C.E.S.S. Working Paper Series

CESS
center for experimental social science

Working Paper #2007-13

# Endogenous Cognitive Types: An Experimental Study

Elizabeth Potamites and Andrew Schotter

Key Words: Beauty Contest; Cognitive Types; Cognitive Hierarchies
JEL Classification: C72, C91

19 West Fourth Street, New York, New York 10003, 212-992-8660

# The Center for Experimental Social Science

*The Center for Experimental Social Science (C.E.S.S.) is an inter-disciplinary research Center at New York University engaged in laboratory experimental work in the social sciences that combines high quality economic theory, psychology and political science. The over arching research tool used in the Center is laboratory experiments in which human subjects engage in decision-making tasks whose purpose is to test both economic theory and/or the properties of proposed or existing economic, political, or social institutions. The working papers in this series reflect these interests.*

**Director:** Andrew Schotter (Economics)

**CoDirectors:** Andrew Caplin (Economics)
Rebecca Morton (Politics)
Tom Tyler (Psychology)

## Staff:

**Director of Operations/Research Scientist:** Rajeev G Advani_
**Technical Team:** Anwar Ruff

## Faculty

Andrew Caplin (Economics)
Kfir Eliaz (Economics)
Eric Dickson (Politics)
Guillaume Frechette (Economics)
Yaw Nyarko (Economics)
Andrew Schotter (Economics)
Tom Tyler (Psychology)

## Faculty Affiliates

Jess Benhabib (Economics, Faculty of Arts and Science)
Alberto Bisin (Economics, Faculty of Arts and Science)
Sally Blount-Lyon (Management and Organizations, Stern)
Luis Cabral (Management/OB, and Economics, NYU Stern)

Peter Carnevale (Psychology, Faculty of Arts and Science)
Marisa Carrasco (Psychology, Faculty of Arts and Science)
Douglas Gale (Economics, Faculty of Arts and Science)
Paul Glimcher (Neuroscience)
Peter Gollwitzer (Psychology, Faculty of Arts and Science)
John Jost (Psychology, Faculty of Arts and Science)
John Leahy (Economics, Faculty of Arts and Science)
Larry Maloney (Psychology, Faculty of Arts and Science)
Efe Ok (Economics, Faculty of Arts and Science)
Ariel Rubinstein (Economics, Faculty of Arts and Science)
Zur Shapira (Management and Organizations, Stern)
Charles Wilson (Economics, Faculty of Arts and Science)

For Information about C.E.S.S. contact Anwar Ruff at 212-992-8660

# Endogenous Cognitive Types: An Experimental Study

Elizabeth Potamites and Andrew Schotter*

New York University
Department of Economics

April 2007

**Abstract**

A reading of the literature on cognitive hierarchies leaves the impression that a subject's type is predetermined before he comes into the lab so that the distribution of types is exogenous and immutable across games. In this paper we view the choice of a person's cognitive level as endogenous and explain it by focusing on subjects' expectations about the cognitive levels endogenously chosen by others. We run a set of experiments using the $2/3^{rd}$'s guessing game where subjects receive public advice offered by a set of advisors. We discover that certain types of public advice, those that are commonly interpreted as meaningful, are capable of shifting the distribution of observed cognitive types indicating that the distribution is endogenous.

Key Words: Beauty Contest, Cognitive Types, Cognitive Hierarchies

JEL Classification: C72, C91

*We thank Colin Camerer, Syngjoo Choi, Guillaume Fréchette, and Brian Rogers for their comments and gratefully acknowledge financial support from the Center for Experimental Social Science (CESS) at New York University. This work was also supported by the NSF, grant number SES-0425118. All errors are our own.

# 1 Introduction

Experimental data (i.e. data generated by controlled laboratory experiments) many times fail to be explained by models positing a homogeneous population of rational agents. This point has been made in dramatic fashion by Nagel (1995), Stahl and Wilson (1995), Ho, Camerer, and Weigelt (1998), Costa-Gomes and Crawford (2006, 2007), and many others following their lead. These models posit the existence of a set of heterogenous types whose levels of rationality form a hierarchy from the lowest type (typically those who choose randomly) to those who either best respond to the type immediately below them (see Stahl and Wilson (1995)) or to some pre-specified distribution of types of lower rationality (see Camerer, Ho, and Chung (2004)). A careful reading of these papers leaves the impression that a subject's type is predetermined before he comes into the lab so that the distribution of types is exogenous (perhaps fixed in the population) and also immutable across games.

A recent model by Choi (2006) relaxes this assumption by positing that agents differ by the cognitive costs they have in functioning at different rationality levels. How high a type one appears to be is determined by a comparison between the marginal benefits one can expect to receive by choosing a different (higher) rationality level as compared to the marginal cost of doing so. In this analysis agents may wind up acting like different types in different games but in fact they are of one type determined exogenously by their cognitive cost function. So in Choi's model, while people cannot act as if they were smarter than they actually are, they can choose to behave in various ways depending on their comparison between the marginal benefit and marginal cost of increasing their rationality level.[1]

In this paper we also view the choice of a person's cognitive level as endogenous but explain it by focusing on their expectations about the cognitive levels endogenously chosen by others. We posit two distributions of types as being relevant. One, which we call the "objective type distribution", is the actual cognitively-correct distribution of types which accurately describes how many steps of logic a given player is capable of in the context of the game they are playing. Note that this distribution is unobservable. The other is what we will call the “observed type distribution” which measures the distribution

[1] Choi's (2006) analysis is an equilibrium analysis so the marginal benefit to increasing one's rationality is definable under the assumption that all other agents are in equilibrium with respect to their level of rationality.

of revealed cognitive types. The two may differ because agents may have a subjective distribution over the type set which differs from the objective one.

For example, take the $2/3^{rd}$'s-guessing game as discussed by Nagel (1995). In this game a cognitive hierarchy model typically assumes that type-zero agents choose randomly, type one agents best respond to type zeros, type 2's best respond to type 1's etc. If one could look inside each person's brain one could, theoretically, measure how many levels each person is capable of and construct the "objective type distribution". Clearly such a distribution is an upper bound on what people in a population are capable of acting like. They may choose to behave as if they were a lower type, however, if they assume that others in their group are distributed in a manner different from the objective distribution, for example if the objective distribution stochastically dominates their subjective guess of it.

In the experiment below we study behavior in an eight-person $2/3^{rd}$'s guessing game where subjects play after receiving advice from others. More precisely, in these experiments we have subjects participate in the standard $2/3^{rd}$'s guessing game experiment except that before they choose they are given advice from another player. This advice is either private and offered to him or her alone (knowing that each other subject is being offered similar private advice from another subject) or public, in which case the advice is announced for all to hear, and is therefore common knowledge. In this context we offer two hypothesis concerning how and why subjects may exhibit behavior that is different from their objective cognitive ability. Hypothesis 1 we call the "focus hypothesis" since it suggests that a subject will exhibit behavior that is of a higher cognitive level if some event occurs which leads him to focus his attention more on the task at hand than he would have in a control treatment where such an event is absent. In our experiment such an event might be receiving a piece of private advice from another subject on how to behave. For example, say the advice was to choose very high, 88 out of the interval [0,100]. Receiving such advice should lead the subject to ask why is 88 either good or bad. This could then lead them to understand that 88 is dominated and might even lead the subject to realize that backward induction leads them to zero. Our other hypothesis, which we call the "common knowledge" hypothesis, suggests that a subject may decide to behave at a higher cognitive level if an event occurs which is common knowledge for all subjects, and therefore leads them all to think that this event has raised the level of cognitive focus on the task at hand. Such an event here might be the issuing of a piece of public advice or some other commonly observed public event.

The implicit model behind these hypotheses works as follows. Assume a subject comes into the lab and reads the instructions of the $2/3^{rd}$ guessing game experiment. Based on his or her cognitive ability and other mental states (how tired he or she is etc.) he or she decides how hard to think about the problem at hand. The behavior exhibited here would be that observed by Nagel (1995) and others who have played this game without advice and defines the observed type distribution. Now say that in another treatment (the focus treatment) the subjects are offered some private advice before they make their choice. This advice may lower their cost of thinking and hence lead them to act more rationally or give them an insight into the problem they had not seen before, in which case they will walk up the cognitive hierarchy. Alternatively, we could run a "common knowledge" treatment in which the advice is public and publicly announced. Now subjects may say that while they were not affected by the announced advice they believe that others are and that others believe that others are etc. In this case it might be advantageous to behave more rationally because you believe that the type-distribution has shifted. So people here climb the cognitive hierarchy because they believe that others are climbing it as well and they must keep up with them (best respond).

Our experiments are designed to separate and test these two hypotheses. Our results provide more support for the common knowledge hypothesis than for the focus hypothesis. More precisely, providing subjects with private advice as to what to choose leads to only a minimal shift in the distribution of observed types. This would indicate that their revealed behavior was not due to a lack of cognitive ability on their part but rather diminished expectation about the cognitive abilities of their cohorts. Providing them with public advice shifts the distribution significantly but only if the subjects believe that the advice offered to them is meaningful (or believe that others will believe it is meaningful). This supports the common knowledge hypothesis because it demonstrates that not all advice revealed publicly can change the observed distribution of types. The advice must be public and credible and commonly assumed to be credible for others.

These results tend to support the view that in any strategic situation people are not functioning at their highest cognitive abilities but rather are poised to if they believe that others will do so. To get them to believe this it is not enough to get them to think about the problem more deeply, as in the private advice case, but rather convince them that others are thinking more deeply about the problem and to make this fact common knowledge, as in our public advice treatment.

In this paper we will proceed as follows. In Section 2 we will present our experiments and discuss our experimental design. In Section 3 we test a set of seven hypotheses that follow from our experimental design. In Section 4 we estimate a modified version of the Camerer, Ho, and Chong (2004) cognitive hierarchy model in an effort to demonstrate how the cognitive type distribution changes as we alter the type of advice presented to subjects. In Section 5 we offer some conclusions.

## 2 Experimental Design and Procedures

Subjects were recruited from the undergraduate student population at New York University and brought into the lab of the Center for Experimental Social Science in groups of 16 where they engaged in a pen and paper experiment. Once in the lab they were offered written instructions describing the $2/3^{rd}$s guessing game (we modified the instructions used by Rosemarie Nagel (1995)) and these instructions were read out loud and questions about them answered.[2] In this game eight subjects are asked to choose a number between 0 and 100. The eight numbers are then averaged and $2/3^{rd}$s of the resulting mean is calculated $(2\overline{x}/3)$ That subject whose number was closest to $2\overline{x}/3$ receives a prize of \$10 while all other subjects receive nothing. If two or more subjects are equally close, they split the \$10 equally. In all treatments but the control, before the instructions were read, subjects were randomly divided into two groups of eight with one group being "advisors" and the other "players". As the instructions indicated, the role of the players was simply to play the $2/3^{rd}$s guessing game. They were paid according to the above description and played the game for 10 rounds so their payoffs were the sum of their payoffs in the 10 rounds of the game.

The other groups of advisors had the task of offering advice to the players. This was done in several ways. In the private advice treatment each advisor was paired with a player and offered him their advice by writing a number to choose on a piece of paper which was anonymously handed to the player's pair member. No one else could see this advice but all subjects knew that such advice was being offered. The advisors were paid in an identical manner as the players. The advisor whose advice was closest to $2\overline{x}/3$ received \$10 while all

[2] See Appendix A for a sample of the instructions offered.

others received zero. Note, however, that the advice offered by advisors did not enter the computation of $2\overline{x}/3$ while the players' choice did. We call this treatment the Low Information Private Advice Treatment (LPri) to distinguish it from another private advice treatment which was identical to this one except that in addition to advisors offering advised numbers to choose, they also offered an unstructured text to accompany it explaining why they thought their advice was meaningful. We call this the High Information Private Advice Treatment (HPri).

In addition to these two private advice treatments we also ran two public advice treatments. In one, called the Mean Public Advice treatment (MPub), before each round we collected the eight pieces of advice just as we did for the private advice treatments but instead of individually handing them out to subjects we calculated the mean and announced it publicly to the players so all could hear. In another public advice treatment, called the Random Public Advice Treatment (1Pub), we did the same thing but rather than announce the mean of the advisors' advice we randomly chose one piece of advice and announced it to all the players. The players were aware of which public advice treatment they were in so they knew if the announced advice was the mean or a randomly chosen piece of advice.

Finally, in the first round of all private and public advice treatments, after the advisors gave advice to their advisees, they were asked to play the $2/3^{rd}$s Guessing Game among themselves. Here they chose a number between 0 and 100 and that player whose choice was closest to $2\overline{x}/3$ of the choices of the advisor group, received $10. Hence, in the first round of the experiment, and only in the first round, advisors could earn money two different ways. One was to give advice which was closest to $2\overline{x}/3$ of the players' group while the other was to make a choice which was closest to $2\overline{x}/3$ of the choices in their group. The results from the advisors game were announced at the end of the experiment.

In our final treatment called the Control Guessing (Control) treatment we again recruited 16 subjects and divided them into two groups. One of the groups were designated as the players and they played a standard $2/3^{rd}$s Guessing Game without advice among themselves. The task given to members of the other group was to make a guess or forecast in each round about the other group of eight subjects. The subject whose guess was closest to $2\overline{x}/3$ of the players won $10 and everyone else won zero. This 'guessers' group was in a separate room and the other players did not know what experiment the 'guessers' were doing so from their perspective they were simply partaking in a conventional guessing game experiment.

In all treatments after each round was over we placed on the blackboard $2\overline{x}/3$ of the players and the ID numbers of the player and advisor (guesser) who won the $10 prize.

Our experimental design is described in Table 1:

**Table 1: Experimental Design**

| Treatment | Advice | Form of Advice | Type of Advice | # of Groups | # of Subjects |
|---|---|---|---|---|---|
| **LPri** | yes | number | Private | 4 | 64 |
| **HPri** | yes | # plus text | Private | 4 | 64 |
| **MPub** | yes | mean of advisors | Public | 4 | 64 |
| **1Pub** | yes | # from 1 advisor | Public | 4 | 63 |
| **Control** | no | NA | NA | 4 | 62 |

## 2.1 Hypotheses

Our Experimental design helps us test our two main hypotheses. If subject behavior is as described by Choi (2005), then behavior in our control treatment should be determined by the trade-off between the marginal benefits of higher cognitive processing and its marginal cost. Providing subjects with private advice should lower these costs by offering some anchor with which to measure one's prior thinking as well as allowing a second round of thinking about the problem. This should cause the subject to focus more on the problem and think at a higher level. Hence we can test our focus hypothesis by comparing behavior in our control treatment with that in both the low and high information private advice treatments (LPri and HPri). If the distribution of types shifts in the direction of more sophistication as a result of private advice, then this shift would offer support for that hypothesis. The null hypothesis can be stated as follows. (We test all of our hypotheses using only the first round data in each of our games since that data is independent of learning which occurs in later rounds. We discuss learning separately.):

**Hypothesis 1: Focus Hypothesis**

**The observed distribution of types in the first round of the $2/3^{rd}$s Guessing Game is not significantly different when comparing the control treatment with either the Low or High information private advice**

**treatments.**

In our experiment advice interacts with thinking in two ways. Advisors play the $2/3^{rd}$s Guessing Game after giving advice (in the first round of our experiments) while advisees play the $2/3^{rd}$s Guessing Game after receiving advice. This allows us to ask which activity, giving or receiving advice, helps to focus attention of subjects most on the game at hand and hence is most likely to shift the observed distribution of types. This yields the following null hypothesis about the relative impacts of these two activities.

**Hypothesis 2: Relative Focus Hypothesis: In the first round of LPri and HPri, there is no difference between the choices of advisors after giving advice and advisees after receiving advice.**

Note that even if private advice allows a subject to focus more on the task, it does not follow that he or she will behave in a more sophisticated manner if he or she does not think that his other cohorts will climb up the cognitive ladder with them. Hence any advice that does not change a subject's expectations about the sophistication of the other players is unlikely to change the observed distribution of cognitive types. This is the motivation for our Common Knowledge hypothesis. According to this hypothesis we would not expect private advice to change the subjects' behavior because even if they benefits from it, our subjects don't know the private advice that was offered others and hence there is no commonly observed event that can lead them to jointly change their expectations about each other. This is not the case when advice is public and common knowledge so we expect under the common knowledge hypothesis that the observed distribution of types should shift when advice is public and lead to more sophisticated behavior no matter what that advice is. This leads to the following null hypothesis.

**Hypothesis 3: Common Knowledge Hypothesis**

**The observed distribution of types in the first round of the $2/3^{rd}$'s Guessing Game is not significantly different when comparing the control treatment with either of the public advice treatments (MPub or 1Pub).**

In addition to these hypotheses our design allows us to test a few more conjectures. For example, our design allows us to investigate the process of advice giving and following. One question that arises is whether in the private advice experiment the advice given is simply their best guess of what the mean in the players groups will be. This can easily be tested since we can compare the advice given in the LPri (HPri) with the guesses made in the control. If advice is simply equal a subjects best guess as to what the mean will be in the players' game, then the distribution of advice offered in the first round of the LPri (HPri) treatment should be identical to the distribution of guesses in the control treatment. This yields the following null hypothesis

**Hypothesis 4: Advice Giving - - Private Advice**

**The distribution of first-round advice in the LPri or HPri treatments should be identical to the distribution of guesses in the control treatment.**

In our experiment we had subjects give advice under two distinctly different conditions – private and public information. This leads us naturally to ask whether the advice offered was different in these two situations. In a previous paper by Chaudhuri, Sopher, and Schotter (2005) advice offered when it was known to be public was quite different from that offered when it was known to be private. This leads us to ask a similar question here and formulate the following null hypothesis.

**Hypothesis 5: Advice Giving - - Public Advice**

**The distribution of first-round advice in the LPri treatment should be identical to the distribution of advice in either the MPub or 1Pub treatments.**

Advice giving is only one half of the equation. We might like to know something about the process of advice following. For this purpose it might be informative to compare the frequency of advice following and the manner in which advice was followed. It is our conjecture that advice is more likely to be followed if its logic is spelled out rather than simply receiving a raw number without any explanation. If this is so we would expect that advice would be followed more in the HPri than in the LPri treatment. This leads to our next null hypothesis.

**Hypothesis 6: Advice Following**

**There is no difference between the frequency or manner in which advice is followed in the first round HPri and LPri treatment**.

Our final hypothesis concerns itself with convergence. Even if advice did not lead our subjects to choose differently in the first round of the experiment, it is still possible that it might have hastened the rate of convergence toward zero as rounds progressed. This may be true for either private or public advice or both when these treatments are compared to the control. In addition we can ask which type of advice speeds up convergence the most. Giving us the following null hypotheses:

**Hypothesis 7: Convergence**

**There is no difference in the rates of convergence toward zero between the control treatment and any private or public advice treatment. In addition, there is not difference in the rate of convergence when we compare private with public advice treatments.**

## 3 Results

We will present our results by testing the seven hypotheses above. In our next section we will use our data to estimate the type distribution observed in each treatment using the Camerer, Ho and Chung (2004) procedure and a modified version of it.

**Hypothesis 1: Focus Hypothesis**

**The observed distribution of types in the first round of the $2/3^{rd}$s Guessing Game is not significantly different when comparing the control experiment control with either the Low or High-information private-advice treatments.**

In order to discuss Hypothesis 1 consider Table 2 and Figure 1. In Table 2 we present the mean choices in round 1 for all of our private advice experiments as well as our control treatment. As you can see, while the mean of control was 29.2 that of the LPri, HPri, were 27.1, 28.4 respectively. When tested using a Kolmogorov-Smirnov test none of the distributions of first round choices were different across these three treatments. This can clearly be seen in Figure 1 where we present the empirical cumulative distributions of these first round choices.

**Figure 1 here**

**Table 2: Private Advice and Control**

| Treatment | $1^{st}$ Round Mean Advice | $1^{st}$ Round Mean Choice | Std dev. of $1^{st}$ Round Choices |
|---|---|---|---|
| **LPri** | 34.5 | 27.1 | 10.9 |
| **HPri** | 37.0 | 28.4 | 16.4 |
| **Control** | NA | 29.2 | 15.6 |

The main import of these tables and figures is to demonstrate that providing subjects with private advice does not focus their attention on the task at hand more than it already is in the control treatment. In other words, it appears as if subjects had already chosen their level of rationality optimally and merely presenting them with one piece of private advice does nothing to lead them to want to change their choice (or what their choice would have been had they not received advice). This suggests that we can not reject Hypothesis 1.

Hypothesis 1 is incomplete since it leaves open the possibility that while *receiving* private advice does not lead to a change in the choices made by subjects, *offering* advice may in fact focus the attention of subjects in a manner that causes them to act at a higher cognitive level. For example, if one struggles to give good advice it very well may be that the process of thinking what this might be could lead a subject to realizations about the game he would not have had he not been forced to give advice or passively receive it. This leads us to consider Hypothesis 2:

**Hypothesis 2: Relative Focus Hypothesis: In the first round of LPri and HPri, there is no difference between the choices of advisors after giving advice and advisees after receiving advice.**

The mean choice of advisors after they gave advice in LPri and HPri was 35.1 and 32.2, compared the mean choices of the receivers, 27.1 and 28.4, respectively. The means in LPri are significantly different when tested using a Mann Whitney U-test (z =2.179, p = .029). In the cdf's of the choices in both treatments, the advisees' choices are shifted to the left relative to the advisors' choices (see Figure 2) and this difference is significant when tested using the Kolmogorov-Smirnov test in the LPri case. This result is somewhat surprising since considerably more thinking is involved in offering advice than passively receiving it (and perhaps ignoring it). Hence we originally thought that it might have an impact on how advisor subjects made their first round choices. Here again we can only conjecture that since such advice does not change a subjects's expectations about his cohorts, it has no impact on behavior.

**Figure 2 here**

The same should not be true for public advice, however, since, being common knowledge, once all subsets hear the announced advice they may start to conjecture that others will respond to it. For example, once a piece of public advice is announced it might make little sense to choose a number above it since a subject knows that others know that others know etc. that they all heard the same announcement. Hence public advice may provide a cognitive anchor from which all subjects commonly adjust their choice downward and hence shift the distribution of types upward. This leads us to investigate Hypothesis 3.

**Hypothesis 3: Common Knowledge Hypothesis**

**The observed distribution of types in the first round of the 2/3$^{rd}$s Guessing Game is not significantly different when comparing the control experiment control with either of the public advice treatments (MPub or 1Pub).**

Table 3 presents the mean results from the first round of our public advice treatments while Figure 3 presents the cdf's of the choices. As we can see, there appears to be no difference between first round choices in the control treatment and the Random Public Advice Treatment (1Pub) (29.2 vs 30.7) while the first round choices in the MPub treatment are much lower (23.7). Using a Mann Whitney U-test the difference between the control and MPub is significant (z = 1.767, p = .077).

**Figure 3 here**

**Table 3: Public Advice and Control**

| **Treatment** | **1$^{st}$ Round Mean Advice** | **1$^{st}$ Round Mean Choice** | **Std dev. of 1$^{st}$ Round Choices** |
|---|---|---|---|
| **MPub** | 36.1 | 23.7 | 9.3 |
| **1Pub** | 33.0 | 30.7 | 15.5 |
| **Control** | NA | 29.2 | 15.6 |

This straightforward result furnishes us with the main point of our paper. Put simply, it appears that the only type of events (advice) that lead to a significant shift in the level of rationality are events which are not only common knowledge but are perceived by subjects to be meaningful to others. More precisely, the only difference between the 1Pub and MPub experiments is that in the MPub experiment subjects are offered public advice which is the mean of the advice from the pool of advisors while in the 1Pub experiment they are offered advice which, while public and common knowledge, is merely a random choice of one piece of advice from the pool of eight advisors. We consider this to mean that while the random public advice was commonly observed it was not commonly considered to be meaningful to all others by the subjects and hence failed to serve as a common cognitive anchor for their thinking. On the other hand, public advice that was the mean of the advisors group in the MPub treatment, appears to have struck our subjects as something to consider as meaningful and which others will take into account when they make their choice of a number. Under these circumstances, a best response is to respond to that advice as well.

One explanation of why public advice of the mean has such a big impact on behavior where other types of advice do not may be that seeing it was equivalent to having seen a round of play of the game you are engaging in (since you learn the mean of the advisors' choices) so what we observe in the first round of our MPub treatment is actually the second round of play of other games. In fact this was far from the truth. In the control treatment the second round mean was 16 (std dev. 5.98) while in the MPub treatment the first round mean was 23.7. We can reject the hypothesis that these means came from the same population at the 1% level using a Mann Whitney U-test. This again supports our Common Knowledge hypothesis as the major explanation.

Our discussion up until now has involved choices and not advice. One question that is natural to ask is whether the advice offered in the LPri and HPri treatments differed significantly from merely the best guesses of the advisors as to what the mean of the players would be. To investigate this we compared the guesses made by our guessing subjects in the control treatment with he advice given by advisors in our two private advice treatments. This yields Hypothesis 4:

**Hypothesis 4: Advice Giving - - Private Advice**

**The distribution of first-round advice in the LPri or HPri treatments should be identical to the distribution of guesses in the control treatment.**

As we can see in Figure 4 the guesses in the control appear be shifted to the left (toward lower numbers) compared to the advice given in private advice treatments. While the mean guess in the control treatment (31.4) is not significantly lower than the mean advice in LPri and HPri treatments (34.5 and 37, respectively) using the Mann-Whitney U-test, a Kolmogorov-Smirnov test confirms that the distribution of guesses or forecasts is significantly different than the distribution of advice in the LPri case (p=.057).

**Figure 4 here**

Hypothesis 4 concerned itself with advice giving in private advice treatments. However, it may be that the advice offered in public advice experiments differed from that offered in the private advice experiment because it was known that such advice would be made public. This conjecture furnishes us with hypothesis 5.

**Hypothesis 5: Advice Giving - - Public Advice**

**The distribution of first-round advice in the LPri and HPri treatments should be identical to the distribution of advice in either the MPub or 1Pub treatments.**

**Table 4: Private and Public Advice**

| **Treatment** | **$1^{st}$ Round Mean Advice** | **Std dev. of $1^{st}$ Round Advice** |
|---|---|---|
| **LPri** | 34.5 | 12.0 |
| **HPri** | 37.0 | 22.1 |
| **MPub** | 36.1 | 17.7 |
| **1Pub** | 33.0 | 15.8 |

Table 4 presents the means (and standard deviations) of the advice given in each of our two private and public advice treatments. Figure 5 presents the disaggregated information behind these means as histograms. As we can see the advice given did not vary much by treatment.

**Figure 5 here**

Up until now we have not commented on how influential advice was in determining the actions of subjects. In other words, did players follow the advice of their advisors? Did they do it more often as treatments changed. Did they do it more often in public rather than private treatments. All of the questions can be tested by comparing the relationship between the advice given and action taken in our private and public advice treatments.

**Hypothesis 6: Advice Following**

**There is no difference between the frequency or manner in which advice is followed in the first round HPri and LPri treatment or in the MPub and 1Pub treatments.**

To get a first insight into testing this hypothesis consider Table 5 which indicates, treatment by treatment, how subjects reacted to the advice offered them. More precisely, it presents the percentage of subjects who chose more than, less than, and approximately equal (+/- 2) the advice offered them.

**Table 5: Advice Following**

| Treatment | Above Advice | Follow Advice | Below Advice |
|---|---|---|---|
| **LPri** | 28% | 28% | 44% |
| **HPri** | 16% | 28% | 56% |
| **1Pub** | 38% | 0% | 62% |
| **MPub** | 9% | 3% | 88% |

Some of these numbers are informative. For example compare the first and last rows of the table which present the following behavior of subjects in our MPub and LPri treatments. As should be obvious, it appears that subjects in the MPub treatment tend to use the advice given them far more often as a cognitive anchor and make choices below it. In fact, 88% of choices are below the announced mean. For the LPri only 44% of subjects chose below the advice they were given. Further, note that practically no subjects followed the advice offered in the MPub treatment (only 3%) while 28% did so in the LPri. This all is consistent with our main findings that in order to raise the observed distribution of cognitive types an event (like advice giving) must be common knowledge and meaningful. To further strengthen this point consider the 1Pub treatment and the fraction of subjects (38%) who choice numbers that were greater than advised to (note that 28% did so in the LPri). The fact that only 9% of subjects did this in the MPub treatment indicates subjects in the 1Pub treatment did not consider the advice announced to be meaningful to all the players since a logical consequence of common knowledge and backward induction in this game is to choose below a number announced if you believe that all others heard that number and that their cohorts will treat it seriously. This is what we mean by a cognitive anchor. Finally notice that while 84% of subjects made choices equal to or less than the advice they were given in the HPri experiment, only 72% did so in the LPri experiment indicating that advice was more persuasive in that treatment or more suggestive about the backward induction needed to think oneself to the Nash equilibrium

Since we repeated experimental game 10 times we can comment on the rate of convergence across our five treatments. We expected that the presence of advice would lead to faster convergence than observed in the control treatment simply because we thought it would hasten the learning process.

**Hypothesis 7: Convergence**

**There is no difference in the rates of convergence toward zero between the control treatment and any private or public advice treatment. In addition, there is no difference in the rate of convergence when we compare private with public advice treatments.**

Figure 6 shows that the rates of the convergence look relatively similar across all treatments. Similar rates of convergence were also found in Nagel's original paper. The main visible difference is the shift in levels.

**Figure 6 here**

Table 6 shows the results of a regression that pools the choices of subjects across all 5 treatments and all 10 rounds. The left hand variable in the regression, subject choice, is regressed on a constant, dummy variables for whether advice was private or public (the omitted category is the control treatment), the round, the round squared, and interaction terms of all of the above. Significant results are indicated. Note that convergence when advice is private (as measured by the coefficient on the variable "round") is not significantly different than the control case. When advice is public, however, the average choices start at a significantly lower level and then decrease over the rounds at a slightly slower rate. Figure 6 shows the choices predicted round by round given these coefficients as well as the mean choice in each round by treatment.

**Table 6: Convergence**

| Variable | Coefficient | (Std. Err.) |
|---|---|---|
| Intercept | 33.98** | (2.85) |
| Round | -9.13** | (1.20) |
| $\text{Round}^2$ | 0.64** | (0.10) |
| Public | -6.76† | (3.56) |
| Pub X Round | 3.35* | (1.50) |
| Pub X $\text{Round}^2$ | -0.26† | (0.13) |
| Private | -3.77 | (3.68) |
| Pri X Round | 2.28 | (1.58) |
| Pri X $\text{Round}^2$ | -0.17 | (0.14) |
| N | 1600 | |
| $R^2$ | 0.133 | |
| $F_{(8,159)}$ | 37.45 | |

Std. errs. account for clustering on Individual level.

Significance levels: † : 10% ∗ : 5% ∗∗ : 1%

## 4 Cognitive Hierarchies

Since our main focus in this paper is on how the distribution of cognitive types changes as we vary the type of advice subjects receive, we need to have a method to measure or at least describe these distributions in order to meaningfully discuss their change. To do this we employ the cognitive hierarchy model of Camerer, Ho, and Chung (2004), CHC. This model, posits that the distribution of types is a Poisson distribution, $f(k) = e^{-\tau} \cdot \frac{\tau^k}{k}$, where $k$ is the number of steps of thinking players do and $\tau$ is a parameter indexing the mean number of thinking steps done by players in the game. Hence, higher $\tau' s$ would imply that subjects in the game are behaving according to a higher level of rationality on average.

According to this model, a player of type $k$, i.e. a player who does $k$ steps of iterative thinking, best responds to the belief that the other players are distributed by this same Poisson distribution truncated (and normalized) over all steps of thinking up to $k-1$. Given a set of choices in the $2/3^{rd}$'s Guessing Game one can estimate the distribution of types by simply finding the $\tau$ which, when used in the Poisson, predicts a mean choice that best matches the mean choice in the data. Our estimate of $\tau$ describes the predicted distribution of types parameterized as a Poisson.[3]

[3]CHC offer an detailed explanation of why a Poisson is a good choice of functional form

To fully specify the model, the behavior of the level zero players must be described. In CHC, level zero players are assumed to choose randomly between 0 and 100. Level one players who by definition believe that everybody else is level zero, think that the average choice will be 50 and respond accordingly. Level two players best respond to the distribution of level zeros and ones that they believe exist in the population. The expected actions of the level zeros define the responses of the rest of the hierarchy.

To get a first approximation to how our treatments change the distribution of types, let us estimate the CHC model as they do by specifying the level zero type as one who chooses actions uniformly over the interval [0,100]. Table 7 shows our estimates of $\tau$ across each treatment. Several things are of interest. Figure 7 compares the distributions of thinking steps predicted by these different estimates of $\tau$.

**Figure 7 here**

**Table 7: Camerer, Chung and Ho model**

| Treatment | Session | Mean Choice | Std. Dev. of Choices | $\tau$ | Predicted Mean | Predicted Std. Dev. |
|---|---|---|---|---|---|---|
| **MPub** | 1 | 19.2 | 6.5 | 3.59 | 19.2 | 9.0 |
| | 2 | 23.2 | 10.8 | 2.80 | 23.2 | 11.3 |
| | 3 | 26.2 | 7.6 | 2.31 | 26.2 | 13.1 |
| | 4 | 26.3 | 11.4 | 2.29 | 26.3 | 13.2 |
| **MPub Treatment Avg.** | | 23.7 | 9.3 | 2.71 | 23.7 | 11.6 |
| **1Pub** | 5 | 25.9 | 9.5 | 2.36 | 25.9 | 12.9 |
| | 6 | 25.9 | 12.0 | 2.35 | 25.9 | 12.9 |
| | 7 | 27.1 | 17.1 | 2.18 | 27.0 | 13.6 |
| | 8 | 44 | 16.3 | 0.39 | 43.9 | 25.3 |
| **1Pub Treatment Avg.** | | 30.7 | 15.5 | 1.68 | 30.7 | 16.0 |
| **LPri** | | 27.1 | 10.9 | 2.17 | 27.1 | 13.7 |
| **HPri** | | 28.4 | 16.4 | 1.99 | 28.4 | 14.5 |
| **Control** | | 29.2 | 15.6 | 2.02 | 28.2 | 14.3 |

and also present a set of desiderata for such a function which they demonstrate are satisfied by the Poisson.

First note that in terms of $\tau$ the biggest impact on our estimate of the type-distribution occurs in our MPub treatment where advice is the mean advice of the advisors and it is made public. While the estimated $\tau$ in the control is 2.02 and in the private advice treatment it is 1.99 and 2.17 in the HPri and LPri treatment respectively, it is 2.71 in the MPub treatment. (It actually falls to 1.68 in the 1Pub treatment but that fall is due to one session where it is 0.39.) Put differently, the addition of credible public advice leads subjects to take .69 more thinking steps on average compared to the control group. Note also that the standard deviation of the choices made is smaller in the MPub treatment than in any other treatment.

To calibrate the impact of various types of advice we can relate our $\tau$'s to those estimated on data from various different subjects pools. Consider table 8 (modified from CHC, p. 36) which presents the estimates of $\tau$ for various guessing games run with different values of $N$ (the number of subjects playing).

**Table 8: Previous Beauty Contest Results**

| Subject Pool | Source | $N$ | Sample Size | Mean | $\tau$ |
|---|---|---|---|---|---|
| Pasadena City College | CHC | 3 | 24 | 47.5 | 0.10 |
| Caltech board | Camerer | 73 | 73 | 42.6 | 0.50 |
| CEO's | Camerer | 20 | 20 | 37.9 | 1.00 |
| German students | Nagel (1995) | 14-16 | 66 | 37.2 | 1.10 |
| 70 year olds | Kovalchik | 33 | 33 | 37 | 1.10 |
| US high school | Camerer | 20-32 | 52 | 32.5 | 1.60 |
| Econ PhD's | Camerer | 16 | 16 | 27.4 | 2.30 |
| Portfolio Managers | Camerer | 26 | 26 | 24.3 | 2.80 |
| Caltech students | Camerer | 17-25 | 42 | 23.0 | 3.00 |
| Newspaper | Nagel (1998) | 1460-3696 | 7884 | 23.0 | 3.30 |
| Game theorists | Nagel (1998) | 27-54 | 136 | 19.1 | 3.70 |

Note that our control group's $\tau$ of 2.02 puts the NYU students between U.S. high school (1.60) and PhD students in Economics (2.30), but after receiving credible public advice in the MPub treatment, NYU students look more like portfolio managers (2.80).

The estimates above are done under the assumption that the level zero types choose uniformly between [0, 100]. While this may be a correct assumption in the informational poor environments that the $2/3^{rds}$-guessing games are typically run in, when public advice is added it may be that those same zero level types decide to use the announced advice as an anchor and adjust their choices accordingly, especially when the advice is common knowledge and credible. In fact, in such circumstances level zero types may decide to blindly follow advice and choose what is announced. Since everyone knows that everyone heard the announced number, the assumption of random choices by the level zeros is less plausible.[4] However, there still may be some fraction of zero level types who do choose randomly despite the existence of advice.

To capture this possibility we estimate a generalized (two-parameter) CHC model where we add a second parameter, $p.p$ is the percentage of level zeros who ignore advice and choose randomly. The remaining level zero types are assumed to follow the announced advice (or at least have a mean choice equal to the advice announced). Estimating $p$ is a natural way to measure one of our main contentions, that people paid more attention to the advice when the mean was announced than when one advisor's suggestion was announced since we would expect that the observed $p$ for the MPub treatment would be lower than that of the 1Pub treatment where, although advice is public it is not necessarily credible.

This conveniently nests both the CHC model and our advice-following modification of it. When p = 1 we are back in the CHC model while when p < 1 we have a mixture model. Since the behavior of the level zeros is crucial for the actions of the rest of the distribution, we feel that the addition of an extra parameter to the model is more than justified.[5] Further, the addition of $p$ is necessary in those cases when a piece of advice is announced and the mean choice of subjects is greater than the announced mean since if all zero level types simply followed advice, then all higher level types would have to choose numbers which were lower than the announced mean which would be impossible if the actual mean were above that which was announced. (This happened in one of the 1Pub treatment sessions where the announced advice was 13 and the mean chosen was 25.9. Obviously forcing all zero level types to choose 13 , and hence all higher types to choose below, could not explain this observation).

[4] We apply this modification only to our public treatments because level ones can only respond to level zeros if they know the advice that they receive. While one could assign beliefs about the advice received by other players in the private treatments (for example, we could assume all players know at least the mean of the advice distribution), we choose to focus on the public advice treatments where the advice received by all is directly observed.

[5] We would like to thank Colin Camerer for suggesting adding $p$ to the model in order to address this problem.

To estimate $p$ and $\tau$ in our modified model, we search over $\tau$ and $p$ in an effort to match both the mean choice and the standard error in each public advice session. While there is a large degree of heterogeneity across sessions, $p$, the percentage of level zeros that ignore the advice is much lower when the mean advice was announced (in the MPub treatment) 12%, as compared to the 51% who ignored advice in the 1Pub treatment. This fact reinforces our interpretation of the data since it confirms that mean public announcements are not only common knowledge but also perceived as credible announcement since they appear worthy of being followed at least by zero-level types. The fact that $p$ averaged 51% in the 1Pub treatments indicates that such announcement were often ignored.

**Table 9: Potamites and Schotter modification of CHC**

| | **Public Advice** | **Mean Choice** | **Std. Dev. of Choices** | **p** | $\tau$ | **Predicted Mean** | **Predicted Std. Dev.** |
|---|---|---|---|---|---|---|---|
| MPub | 31.7 | 19.2 | 6.5 | 0.01 | 1.76 | 19.2 | 6.5 |
| | 28.6 | 23.2 | 10.8 | 0.18 | 1.12 | 23.14 | 10.8 |
| | 46.5 | 26.2 | 7.6 | 0.00 | 2.02 | 26.2 | 9.1 |
| | 37.6 | 26.3 | 11.4 | 0.29 | 1.53 | 26.3 | 11.4 |
| 1Pub | 30 | 25.9 | 9.5 | 0.08 | 0.62 | 25.9 | 9.5 |
| | 13 | 25.9 | 12.0 | 1.00 | 2.59 | 24.4 | 12.0 |
| | 33 | 27.1 | 17.1 | 0.72 | 1.18 | 31.7 | 17.1 |
| | 73 | 44 | 16.3 | 0.24 | 1.45 | 44.0 | 16.2 |

# 5 Conclusions

In this paper we have attempted to explain the endogenous choice of rationality levels for subjects in $2/3^{rd}$'s guessing games. We have argued that the distribution of cognitive types revealed in an experiment may not be identical with the true distribution of types that the subjects walk into the lab with. What differentiates the two is the belief on the parts of the subjects of how rationally their cohorts are likely to behave. If these beliefs can be altered by having subjects observe some outside event, offering them advice in various ways, then we would expect that the distribution of revealed types could change. We hope to have demonstrated that a necessary condition for such a distributional change is that the external event (advice) be public, common knowledge, and credible.

If either of these elements is missing then it is unlikely that we will observe a change in the observed cognitive hierarchy.

We feel that this result is important because it leads us to think twice about interpreting models which attempt to explain the unobserved heterogeneity of cognitive types by assuming that these types are exogenous and immutable. Our results suggest that they are, in fact, endogenous but that this fact can be observed only if we attempt to alter them in specific ways.

# 6 Bibliography

## References

[1] Camerer, Colin, Teck-Hua Ho and Juin-Kuan Chung. “A Cognitive Hierarchy Model of Games.” Working Paper, 2004.

[2] Choi, Syngjoo. “A Cognitive Hierarchy Model of Learning in Networks.” Working Paper, 2006.

[3] Costa-Gomes, Miguel, and Vincent Crawford. “Cognition and Behavior in Two-Person Guessing Games: An Experimental Study,” *American Economic Review* December 2006, 96(5), pp. 1737-1768.

[4] Costa-Gomes, Miguel, and Vincent Crawford. “Studying Cognition via Information Search in Two-Person Guessing Game Experiments,” manuscript in preparation, 2007.

[5] Costa-Gomes, Miguel, Vincent Crawford, and Bruno Broseta. 2001. “Cognition and Behavior in Normal-Form Games: An Experimental Study,” *Econometrica*, September 2001, 69(5), pp. 1193-1235.

[6] Duffy John, and Rosemarie Nagel. “On the Robustness of Behavior in Experimental ‘Beauty Contests.’ ” *The Economic Journal*, November 1997, 107(445), pp. 1684-700.

[7] Ho, Teck-Hua, Colin Camerer, and Keith Weigelt. “Iterated Dominance and Itera Best Response in Experimental ‘$p$-Beauty Contests.’ ” *American Economic Review*, September 1998, 88(4), pp. 947-69.

[8] Kocher, Martin G., Matthias Sutter, and Florian Wakolbinger. “The Impact of Naive Advice and Observational Learning in Beauty-contest Games,” Tinbergen Institute Discussion Papers, 2007.

[9] Nagel, Rosemarie. “Unravelling in Guessing Games: An Experimental Study.” *American Economic Review*, December 1995, 85(5), pp. 1313-26.

[10] Stahl, Dale and Paul Wilson. “On Players’ Models of Other Players: Theory and Experimental Evidence.” *Games and Economic Behavior*, July 1995, 10(1), pp. 218-54

[11] Chaudhuri, Ananish, Andrew Schotter, and Barry Sopher, "Talking Ourselves to Efficiency: Coordination in Inter-Generational Minimum Effort Games with, Almost Common and Common Knowledge of Advice." Working paper, 2005.