Wissenschaftszentrum Berlin
für Sozialforschung

Dietmar Fehr
Steffen Huck

Who knows it is a game? On strategic awareness and cognitive ability

Discussion Paper

SP II 2013–306r

July 2013 (revised November 2014)

Research Area
Markets and Choice

Research Unit
Economics of Change

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Affiliation of the authors:

Dietmar Fehr, WZB
Steffen Huck, WZB and University College London

Abstract

Who knows it is a game? On strategic awareness and cognitive ability*

We introduce the notion of strategic awareness in experimental games which captures the idea that subjects realize they are playing a game and thus have to form beliefs about others' actions. The concept differs from both, rule understanding and rationality. We then turn to experimental evidence from a beauty contest game where we elicit measures of cognitive ability and beliefs about others' cognitive ability. We show that the effect of cognitive ability is highly non-linear. Subjects' behavior below a certain threshold choose numbers in the whole interval and does not correlate with beliefs about others ability. In contrast, choices of subjects who exceed the threshold avoid choices above 50 and react very sensitively to beliefs about others' cognitive ability.

In diesem kurzen Artikel führen wir das Konzept von *strategic awareness* in Experimenten ein. Dieses neue Konzept beschreibt die Fähigkeit von Experimentteilnehmer, strategische Situationen zu erkennen und daher Erwartungen über das Verhalten von anderen zu bilden. Das Konzept unterscheidet sich sowohl von Rationalität als auch vom bloßen Verstehen von den Regeln eines Experiments. Wir demonstrieren das Konzept empirisch mit Hilfe von Daten eines Beauty Contest Games, in dem wir die kognitiven Fähigkeiten der Teilnehmer und ihre Einschätzungen über die kognitiven Fähigkeiten der anderen Teilnehmer erheben. Die Resultate zeigen, dass kognitive Fähigkeiten einen starken nicht-linearen Effekt auf die Entscheidungen in dem Beauty Contest Game haben. Das Verhalten von Experimentteilnehmer, die unter einer bestimmten Schwelle liegen, kann nicht von zufälligen Entscheidungen unterschieden werden und korreliert auch nicht mit deren Einschätzung über die kognitiven Fähigkeiten der anderen Teilnehmer. Im Gegensatz dazu vermeiden Teilnehmer, die über dieser Schwelle liegen, dominierte Entscheidungen und basieren ihre Entscheidungen auf ihrer Einschätzung über die kognitiven Fähigkeiten der anderen Teilnehmer.

Keywords: Strategic awareness, cognitive ability, beauty contest

JEL classification: C7, C9, D0

---

* Financial support from the Deutsche Forschungsgemeinschaft (DFG) through the SFB 649 "Economic Risk" is gratefully acknowledged.
E-mail: dietmar.fehr@wzb.eu and *steffen.huck@wzb.eu*

1. Introduction

Life outside the laboratory is often characterized by rather subtle strategic structures that agents might easily overlook. For example, a worker in a firm might think that his wage (or a future promotion) just depends on his own effort while, in reality, it is linked to the competition between different units and their managers. Understanding such strategic ripple effects can be taxing for even the most sophisticated game theorists.

In laboratory experiments strategic reasoning is often promoted through framing that vividly illustrates strategic interdependencies. Game matrices are an obvious example. Researchers typically spend much care on the writing of their instructions in order to make sure subjects understand them well. Often they employ pre-experimental questionnaires with understanding tests and some only proceed to the experiment proper once all subjects have answered all questions correctly. But what does it actually mean to *understand* instructions for an economic experiment? For game experiments it appears natural to equate understanding of the experiment with the understanding of the precise rules of the game. That is, subjects should be aware of all choice possibilities (in case of extensive form games, they should be aware of the entire tree structure) and they should know all players' payoff functions that map choices into outcomes. In other words, subjects should be able to answer questions like "What are the available actions for player Y if player X has chosen action *s*?" and "What will be player Y's outcome if X chooses *s* and she plays *t*?" If a subject can answer all these questions correctly, we say that she has *rule understanding*.

As such rule understanding is a prerequisite for rationality which demands that players best respond to their beliefs. While data analysis in experimental economics does not necessarily assume rationality (although it frequently does, for example, when researchers try to infer preferences from belief and choice data), it almost invariably assumes rule understanding.

In this note we explore a concept – *strategic awareness* – that is weaker than rationality and related to rule understanding. We say that subjects have strategic awareness if they understand that the decision situation they face is interactive, i.e., that they are, in fact, playing a game. Therefore, strategic awareness requires that subjects realize that reasoning about how to play the game requires reasoning about others. While forming beliefs about others is an essential requirement for strategic awareness, beliefs need not be based on equilibrium models of others players' behavior. Like in alternative models of strategic thinking, beliefs can be based on non-equilibrium behavior, heuristics, etc.

According to our definition of strategic awareness, rule understanding does not necessarily imply strategic awareness. In principle, subjects might be able to enumerate all payoff consequences of all sequences of actions and yet fail to understand that, in order to play the game, they have to think about what others will do. Similarly, strategic awareness does not require full rule understanding. Subjects can be aware of strategic considerations without understanding all details of the payoff functions. For example, in a principal-agent setting the agent might be aware that the principal can monitor his actions, but may overlook some subtle details of the payoff function or monitoring technology. A good example from the experimental economics literature comes from two-person guessing games where the choice of zero is weakly dominant (see, for example, Grosskopf and Nagel, 2008). While standard descriptions of the payoff function induce strategic awareness in at least some players, many fail to choose zero. This can be attributed to a failure of rule understanding by comparing the standard instruction ("whoever is closest to two thirds of the average wins") with instructions that simply state "whoever chooses the lowest number wins", which vastly enhances compliance with weak dominance.[1]

As a vehicle to study the presence or absence of strategic awareness we employ Nagel's (1995) classic beauty contest game (BCG). In the BCG, subjects have to pick a number $a \in [0, 100]$ and the subject who is closest to a fraction *p* of the group average wins a prize m. Given that $p \in (0, 1)$, the unique Nash equilibrium is zero. In the last decade BCGs have been

[1] See, for example, Chou *et al.*'s (2009) "battle protocol". Chou et al. argue that violations of game theory may be traced back to a failure of "game form recognition", i.e., the failure to understand the rules of the game. Their concept of game form recognition, however, differs from strategic awareness since strategic awareness does not explicitly rule out that subjects act strategically despite a lack of full understanding the game. Chou *et al.* vary their instructions in several aspects, such as emphasizing strategic thinking, giving subjects a hint on how to play the game or paraphrasing subjects' task as a battle, in order to reduce violations of weak dominance. Indeed, these variations led to fewer violations of weak dominance. However, it is not clear whether subjects followed the advice in the instructions, or whether these variations induced subjects to reason strategically.

extensively used to examine rationality and belief formation, in particular, in the context of level-*k* models (see, for example Nagel, 1995, Ho, Camerer and Weigelt, 1998, Costa-Gomes and Crawford, 2006, and for a review on recent evidence on strategic thinking Crawford, Costa-Gomes and Iriberri, 2013). The typical level-*k* model assumes that higher level players anchor their best responses on the behavior of non-strategic level-0 players and that belief formation proceeds in iterative steps. Thus choices of sophisticated players depend on their model of level-0 play, their own level of reasoning and their expectation about the level of reasoning of others. From the perspective of the vast literature on level-*k* thinking, we take one step back and pose the question which subjects have strategic awareness. Importantly, the answer to this question does not depend on the specification of level-0 players; it only requires that subjects base their decisions on their beliefs about others' behavior.

We present experimental evidence that links strategic awareness to cognitive ability and beliefs about others' cognitive ability in the BCG. Intuitively, cognitive ability might be an important determinant of the reasoning process and recently a few studies linked subjects' behavior in the BCG directly to their cognitive ability (e.g., Burnham et al., 2009, Branas-Garza, Garcia-Munoz and Gonzalez, 2012, Georganas, Healy and Weber, 2012 and Gill and Prowse, 2013) or brain activity (Coricelli and Nagel, 2009). If a higher level of reasoning is positively linked to cognitive ability, then we may expect that those with higher cognitive ability will not only think about the levels of reasoning of their opponents but also about their opponents' cognitive ability. In other words, strategically aware subjects will base their choices in the BCG on their expectation about the cognitive ability of others. In order to identify the link between strategic awareness and cognitive ability we rely on a simple, but reliable measure of cognitive ability and measure subjects' expectations about others' cognitive ability. Thus, if our hypothesis about strategic awareness is correct, we should observe that beliefs about the cognitive ability of opponents affect choices in the BCG for subjects with high cognitive ability, but not for others. While our approach to identify strategic awareness is indirect, an important advantage is that it does not conflict with behavior in the game itself as would be the case with direct belief elicitation and targeted questions.

We find a strong interesting link between cognitive ability, strategic awareness and choices in the BCG. Essentially, our data shows that subjects need to have some minimal cognitive ability in order to acquire strategic awareness. Subjects below a certain threshold play the BCG as if it was a game of luck: there is no evidence for reasoning about others and their choices appear to be randomly distributed over the whole interval.[2] In contrast, subjects above the threshold avoid choices above 50 and do reason about others. Specifically, their choices depend on their beliefs about others' cognitive ability.

Our paper complements a recent literature that links subjects' choices in the BCG to cognitive ability by providing evidence that choices of strategic subjects also depend on their expectations about the cognitive ability of opponents. While Burnham et al. (2009) document that lower choices in a one-shot $p = 1/2$ BCG are associated with higher cognitive ability as measured through a comprehensive IQ, other studies find no or little evidence for such a relationship when looking at initial choices (Branas-Garza, Garcia-Munoz and Gonzalez, 2012, Georganas, Healy and Weber, 2012 and Gill and Prowse, 2013).[3] Coricelli and Nagel (2009) use fMRI data to relate brain activity to subjects' levels of reasoning in several BCGs. They find substantial evidence that higher levels of reasoning trigger activity in certain brain areas associated with mentalizing. The findings of our study suggest that cognitive ability is an important ingredient for strategic thinking in novel situations and more importantly that strategic reasoning is endogenously determined by expectations about others' cognitive skills.

Two closely related studies attempt to manipulate subjects' beliefs about their opponents' sophistication without measuring cognitive ability directly. Agranov *et al.* (2012) show that behavior shifts to higher observed levels of reasoning, when undergraduate subjects know

[2] This finding is in line with other recent evidence on the presence of non-strategic players in the BCG. For example, a study by Agranov, Caplin and Tergiman (2012) identifies a large proportion of $k = 0$ players (43 percent) by observing players' provisional BCG choices within a three-minute time frame. Related, Burchardi and Penczynski (2014) use communication protocols among team members playing a BCG to classify strategic and non-strategic reasoning (about 33 percent).

[3] Branas-Garza, Garcia-Munoz and Gonzalez (2012) used the Raven test and the cognitive reflection test (CRT) to measure subjects' cognitive ability and related the test scores to behavior in six BCGs with varying p. They find no effect for cognitive ability in the first two beauty contests (with $p = 2/3$ and $p = 1/8$), but in the remaining BCGs ($p = 1/5$, $p = 1/3$, $p = 1/2$, $p = 3/4$) they observe a negative correlation of CRT and chosen numbers. Similarly, Gill and Prowse (2013) find no relationship between Raven's test scores and choices in the first round of a repeated BCG ($p = 7/10$), but they observe that more cognitively able subjects converge more frequently to equilibrium over time. Georganas, Healy and Weber (2012) find that CRT scores are related to level-2 play in a variety of 2-person BCGs.

they are playing the BCG against experienced graduate students instead of other undergraduates. In a similar vein Alaoui and Penta (2013) show that subjects' level of reasoning in an 11-20 game varies with their knowledge about the inferred sophistication of opponents.[4]

More generally, our findings add to an emerging literature that explores the impact of cognitive ability on economic behavior. For example, the studies of Benjamin, Brown and Shapiro (2013), Burks et al. (2009), Dohmen et al. (2010), Frederick (2005) or Huck and Weizsäcker (1999) demonstrate that higher cognitive ability is associated with less biased risk-taking and time discounting behavior.[5] In light of this evidence, our results emphasize the importance of cognitive ability for economic behavior in strategic contexts.

2. Setup

We have data from 240 subjects playing BCGs ($p = 2/3$) in 15 computerized sessions, which were conducted at the WZB-TU laboratory in Berlin between May 2010 and June 2013.[6] Subjects were randomly matched into groups of six and the winner's prize was 5 euros, which was equally shared in case of a tie.[7] The instructions of the game, including an illustrative example, were presented on the computer screen. Before subjects proceeded to the game they had to indicate whether they understood the game and, if necessary, they could clarify any open issues. Subjects learned the outcome of the BCG at the end of a session.

We measured subjects' cognitive ability with a simple test, the cognitive reflection test (CRT), introduced by Frederick (2005). The CRT is a three-question test that builds upon the presumption that our brain uses two types of cognitive processes, a more intuitive or emotional

---

[4] Alaoui and Penta (2013) show theoretically that their experimental results are consistent with individuals engaging in a cost-benefit analysis of applying additional rounds of reasoning. See also Strzalecki (2010) for an alternative theoretical approach in which behavior also depends on own bounded rationality and beliefs about opponents' bounds.

[5] Other studies find that higher cognitive ability is, for example, related to better financial decision making (Agarwal and Mazumder, 2013) and stock market participation (Christelis, Jappelli and Padula, 2010) or health insurance take up (Fang, Keane and Silverman, 2008).

[6] The BCG was conducted alongside other experimental modules (for details and subjects characteristics, see Appendix A and Fehr, 2013). Subjects were recruited through ORSEE (Greiner, 2004) and the experiment was run with z-tree (Fischbacher, 2007).

[7] In four sessions subjects were unintentional matched in groups of twelve (due to a programming mistake) and thus in total there are 36 groups in the BCG. While the average choice turned out to be similar in groups of twelve (42.3) and six (45.8), the regression analysis controls for group-size effects.

process and a more deliberate and cognitively demanding process. The test measures subjects' ability to inhibit intuitive (but wrong) responses and therefore measures one particular component of cognitive ability. The CRT is known to correlate with various behavioral biases, such as the conjunction fallacy or the base-rate fallacy (Oechssler, Roider and Schmitz, 2009, Hoppe and Kusterer, 2011), and with the incidence of Nash equilibrium play (Grimm and Mengel, 2012, Altmann and Falk, 2011).

The CRT was administered before subjects learned about and played the BCG. After the test, subjects had to estimate for each of the three questions the proportion of correct answers among the subjects in their session. This provides us for each subject with a measure for the perceived difficulty of each question as well as with a measure of how they assess others' cognitive ability. We will use this measure in our analysis in order to identify the link between strategic awareness and cognitive ability. Subjects were rewarded with 1 Euro for each correct answer in the CRT and with 25 cents for each correct prediction of the proportion of a correct answer to a CRT question, i.e., they could earn a modest 75 cents if all three predictions were correct.

We complemented the CRT with a five-minute 20-question variant of the Wonderlic Personnel Test (WPT). This test provides a measure for general cognitive ability and we use it mainly to check the robustness of the CRT. Therefore, we did not elicit subjects' beliefs about other subjects' WPT scores. The WPT was administered at the end of a session and subjects were rewarded with 25 Cents per correct answer.[8]

3. Results

Figure 1 shows the distribution of choices in the BCG. While choices are spread out over the whole interval, there are two modes at 50 and 33 (about 68 percent of choices are below 51). The average choice was 45.1, which is higher than in previous experiments. For example, Nagel (1995) and Agranov et al. (2012) report an average of 36.7 and 35.1 for initial choices in their $p = 2/3$ BCG, whereas Gill and Prowse (2013) report an average of 44.2 for first-round choices in a $p = 7/10$ BCG. About 19 percent of choices were in the range between 67 and 100, which

[8] In one session there was a false fire alarm during the WPT and hence the test scores are not used in the analysis. The other tasks were not affected because they were completed before the false alarm.

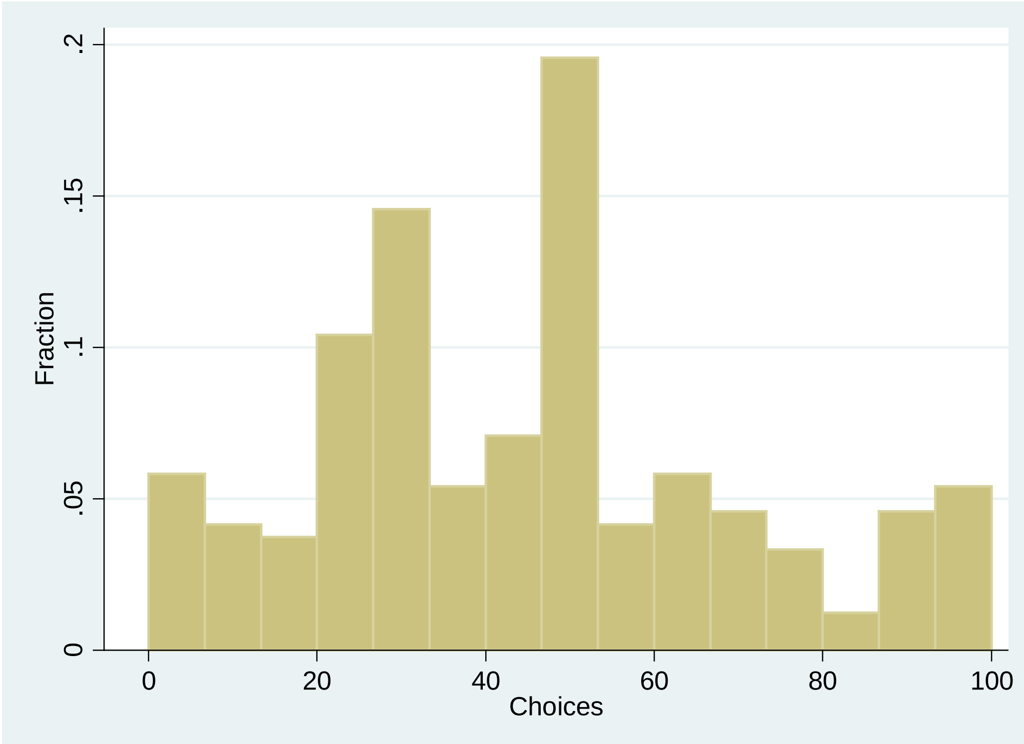

Figure 1: Distribution of Choices in the BCG

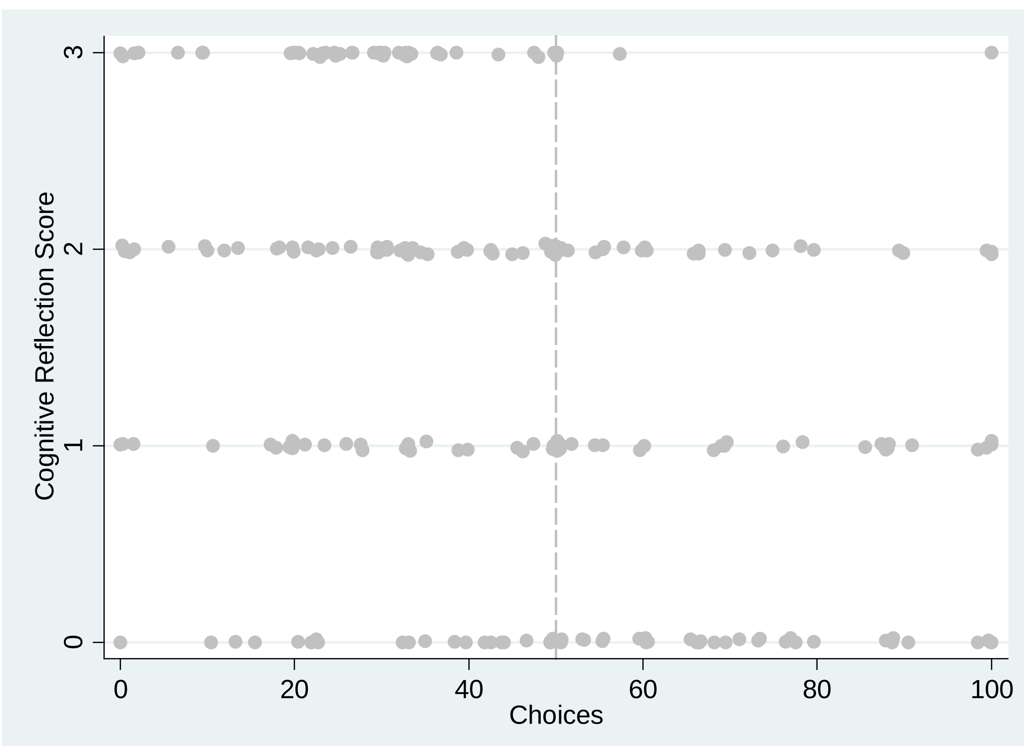

Figure 2: Choices in the BCG and relationship to cognitive reflection (CRT score).

Table 1: Distribution of answers in the CRT.

Question	Answer correct	intuitive	other	expected correct answers
Bat and Ball (BBQ)	32%	60%	8%	65%
Widgets	59%	29%	12%	59%
Lake	58%	21%	21%	52%

Notes: *BBQ*: A bat and a ball cost 1.10 Euro in total. The bat costs 1.00 Euro more than the ball. How much does the ball cost? *Widgets:* If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? *Lake:* In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

is clear evidence for a violation of the dominance principle.[9] Choices in this range are associated with lower cognitive reflection: the average CRT score of subjects choosing a number between 67 and 100 is 0.89, which is substantially and significantly lower than the average CRT score of 1.63 for subjects choosing numbers below 67 ($p < 0.01$, two-sided t-test).

Figure 2 shows the relationship of BCG choices and CRT score. The figure reveals a striking non-linear relationship. As can be seen, subjects who answered two or less questions correctly chose numbers over the whole interval as if they choosing randomly. In contrast, subjects who answered all three questions correctly systematically avoid numbers above 50. We can reject that the median choice is the same for more cognitively able subjects (3 correct answers) and less cognitively able subjects (2 or less correct answers) (Mann-Whitney test, z = 5.351, p<0.01).

Table 1 shows the distribution of correct answers for the three CRT questions as well as subjects' expectations about how many others would answer the question correctly.[10] Our subjects found the „Bat and Ball" question (henceforth BBQ) more difficult than the other two

[9] Recent evidence by Agranov, Caplin and Tergiman (2012) shows that more than 50 percent of subjects choose at least once a dominated action when subjects can modify their choice as often as they want within a three-minute time frame. In their setup, non-strategic subjects switch their choices more than five times as often as strategic subjects and the average choices of non-strategic subjects are remarkably close to 50. Similarly, we find that, in particular, the choices of subjects with a CRT score of 0 or 1 are close to 50 on average (about 48 percent).

[10] The average score in the 3-item CRT is 1.49. About 48% of the subjects have one or less answer correct and 52% of the subjects have 2 or 3 answers correct. Other studies find similar average scores. For instance, the mean score of the whole sample in Frederick (2005) was 1.24 (with a maximum of 2.18 and a minimum of 0.57), whereas Hoppe and Kusterer (2009) and Oechssler, Roider and Schmitz (2009) report slightly higher scores of 1.84 and 2.05 among German students, respectively.

questions.[11] Only 32 percent of subjects gave the correct answer to the BBQ, whereas almost twice as many gave correct answers to the other two questions. Almost two-thirds of those subjects, who answered the BBQ correctly (62 percent), did answer all three questions correctly. The average choice in the BCG of subjects answering all three questions correctly was 29.1, whereas the average choice was 38.3 for subjects with a correct BBQ but less than the maximum CRT score.

It is only superficially surprising that subjects estimates of others' performance are inversely related to the actual performance. Those who believe in the "intuitive" but wrong answer to the BBQ must be convinced that the answer is stunningly simple. Indeed, they expect that on average 73 percent of other subjects gave the correct answer to the BBQ. In contrast, we find that subjects who got the correct answer to the BBQ have a better assessment of its difficulty. They expect that 52 percent of subjects gave the correct answer to the BBQ ($p < 0.01$, two-sided t-test).[12]

Table 2 presents estimation results from Tobit regressions that examine the relationship between choices in the BCG and cognitive ability controlling for individual characteristics such as gender, major and time enrolled. The first specification in column (1) indicates that subjects with a technical background choose lower numbers than subjects from non-technical fields. Once we include our cognitive measures, this effect is smaller. In the regression in column (2) we include a dummy variable for subjects with the maximum CRT score of three to account for the observed non-linear effect of cognitive ability. The result confirms that those subjects with the maximum score in the CRT choose significantly lower numbers in the BCG. As shown above, the questions in the CRT varied in their difficulty, and the BBQ on its own appears to be a good proxy for the maximum CRT score. In column (3) we use the BBQ as an explanatory variable

[11] The "Bat and Ball" question (BBQ) is the following: A bat and a ball cost 1.10 Euro in total. The bat costs 1.00 Euro more than the ball. How much does the ball cost? The other questions are: If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

[12] It is also the case that subjects who answered the BBQ correctly were more likely to win the prize ($p = 0.04$, χ^2-test).

Table 2: Regressions -- Cognitive reflection and choices.

	Dependent variable: choice in beauty-contest game			
	(1)	(2)	(3)	(4)
CRT score = 3 (d)		-19.534***		
		(3.792)		
BBQ (d)			-18.573***	
			(2.811)	
WPT Score				-3.522***
				(0.690)
Female (d)	4.841	5.126	4.624	4.977*
	(3.265)	(3.165)	(3.038)	(2.775)
Major: engineering (d)	-9.625**	-7.173**	-6.188*	-6.842
	(3.717)	(3.299)	(3.178)	(4.187)
# of terms enrolled	4.091*	4.535**	4.994**	5.347**
	(2.339)	(2.243)	(2.384)	(2.285)
Instruction time	3.830	1.308	0.900	0.591
	(4.414)	(4.257)	(4.536)	(3.577)
Group size (d)	-5.595	-4.441	-6.683**	-6.775*
	(3.426)	(3.203)	(3.182)	(3.863)
Constant	27.913	40.307**	43.452**	78.389***
	(19.884)	(19.624)	(20.333)	(14.297)
N	240	240	222	240

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tobit regressions with standard errors clustered on the session level in parentheses. “CRT score = 3 (d)” is a dummy variable, which equals one if a subject answered all three questions in the CRT correct. “WPT Score” is the number of correct answers in the WPT, ranging from 6 to 17. “BBQ (d)” is a dummy variable and equals 1 for a correct answer to the Bat & Ball question (BBQ). “Instruction time” is the time taken for reading the instructions and questions in log(seconds). The number of terms enrolled denotes the time enrolled at the university in log(semester). “Group size (d)” indicates sessions were the group size was n = 12 instead of n = 6.

instead of the maximum CRT score. The resulting coefficient is similar to the previous estimate and highly significant. This lends further support to the observation that the BBQ on its own is an important indicator for cognitive ability and subsequently for choices in the BCG.[13]

The CRT measures subjects' ability to reflect on a problem and to suppress intuitive responses as well as mathematical skills. This is only one particular aspect of cognitive ability and we want to see whether the observed relationship also holds for a more general measure of cognitive ability such as the WPT score. The correlation between the CRT score and the WPT score is about 0.47 and statistically significant ($p<0.01$), which indicates that the two tests indeed capture a similar underlying cognitive trait. Therefore it is not surprising to see that the association between CRT score and choices in the BCG is robust to using the WPT score as an alternative measure for cognitive ability (column 4). This result also suggests that even though the CRT is arguably a coarse measure for cognitive ability (only three questions), the questions (in fact, the BBQ alone) are powerful enough to detect substantial differences in cognitive ability in the context of our experiment.

We hypothesized in our introduction that cognitive ability is related to strategic awareness, i.e., subjects' understanding that they have to think about others. While we have seen that those who fail on at least one of the CRT questions behave in a way that is indistinguishable from uniform random choice, we have not yet seen direct evidence that those who did answer all questions correctly do reason about others. For that purpose we will now utilize subjects' beliefs about others' cognitive ability. Note that eliciting beliefs about others' choices in a BCG with standard instructions would in essence be equivalent to asking for their strategic choice (albeit multiplied by *p*). Moreover, eliciting beliefs directly may induce strategic reasoning as it makes the strategic aspect of the BCG more salient.[14] Thus we prefer to examine whether subjects' beliefs about others' cognitive ability correlates with their choice in the BCG.

[13] That the BBQ is an important indicator for choices in the BCG is also confirmed by testing the predictive power that each of the three CRT questions provides conditional on the other questions. That is, including a dummy variable for each correctly answered question reveals that the BBQ has by far the largest the impact on choices in the BCG and we can reject the hypothesis that all three questions are equally predictive (Wald test, $p = 0.058$, two sided).

[14] A similar problem may arise by using post-experimental questions. Even though such questions cannot influence behavior in the game if presented afterwards, they can induce biased responses because subjects may realize ex-

In order to investigate the relationship between cognitive ability and strategic awareness we focus on subjects' beliefs about how others will do in the BBQ. We concentrate on this question because previous studies revealed that it is the most difficult question (Oechssler, Roider and Schmitz, 2009, Hoppe and Kusterer, 2011, Meyer, Spunt and Frederick, 2013) and thus it may give us the best assessment of what subjects think about other subjects' cognitive ability.[15] If our hypothesis about strategic awareness is correct, we should find that beliefs about others' performance in the BBQ matter for subjects with the maximum CRT score of three but not for subjects with a lower CRT score.

The Tobit regressions in columns 1 and 2 of Table 3 present our main evidence regarding the role of cognitive ability for strategic awareness. In the first column we focus only on subjects that did achieve the maximum CRT score, whereas in column 2 we only examine subjects with a lower score. The difference between both columns is pronounced. While the BCG choices of subjects who answered all three CRT questions correctly are highly sensitive to their beliefs about how other subjects answered the BBQ, other subjects' choices do not correlate at all with their beliefs about others' cognitive ability.[16] We should note that our results not necessarily imply that subjects actively think about others' ability to answer the BBQ when choosing their number. It rather suggests that subjects who are perfect in the CRT not only have a better understanding of the task but are also inclined to reason about others and behave strategically.

This result is robust to using an alternative belief measure and sample split. Instead of using only a subject's expectation about correct answers to the BBQ, we use all three belief statements of a subject. That is, we consider a subject's average expectation of correct answers to the three questions as an explanatory variable (columns 3 and 4). However, we have to keep in mind that this is only a coarse measure for beliefs about the overall cognitive ability of others

---

post that the decision environment was strategic. Note that in the BCG it is not possible to distinguish cleanly between non-strategic and strategic behavior, since low numbers might be evidence for both.

[15] Relatedly, the BBQ (and more generally the CRT) is found to be a good predictor for rational thinking, i.e., individuals' immunity to cognitive biases (see e.g., Toplak, West and Stanovich, 2011).

[16] Gill and Prowse (2013) find a similar result in an environment that facilitates learning. Using data from a repeated BCG they find that subjects with higher cognitive ability vary their level of reasoning depending on the known cognitive ability of their opponents, whereas low ability subjects do not react at all to the cognitive ability of their opponents. However, they do not find differences between high and low ability subject for initial responses in the BCG.

Table 3: Regressions – Beliefs about cognitive ability and choices.

	Dependent variable: choice in beauty-contest game					
	(1)	(2)	(3)	(4)	(5)	(6)
	CRT = 3	CRT < 3	CRT = 3	CRT < 3	BBQ correct	BBQ wrong
Expected proportion BBQ correct	-15.693**	-6.333			-11.748**	-7.309
	(7.498)	(6.363)			(5.519)	(6.709)
Avg. expected proportion correct			-12.747	-5.032		
			(8.321)	(10.052)		
Female (d)	0.715	6.207	-0.641	6.253	-1.941	7.744*
	(5.401)	(3.762)	(5.623)	(3.809)	(3.964)	(4.386)
Major: engineering (d)	0.914	-8.584**	1.112	-8.896**	-4.001	-6.372
	(8.882)	(4.105)	(8.261)	(4.123)	(8.707)	(4.224)
# of terms enrolled	0.125	5.641**	0.103	5.441**	3.388	5.782**
	(5.908)	(2.343)	(5.647)	(2.223)	(4.964)	(2.351)
Decision time	-9.898*	4.105	-8.986*	4.179	-2.536	1.657
	(5.072)	(5.280)	(5.168)	(5.353)	(4.595)	(6.131)
Group size (d)	-8.887**	-6.170	-8.732*	-5.844	-10.934***	-7.636**
	(3.916)	(3.900)	(4.363)	(3.826)	(3.455)	(3.857)
Constant	81.275***	31.261	76.680***	30.074	50.400**	42.945
	(23.501)	(24.479)	(23.094)	(27.005)	(20.494)	(28.834)
N	48	192	48	192	77	163

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tobit regressions with standard errors clustered on the session level in parentheses. "Expected proportion BBQ correct" is the expected proportion of correct answers to the BBQ in a session (ranging from 0 to 1) and "Average expected proportion correct" is the average expected proportion of correct answers to each CRT question. "Instruction time" is the time taken for reading the instructions and questions in log(seconds). The number of terms enrolled denotes the time enrolled at the university in log(semester). "Group size (d)" indicates sessions were the group size was n = 12 instead of n = 6.

and also does not account for the observed non-linearity. (Ideally we would have liked to include information about the distribution of test scores within a session.) Nevertheless, the coefficient estimates are remarkably close to the estimates of our main specification in both subsamples. They are less precisely estimated and for subjects with the maximum score in the CRT the coefficient is at the margin of being statistically significant (column 3). In the last two columns of Table 3, we use an alternative sample split, which allows us to focus directly on the beliefs of those who answered the BBQ correctly. Accordingly, we split the sample into subjects with a correct answer to the BBQ and into subjects with a wrong answer to that question and focus on the estimated proportion of correct answers to the BBQ within a session (as in our main specification). Again, we find that the corresponding coefficient for "Expectation BBQ correct" is significant at the 5-percent level for subjects with a correct BBQ (column 5), but not for subjects who fail on the BBQ (column 6). In summary, we find support for our hypothesis. Subjects above a certain threshold for cognitive ability, i.e., those with a perfect score in the CRT, adjust their choices in the BCG based on their expectations about the cognitive ability of others.

4. Concluding Remarks

In this note we introduced the new concept of strategic awareness, which describes subjects' understanding that reasoning about others is a prerequisite for reasoning about how to play the game. We presented evidence that subjects with higher cognitive ability are more likely to be aware of strategic aspects than less cognitively able subjects.

Strategic awareness will often convey important advantages in life. For example, in retail finance markets consumers' failure to understand the financial incentives of advisors can cause serious harm (see, e.g., Inderst and Ottaviani, 2012 and Chater, Huck and Inderst, 2010) and disclosure rules that create transparency and promote the salience of strategic situation may be an important instrument for consumer protection. There are, of course, also situations where a lack of strategic awareness can be advantageous by inducing *de facto* commitment power. Cournot experiments are a good example, where subjects that simply choose quantities somewhere in the middle of the action space and do not (best) respond to others unwittingly become

Stackelberg leaders outperforming their more sophisticated competitors, see e.g., Huck, Normann, and Oechssler (1999).

While we found evidence for strategic awareness in a specific environment, it is an open question how strategic awareness depends on context. More research in that direction as well as research that explicitly studies how strategic awareness is acquired and how it can be promoted through hints or nudges appear fruitful for both, fundamental and policy-related applied research.

References

Alaoui, Larbi and Antonio Penta. 2013. Endogenous Depth of Reasoning. mimeo.

Agarwal, Sumit, and Bhashkar Mazumder. 2013. Cognitive Abilities and Household Financial Decision Making. *American Economic Journal: Applied Economics*. 5 (1): 193-207.

Agranov, Marina, Andrew Caplin, and Chloe Tergiman. 2012. Naive Play and the Process of Choice in Guessing Games. mimeo.

Agranov, Marina, Elizabeth Potamites, Andrew Schotter, and Chloe Tergiman. 2012. Beliefs and Endogenous Cognitive Levels: An Experimental Study. *Games and Economic Behavior*. 75 (2): 449-463.

Altmann, Steffen, and Armin Falk. 2011. The Impact of Cooperation Defaults on Voluntary Contributions to Public Goods. mimeo.

Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro. 2013. Who is Behavioral? Cognitive Ability and Anomalous Preferences. *Journal of the European Economic Association*. forthcoming

Branas-Garza, Pablo, Teresa Garcia-Munoz, and Roberto H. Gonzalez. 2012. Cognitive Effort in the Beauty Contest Game. *Journal of Economic Behavior & Organization*. 83 (2): 254-260.

Burchardi, Konrad, and Stefan Penczynski. 2014. Out of Your Mind: Eliciting Individual Reasoning in One Shot Games. *Games and Economic Behavior*. forthcoming.

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

Burnham, Terence C., David Cesarini, Magnus Johannesson, Paul Lichtenstein, and Björn Wallace. 2009. Higher Cognitive Ability is Associated with Lower Entries in a p-Beauty Contest. *Journal of Economic Behavior and Organization*. 72 (1): 171-175.

Camerer, Colin. 2003. *Behavioral game theory: Experiments on strategic interaction.* Princeton: Princeton University Press.

Chater, Nick, Steffen, Huck, and Roman Inderst. 2010. Consumer Decision Making in Retail Financial Services. Report prepared for SANCO/EC. 2010.

Chou, Eileen, Margaret, McConnell, Rosemarie, Nagel and Charles Plott. 2009. The Control of Game Form Recognition in Experiments: Understanding Dominant Strategy Failures in a Simple Two Person "Guessing" Game. *Experimental Economics*. 12 (2): 159-179.

Christelis, Dimitris, Tullio, Jappelli, Mario Padula. 2010. Cognitive Abilities and Portfolio Choice. *European Economic Review*. 54 (1): 18-38.

Coricelli, Giorgio, and Rosemarie Nagel. 2009. Neural Correlates of Depth of Strategic Reasoning in Medial Prefrontal Cortex. *Proceedings of the National Academy of Sciences: Economic Sciences*. 106 (23): 9163-9168.

Costa-Gomes, Miguel A., and Vincent P. Crawford. 2006. Cognition and Behavior in Two-Person Guessing Games: An Experimental Study. *American Economic Review* 96 (5): 1737-1768.

Crawford, Vincent P., Miguel A. Costa-Gomes, and Nagore Iriberri. 2013. Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications. *Journal of Economic Literature* 51 (1): 5-62.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2010. Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review*. 100 (3): 1238-1260.

Fang, Hanming, Michael P. Keane, and Dan Silverman. 2008. Sources of Advantageous Selection: Evidence from the Medigap Insurance Market. *Journal of Political Economy* 116 (2): 303-350.

Fehr, Dietmar. 2013. Coordination Failure and Forgone Communication. mimeo.

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

Frederick, Shane. 2005. Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*. 19(4): 25-42.

Georganas, Sotiris, Healy, Paul and Roberto Weber. 2012 “On the Persistence of Strategic Sophistication”. mimeo.

Gill, David, and Victoria Prowse. 2013. Cognitive Ability and Learning to Play Equilibrium: A Level-*k* Analysis. mimeo.

Greiner, Ben. 2004. An Online Recruitment System for Economic Experiments. In Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63. ed. Kurt Kremer and Volker Macho. 79-93. Ges. fuer Wiss. Datenverarbeitung, Goettingen.

Grimm, Veronika, and Friederike Mengel. 2012. An Experiment on Learning in a Multiple Games Environment. *Journal of Economic Theory*. 147 (6): 2220-2259.

Grosskopf, Brit, and Rosemarie Nagel. 2008. The Two-Person Beauty Contest. *Games and Economic Behavior.* 62 (1): 93-99.

Ho, Teck-Hua, Colin Camerer, and Keith Weigelt. 1998. Iterated Dominance and Iterated Best Response in Experimental p-beauty Contests. *American Economic Review.* 88 (4): 947-969.

Hoppe, Eva, and David J. Kusterer. 2011. Behavioral Biases and Cognitive Reflection. *Economics Letters*. 110 (2): 97-100.

Huck, Steffen, and Georg Weizsäcker. 1999. Risk, Complexity, and Deviations from Expected-Value Maximization, *Journal of Economic Psychology*. 20 (6): 699-715.

Huck, Steffen, Hans-Theo Normann, and Jörg Oechssler. 1999. Learning in Cournot Oligopoly: An Experiment. *Economic Journal*. 109 (454): C80-C95.

Inderst, Roman, and Marco Ottaviani. 2012. How (Not) to Pay for Advice: A Framework for Consumer Financial Protection. *Journal of Financial Economics*. 105 (2): 393-411.

Meyer, Andrew, Bob Spunt, and Shane Frederick. 2013. The Bat and Ball Problem. mimeo.

Nagel, Rosemarie. 1995. Unraveling in Guessing Games: An Experimental Study. *American Economic Review*. 85(5): 1313-1326.

Oechssler, Joerg, Andreas Roider, and Patrick W. Schmitz. 2009. Cognitive Abilities and Behavioral Biases. *Journal of Economic Behavior and Organization*. 72 (1): 147-152.

Stahl, Dale O., and Paul W. Wilson. 1995. On Players Models of Other Players: Theory and Experimental Evidence. *Games and Economic Behavior*. 10 (1): 218-254.

Strzalecki, Tomasz. 2010. Depth of Reasoning and Higher-Order Beliefs. mimeo.

Toplak, Maggie E., Richard F. West and Keith E. Stanovich. 2011. The Cognitive Reflection Test as a Predictor of Performance on Heuristics-and-Biases Tasks. *Memory and Cognition*. 39 (7): 1275-1289.

Appendix (not for publication)

A) Experimental Procedures

Table A1: Subject characteristics

Variable	N	Mean	SD
Female	240	0.40	0.49
Age	240	24.41	4.83
Major: engineering	240	0.70	0.46
# of terms enrolled (semesters)	240	6.02	4.41
Reading BCG instructions (in seconds)	240	92.83	38.70

The beauty-contest game (BCG) was conducted alongside other experimental modules (for more details, see Fehr 2013). The sequence of these modules was as follows. Subjects participated first in a repeated minimum-effort game (ten rounds) and completed the cognitive reflection test (CRT) – see Appendix C – as well as questions about how many other subjects answered each of the CRT questions correctly afterwards. Then they were informed about the BCG and that they were matched into new groups of six. After they played the BCG, subjects had to complete a risk elicitation task and had to answer several questions regarding hypothetical sooner or later payments. Finally, they completed a five-minute 20-question variant of the Wonderlic Personnel Test (WPT). Subjects were aware that a session would consist of several different modules, but they had no a priori information on the nature of these modules. There was no feedback about outcomes between modules, but subjects received feedback about their payoff after each of ten rounds in the minimum-effort game. Although only 5 of 10 rounds of the minimum-effort game were selected randomly for payment at the end of a session, subjects could infer the approximate range of their payment from their feedback after each round. However, we do not find evidence that subjects' payoffs in the minimum-effort game affected their subsequent play in the BCG ($p > 0.87$, two-sided t-test, standard errors clustered on the group level) and also not their performance in the CRT ($p > 0.53$, two-sided t-test, standard errors clustered on the group level) or WPT ($p > 0.54$, two-sided t-test, standard errors clustered on the group level).

B) Instructions (Translation from German)

Task:

You will be randomly matched into a group with 5 other participants. The group will not be the same as in the previous part.

Your task is to choose a number between 0 and 100. The number 0 and 100 are also possible and the number you choose can have up to two decimals.

Payoff:

First, we will calculate a target number. The participant in a group, who is closest to this target number, will get 5 Euro. All other participants in a group will get nothing. If two or more participants in a group are equally close to the target number, the prize will be shared among those participants.

How we determine the target number:

First, we will calculate the average of all chosen numbers in your group. That means that we add up all numbers and then divide the sum by six. Then we multiply this average by 2/3 (two-thirds), that is the target number = 2/3 x the average of chosen numbers in your group.

Here is an example:

Suppose the other five participants in your group all choose 100 and you will choose 99. In this case, the average of the chosen numbers in your group is 99.83 [=(5 x 100 +99)/6] and thus the target number will be 66,55 [=2/3*99,83]. Therefore your chosen number will be closest to the target number and you would get the prize of 5 Euro.

If you have any questions, please raise your hand and the experimenter will clarify your question in private.

C) CRT questions:

1. **BBQ:** A bat and a ball cost 1.10 Euro in total. The bat costs 1.00 Euro more than the ball. How much does the ball cost?

2. **Widgets:** If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

3. **Lake:** In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

Discussion Papers of the Research Area Markets and Choice 2013

Research Unit: **Market Behavior**

Nadja Dwenger, Dorothea Kübler, Georg Weizsäcker Preference for Randomization: Empirical and Experimental Evidence	SP II 2013-201
Kai A. Konrad, Thomas R. Cusack Hanging Together or Being Hung Separately: The Strategic Power of Coalitions where Bargaining Occurs with Incomplete Information	SP II 2013-202
David Danz, Frank Hüber, Dorothea Kübler, Lydia Mechtenberg, Julia Schmid 'I'll do it by myself as I knew it all along': On the failure of hindsight-biased principals to delegate optimally	SP II 2013-203
David Hugh-Jones, Morimitsu Kurino, Christoph Vanberg An Experimental Study on the Incentives of the Probabilistic Serial Mechanism	SP II 2013-204
Yan Chen, Onur Kesten From Boston to Chinese Parallel to Deferred Acceptance: Theory and Experiments on a Family of School Choice Mechanisms	SP II 2013-205
Thomas de Haan, Roel van Veldhuizen Willpower Depletion and Framing Effects	SP II 2013-206
Dietmar Fehr, Steffen Huck Who knows it is a game? On rule understanding, strategic awareness and cognitive ability	SP II 2013-306
Christine Binzel, Dietmar Fehr Giving and sorting among friends: evidence from a lab-in-the field experiment	SP II 2013-207

Research Unit: **Economics of Change**

Luisa Herbst, Kai A. Konrad, Florian Morath Endogenous Group Formation in Experimental Contests	SP II 2013-301
Kai A. Konrad, Florian Morath Evolutionary Determinants of War	SP II 2013-302
Armin Falk, Nora Szech Organizations, Diffused Pivotality and Immoral Outcomes	SP II 2013-303
Maja Adena, Steffen Huck, Imran Rasul Charitable Giving and Nonbinding Contribution-Level Suggestions. Evidence from a Field Experiment	SP II 2013-304

All discussion papers are downloadable:
http://www.wzb.eu/en/publications/discussion-papers/markets-and-choice

Dominik Rothenhäusler, Nikolaus Schweizer, Nora Szech SP II 2013-305
Institutions, Shared Guilt, and
Moral Transgression

Dietmar Fehr, Steffen Huck SP II 2013-306
Who knows it is a game? On rule understanding, strategic awareness
and cognitive ability

All discussion papers are downloadable:
http://www.wzb.eu/en/publications/discussion-papers/markets-and-choice