Cognitive load in the multi-player prisoner’s dilemma game: Are there brains in games?

Duffy, Sean and Smith, John
Rutgers University-Camden

15 May 2012

Online at https://mpra.ub.uni-muenchen.de/53123/
MPRA Paper No. 53123, posted 23 Jan 2014 03:08 UTC
Cognitive load in the multi-player prisoner’s dilemma game: Are there brains in games?

Sean Duffy* and John Smith†

January 17, 2014

Abstract

We find that differences in the ability to devote cognitive resources to a strategic interaction imply differences in strategic behavior. In our experiment, we manipulated the availability of cognitive resources by applying a differential cognitive load. In cognitive load experiments, subjects are directed to perform a task which occupies cognitive resources, in addition to making a choice in another domain. The greater the cognitive resources required for the task implies that fewer such resources are available for deliberation on the choice. In our experiment, subjects played a finitely repeated multi-player prisoner’s dilemma game under two cognitive load treatments. In one treatment, subjects were placed under a high cognitive load (given a 7 digit number to recall) and subjects in the other were placed under a low cognitive load (given a 2 digit number). According to two different measures, we find evidence that the low load subjects behaved more strategically. First, the low load subjects exhibited more strategic defection near the end of play than the high load subjects. Second, we find evidence that low load subjects were better able to condition their behavior on the outcomes of previous periods.

Keywords: bounded rationality, experimental economics, experimental game theory, public goods game, strategic sophistication, rational inattention

JEL: C72, C91

*Rutgers University-Camden, Department of Psychology, 311 N. 5th Street, Camden, New Jersey, USA, 08102.
†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

There have been advancements in the understanding of play in games based on the conceptualization that players exhibit heterogenous levels of deliberation on their strategy (Stahl and Wilson, 1994, 1995; Nagel, 1995; Costa-Gomes, Crawford, and Broseta, 2001; Camerer, Ho, and Chong, 2004). This conceptualization is often supported by observations of play in a game and determining whether these models improve the fit. In addition to comparing the predictions with the observations, these models are also supported by the measurement of data related to the level of cognition. For instance, studies measuring the decision to lookup relevant and available information,\(^1\) eyetracking studies which measure the location of the attention of the subjects,\(^2\) studies which examine the decision time of subjects,\(^3\) studies which use novel methods in order to learn the distribution of strategic sophistication,\(^4\) studies which employ measures of the cognitive ability of the subjects,\(^5\) and neurological studies\(^6\) have improved our understanding of play in games.

In a rough sense, these experimental papers ask whether one can observe the effects of cognition on strategic behavior. In these studies, researchers perform a measure of cognition or a measure related to the level of cognition and compare this with the observed behavior in games. In this paper, we take a complementary approach. Rather than measure the level of cognition or perform a measure related to the level of cognition, we manipulate the level of cognition. This procedure has the advantage that, since we can randomly assign subjects to a cognitive load treatment, we can mitigate the effects of the heterogeneity of the subjects.\(^7\) Although we do not study behavior which would provide direct evidence on cognitive hierarchy models, similar to these papers, we are interested in examining the role

---


\(^4\)See Agranov, Caplin, and Tergiman (2013) and Burchardi and Penczynski (2013).


\(^7\)We note that previous research has found a relationship between the efficacy of the cognitive load manipulation and the cognitive ability of the subject (Carpenter, Graham, and Wolf, 2013).
of cognition in strategic outcomes. In this sense, our study is complimentary way of asking, "Are there brains in games?"

In the experiment described below, we find a relationship between the heterogenous ability to devote cognitive resources to a strategic interaction and behavior in the interaction. This heterogeneity arises because we apply a differential cognitive load on subjects who are playing the game. In cognitive load experiments, subjects are directed to perform a memorization task in parallel to making a choice in another domain. This additional memorization task occupies cognitive resources which cannot be devoted to deliberation on the choice. In this sense, the condition of subjects under a larger cognitive load could be thought of as similar to the condition of subjects with a diminished ability to reason.

Much is known about the behavior of subjects under a cognitive load. For instance, the literature finds that subjects under a larger cognitive load tend to be more impulsive and less analytical. However, there are only a few studies which examine the effects of cognitive load on behavior in strategic games.\(^8\) \(^9\) One objective of this study is to contribute to the literature which examines the relationship between cognitive load and behavior in games, by studying behavior in the repeated multi-player prisoner’s dilemma.

In our experiment, we imposed a cognitive load on subjects while they were playing a finitely repeated four-player prisoner’s dilemma game.\(^10\) In each period, subjects were told to memorize a number. In the low load treatment, this was a small number and therefore relatively easy to remember. In the high load treatment, this was a large number and therefore relatively difficult to remember. The subjects then played the game. After the subjects made their choice in the game, they were asked to recall the number. As suggested above, subjects in the low load condition were better able to commit cognitive resources to deliberation on

---

\(^8\)See Allred, Duffy, and Smith (2013), Cappelletti, Gütth, and Ploner (2011), Carpenter, Graham, and Wolf (2013), Milinski and Wedekind (1998), and Roch et al. (2000). Below, we discuss these in greater detail.

\(^9\)Researchers have also studied the effects of the contraints on the complexity of strategies on outcomes in the finitely repeated prisoner’s dilemma game. For instance, see Neyman (1985, 1998). Also see Béal (2010) for a more recent reference. Our study can be regarded as a similar exercise, but in the laboratory.

\(^10\)See Komorita et al. (1980). We employ this version because the game is relatively simple, as the decision is binary and the game is linear. On the other hand, the four-player version requires more thought than the two-player version because outcomes depend on the actions of three opponents, rather than just one opponent. Additionally, we were concerned that the subjects could be familiar with the two-player version and would import prior experience into the experiment. The four-player version seems to strike the appropriate balance among these concerns.
their action in the game.

We find that the subjects in the low load condition exhibited more strategic defection near the end of play than those in the high load treatment.\textsuperscript{11} We also find that low load subjects were better able to condition their strategy on previous outcomes. Our paper contributes to the cognitive load literature in that we find that subjects under a high cognitive load are less strategic according to these two measures. Our paper also contributes to the literature on measures of cognitive ability and strategic behavior. To the extent that the cognitive load manipulation affects the cognitive ability of the subject, our results would suggest a negative relationship between measures of cognitive ability and both end-of-game defection and the ability to condition play on previous outcomes in the finitely repeated multi-player prisoner’s dilemma.

1.1 Related literature

The cognitive load literature finds that subjects under a larger cognitive load tend to be more impulsive and less analytical. These differences in behavior stem from the fact that those under a larger cognitive load are less able to devote cognitive resources to reflect on their decision. For instance, Shiv and Fedorikhin (1999) described an experiment in which subjects were given an option of eating an unhealthy cake or a healthy serving of fruit. The authors found that the subjects were more likely to select the cake when they were under a high cognitive load.

Much is known about how the cognitive load manipulation affects subjects in nonstrategic settings. In addition to being more impulsive and less analytical (Hinson, Jameson, and Whitney, 2003) it has been found that subjects under a cognitive load tend to be more risk averse and exhibit a higher degree of time impatience (Benjamin, Brown, and Shapiro, 2013), make more mistakes (Rydval, 2011), have less self control (Shiv and Fedorikhin, 1999; Ward and Mann, 2000), fail to process available information (Gilbert, Pelham, and Krull, 1988; Swann et al., 1990), perform worse on gambling tasks (Hinson, Jameson, and Whitney, 2002),

\textsuperscript{11}These results near the end of the game were also found by Andreoni and Miller (1993) and Keser and van Winden (2000).
are more susceptible to a social label (Cornelissen et al., 2007), make different choices in allocation decisions (Cornelissen, Dewitte, and Warlop, 2011; Hauge et al., 2009, Schulz et al., 2012), and have different evaluations of the fairness of outcomes (van den Bos et al., 2006).

The cognitive load manipulation is effective because it occupies a portion of the working memory of a subject. Working memory is the capacity to temporarily store information so that it can be processed or manipulated. A strong relationship has been found between working memory capacity and measures of cognitive ability (Burgess et al., 2011; Conway, Kane, and Engle, 2003; Süß et al., 2002). In fact, research suggests that training designed to improve working memory can improve the cognitive ability of both humans and mice. Therefore, a reduction of the available working memory capacity of the subject can be thought of as similar to the condition of having a diminished ability to reason.

To our knowledge, there are only a few papers which investigate the relationship between the manipulation of cognitive load and behavior in games. Roch et al. (2000) found that subjects in the low cognitive load condition requested more resources in a common resource game. However, in Roch et al. the subjects were not told the penalty if the sum of the group’s requests exceeded the amount to be divided. As a result, one cannot determine whether the cognitive load manipulation implied differences in strategic behavior or differences in the regard for instructions which are not incentivized.

Milinski and Wedekind (1998) studied the effect of the cognitive load manipulation on behavior in the repeated prisoners dilemma game, without a definite last period. The authors found that most subjects under a low cognitive load conditioned their play on the opponent’s action in the previous period, in addition to their own action in the previous period. In contrast, many high load subjects simply conditioned their play on their opponent’s previous action. The authors also found that low load subjects were better able than high load subjects to remember the outcome of the game two periods in the past. We also find that low load

---

12 Although most cognitive load studies investigate a single effect at a time, Deck and Jahedi (2013) examine several effects at a time. The authors found that subjects under a cognitive load are less patient, more risk averse, perform worse on arithmetic tasks, and are more prone to anchoring effects.

13 See Conway and Getz (2010).

14 Although we note that Milinski and Wedekind (1998) did not investigate the extent to which behavior was affected by outcomes two periods in the past.
subjects condition their play on a larger number of variables related to previous outcomes. While we do not find a difference between the ability to condition on outcomes one period in the past, we do find that low load subjects are better able to condition their play on outcomes two periods in the past.

Cappelletti, Güth, and Ploner (2011) studied behavior in the ultimatum game and varied the ability of subjects to deliberate, by manipulating both time pressure and cognitive load. The authors found that cognitive load did not affect behavior as either a proposer or responder.\textsuperscript{15} In contrast, we find that cognitive load affected strategic behavior. The difference in the efficacy of the cognitive load manipulation is likely due to the differences in its incentivization. We further discuss this issue below.

Additionally, Allred, Duffy, and Smith (2013) performed a within-subject analysis of the effect of cognitive load on various simultaneous games which were designed to measure the strategic sophistication of the subjects. Similarly, Carpenter, Graham, and Wolf (2013) induced a differential cognitive load in subjects then observed play in a pair of strategic games. Carpenter et al. found that subjects under a high cognitive load were less strategic, in general, and less able to perform backwards induction, in particular. We find a similar result in that subjects under a high cognitive load exhibited less strategic defection near the end of play.

There is a recent interest in the relationship between cognitive ability and preferences.\textsuperscript{16} This literature largely finds a negative relationship between cognitive ability and both risk aversion and time impatience. We note the similarities between the findings in the cognitive ability literature and those in the cognitive load literature. There is also a recent interest in the relationship between measures of cognitive ability and behavior in games.\textsuperscript{17} For instance, Burnham et al. (2009) found a relationship between a measure of cognitive ability and strategic

\textsuperscript{15}In a related paper, Buckert, Oechssler, and Schwieren (2013) found that both time pressure and cognitive load affect behavior in a repeated Cournot oligopoly game.


behavior in a beauty contest game. Jones (2008) found a relationship between cooperation in the repeated prisoner’s dilemma and the average SAT scores at the university where the experiment was conducted. Devetag and Warglien (2003) found a relationship between the measured working memory capacity of a subject and the congruence of play to equilibrium behavior. We contribute to this literature by manipulating cognition, rather than by measuring cognition, and examining its affect on strategic behavior.

2 Method

2.1 Discussion of the experimental design

Although the cognitive load manipulation is common, most cognitive load manipulations are not repeated. As a result, it was not obvious to us whether we should balance the experiment so that each subject would undergo the high and low loads an equal number of times. However, we decided to keep the subjects in a single treatment throughout the experiment. In part, this decision was due to the results in Dewitte et al. (2005) which report that the effects of the cognitive load manipulation can be lasting. Also note that we decided to use a 7 digit number as the high load manipulation because it is standard in the literature and because Miller (1956) found that this tends to be near the limit of the memory of subjects.

The bulk of the cognitive load literature does not incentivise the memorization task. Benjamin et al. (2012), Cappelletti et al. (2011), Carpenter et al. (2013), and Schulz et al. (2012) are examples of experiments with such material incentives. Cappelletti et al. (2011) paid the subjects per correct digit. On the other hand, we paid the full amount earned in the game for correct recall and we paid nothing for incorrect recall. However, like Cappelletti et al. (2011), we did not provide feedback regarding the accuracy of the memorization task. We

---

18 See Rydval and Ortmann (2004) for a similar result.
19 Somewhat related to our approach, Bednar et al. (2012) describe an experiment in which subjects simultaneously played two distinct games with different opponents. The authors found that behavior in a particular game was affected by the corresponding paired game. Also see Savikhina and Sheremeta (2013).
20 See Hinson et al. (2002), Hinson et al. (2003), and Schulz et al. (2012) for exceptions.
21 Also, see Cowan (2001) for more recent view on the memory capacity literature.
22 Although there is evidence that subjects perform better on tasks which require attention when the tasks are incentivized. See Camerer and Hogarth (1999).
made these two design decisions in order to reduce the ability of the subjects to strategically allocate cognitive resources.\footnote{Another means of incentivising the cognitive load, without inducing possible differences in payment, is to pay the subjects based on the rank of correct answers within their treatment. While this procedure has the advantage that payments across treatments would be equal, in our view this is less satisfactory than our design. First, in order to make these instructions comprehensible in a session with both treatments, we would have to explain to the subjects that there are different cognitive load treatments. Second, we could avoid informing subjects of the different treatments by running sessions with only a single treatment. However, given the between-subject nature of the experimental design, we worried that this might introduce additional heterogeneity between the treatments.}

Also note that we designed the experiment so that the subject would only enter the following stage when each player completed the preceding stage. This was done in order to mitigate the ability of the subjects to strategically decide the timing of their decisions. Due to our design, there was little incentive for the subjects in the low load condition to quickly leave the stage where they were given the number. Additionally, the subjects in the high load condition could not quickly make their decision in the prisoner’s dilemma game, in order to spill their number in the memorization task. We suspect that our results would be stronger if we allowed subjects to immediately proceed to the subsequent stage.

### 2.2 Experimental design

A total of 48 subjects participated in the experiment. The subjects were graduate and undergraduate students at Rutgers University-Camden. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

Subjects were matched with three other subjects and they played a repeated prisoner’s dilemma game. The subjects were told that the group would remain fixed throughout the experiment.\footnote{The instructions were given via power point slides. The slides, along with any experimental material, are available from the corresponding author upon request.} The subjects were given no information about the composition of their group.

The individual decision was to select $X$ (the cooperative action) or $Y$ (the uncooperative action). Of the four subjects in the group, if $x$ subjects play $X$ then selecting $X$ yields a payoff of $20x$ points whereas selecting $Y$ yields $20x + 40$. The exchange rate was $1$ for every $150$ points. Additionally, the subjects were paid a $5$ show-up fee. While making a decision in the game, the subjects were provided with the payoffs in two logically equivalent formats.\footnote{See the appendix for the screen shown to the subjects during their decision in the game.}
Before play in each period, the subjects were given 15 seconds to commit a number to memory. The subjects were aware that they would be asked to recall the number after their choice was made in the game. There were two cognitive load treatments: in the low load treatment, subjects were directed to memorize a 2 digit number, and in the high load treatment, subjects were directed to memorize a 7 digit number. There were 20 subjects in the low load treatment and 28 in the high load treatment. The groups were homogenous in that they contained only a single load treatment. The subjects were told that they would only receive payment in the periods in which they correctly recalled the number. After each period, subjects were given feedback regarding play in the game, however they were not given feedback about their performance on the memorization task.

To summarize the timing in each period, subjects were given the number (7 digits or 2 digits), they made their choice in the game, they were asked to recall the number, and they were given feedback on the game outcome but not on the memorization task outcome. This procedure was repeated for 30 periods, with a new number in each period. The average amount earned was $14.86.

At the conclusion of period 30, the subjects answered the following manipulation check questions on a scale of 1 to 7: Which featured into your decisions between X and Y, your prudent side or your impulsive side (1 prudent, 7 impulsive)? How difficult was it for you to recall your numbers (1 very difficult, 7 not very difficult)? How difficult was it for you to decide between X and Y (1 very difficult, 7 not very difficult)? How distracting was the memorization task (1 very distracting, 7 not very distracting)? and How many of the memorization tasks do you expect that you correctly answered (1 none correct, 7 all correct)?

3 Results

3.1 Manipulation checks and an overview of the data

The manipulation check questions demonstrated some evidence of differences between the high and low load treatments. Specifically, those in the high load treatment reported hav-
ing significantly more difficulty in recalling the number \((Z = 3.75, p < 0.001)\), found the memorization task to be significantly more distracting \((Z = 3.79, p < 0.001)\), and expected to correctly recall the number with a significantly lower precision \((Z = 2.74, p = 0.006)\) than those in the low load treatment. There is not a significant difference in the reported difficulty in deciding on an action in the game \((Z = 1.10, p = 0.27)\) or in the reported impulsiveness \((Z = -0.94, p = 0.34)\). The subjects in the high load treatment spent a significantly longer time committing the number to memory \((M = 9.08, SD = 4.99)\) than the subjects in the low load treatment \((M = 1.31, SD = 2.28)\), \(Z(840,600) = 28.35, p < 0.001\).

Despite its difficulty, we are surprised by the success of the high load subjects on the memorization task. In the high load treatment, 676 of the 840 (80.5\%) of the memorization tasks were performed correctly. By comparison, 592 of 600 (98.7\%) of the memorization tasks in the low load were performed correctly.

Finally, we provide an overview of the rates of cooperation in the experiment. In Figure 1, we show the average cooperation rates by cognitive load treatment across periods. In Figure 2, we show the cooperation rates of each of the 12 groups across periods. In Table 1, we list the rates of cooperation by treatment in blocks of periods and the aggregate data.

<<Figures 1 and 2 about here>>

<table>
<thead>
<tr>
<th>Table 1. Cooperation rates by treatment and period</th>
</tr>
</thead>
<tbody>
<tr>
<td>---------</td>
</tr>
<tr>
<td>High load treatment</td>
</tr>
<tr>
<td>Low load treatment</td>
</tr>
<tr>
<td>Z-statistic</td>
</tr>
<tr>
<td>p-value</td>
</tr>
</tbody>
</table>

We report the results of Mann-Whitney tests for the difference between the cooperation rates for the high and low load treatments. We perform these tests on blocks of 5 periods and also on the aggregate data. The former tests have \(n_1 = 140\) and \(n_2 = 100\), whereas the latter test has \(n_1 = 840\) and \(n_2 = 600\).

\(^{26}\)These are the results of two-sided Mann-Whitney tests between the high and low load subjects, where \(n_1 = 28\) and \(n_2 = 20\).

\(^{27}\)The z-Tree output specified the time remaining when the Click to Proceed button was pressed. However, there were instances where the output suggested that the decision was made with 99999 seconds remaining. This output seems to have occurred if the "Click to Proceed" button was pressed before the clock could begin. In the stage in which the number was given to the subjects, we recoded the 56 instances of the 99999 output as 16, because 15 seconds were allotted.
Table 1 suggests that there does not exist a large difference between the overall rates of cooperation of the high and low load subjects. When considering the periods in blocks of 5 or the aggregate data, we only find significant differences between the treatments in the final 5 periods, where the low load subjects cooperated less, and between periods 11 – 15, where the low load subjects cooperated more. The former suggests that, as the end of the game approached, the low load subjects played more strategically than the high load subjects. The latter suggests that, despite that they were to be less cooperative near the end of the game, low load subjects exhibited an elevated rate of cooperation for intermediate periods. In other words, there is evidence of rational cooperation by the low load subjects. To investigate this further, we test for evidence of rational cooperation in periods 1 – 25. We find that when excluding the last 5 periods, the low load subjects ($M = 0.452, SD = 0.498$) exhibited more cooperation than the high load subjects ($M = 0.389, SD = 0.488$), $Z(700,500) = 2.198$, $p = 0.028$.

3.2 Differences in behavior

Here, our dependent variable obtains a value of 1 if the cooperative action ($X$) was selected and 0 otherwise. We use a dummy variable where 1 indicates that the subject was in the low load treatment and 0 otherwise. We also use a dummy variable indicating whether the period was within the last 5 periods.

We note the panel nature of our data, consisting of repeated binary decisions of subjects in fixed groups. In order to account for this, we perform random-effects repeated measures logistic regressions. Our analysis assumes an exchangeable log odds ratio, clustered by group and subclustered by subject. In other words, we assume a constant log odds ratio relating any two observations involving a particular subject. We also assume a constant log odds ratio involving two different subjects in the same group. However, observations involving two different subjects in different groups are considered to be 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).\textsuperscript{28} See Table 2 for the results of these regressions.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Period</strong></td>
<td>-0.0327***</td>
<td>-0.0226**</td>
<td>-0.0315***</td>
<td>-0.0345***</td>
</tr>
<tr>
<td></td>
<td>(0.0066)</td>
<td>(0.0105)</td>
<td>(0.0066)</td>
<td>(0.0103)</td>
</tr>
<tr>
<td><strong>Last 5</strong></td>
<td></td>
<td>-0.0498</td>
<td>0.344***</td>
<td>0.390***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.181)</td>
<td>(0.0861)</td>
<td>(0.0895)</td>
</tr>
<tr>
<td><strong>Low load</strong></td>
<td></td>
<td>0.436**</td>
<td>0.263</td>
<td>0.175</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.2198)</td>
<td>(0.202)</td>
<td>(0.239)</td>
</tr>
<tr>
<td><strong>Last 5-Low load Interaction</strong></td>
<td></td>
<td></td>
<td>-1.003***</td>
<td>-1.108***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.233)</td>
<td>(0.261)</td>
</tr>
<tr>
<td><strong>Period-Low load Interaction</strong></td>
<td></td>
<td></td>
<td>0.0210**</td>
<td>0.0070</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0107)</td>
<td>(0.0123)</td>
</tr>
<tr>
<td><strong>QIC</strong></td>
<td>1915.84</td>
<td>1921.73</td>
<td>1913.09</td>
<td>1914.05</td>
</tr>
</tbody>
</table>

We provide the coefficient estimates with standard errors in parentheses, where ** indicates significance at 0.05, and *** indicates significance at 0.01. Each regression accounts for the repeated observations by assuming an exchangeable log odds ratio clustered by group and subclustered by subject. Each regression has 1440 observations from 48 subjects in 30 periods. QIC refers to the Quasi-likelihood information criterion.

There is evidence of convergence to the Subgame Perfect Nash Equilibrium (SPNE) prediction. In every specification involving the period, our results indicate that subjects played less cooperatively across time. In regression (2), the significance of the Low load variable suggests evidence of rational cooperation by the low load subjects. However, this relationship is not significant in regressions (3) and (4). This lack of significance stands in contrast to the rational cooperation results of the Mann-Whitney test over periods 1 – 25. We also find that the actions of the subjects in the low load treatment exhibited significantly more strategic defection near the end of play. We summarize this analysis with the following result.

**Result 1** Across both treatments, behavior converged to the SPNE behavior. Additionally, the subjects in the low load treatment exhibited significantly more strategic defection near the end of play.

\textsuperscript{28}For more on QIC, see Pan (2001).
3.3 Differences in cognitive resources or differences in expected payments?

One alternate explanation for the difference in the behavior of the subjects in the high and low load treatments relates to the possible difference in the expected payments across treatments. Although it would seem difficult to argue that Result 1 was driven by a difference in payment expectations, we nonetheless explore this explanation. While it is not possible to determine the precise difference in the payment expectations, it is possible to look for evidence that the difference in behavior was motivated by the income effect rather than the cognitive load.

One possibility is that the subjects in the high load treatment completely forgot the number, and therefore selected the action in the game with the knowledge that they would not receive payment in that period. If this was the case then we would expect to see subjects having quickly entered an incorrect number so that they could use this additional time to rest and therefore perform better in the subsequent period. Here we look for evidence that high load subjects quickly entered incorrect responses on the memorization task. In Table 3, we summarize the relationship between the memorization task and the time remaining when the stage was exited. We provide the number of correct responses, the number of total responses, and the percent correct by the time remaining when the stage was exited. Recall that subjects were given 15 seconds in which to provide the number.29

<table>
<thead>
<tr>
<th>Time Remaining</th>
<th>14 or more</th>
<th>13 or 12</th>
<th>11 or 10</th>
<th>9 or 8</th>
<th>7 or 6</th>
<th>5 or less</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct High load treatment</td>
<td>21</td>
<td>302</td>
<td>227</td>
<td>72</td>
<td>36</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>331</td>
<td>281</td>
<td>102</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>Percent</td>
<td>95.5%</td>
<td>91.2%</td>
<td>80.8%</td>
<td>70.6%</td>
<td>69.2%</td>
<td>34.6%</td>
</tr>
<tr>
<td>Correct Low load treatment</td>
<td>337</td>
<td>214</td>
<td>28</td>
<td>11</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>342</td>
<td>214</td>
<td>29</td>
<td>12</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Percent</td>
<td>98.5%</td>
<td>100%</td>
<td>96.6%</td>
<td>91.7%</td>
<td>100%</td>
<td>50.0%</td>
</tr>
</tbody>
</table>

In Table 3, we observe that relatively few incorrect responses to the memorization task occurred early in the stage. This suggests that it was not common for the subject to leave the

29As above, we recorded 5 instances of the time remaining output of 99999. This output seems to have occurred if the "Click to Proceed" button was pressed before the clock could begin. We recoded these as 16, because 15 seconds were allotted.
game stage having forgotten the number because there is evidence that the subjects exerted effort to correctly perform the memorization task. The data summarized in Table 3 seems to be consistent with the hypothesis that the subjects in both treatments attempted to correctly perform the memorization task, although the high load subjects took longer and did so with less success.

While the results of Table 3 suggest that the subjects attempted to correctly respond to the memorization task, it is possible that response times would not capture the perceived likelihood of payment. To account for this possibility, we employ a different measure of the subject’s expectation of payment in that period: whether the subject correctly responded to the memorization task in that period. Here we perform an analysis, similar to that summarized in Table 2, with the exception that we include a variable Correct, which assumes a value of 1 if the memorization task in that period was performed correctly, and 0 otherwise. We present a summary of this analysis in Table 4.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Period</strong></td>
<td>-0.0326***</td>
<td>-0.0226**</td>
<td>-0.0315***</td>
<td>-0.0345***</td>
</tr>
<tr>
<td></td>
<td>(0.0066)</td>
<td>(0.0104)</td>
<td>(0.0066)</td>
<td>(0.0102)</td>
</tr>
<tr>
<td><strong>Last 5</strong></td>
<td></td>
<td>-0.0497</td>
<td>0.344***</td>
<td>0.389***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.1809)</td>
<td>(0.0859)</td>
<td>(0.0924)</td>
</tr>
<tr>
<td><strong>Low load</strong></td>
<td></td>
<td>0.452**</td>
<td>0.277</td>
<td>0.190</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.226)</td>
<td>(0.221)</td>
<td>(0.248)</td>
</tr>
<tr>
<td><strong>Last 5-Low load Interaction</strong></td>
<td></td>
<td></td>
<td>-1.003***</td>
<td>-1.106***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.234)</td>
<td>(0.264)</td>
</tr>
<tr>
<td><strong>Period-Low load Interaction</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.0069</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0106)</td>
</tr>
<tr>
<td><strong>Correct</strong></td>
<td>-0.0594</td>
<td>-0.0813</td>
<td>-0.0733</td>
<td>-0.0714</td>
</tr>
<tr>
<td></td>
<td>(0.231)</td>
<td>(0.244)</td>
<td>(0.243)</td>
<td>(0.242)</td>
</tr>
<tr>
<td><strong>Correct p-value</strong></td>
<td>0.80</td>
<td>0.74</td>
<td>0.76</td>
<td>0.77</td>
</tr>
<tr>
<td><strong>QIC</strong></td>
<td>1919.0426</td>
<td>1924.44</td>
<td>1915.88</td>
<td>1916.84</td>
</tr>
</tbody>
</table>

We provide the coefficient estimates with standard errors in parentheses, where ** indicates significance at 0.05, and *** indicates significance at 0.01. Each regression accounts for the repeated observations by assuming an exchangeable log odds ratio clustered by group and subclustered by subject. Each regression has 1440 observations from 48 subjects in 30 periods. QIC refers to the Quasi-likelihood information criterion.
First, we note that Result 1 is not affected by the presence of the Correct variable. Given our measure of the confidence that the subject would correctly perform the memorization task in that period, we still observe more strategic defection near the end of play from the low load subjects. Second, we note that the Correct variable is not significant in any of the regressions. Hence, there does not appear to be a relationship between cooperation and successful performance of the memorization task in that period. Finally, we note that the evidence of rational cooperation, based on the significance of the Low load variable, is similar to that found in Table 2.

Alternatively, we could account for the possibility of differences in the expectations of payment by excluding observations in which the subject incorrectly performed the memorization task. Consider an analysis similar to that summarized in Table 2, with the exception that we only include the 1268 observations in which the memorization task was performed correctly in that period. These results are qualitatively similar to that summarized in Table 4. In light of this analysis, we offer the following result.

**Result 2** We do not find evidence that the subjects were motivated by a difference in payment rather than a difference in cognitive load.

### 3.4 Differences in ability to condition on previous outcomes

Now we explore another measure of strategic behavior: whether the low load subjects were better able to condition their play on past outcomes. In order to investigate this possibility, we offer a model of cooperation which is possibly dependent on previous outcomes. In the analysis described below, we assume that the subject considered features of these previous outcomes to be state variables upon which play could be conditioned.

We describe the variables upon which the subject could condition. One possibility is that the subject could condition play in period $t$ on the number of other players in the group who played cooperatively in the previous period. We refer to this variable as $Others_X in t - 1$. Note that this variable can range from 0 to 3. Another possibility is that the subject could

---

30 This analysis is available from the corresponding author upon request.
condition play in period $t$ on their own action in the previous period. We refer to this variable as $Subject X in t - 1$. This variable obtains a 1 if the subject cooperated in the previous period and a 0 otherwise. We also include the interaction between Others X in $t - 1$ and Subject X in $t - 1$. In addition to these variables with a one period lag, we also include the corresponding variables with a two period lag, which are denoted by $t - 2$.

Similar to Milinski and Wedekind (1998) we investigate the extent to which the other cooperators in the previous period and the subject’s own behavior in the previous period can affect behavior. However unlike Milinski and Wedekind (1998), we also study the extent to which the corresponding variables with a two period lag affect behavior.

As above, we employ a random-effects repeated measures analysis clustered by group and subclustered by subject with an exchangeable log odds ratio. In regressions (1)-(3), we restrict attention to the high load subjects. In regressions (4)-(6), we restrict attention to the low load subjects. In order to investigate the possibility that end-of-game behavior is affected by past outcomes, we offer three specifications for both treatments. In regressions (1) and (4), we analyze data from periods 3 through 30.\(^{31}\) In regressions (2) and (5), we analyze the same data but we include a dummy variable indicating whether the observation came from the last 5 periods. In regressions (3) and (6), we exclude data obtained from the last 5 periods. This analysis is summarized in Table 5.

\(^{31}\)Periods 1 and 2 do not contain observations with a two period lag.
in other hand, none of the three high load specifications exhibit a significant
This suggests that the low load subjects were able to condition their play on a greater number

<table>
<thead>
<tr>
<th>Others X in $t - 1$</th>
<th>High load treatment</th>
<th>Low load treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>$-0.106$</td>
<td>$-0.114$</td>
<td>$-0.0118$</td>
</tr>
<tr>
<td>$(0.111)$</td>
<td>$(0.111)$</td>
<td>$(0.101)$</td>
</tr>
<tr>
<td>$-0.314$</td>
<td>$-0.334$</td>
<td>$-0.173$</td>
</tr>
<tr>
<td>$(0.338)$</td>
<td>$(0.323)$</td>
<td>$(0.383)$</td>
</tr>
<tr>
<td>$0.425^*$</td>
<td>$0.440^*$</td>
<td>$0.382^*$</td>
</tr>
<tr>
<td>Subject X in $t - 1$</td>
<td>$(0.179)$</td>
<td>$(0.181)$</td>
</tr>
<tr>
<td></td>
<td>$(0.188)$</td>
<td>$(0.158)$</td>
</tr>
<tr>
<td></td>
<td>$(0.345)$</td>
<td>$(0.331)$</td>
</tr>
<tr>
<td>Others X in $t - 1^*$</td>
<td>$0.375^*$</td>
<td>$0.401^*$</td>
</tr>
<tr>
<td>Subject X in $t - 1$</td>
<td>$(0.179)$</td>
<td>$(0.181)$</td>
</tr>
<tr>
<td></td>
<td>$(0.188)$</td>
<td>$(0.158)$</td>
</tr>
<tr>
<td></td>
<td>$(0.345)$</td>
<td>$(0.331)$</td>
</tr>
<tr>
<td>Others X in $t - 2$</td>
<td>$0.0147$</td>
<td>$-0.156$</td>
</tr>
<tr>
<td></td>
<td>$-0.131$</td>
<td>$-0.0089$</td>
</tr>
<tr>
<td></td>
<td>$(0.170)$</td>
<td>$(0.174)$</td>
</tr>
<tr>
<td></td>
<td>$(0.184)$</td>
<td>$(0.0880)$</td>
</tr>
<tr>
<td>Subject X in $t - 2$</td>
<td>$0.0617$</td>
<td>$0.046$</td>
</tr>
<tr>
<td></td>
<td>$0.169$</td>
<td>$-0.0232$</td>
</tr>
<tr>
<td></td>
<td>$(0.384)$</td>
<td>$(0.385)$</td>
</tr>
<tr>
<td></td>
<td>$(0.390)$</td>
<td>$(0.211)$</td>
</tr>
<tr>
<td>Others X in $t - 2^*$</td>
<td>$0.0139$</td>
<td>$0.023$</td>
</tr>
<tr>
<td></td>
<td>$-0.0171$</td>
<td>$0.178^*$</td>
</tr>
<tr>
<td></td>
<td>$(0.227)$</td>
<td>$(0.229)$</td>
</tr>
<tr>
<td></td>
<td>$(0.218)$</td>
<td>$(0.0822)$</td>
</tr>
<tr>
<td>Subject X in $t - 2$</td>
<td>$-0.148$</td>
<td>$-1.062^{***}$</td>
</tr>
<tr>
<td>Last 5</td>
<td>$(0.129)$</td>
<td>$(0.201)$</td>
</tr>
<tr>
<td>$QIC$</td>
<td>$1044.30$</td>
<td>$1044.85$</td>
</tr>
<tr>
<td>Observations</td>
<td>$784$</td>
<td>$784$</td>
</tr>
<tr>
<td></td>
<td>$644$</td>
<td>$560$</td>
</tr>
<tr>
<td></td>
<td>$560$</td>
<td>$460$</td>
</tr>
</tbody>
</table>

We provide the coefficient estimates with standard errors in parentheses, where * indicates significance at 0.1, ** indicates significance at 0.05, and *** indicates significance at 0.01. Each regression accounts for the repeated observations by assuming an exchangeable log odds ratio clustered by group and subclustered by subject. Regressions (1) and (2) have 784 observations from 28 subjects in 28 periods. Regression (3) has 644 observations from 28 subjects in 23 periods. Regressions (4) and (5) have 560 observations from 20 subjects in 28 periods. Regression (6) has 460 observations from 20 subjects in 23 periods. QIC refers to the Quasi-likelihood information criterion.

First, we note that the $t - 1$ interaction term is significant in every specification for both cognitive load treatments. However, we note that neither of the other $t - 1$ variables are significant in any of the regressions. Here we find evidence that the high load and the low load subjects conditioned their play on both the number of other cooperators in the previous period and their own play in the previous period. Second, we note that in every specification of the low load regressions, at least one $t - 2$ variable is significant. In regressions (4) and (5) the $t - 2$ interaction variable is significant. Additionally, in regression (5) the Others X in $t - 2$ variable is significant and in regression (6), Subject X in $t - 2$ is significant. On the other hand, none of the three high load specifications exhibit a significant $t - 2$ coefficient. This suggests that the low load subjects were able to condition their play on a greater number

17
of variables related to previous outcomes, even when controlling for the end of play. Third, as expected from the previous analysis, the Last 5 coefficient is significant in the Low load regressions but not in the High load regressions. We summarize the new insights from this analysis with the following result.

**Result 3:** There is evidence that the low load subjects were better able than high load subjects to condition their play on variables related to previous outcomes.

### 4 Conclusion

So are there brains in games? Our results suggest a qualified "yes." Given our manipulation of the availability of cognitive resources in our particular strategic environment, we found that differences in cognitive resources implied differences in strategic behavior. We found that behavior of both high and low load subjects in the multi-player prisoner’s dilemma converged to the SPNE behavior. We also found that the low load subjects exhibited more strategic defection near the end of play. Additionally, we found evidence that the low load subjects were better able to condition their behavior on previous outcomes.

We note that previous research (Jones, 2008) has found a positive relationship between measures of cognitive ability and cooperation in repeated prisoner’s dilemma games, played without a definite last period. However, we do not find strong evidence that the cognitive load treatment is significantly related to cooperative behavior in our regressions. Perhaps we do not find such a relationship since we studied behavior in the finitely repeated prisoner’s dilemma game. It is possible that if we studied behavior in the prisoner’s dilemma game, without a definite last period, then we would find a significant treatment variable.

The relationship between cognitive resources and play in games is also of interest to researchers who study nonequilibrium models. In response to the mounting evidence that subjects rarely play according to the equilibrium predictions, researchers have been turning their attention to nonequilibrium models which can account for hierarchical levels of thinking (Camerer et al., 2004; Costa-Gomes, et al. 2001). It would seem natural to expect that the cognitive ability of the subject would be related to the level of strategic sophistication of
the subject. However, Georganas, Healy, and Weber (2010) found that the mapping of measures of cognitive ability to the estimated hierarchical level of thinking varied across games. While there could be other reasons for this negative result, evidence of this kind is crucial in supporting existing nonequilibrium models or in suggesting modifications to existing models. Whereas the repeated nature of our experiment does not provide direct evidence related to the cognitive hierarchy literature, our paper suggests that it could be fruitful to investigate the relationship between the nonequilibrium models and the cognitive ability of subjects, through the application of a differential cognitive load.

There remain several interesting and unanswered questions. For instance, it is unclear how our results would be affected by an increase (i.e., a public goods game or auction) or a decrease (i.e., a two-player prisoner’s dilemma) in the complexity of the game. Another unanswered question relates to the significance of the incentives regarding the memorization task. While our cognitive load manipulation was successful, and we did not find evidence of an income effect, it is possible that the subjects were motivated by differences in payments across treatments. Also note that we only applied a cognitive load during the stage in which the subjects selected an action in the game. We conjecture that our results would be strengthened if the load was applied during both the game decision stage and the feedback stage. We hope that future work can address these matters.

Finally, we note that the cognitive load manipulation is possibly helpful in any setting in which cognitive constraints affect behavior. Consider the rational inattention literature. These models assume that decision makers are unable to process all of the available information, however they optimally allocate their attention in order to make decisions. In our experiment, we provide evidence of the effects of these constraints, in that subjects under a smaller cognitive load were better able to condition their play on the outcomes of previous periods. The results of our experiment suggest that manipulating the ability to process information via cognitive load could be a productive supplement to efforts in observing behavior.

---

consistent with rational inattention.$^{34,35}$

**Acknowledgements**

We thank Hans Czap, Natalia Czap, Tyson Hartwig, Hrvoje Stojic, Roel Van Veldhuizer, Jack Worrall, and participants at the 7th IMEBE in Barcelona, the IAREP/SABE conference in Exeter UK, and seminar participants at Drexel University and Rutgers University for helpful comments. This research was supported by Rutgers University Research Council Grant #202171.

---

$^{34}$We also note that the predictions of rational inattention have been studied in the laboratory (Cheremukhin, Popova, and Tutino, 2012).

$^{35}$The cognitive load manipulation is also possibly of helpful in the effort to investigate the deviations from optimal search (Sanjurjo, 2012a, 2012b).
## Appendix

The screen during the game decision:

### Period

1 Of 3

<table>
<thead>
<tr>
<th>Period</th>
<th>1 Of 39</th>
<th>Time Remaining</th>
<th>30</th>
</tr>
</thead>
</table>

### Others play XXX

<table>
<thead>
<tr>
<th>Others play XXX</th>
<th>Others play XXX</th>
<th>Others play XXX</th>
<th>Others play XXX</th>
</tr>
</thead>
<tbody>
<tr>
<td>X: 60 Y: 40</td>
<td>Y: 100</td>
<td>X: 40 Y: 80</td>
<td>X: 20 Y: 60</td>
</tr>
</tbody>
</table>

### Others play YYY

<table>
<thead>
<tr>
<th>Others play YYY</th>
<th>Others play YYY</th>
<th>Others play YYY</th>
<th>Others play YYY</th>
</tr>
</thead>
<tbody>
<tr>
<td>X: 00 Y: 100</td>
<td>X: 40 Y: 80</td>
<td>X: 20 Y: 60</td>
<td>X: 40 Y: 40</td>
</tr>
</tbody>
</table>

### Your move

<table>
<thead>
<tr>
<th>Your move</th>
<th>X: 00 Y: 100</th>
<th>X: 40 Y: 80</th>
<th>X: 20 Y: 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X: 40 Y: 80</td>
<td>X: 20 Y: 60</td>
<td>X: 40 Y: 40</td>
</tr>
<tr>
<td>Y</td>
<td>X: 100 Y: 0</td>
<td>X: 60 Y: 0</td>
<td>X: 40 Y: 40</td>
</tr>
</tbody>
</table>

### Payoff Matrix

<table>
<thead>
<tr>
<th>Payoff</th>
<th>0 play X</th>
<th>1 play X</th>
<th>2 play X</th>
<th>3 play X</th>
<th>4 play X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Play X</td>
<td>--</td>
<td>Eams 20</td>
<td>Eams 40</td>
<td>Eams 60</td>
<td>Eams 80</td>
</tr>
<tr>
<td>Play Y</td>
<td>Eams 40</td>
<td>Eams 60</td>
<td>Eams 80</td>
<td>Eams 100</td>
<td>--</td>
</tr>
</tbody>
</table>

Select an action: |

---

Click to Proceed
References


Baghestanian, Sascha and Frey, Seth, 2012. "GO Figure: Analytic and Strategic Skills are Separable." working paper, Indiana University.


Dewitte, Siegfried, Pandelaere, Mario, Briers, Barbara, and Warlop, Luk, 2005. "Cognitive load has negative after effects on consumer decision making." working paper Katholieke Universiteit Leuven, Belgium.


Figure 1-Average cooperation rates by treatment across periods
Figure 2 - Cooperation rates for each of the 12 groups across periods