	0.003
19	20	1.00	0.317	33	0.995	0.988	0.990
21	40	49.00	< 0.001	16	0.368	0.260	0.349
23	35	30.25	< 0.001	23	0.699	0.551	0.581
24	32	21.16	< 0.001	17	0.021	0.001	0.014
25	37	37.21	< 0.001	24	0.942	0.902	0.916
26	29	13.69	< 0.001	32	0.983	0.966	0.983
28	38	40.96	< 0.001	17	0.254	0.135	0.208
29	21	1.69	0.194	25	0.516	0.399	0.496
32	32	21.16	< 0.001	28	0.916	0.862	0.913
35	25	6.25	0.012	23	0.237	0.159	0.163
36	26	7.84	0.005	37	0.999	0.998	0.999
37	20	1.00	0.317	30	0.950	0.912	0.947
38	28	11.56	< 0.001	25	0.483	0.371	0.440
39	28	11.56	< 0.001	30	0.921	0.869	0.869

Camden subjects are labeled $1-78$. Haverford subjects are labeled $101-152$. Among the 50 decisions against an Exploitative opponent, we report the number of down actions, the χ^2 statistic and the corresponding p-value as discussed in subsection 4.5. We report the number of runs, the two CDF statistics, and the draw of the uniform between these, as discussed in subsection 4.6.

High Load against Exploitative opponent

Subject	#Down	χ^2	p-value	Runs	$F(r)$	$F(r-1)$	U draw
40	23	3.61	0.057	23	0.250	0.169	0.193
41	24	4.84	0.028	27	0.669	0.562	0.610
42	21	1.69	0.194	27	0.735	0.629	0.673
44	49	94.09	< 0.001	3	1	0.04	0.929
45	18	0.16	0.689	26	0.772	0.676	0.676
46	19	0.49	0.484	22	0.264	0.177	0.185
48	28	11.56	< 0.001	34	0.995	0.989	0.990
50	31	18.49	< 0.001	16	0.007	0.003	0.005
51	33	24.01	< 0.001	21	0.270	0.172	0.230
54	31	18.49	< 0.001	25	0.612	0.488	0.539
55	22	2.56	0.110	27	0.704	0.598	0.655
56	29	13.69	< 0.001	20	0.077	0.043	0.077
57	23	3.61	0.057	27	0.682	0.576	0.659
58	23	3.61	0.057	22	0.169	0.105	0.120
59	25	6.25	0.012	32	0.969	0.942	0.948
60	28	11.56	< 0.001	29	0.869	0.796	0.851
61	32	21.16	< 0.001	21	0.217	0.135	0.193
63	23	3.61	0.057	28	0.778	0.682	0.744
66	30	16.00	< 0.001	30	0.950	0.912	0.946
69	23	3.61	0.057	25	0.460	0.351	0.368
72	24	4.84	0.028	33	0.985	0.970	0.984
73	20	1.00	0.317	20	0.090	0.051	0.051
74	22	2.56	0.110	36	0.999	0.998	0.999
75	33	24.01	< 0.001	7	0.000	0.000	0.000
76	49	94.09	< 0.001	3	1	0.04	0.296

Camden subjects are labeled $1-78$. Haverford subjects are labeled $101-152$. Among the 50 decisions against an Exploitative opponent, we report the number of down actions, the χ^2 statistic and the corresponding p-value as discussed in subsection 4.5. We report the number of runs, the two CDF statistics, and the draw of the uniform between these, as discussed in subsection 4.6.

High Load against Exploitative opponent

Subject	#Down	χ^2	p-value	Runs	$F(r)$	$F(r-1)$	U draw
101	31	18.49	< 0.001	30	0.965	0.936	0.951
104	31	18.49	< 0.001	27	0.815	0.718	0.727
106	27	9.61	0.002	35	0.998	0.994	0.997
107	28	11.56	< 0.001	31	0.956	0.921	0.952
108	24	4.84	0.028	24	0.339	0.240	0.271
109	23	3.61	0.057	23	0.250	0.169	0.224
111	21	1.69	0.194	26	0.629	0.516	0.549
113	24	4.84	0.028	29	0.844	0.767	0.816
114	20	1.00	0.317	31	0.975	0.950	0.952
115	20	1.00	0.317	24	0.438	0.327	0.372
116	24	4.84	0.028	29	0.844	0.767	0.832
117	24	4.84	0.028	33	0.985	0.970	0.984
118	24	4.84	0.028	32	0.970	0.944	0.953
122	19	0.49	0.484	16	0.007	0.003	0.004
123	29	13.69	< 0.001	18	0.022	0.010	0.017
124	27	9.61	0.002	27	0.682	0.576	0.612
126	24	4.84	0.028	30	0.904	0.844	0.861
127	30	16.00	< 0.001	31	0.975	0.950	0.958
129	42	57.76	< 0.001	17	1	0.822	0.937
132	24	4.84	0.028	40	0.999	0.999	0.999
133	29	13.69	< 0.001	30	0.935	0.889	0.891
135	24	4.84	0.028	30	0.904	0.844	0.858
138	26	7.84	0.005	28	0.767	0.669	0.691
139	26	7.84	0.005	30	0.904	0.844	0.889
143	26	7.84	0.005	21	0.100	0.059	0.071
148	31	18.49	< 0.001	29	0.936	0.883	0.916

Camden subjects are labeled $1-78$. Haverford subjects are labeled $101-152$. Among the 50 decisions against an Exploitative opponent, we report the number of down actions, the χ^2 statistic and the corresponding p-value as discussed in subsection 4.5. We report the number of runs, the two CDF statistics, and the draw of the uniform between these, as discussed in subsection 4.6.