Strategic complexity and the value of thinking [*]

David Gill [†]

Victoria Prowse [‡]

This version: February 16, 2018

First version: January 19, 2017

Abstract

We measure the strategic complexity of a situation by how long people think on average when they face that situation (where we categorize situations according to the characteristics of play in the previous round). We find that strategic complexity varies significantly across situations, and we find considerable heterogeneity in how responsive subjects' thinking times are to complexity. We also study how variation in response times at the individual level affects strategic behavior and success. 'Overthinking' is detrimental: when a subject thinks for longer than she would normally do in a particular strategic situation, she wins less frequently and earns less.

Keywords: Response time; decision time; thinking time; strategic complexity; game theory; strategic games; repeated games; beauty contest; cognitive ability; personality.

JEL Classification: C72; C91.

1 Introduction

Psychologists have long been interested in choice response or decision times (e.g., Stone, 1960). Recently economists have become interested in how the speed of decision making affects behavior and outcomes. Response times are thought to be connected to decision-making style: fast thinking is linked to intuitive or instinctive decision making, while slower thinking is linked to a more deliberate or contemplative mode of thought (e.g., Rubinstein, 2016).[1] Following the psychology literature, one stream of research in economics looks at response times in non-strategic environments: see Spiliopoulos and Ortmann (2016)'s comprehensive survey.[2]

In contrast to the literature that studies response times in non-strategic settings, our focus is on response times in strategic settings. Speed of decision making is of particular importance in strategic settings because of the complexity of the environment: players need to think about the payoff structure and then form beliefs about how others will behave. For a sophisticated agent, forming such beliefs involves complex iterative thinking. The type of reasoning required to perform well in games requires substantial cognitive effort, but standard theory is silent about how much cognitive effort people exert in strategic settings.[3] In reality, agents vary in their willingness to exert cognitive effort and in how useful cognitive effort is for their performance in strategic settings. According to Rubinstein (2007), response time is a simple low-cost indicator of the process of reasoning in strategic games, with more 'cognitive' choices taking longer than 'instinctive' ones. The speed of decision making in strategic environments has received some recent interest. Experiments have considered how response times relate to: the type of game being played; rates of cooperation in public goods games; behavior in contests; strategies designed to persuade; threshold strategies in global games; and private information in auctions and environments with social learning.[4]

---

[1] As noted by Rubinstein (2016), this distinction between instinctive and contemplative decision making is in some sense consistent with Kahneman (2011)'s distinction between 'system I' and 'system II' decision processes, although economists generally study response times measured in seconds while psychologists generally study response times measured in fractions of seconds.

[2] Wilcox (1993) finds that lottery choice response times respond positively to incentives. Chabris et al. (2009) find that response times increase with the similarity in expected values of binary choices. Piovesan and Wengström (2009) and Lohse et al. (2014) find that generosity is associated with longer response times (both across and within-individuals). Achtziger and Alós-Ferrer (2013) link response times to Bayesian updating. Clithero and Rangel (2013) use response times together with choice data to predict choice out of sample. Rubinstein (2013) finds that shorter response times are associated with more mistakes but not with the incidence of behavior inconsistent with standard models of decision making. In a single-agent forecasting task, Moritz et al. (2014) find that very fast or very slow decisions compared to the prediction from a learning model with individual-level heterogeneity perform worse. Hutcherson et al. (2015), Cappelen et al. (2016), Chen and Fischbacher (2016) and Konovalov and Krajbich (2016) show that response times are associated with risk, time and social preferences. Caplin and Martin (2016) find that quick automatic decisions are of lower quality.

[3] A small literature in behavioral economics addresses this issue: see, e.g., Alaoui and Penta (2016b) who present a cost-benefit analysis of the endogenous depth of reasoning in games. Alaoui and Penta (2016a) extend the model, and Alaoui and Penta (2016c) apply the extended model to Avoyan and Schotter (2016)'s data on the allocation of time across games.

[4] Kuo et al. (2009) and Polonio et al. (2015) find faster response times in coordination games than in dominance-solvable games, while Di Guida and Devetag (2013) find shorter response times in games with focal points and Rand et al. (2015) find that response times vary according to whether decisions are implemented with error and intentions are observable. Gneezy et al. (2010) find that response times in the Race Game are longer in losing positions. Arad and Rubinstein (2012) study the relationship between response times and behavior in the Colonel Blotto contest. Glazer and Rubinstein (2012) study the association between response times and behavioral types in a game of persuasion. Rand et al. (2012), Lotito et al. (2013) and Nielsen et al. (2014) find that shorter response times are associated with more cooperation in public goods games, although Evans et al. (2015) find a U-shaped relationship, Krajbich et al. (2015) argue that the direction of the correlation is not robust to changing the relative attractiveness of the selfish and cooperative actions, Recalde et al. (2015) argue that the relationship may reflect mistakes and Nishi et al. (2017) find cross-cultural differences. Schotter and Trevino (2014) look at the relationship between response times and threshold strategies in a global game. Agranov et al. (2015) elicit incentivized choices at multiple points in time and find that 'sophisticated' players decrease their choice with thinking time in a p-beauty contest variant. Turocy and Cason (2015) consider the relationship between signals and response times in auctions. Frydman and Krajbich (2016) find that in

In particular, we aim to study response times in *repeated* strategic interactions. Strategic interactions of economic interest are often repeated: examples include repeated rounds of job hiring and searching, markets with repeated price competition, repeated selling of goods via auction and multiple rounds of competition for promotions within firms. We leverage the dynamic nature of response-time data from repeated strategic interactions to: (i) measure the strategic complexity of a situation by how long people think on average when they face that situation (where we categorize situations according to the characteristics of play in the previous round); and (ii) discover whether variation in thinking time across rounds in which an individual faces the same strategic situation affects strategic behavior and success.

To study response times in repeated strategic interactions we use the experimental data collected by us in Gill and Prowse (2016).[5] In the experiment, 780 student subjects were matched into groups of three players. Each group played the p-beauty contest for ten rounds with feedback and no rematching. In each round, the subject whose chosen number was closest to seventy percent of the average of the three numbers chosen by the group members (i.e., $p = 0.7$) won six dollars. The subjects had ninety seconds to make their choice in each round. Gill and Prowse (2016) also measured the subjects' cognitive ability and personality.

In the p-beauty contest the incentive to undercut the average of the choices drives the equilibrium to zero. With repetition, choices in beauty contest experiments move toward the equilibrium (e.g., Nagel, 1995, Ho et al., 1998, Gill and Prowse, 2016). However, the game is well-suited to studying strategic thinking: players who expect others to select non-equilibrium actions often have an incentive to choose away from equilibrium themselves. Real-world parallels of the p-beauty contest include timing games in financial and labor markets: during a bubble or in a job market, there is an advantage to trading or making job offers a little earlier than competitors, but moving too early is costly.

A few papers find a positive between-subject relationship between response times and success in strategic games.[6] We replicate this finding for our beauty contest game by showing that subjects who think for longer on average win more rounds and choose numbers closer to equilibrium. We find no statistically significant relationship between cognitive ability or personality and average response times.[7]

Our first substantive contribution is to leverage the dynamic nature of repeated-game response times to develop a measure of strategic complexity. In our repeated-game setting with fixed groups, the subjects may perceive that the strategic complexity of the situation that they face varies with the characteristics of play in the previous round. Motivated by this observation, we categorize eleven different situations according to the particular subject's earnings in the previous round, the rank-order of the choices of the three group members in the previous round, and whether the group played the

---

a social learning environment people can learn about others' private information from observing their response time. Nishi et al. (2016) and Nishi et al. (2017) find that in repeated social dilemmas response times correlate with the level of previous group cooperation. Spiliopoulos (2016) studies a repeated 2×2 constant-sum game and finds that the win-stay-lose-shift heuristic is associated with faster choices. Brañas-Garza et al. (forthcoming) study the response times of ultimatum-game proposers.

[5]Gill and Prowse (2016) investigate how cognitive ability and personality influence the evolution of play toward Nash equilibrium. The paper does not study response times.

[6]Arad and Rubinstein (2012) find that longer response times are associated with winning more battles in the Colonel Blotto game. Brañas-Garza et al. (forthcoming) find that ultimatum-game proposers who think longer earn more. With hypothetical payoffs, Rubinstein (2016) finds that in a 2×2 zero-sum game, the more contemplative action (that is, the action associated with more thought on average) yields a higher expected payoff (in non-zero-sum games the relationship between contemplative actions and payoffs is not as clear).

[7]We are not aware of any other work that measures the relationship between personality and response times in strategic games. With respect to cognitive ability, in a public goods game Nielsen et al. (2014) find mixed results.

Nash equilibrium in that round.[8] We then measure the strategic complexity of a situation by how long subjects think on average when they face that particular situation. Using average thinking time to measure the strategic complexity of a situation in our repeated-game setting is related to Rubinstein (2016)'s distinction between 'contemplative' and 'instinctive' actions in one-shot games according to how long subjects think on average before choosing the action: according to Rubinstein (2016), more contemplative actions require more strategic reasoning. Having developed our measure of strategic complexity, we show that strategic complexity varies significantly across situations. We also find considerable between-subject heterogeneity in how responsive subjects' thinking times are to changes in strategic complexity: we estimate a two-type mixture regression model and find that one type of subject varies her response times substantially with the strategic complexity of the situation that she faces, while the other type hardly varies her response times at all.

Our second substantive contribution is to study how within-subject variation in response times across rounds affects behavior and success in a repeated-game strategic setting. Specifically, we often observe the same subject facing the same strategic situation more than once, and we can measure whether thinking for longer or for less long than the subject would normally do in that situation affects the subject's choices and her probability of winning the round.[9,10] We find that thinking more than normal results in worse performance: when a subject thinks for longer than she would normally do in a particular situation, she wins less frequently and earns less. The behavioral mechanism that drives the reduction in performance is a tendency to move away from Nash equilibrium behavior: when the subject thinks for longer than normal she is more likely to increase her choice relative the previous round and she is less likely to choose the equilibrium number. We conclude that 'overthinking' is detrimental to performance in our strategic setting. Interestingly, these results based on within-subject variation show that overthinking is detrimental even though, as noted above, between-subject variation reveals that subjects who think for longer on average (across situations and rounds) perform better on average. However, this between-subject variation fails to capture the effect of overthinking at the individual level, and the between-subject correlation between average response time and average performance might be driven by the effect of unobserved individual characteristics on both response times and performance.

Together, our findings shed new light on how subjects allocate cognitive resources in games and the consequences of doing so. Our finding that, in repeated games, thinking time varies across situations provides evidence that subjects respond to the characteristics of the situation that they face when deciding how much cognitive effort to allocate to the situation. However, only one type of subject

---

[8]To give a flavor, we define three of the eleven situations here. Situation 1: the subject won in the previous round with the lowest choice, and the two other subjects chose the same higher number. Situation 2: the subject won in the previous round with an 'intermediate' choice (that is, one of the other subjects lost with a lower number and the other lost with a higher number). Situation 3: the subject lost in the previous round with the highest choice, and the other two subjects chose lower numbers different from each other.

[9]Recall from the previous paragraph that we categorize situations according to the characteristics of play in the previous round.

[10]Some existing papers constrain thinking time (see the survey by Spiliopoulos and Ortmann, 2016). However, in the context of strategic games, constraining players' thinking time does not give a clean measure of how a player's strategic behavior varies in her own thinking time because a player's behavior could also change in anticipation of the time constraint on the behavior of the other players. Furthermore, under a time constraint experimental subjects are not choosing how long to think and the imposed time constraint is often very short. Most closely related to our competitive strategic setting: (i) with a time limit of just fifteen seconds Kocher and Sutter (2006) find slower convergence to equilibrium in a modified beauty contest; (ii) with a time limit of fifteen seconds to both read the instructions and then decide Lindner and Sutter (2013) find fewer sophisticated choices in the 11-20 game (but also choices closer to the mixed Nash equilibrium); and (iii) in a design where subjects have just twenty seconds to search payoff boxes using their mouse and then decide Spiliopoulos et al. (2015) find that time pressure reduces payoff search and induces simpler heuristics in normal-form games.

responds to strategic complexity, which highlights the importance of taking seriously across-subject heterogeneity. Our finding that overthinking can be detrimental to performance suggests that allocating too many cognitive resources in games can be counterproductive relative to quicker more instinctive thinking.

We hope that our findings will help to refine existing models of boundedly-rational thinking in games and help to build new ones. In the context of the level-k model (Stahl and Wilson, 1995), our results suggest that the level of strategic reasoning that people use might vary predictably with the particular characteristics of the strategic situation that they face. Furthermore, our findings suggest that new models that incorporate explicitly the choice of how hard to think in strategic interactions could yield better predictive power.

2 Experimental design

We use the experimental data collected by us in Gill and Prowse (2016). In total, 780 student subjects participated in our experiment at the University of Arizona's Experimental Science Laboratory. On average subjects earned twenty-five dollars. Online Appendix I provides the experimental instructions.

Subjects played ten rounds of the p-beauty contest game in fixed groups of three without rematching. In every round each group member chose an integer $x \in \{0, 1, ..., 100\}$. We implemented the beauty contest with $p = 0.7$: the group member whose chosen number was closest to seventy percent of the mean of the three numbers chosen by the group members was paid six dollars and the other group members received nothing. In the case of ties, the six dollars was split equally among the subjects who tied. The unique Nash equilibrium is for all players to choose zero.

The subjects had ninety seconds to make their choice in each round. If a subject failed to make a choice within ninety seconds, then a flashing request prompted an immediate choice.[11] At the end of the ninety seconds, all groups advanced together to a feedback stage that lasted thirty seconds. We provided feedback about the group members' choices in that round, seventy percent of the mean of the choices, and the earnings of the group members in that round.

We measured subjects' cognitive ability using a Raven test, and subjects were matched into groups of three to play the p-beauty contest according to their Raven test score. We also measured the personality (Big Five, grit and a measure of future orientation) of 270 of our subjects. Gill and Prowse (2016) provide further details, including details of the subject matching by Raven test score.

3 Descriptive statistics

We observe 780 subjects for ten rounds, giving 7,800 subject-round observations. Figure 1(a) shows that the average response time varies little across rounds.[12] Figure 1(b) shows the distribution of the subject-round observations of response time (this captures both within-subject heterogeneity, which arises when a subject varies her response time from one round to the next, and across-subject heterogeneity, which originates from systematic differences in average response times across subjects). We see considerable heterogeneity in response time: while around one-quarter of responses occur within the first three seconds, around five percent of responses occur during the final five seconds of the ninety-second-long response window.

---

[11]We code choices that occurred during the flashing request as having taken exactly ninety seconds. Such choices make up just 0.8% of our observations.

[12]We test the significance of the trend by regressing response time on a linear round variable. The two-sided p-value is 0.656.
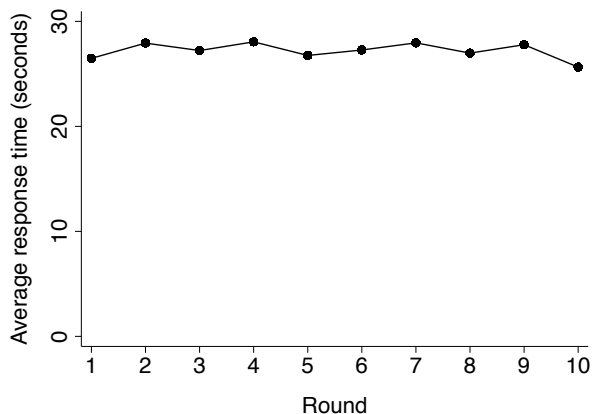
(a) Average response time in rounds 1–10
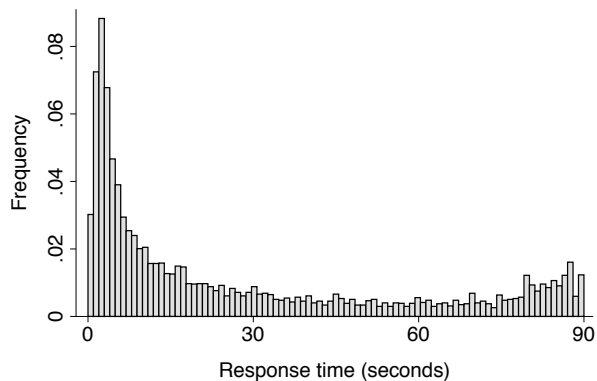
(b) Distribution of subject-round observations of response time

Figure 1

As discussed in footnote 6, a few papers find a positive between-subject relationship between response times and success in strategic games. We replicate this finding. Specifically, we run a between-subject analysis in which we regress measures of success on the subject-level average response time. The first three columns of Table 1 show that a subject's average response time is a predictor of her success: subjects who think for longer on average are more likely to win and earn more. To understand the behavior that underlines these patterns, we regress choices on the subject-level average response time. The last column of Table 1 shows that subjects who think for longer on average choose lower numbers in the beauty contest, i.e., they choose numbers closer to Nash equilibrium. Table SWA.1 in Online Appendix II shows that we find no significant relationship between thinking time and cognitive ability or personality.

	Fraction of rounds won	Earnings per round (cents)	Log earnings per round (cents)	Average choice
Average response time (minutes)	0.060***	38.118***	0.160***	-2.002**
	(0.019)	(9.027)	(0.042)	(0.872)
Intercept	0.402***	182.716***	4.811***	19.929***
	(0.011)	(4.153)	(0.021)	(0.645)
Subjects	780	780	780	780

Notes: All averages are taken at the subject level. When calculating the fraction of rounds won, a subject is considered to be a winner if she won all or part of the prize. Log earnings per round is calculated by taking the log of earnings at the round level and then averaging at the subject level over rounds. When taking the log of earnings, we add fifty cents to earnings in each round (the show-up fee of five dollars divided by the number of rounds) to avoid taking the log of zero. Heteroskedasticity-consistent standard errors with clustering at the group level are shown in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels (two-sided tests).

Table 1: Average response time, success and choices

4 Strategic complexity

We now turn to our first substantive contribution: we use the dynamic nature of the repeated response times in our beauty contest game to measure the strategic complexity of the situations that our subjects face. We proceed in three stages: first, we categorize situations according to the characteristics of play in the previous round; second, we measure the strategic complexity of each situation by how long subjects think on average when they face that situation; and finally, we explore heterogeneity in how subjects respond to changes in strategic complexity.

4.1 Categorizing situations

In our repeated-game setting with fixed groups, subjects may perceive that the strategic complexity of the situation that they face in a given round varies with the characteristics of play in the previous round. Motivated by this observation, we categorize situations according to: (i) the subject's earnings in the previous round; (ii) the rank-order of the choices of the three group members in the previous round; and (iii) whether the group played the Nash equilibrium in the previous round. Let n denote the number chosen by the subject in the previous round, and let $\underline{n}^o$ and $\overline{n}^o$ denote the numbers chosen in the previous round by the subject's opponents. Without loss of generality, we order the opponents such that $\underline{n}^o \leq \overline{n}^o$. The subject's situation in a given round is determined by: (i) the subject's earnings in the previous round, (ii) the ordering of n, $\underline{n}^o$ and $\overline{n}^o$; and (iii) whether all three choices were zero (the Nash equilibrium) in the previous round. We focus on the eleven situations described in the first column of Table 2. The second column of Table 2 shows that we have between 123 and 1,739 subject-round observations for each situation, giving a total of 7,012 subject-round observations across the eleven situations.[13] We order the situations by strategic complexity: the third column of Table 2 reports our measure of strategic complexity, which we explain in Section 4.2.

4.2 Measuring strategic complexity

We measure the strategic complexity of a situation by how long subjects think on average when they face that particular situation. To identify the influence of the situation on average response times we cannot simply measure the average response time in each situation because subjects vary systematically in how long they think and subjects who tend to think for longer might be more likely to face certain situations. Thus, we leverage the repeated observations of response times to separate the component of response times that is attributable to the situation from the component of response times that is due to systematic differences in thinking times between subjects. We do this by running the following fixed-effects regression for subject i's response time in seconds in round r:

$$\text{ResponseTime}_{i,r} = \sum_k \theta_k D_{i,r}^k + \alpha_i + \tau_r + \varepsilon_{i,r} \quad \text{for} \quad r = 2, ..., 10, \tag{1}$$

where $D_{i,r}^k$ is an indicator for subject i being in situation k in round r, α_i is a subject fixed effect that absorbs systematic between-subject difference in response times, τ_r is a round fixed effect that

---

[13]These eleven situations represent all feasible situations with two exceptions: the situation where $n < \underline{n}^o < \overline{n}^o$ and the subject earned three dollars (the subject tied with the opponent who chose $\underline{n}^o$); and the situation where $\underline{n}^o < n < \overline{n}^o$ and the subject earned three dollars (again the subject tied with the opponent who chose $\underline{n}^o$). We omit these two situations because we have only four subject-round observations from each of the two situations, giving eight omitted observations. Subtracting these eight observations from the 7,020 subject-round observations from the second round onward gives our total of 7,012. We cannot use the 780 response-time observations from the first round since we categorize situations according to the characteristics of play in the previous round.

captures any trends in response times over the experiment (round five is the reference category), and $\varepsilon_{i,r}$ is an error term.

When estimating (1), one situation is arbitrarily chosen to be the reference category ($k = 1$; $\theta_1 = 0$), and the parameter θ_k then measures the pure effect on response time of being faced with situation k instead of this reference situation. Using estimates from the fixed-effects regression described in (1), we calculate the strategic complexity of situation k as follows:

$$\text{StrategicComplexity}_k \equiv \hat{\theta}_k + \overline{\hat{\alpha}_i}, \tag{2}$$

where $\hat{\theta}_k$ denotes the estimate of θ_k and $\overline{\hat{\alpha}_i}$ is the average of the estimates of the subject fixed effects.

The third column of Table 2 reveals that strategic complexity, measured by average thinking time as detailed in (2), varies substantially across the eleven situations. The most complex situation is where the subject won the entire prize of six dollars in the previous round with a choice that was between that of her opponents; in this situation subjects think for an average of thirty seconds. The least complex situation is where the group was at the Nash equilibrium in the previous round (that is, all three group members chose zero); in this situation subjects think for an average of fifteen seconds. The variation in strategic complexity across situations is statistically significant: the null hypothesis that strategic complexity is constant across the eleven situations is strongly rejected (an F-test returns $p = 0.000$).

Situation	Subject-round observations	Strategic Complexity (Average response time in seconds)
$n = \underline{n}^o = \overline{n}^o = 0$ ($\therefore$ subject earned \$2)	318	15.398
$\underline{n}^o < \overline{n}^o < n$ ($\therefore$ subject earned \$0)	1,739	25.249
$\underline{n}^o = \overline{n}^o < n$ ($\therefore$ subject earned \$0)	273	25.426
$n < \underline{n}^o < \overline{n}^o$ & subject earned \$0	633	26.064
$\underline{n}^o < n < \overline{n}^o$ & subject earned \$0	1,102	26.373
$n = \underline{n}^o = \overline{n}^o > 0$ ($\therefore$ subject earned \$2)	123	26.680
$\underline{n}^o < n = \overline{n}^o$ ($\therefore$ subject earned \$0)	362	26.775
$n < \underline{n}^o = \overline{n}^o$ ($\therefore$ subject earned \$6)	181	27.027
$n < \underline{n}^o < \overline{n}^o$ & subject earned \$6	1,102	27.984
$n = \underline{n}^o < \overline{n}^o$ ($\therefore$ subject earned \$3)	546	28.447
$\underline{n}^o < n < \overline{n}^o$ & subject earned \$6	633	30.284

Notes: n denotes the number chosen by the subject in the previous round, and $\underline{n}^o$ and $\overline{n}^o$ denote the numbers chosen in the previous round by the subject's opponents (with $\underline{n}^o \leq \overline{n}^o$). Earnings refer to the subject's earning in the previous round. We have a total of 7,012 subject-round observations across the eleven situations (see footnote 13). The third column reports strategic complexity as described in (2).

Table 2: Strategic complexity

4.3 Heterogeneity

We now study whether subjects vary in how they adjust thinking time in response to changes in strategic complexity. We postulate that some subjects may be sensitive to changes in strategic complexity and choose to devote more time to thinking in more complex situations. Other subjects may not tailor their thinking time to the complexity of their situation, either because they fail to recognize that situations vary in complexity or because they lack the self-control needed to think for longer.

We explore the empirical support for this reasoning by estimating the following two-type mixture regression model, which captures heterogeneity in how thinking time responds to changes in strategic complexity:

$$\text{ResponseTime}'_{i,r} = \beta_i \text{StrategicComplexity}'_{i,r} + \upsilon_r + \sigma_i \epsilon_{i,r} \quad \text{for} \quad r = 2, ..., 10, \qquad (3)$$

where

$$\text{StrategicComplexity}_{i,r} \equiv \sum_k \text{StrategicComplexity}_k \times D^k_{i,r} \quad \text{for} \quad r = 2, ..., 10, \qquad (4)$$

denotes the strategic complexity of the situation facing subject i in round r, $X'_{i,r}$ denotes $X_{i,r}$ expressed in deviation form (i.e., differenced relative to the average of $X_{i,r}$ for subject i), υ_r is a round fixed effect, and $\epsilon_{i,r}$ is an independent error term with a standard normal distribution (see Section 4.2, and in particular (1) and (2), for the definitions of $\text{StrategicComplexity}_k$ and $D^k_{i,r}$). The parameter β_i describes how the subject's response time changes when strategic complexity deviates from the average strategic complexity faced by the subject, and σ_i measures the standard deviation of the component of the subject's response time that is unresponsive to changes in strategic complexity. We distinguish two types of subjects: type 1 subjects have $[\beta_i, \sigma_i] = [\beta_1, \sigma_1]$ and type 2 subjects have $[\beta_i, \sigma_i] = [\beta_2, \sigma_2]$. The probability of a subject being of type j is π_j for $j = 1, 2$.

Table 3 reports the parameter estimates for this two-type mixture regression model. We find that the thinking times of the two types respond very differently to changes in strategic complexity. Around sixty percent of subjects are type-1 subjects who increase thinking time substantially when strategic complexity increases (specifically, when the strategic complexity of the situation increases by one second, the response time of type-1 subjects increases by around 1.3 seconds). In contrast, type-2 subjects, who make up around forty percent of the subject population, hardly respond at all to changes in the complexity of their situation: the estimated effect size for type-2 subjects is around one-eighth of that for type-1 subjects.

	Type 1	Type 2
β	1.311***	0.165***
	(0.179)	(0.057)
σ	23.093***	5.844***
	(0.419)	(0.326)
π (type probability)	0.634***	0.366***
	(0.023)	(0.023)

Notes: Parameter estimates were obtained by applying Maximum Likelihood estimation to the sample of 7,012 subject-round observations described in Table 2. Heteroskedasticity-consistent standard errors with clustering at the group level are shown in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels (two-sided tests for β and one-sided tests elsewhere).

Table 3: Heterogeneous effect of strategic complexity on response time: two-type model

4.4 Robustness of results on strategic complexity

Table 2 shows that the complexity of the situation in which a subject's group played the Nash equilibrium in the previous round (the 'equilibrium' situation) is substantially lower than that of the other ten situations (the 'non-equilibrium' situations). In Online Appendix II we show that our results on strategic complexity are not driven by differences between the behavior of subjects in the equilibrium situation and subjects in the non-equilibrium situations. In particular, Table SWA.2 shows that our estimates of the complexities of the ten non-equilibrium situations reported in Table 2 are essentially unchanged if we re-estimate these complexities using a sample that excludes the 318 subject-round observations where the subject's group played the equilibrium in the previous round. The notes to Table SWA.2 also report statistically significant variation in strategic complexity across the ten non-equilibrium situations. Tables SWA.3 and SWA.4 show that our results on the heterogeneous effect of strategic complexity on thinking time reported in Section 4.3 are robust to: (i) excluding the 318 subject-round observations where the subject's group played the equilibrium in the previous round; and (ii) including three types in the mixture regression model.

5 The value of thinking more

In Section 4 we categorized eleven situations according to the characteristics of play in the previous round and we measured the strategic complexity of a situation by how long subjects think on average when they face that particular situation. In Section 3 we showed that subjects who think for longer on average (across situations and rounds) are more successful. However, this between-subject correlation between average response time and average performance might be driven by the effect of unobserved individual characteristics on both response times and performance. We now turn to a more subtle question: is a subject more successful when she thinks for longer than she would normally do in a particular situation?

5.1 Estimation strategy and results

We explore whether a subject is more successful when she thinks for longer than she would normally do in a particular situation by exploiting within-subject variation in response times across rounds that is not attributable to variation in the strategic complexity of the situation facing the subject. We often observe the same subject facing the same situation more than once, and we can measure whether thinking for longer or for less long than the subject would normally do in that situation affects the subject's success.

We consider three measures of subject i's success in round r: being a winner of the round; earnings in the round; and the log of earnings in the round. Specifically, for each measure of success we run the following fixed-effects regression, which describes how a subject's success in a particular round depends on her thinking time in that round:

$$\text{Success}_{i,r} = \lambda \text{ResponseTime}_{i,r} + \eta_{i,s} + \gamma_r + e_{i,r}, \quad \text{for} \quad r = 2, ..., 10, \tag{5}$$

where $\eta_{i,s}$ is a subject-situation fixed effect, γ_r is a round fixed effect and $e_{i,r}$ is an error term. The subject-situation fixed effects absorb subject-specific systematic differences in success across situations. In particular, they absorb the effect of systematic differences in thinking times across situations at the subject level that are correlated with the subject's success. The parameter λ is thus identified from

within-subject variation in response times across rounds in which the subject faced the same situation. It follows that we can interpret our estimate of λ as the effect of a subject thinking for longer than she ordinarily does in a particular situation (and $-\lambda$ as the effect of thinking for less long than normal).

The top panel in Table 4 reports parameter estimates from the fixed-effects regression described in (5). The table shows that thinking for longer than normal results in worse performance. In particular, when a subject thinks for one minute longer than she would normally do in a particular situation, the probability that the subject wins the round decreases by almost six percentage points. This translates into a reduction in earnings in the round of around twenty-seven cents or thirteen percent. We conclude that 'overthinking' is detrimental to success in our strategic setting.

We explore the behavioral mechanism that drives the detrimental effect of overthinking on performance by running fixed-effects regressions as described by (5), but with dependent variables based on the subject's choice in the round. The behavioral mechanism that drives the reduction in performance is a tendency to move away from Nash equilibrium behavior: the bottom panel of Table 4 shows that overthinking makes the subject more likely to increase her choice relative to her choice in the previous round and less likely to choose the equilibrium action of zero. In more detail, when a subject thinks for one minute longer than normal, the probability that her choice increases relative to the previous round goes up by six percentage points and the probability that she chooses the equilibrium action falls by three percentage points. Thinking more than normal also increases the subject's choice in the round, but this effect is noisy and therefore insignificant.

	Winner of round	Earnings in round	Log earnings in round
Effect of thinking longer than normal (minutes)	-0.059** (0.028)	-27.262* (15.003)	-0.132** (0.067)

	Choice in round	Choose zero in round	Increased choice in round
Effect of thinking longer than normal (minutes)	0.990 (1.025)	-0.027** (0.013)	0.057*** (0.019)
Subject-round observations	4,774	4,774	4,774

Notes: We start with the sample of 7,012 subject-round observations described in Table 2. From this sample, we then use the subject-round observations for which the subject is observed in the same situation in at least one other round (this ensures that subject-round observations are not fully absorbed by the fixed effects). This gives us 4,774 subject-round observations and 1,763 subject-situation fixed effects. A subject is considered to be a winner if she won all or part of the prize. When taking the log of earnings, we add fifty cents to earnings in each round (the show-up fee of five dollars divided by the number of rounds) to avoid taking the log of zero. 'Choose zero in round' is an indicator for the subject having chosen the equilibrium action of zero. 'Increased choice in round' is an indicator for the subject's choice in the round being greater than her choice in the previous round. Heteroskedasticity-consistent standard errors with clustering at the group level are shown in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels (two-sided tests).

Table 4: Effect of thinking for longer than normal on success and strategic behavior

5.2 Robustness of results on the value of thinking more

A potential concern is that our results could be biased in the presence of a relationship between thinking time and experience. For example, subjects who have little experience of a situation might also think for longer and perform worse. We address this concern by showing that our results are robust to including controls for the subject's experience of her current situation: see Table SWA.5 in Online Appendix II.

Even though we have included round fixed effects that control for trends over rounds, a second potential concern is that our results could be biased in the presence of heterogeneous round trends. For example, some subjects might exhibit declining response times and improving performance over rounds, while other subjects might exhibit constant or increasing response times and worsening performance over rounds. We address this concern by showing that our results are robust to including controls for heterogeneous round trends: see Table SWA.6 in Online Appendix II.

In Section 4.4 we showed that our results on strategic complexity are not driven by differences between behavior in the equilibrium situation (that is, after the subject's group played equilibrium in the previous round) and behavior in the ten non-equilibrium situations. Table SWA.7 in Online Appendix II show that our results on the effects of thinking for longer than normal are also robust to excluding the equilibrium situation.

6 Conclusion

Response times in games allow researchers to gain insight into subjects' reasoning processes that choices alone cannot (Rubinstein, 2016). We use experimental data on response times from repeated games to develop a measure of the strategic complexity of a situation based on how long subjects think on average when they face that situation, where situations are defined according to the characteristics of play in the previous round. Our finding that strategic complexity varies significantly across situations provides evidence that subjects respond to the characteristics of the situation that they face when deciding how much cognitive effort to allocate to the situation. But not all subjects do this: one type of subject responds strongly to strategic complexity, while another type hardly responds at all.

We also leverage our response-time data from repeated strategic interactions to show that when subjects think for longer than they would normally do in a particular situation, they perform less well. This finding suggests that allocating too many cognitive resources in games can be counterproductive: instead, quicker more instinctive responses can be more successful.

A nascent theoretical literature attempts to model how agents allocate cognitive resources in strategic situations (e.g., Alaoui and Penta, 2016b). Our experimental results suggest that the willingness to engage in strategic reasoning varies systematically with the characteristics of the strategic situation that agents face. We hope that our findings will help to improve the predictive power of existing models of boundedly-rational thinking in games while also inspiring new empirically-grounded models that incorporate explicitly the choice of how hard to think in strategic interactions.

References

Achtziger, A. and **Alós-Ferrer, C.** (2013). Fast or rational? A response-times study of Bayesian updating. *Management Science*, 60(4): 923–938

Agranov, M., **Caplin, A.**, and **Tergiman, C.** (2015). Naive play and the process of choice in guessing games. *Journal of the Economic Science Association*, 1(2): 146–157

Alaoui, L. and **Penta, A.** (2016a). Cost-benefit analysis in reasoning. *Mimeo, University of Wisconsin*

Alaoui, L. and **Penta, A.** (2016b). Endogenous depth of reasoning. *Review of Economic Studies*, 83(4): 1297–1333

Alaoui, L. and **Penta, A.** (2016c). Endogenous depth of reasoning and response time, with an application to the attention allocation task. *Mimeo, University of Wisconsin*

Arad, A. and **Rubinstein, A.** (2012). Multi-dimensional iterative reasoning in action: The case of the Colonel Blotto game. *Journal of Economic Behavior & Organization*, 84(2): 571–585

Avoyan, A. and **Schotter, A.** (2016). Attention in games: An experimental study. *Mimeo, New York University*

Brañas-Garza, P., **Meloso, D.**, and **Miller, L.M.** (forthcoming). Strategic risk and response time across games. *International Journal of Game Theory*

Caplin, A. and **Martin, D.** (2016). The dual-process drift diffusion model: Evidence from response times. *Economic Inquiry*, 54(2): 1274–1282

Cappelen, A.W., **Nielsen, U.H.**, **Tungodden, B.**, **Tyran, J.R.**, and **Wengström, E.** (2016). Fairness is intuitive. *Experimental Economics*, 19(4): 727–740

Chabris, C.F., **Laibson, D.**, **Morris, C.L.**, **Schuldt, J.P.**, and **Taubinsky, D.** (2009). The allocation of time in decision-making. *Journal of the European Economic Association*, 7(2-3): 628–637

Chen, F. and **Fischbacher, U.** (2016). Response time and click position: Cheap indicators of preferences. *Journal of the Economic Science Association*, 2(2): 109–126

Clithero, J.A. and **Rangel, A.** (2013). Combining response times and choice data using a neuroeconomic model of the decision process improves out-of-sample predictions. *Mimeo, California Institute of Technology*

Di Guida, S. and **Devetag, G.** (2013). Feature-based choice and similarity perception in normal-form games: An experimental study. *Games*, 4(4): 776–794

Evans, A.M., **Dillon, K.D.**, and **Rand, D.G.** (2015). Fast but not intuitive, slow but not reflective: Decision conflict drives reaction times in social dilemmas. *Journal of Experimental Psychology: General*, 144(5): 951–966

Frydman, C. and **Krajbich, I.** (2016). Using response times to infer others' beliefs: An application to information cascades. *Working Paper 2817026, SSRN*

Gill, D. and **Prowse, V.** (2016). Cognitive ability, character skills and learning to play equilibrium: A level-k analysis. *Journal of Political Economy*, 124(6): 1619–1676

Glazer, J. and **Rubinstein, A.** (2012). A model of persuasion with boundedly rational agents. *Journal of Political Economy*, 120(6): 1057–1082

Gneezy, U., **Rustichini, A.**, and **Vostroknutov, A.** (2010). Experience and insight in the Race game. *Journal of Economic Behavior & Organization*, 75(2): 144–155

Ho, T.H., **Camerer, C.**, and **Weigelt, K.** (1998). Iterated dominance and iterated best response in experimental "p-beauty contests". *American Economic Review*, 88(4): 947–969

Hutcherson, C.A., **Bushong, B.**, and **Rangel, A.** (2015). A neurocomputational model of altruistic choice and its implications. *Neuron*, 87(2): 451–462

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan

Kocher, M.G. and **Sutter, M.** (2006). Time is money? Time pressure, incentives, and the quality of decision-making. *Journal of Economic Behavior & Organization*, 61(3): 375–392

Konovalov, A. and **Krajbich, I.** (2016). Revealed indifference: Using response times to infer preferences. *Mimeo, Ohio State University*

Krajbich, I., **Bartling, B.**, **Hare, T.**, and **Fehr, E.** (2015). Rethinking fast and slow based on a critique of reaction time reverse inference. *Nature Communications*, 6(7455): 1–9

Kuo, W.J., **Sjöström, T.**, **Chen, Y.P.**, **Wang, Y.H.**, and **Huang, C.Y.** (2009). Intuition and deliberation: Two systems for strategizing in the brain. *Science*, 324(5926): 519–522

Lindner, F. and **Sutter, M.** (2013). Level-k reasoning and time pressure in the 11–20 money request game. *Economics Letters*, 120(3): 542–545

Lohse, J., **Goeschl, T.**, and **Diederich, J.** (2014). Giving is a question of time: Response times and contributions to a real world public good. *Department of Economics Discussion Paper 566, University of Heidelberg*

Lotito, G., **Migheli, M.**, and **Ortona, G.** (2013). Is cooperation instinctive? Evidence from the response times in a public goods game. *Journal of Bioeconomics*, 15(2): 123–133

Moritz, B., **Siemsen, E.**, and **Kremer, M.** (2014). Judgmental forecasting: Cognitive reflection and decision speed. *Production and Operations Management*, 23(7): 1146–1160

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, 85(5): 1313–1326

Nielsen, U.H., **Tyran, J.R.**, and **Wengström, E.** (2014). Second thoughts on free riding. *Economics Letters*, 122(2): 136–139

Nishi, A., **Christakis, N.A.**, **Evans, A.M.**, **O'Malley, A.J.**, and **Rand, D.G.** (2016). Social environment shapes the speed of cooperation. *Scientific Reports*, 6(29622): 1–10

Nishi, A., **Christakis, N.A.**, and **Rand, D.G.** (2017). Cooperation, decision time, and culture: Online experiments with American and Indian participants. *PloS ONE*, 12(2): 1–9

Piovesan, M. and **Wengström, E.** (2009). Fast or fair? A study of response times. *Economics Letters*, 105(2): 193–196

Polonio, L., **Di Guida, S.**, and **Coricelli, G.** (2015). Strategic sophistication and attention in games: An eye-tracking study. *Games and Economic Behavior*, 94: 80–96

Rand, D.G., **Fudenberg, D.**, and **Dreber, A.** (2015). It's the thought that counts: The role of intentions in noisy repeated games. *Journal of Economic Behavior & Organization*, 116: 481–499

Rand, D.G., **Greene, J.D.**, and **Nowak, M.A.** (2012). Spontaneous giving and calculated greed. *Nature*, 489(7416): 427–430

Recalde, M.P., **Riedl, A.**, and **Vesterlund, L.** (2015). Error prone inference from response time: The case of intuitive generosity. *Department of Economics Working Paper 15/004, University of Pittsburgh*

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

Rubinstein, A. (2013). Response time and decision making: An experimental study. *Judgment and Decision Making*, 8(5): 540–551

Rubinstein, A. (2016). A typology of players: Between instinctive and contemplative. *Quarterly Journal of Economics*, 131(2): 859–890

Schotter, A. and **Trevino, I.** (2014). Is response time predictive of choice? An experimental study of threshold strategies. *Discussion Paper SP II 2014-305, WZB*

Spiliopoulos, L. (2016). The determinants of response time in a repeated constant-sum game: A robust Bayesian hierarchical model. *Working Paper 2740989, SSRN*

Spiliopoulos, L. and **Ortmann, A.** (2016). The BCD of response time analysis in experimental economics. *Working Paper 2401325, SSRN*

Spiliopoulos, L., **Ortmann, A.**, and **Zhang, L.** (2015). Attention and choice in games under time constraints: A process analysis. *Working Paper 2620163, SSRN*

Stahl, D.O. and **Wilson, P.W.** (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(1): 218–254

Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, 25(3): 251–260

Turocy, T.L. and **Cason, T.N.** (2015). Bidding in first-price and second-price interdependent-values auctions: A laboratory experiment. *Mimeo, University of East Anglia*

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

Online Appendix

(Intended for Online Publication)

Online Appendix I   Experimental Instructions

Please look at your screen now. I am reading from the instructions displayed on your screen. Please now turn off cell phones and any other electronic devices. These must remain turned off for the duration of this session. Please do not use or place on your desk any personal items, including pens, paper, phones etc. Please do not look into anyone else's booth at any time. Thank you for participating in this experimental session on economic decision-making. You were randomly selected from the Economic Science Laboratory's pool of subjects to be invited to participate in this session. There will be a number of pauses for you to ask questions. During such a pause, please raise your hand if you want to ask a question. Apart from asking questions in this way, you must not communicate with anybody in this room or make any noise.

You will be paid a show-up fee of $5 together with any money you accumulate during this session. The amount of money you accumulate will depend partly on your actions and partly on the actions of other participants. You will be paid privately in cash at the end of the session.

{*Further instructions in the 15 sessions that included questionnaires to measure personality:* Please raise your hand if you have any questions. Before we start the experimental session, I would like you to complete a pre-experimental survey. The survey is made up of 68 questions. There are 17 pages with 4 questions on each page. For each question, please enter your answer in the column to the right of the question. There are no right or wrong answers to the questions. You will have 8 minutes to complete the pre-experimental survey. During the 8 minutes, you can move back and forth between the 17 pages and you can change your previous answers. The top right-hand corner of the screen will display the time remaining (in seconds). Before we start the pre-experimental survey, please raise your hand if you have any questions. During the pre-experimental survey, please raise your hand if you have a problem with your computer. [Subjects complete questionnaire] Thank you for completing the pre-experimental survey. I will now describe the experimental session.}

The session is made up of 2 parts. In the first part you will complete a test. Right at the end of the session you will find out your own test score, but you will not be paid for completing the test. I will describe the second part of the session after you have completed the test. Please raise your hand if you have any questions.

I will now describe the test which makes up the first part of the session. The test is made up of 60 questions, divided into parts A, B, C, D and E. Each of these parts is made up of 12 questions. For every question, there is a pattern with a piece missing and a number of pieces below the pattern. You have to choose which of the pieces below is the right one to complete the pattern. For parts A and B of the test, you will see 6 pieces that might complete the pattern. For parts C, D and E you will see 8 pieces that might complete the pattern. In every case, one and only one of these pieces is the right one to complete the pattern.[14] For each question, please enter your answer in the column to the right of the pattern. You will score 1 point for every right answer. You will not be penalized for wrong answers. You will have 3 minutes to complete each of parts A and B, and you will have 8 minutes to complete each of parts C, D, and E. During each part, you can move back and forth between the 12 questions in that part and you can change your previous answers. The top right-hand corner of the screen will display the time remaining (in seconds). Before we start the test, please raise your hand if you have any questions. During the test, please raise your hand if you have a problem with your computer. [Subjects complete test]

Your screen is now displaying whether your test score was in the top half of the test scores of

---

[14]The wording of this description follows the standard Raven test convention.

all participants in the room or was in the bottom half of the test scores of all participants. [30 second pause] [Example (not read aloud): Your test score was in the top half of the test scores of all participants in the room.] At the end of the session you will find out your own test score.

I will now describe the second and final part of the session. This second part is made up of 10 rounds. You will be anonymously matched into groups of 3 participants. You will stay in the same group for all 10 rounds. In each round, you and your other 2 group members will separately choose a whole number between 0 and 100 (0, 100 or any whole number in between is allowed). The group member whose chosen number is closest to 70% of the average of all 3 chosen numbers will be paid $6 for that round and the other 2 group members will be paid nothing. If more than one group member chooses a number which is closest to 70% of the average of all 3 chosen numbers, the $6 will be split equally among the group members who chose the closest number or numbers. Your total payment will be the sum of your payments in each round together with your show-up fee of $5. In each round you will have 90 seconds to choose your number. If you choose your number early you will still have to wait until the end of the 90 seconds. The top right-hand corner of the screen will display the time remaining (in seconds). The screen will also include a reminder of the rules.

At the end of each round you will discover: (i) the numbers chosen by all your group members; (ii) the average of all 3 chosen numbers; (iii) what 70% of the average of all 3 chosen numbers was; and (iv) how much each group member will be paid for the round. Please raise your hand if you have any questions.

You will stay in the same group of 3 for all 10 rounds. Each group member has been randomly allocated a label, X, Y or Z. Your screen is now displaying your label and whether the test scores of the members of your group were in the top half or the bottom half of the test scores of all participants in the room. [60 second pause] [Example (not read aloud): You are group member Y. Your test score was in the top half of the test scores of all participants in the room. You have been matched with 2 participants (group member X and group member Z). Group member X was randomly selected from those whose test scores were also in the top half. Group member Z was randomly selected from those whose test scores were in the bottom half.] Please raise your hand if you have any questions. There will be no further opportunities for questions.

[10 rounds of beauty contest with feedback as described in Section 2.3]

[Screen asks subjects to report their gender]

[Screen reports the subject's score in the Raven test]

The session has now finished. Your total cash payment, including the show-up fee, is displayed on your screen. Please remain in your seat until you have been paid. Thank-you for participating.

Online Appendix II Additional tables

	Average response time (minutes)		
Raven test score	-0.007		0.012
(cognitive ability)	(0.013)		(0.019)
Personality factor 1		-0.027	-0.026
(conscientiousness, grit and future orientation)		(0.022)	(0.022)
Personality factor 2		0.006	0.005
(agreeableness and emotional stability)		(0.018)	(0.019)
Personality factor 3		-0.015	-0.015
(openness, extraversion and future orientation)		(0.021)	(0.021)
Intercept	0.453***	0.442***	0.442***
	(0.013)	(0.019)	(0.019)
Subjects	780	270	270

Notes: All averages are taken at the subject level. The Raven test score and personality factors 1–3 have been standardized to have means of zero and standard deviations of one (Gill and Prowse, 2016, describe the construction of the personality factors). Personality was measured for 270 of our 780 subjects. Heteroskedasticity-consistent standard errors with clustering at the group level are shown in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels (two-sided tests).

Table SWA.1: Raven test score, personality and average response time

Situation	Subject-round observations	Strategic Complexity (Average response time in seconds)
$\underline{n}^o < \overline{n}^o < n$ ($\therefore$ subject earned \$0)	1,739	25.092
$\underline{n}^o = \overline{n}^o < n$ ($\therefore$ subject earned \$0)	273	25.526
$n < \underline{n}^o < \overline{n}^o$ & subject earned \$0	633	26.028
$\underline{n}^o < n < \overline{n}^o$ & subject earned \$0	1,102	26.286
$n = \underline{n}^o = \overline{n}^o > 0$ ($\therefore$ subject earned \$2)	123	26.295
$\underline{n}^o < n = \overline{n}^o$ ($\therefore$ subject earned \$0)	362	26.806
$n < \underline{n}^o = \overline{n}^o$ ($\therefore$ subject earned \$6)	181	27.333
$n < \underline{n}^o < \overline{n}^o$ & subject earned \$6	1,102	27.998
$n = \underline{n}^o < \overline{n}^o$ ($\therefore$ subject earned \$3)	546	28.277
$\underline{n}^o < n < \overline{n}^o$ & subject earned \$6	633	30.272

Notes: We start with the sample described in Table 2. From this sample, we exclude the equilibrium situation: that is, we exclude the 318 subject-round observations where the subject's group played equilibrium in the previous round. This gives us a total of 6,694 subject-round observations across the ten non-equilibrium situations. The third column reports strategic complexity as described in (2). See the notes to Table 2 for notational definitions. The null hypothesis that strategic complexity is constant across the ten non-equilibrium situations is strongly rejected (an F-test returns $p = 0.001$).

Table SWA.2: Robustness of the results in Table 2 to excluding the equilibrium situation

	Type 1	Type 2
β	1.585***	0.132
	(0.287)	(0.091)
σ	23.099***	5.865***
	(0.443)	(0.337)
π (type probability)	0.627***	0.373***
	(0.023)	(0.023)

Notes: Parameter estimates were obtained by applying Maximum Likelihood estimation to the sample of 6,694 subject-round observations described in Table SWA.2. Heteroskedasticity-consistent standard errors with clustering at the group level are shown in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels (two-sided tests for β and one-sided tests elsewhere).

Table SWA.3: Robustness of the results in Table 3 to excluding the equilibrium situation

	Type 1	Type 2	Type 3
β	1.427***	0.240**	0.022
	(0.207)	(0.108)	0.036
σ	23.868***	7.671***	2.140***
	(0.402)	(0.423)	(0.202)
π (type probability)	0.581***	0.313***	0.106***
	(0.023)	(0.019)	(0.014)

Notes: Parameter estimates were obtained by applying Maximum Likelihood estimation to the sample of 7,012 subject-round observations described in Table 2. Heteroskedasticity-consistent standard errors with clustering at the group level are shown in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels (two-sided tests for β and one-sided tests elsewhere).

Table SWA.4: Heterogeneous effect of strategic complexity on response time: three-type model

	Winner of round	Earnings in round	Log earnings in round
Effect of thinking longer than normal (minutes)	-0.057**	-26.800*	-0.129*
	(0.028)	(15.086)	(0.067)

	Choice in round	Choose zero in round	Increased choice in round
Effect of thinking longer than normal (minutes)	0.908	-0.027**	0.058***
	(1.026)	(0.013)	(0.019)
Subject-round observations	4,774	4,774	4,774

Notes: We run the same regressions on the same sample as in Table 4, adding controls for the subject's experience of her current situation by including the indicator functions $\mathbf{1}_{\{k\}}(\chi_{i,r})$ for $k = 1, ..., 6$, where $\chi_{i,r} \in \{0, 1, ..., 6\}$ is equal to the number of times that subject i experienced her current situation prior to round r (zero experience is the omitted category). No subject experienced her current situation in eight previous rounds (the maximum). One subject experienced her current situation in seven previous rounds: we set $\chi_{i,r}$ equal to six for this observation to ensure that the observation continues to contribute to the estimation. Heteroskedasticity-consistent standard errors with clustering at the group level are shown in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels (two-sided tests).

Table SWA.5: Robustness of the results in Table 4 to controlling for experience

	Winner of round	Earnings in round	Log earnings in round
Effect of thinking longer than normal (minutes)	-0.090** (0.039)	-40.807* (21.069)	-0.201** (0.093)

	Choice in round	Choose zero in round	Increased choice in round
Effect of thinking longer than normal (minutes)	1.197 (1.355)	-0.020 (0.015)	0.080*** (0.025)
Subject-round observations	4,774	4,774	4,774

Notes: We run the same regressions on the same sample as in Table 4, controling for heterogeneous round trends by including a separate linear round trend for each of the 780 subjects in the sample (formally, we augment (5) by adding the term $\rho_i \times r$ and treat ρ_i for $i = 1, ..., 780$, as parameters to be estimated). Heteroskedasticity-consistent standard errors with clustering at the group level are shown in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels (two-sided tests).

Table SWA.6: Robustness of the results in Table 4 to controlling for heterogeneous round trends

	Winner of round	Earnings in round	Log earnings in round
Effect of thinking longer than normal (minutes)	-0.060** (0.029)	-28.407* (15.600)	-0.136** (0.069)

	Choice in round	Choose zero in round	Increased choice in round
Effect of thinking longer than normal (minutes)	1.026 (1.066)	-0.027** (0.013)	0.058*** (0.019)
Subject-round observations	4,510	4,510	4,510

Notes: We start with the sample of 6,694 subject-round observations described in Table SWA.2. From this sample, we then use the subject-round observations for which the subject is observed in the same situation in at least one other round (as explained in the notes to Table 4, this procedure ensures that subject-round observations are not fully absorbed by the fixed effects). This gives us 4,510 subject-round observations and 1,682 subject-situation fixed effects. We then run the same regressions as in Table 4. Heteroskedasticity-consistent standard errors with clustering at the group level are shown in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels (two-sided tests).

Table SWA.7: Robustness of the results in Table 4 to excluding the equilibrium situation