CBESS Discussion Paper 10-16

On the content of focal points

by David Rojo Arjona*

*University of East Anglia

Abstract
This paper presents a methodology to study the identification of the content of focal points (Schelling 1960). This question is important for external validity and operationalising theories of decision making (e.g. team-reasoning and level-k). Choices implied by different concepts are mapped into Pure Coordination Games: open (Mehta et al. 1994) and closed (Bardsley et al. 2010). Results show that focal points are consistent with well-defined concepts, such as prototypicality, typicality and prominence in open sets but not in closed sets. Furthermore, subjects make different choices between open and closed sets leading to surprisingly more success in open than closed sets.

JEL classification codes
C70, C92

Keywords
Behavioural game theory, experiments, coordination, framing effects, salience, focal point, typicality, prototypicality

Acknowledgements
I am extremely grateful to Robert Sugden, Shaun Hargreaves Heap and Graham Loomes for detailed comments and suggestions. I am also grateful to Kei Tsutsui and Andrea Isoni for providing help with the experimental program, and Peter Moffatt, Ganna Pogrebna and Panayiotis Agisilaou for useful comments. This paper has also benefited from suggestion by the participants at the Foundations of Utility and Risk conference 2010, PGR seminars (UEA), CCC workshop and Max-Planck institute (Jena).

CBESS
University of East Anglia
Norwich NR4 7TJ
United Kingdom
www.uea.ac.uk/ssf/cbess

I. Introduction

People coordinate to a surprising degree in one shot pure coordination games (see Schelling 1960; Mehta et al. 1994; Crawford et al. 2008; Bardsley et al. 2010). There are theories of how people reason to achieve such a result in these games: level-k models (Bacharach and Stahl 2000; Crawford et al. 2008; Bardsley et al. 2010), team reasoning (Sugden 1993, 1995; Bacharach 1999, 2006) and rule-rationality (Aumann 2008). Level-k explains coordination by a combination of level 0 and higher order players where level 0 acts non-strategically upon salience and higher order players adopt the best reply to lower choices. Team reasoning explains coordination by imaging that individuals ask themselves what rule if followed by both players would most likely achieve coordination. Rule-rationality explains coordination by rules that usually, but not always, maximize utility. What neither of these theories supply, however, is an account of the actual content of the focal point that emerges from this reasoning process. That is, in level-k what level 0 players actually choose[2] and what rule actually solves the problem best in team reasoning and rule-rationality. This paper is concerned with identifying the content in this sense of the focal points that occur in one shot coordination games; and in particular with whether the content arises from the use of some general principle of focality or is idiosyncratic to the particular coordination problem.

Identification of the content of a focal point and whether it arises idiosyncratically or from the application of some general principle of focality is important for at least two reasons. First, if subjects do use a general principle, one might reasonably assume that such principles are used outside the laboratory too – and so contribute to external validity of the laboratory results.[3] Second, these general principles will operationalise theories of decision making in new settings.

This paper presents a new experimental methodology in order to identify the content of focal points. This methodology uses elicited choices according to different concepts to see which of these concepts are consistent with choices in Pure Coordination Games (PCG). If a

[2] Bacharach and Stahl (2000) is an exception to this. They apply Bacharach's (1993) theory of focal points to level-k models.

[3] The claim here is that the general principles might need to be the same for any population. However, the intuition that two groups of subjects produce different answers must be due to different interpretations of the general principle (e.g. "Choose your co-player's favourite") is a well-defined principle, but implies different choices in different populations. The same happens with other concepts.

concept is consistent with choices across different PCG, that concept must be a general principle.

Candidate general principles are evaluated along three dimensions. The first dimension is the concept used. The literature on focality is prolific in proposing different concepts: prominence, conspicuousness, uniqueness, simplicity, unambiguousness, precendent, some rationale that makes any strategy qualitatively different from the continuum of possible alternative (Schelling 1960), salience (Lewis 1969; Mehta et al. 1994; Bardsley et al. 2010), rarity preference (Bacharach 1993; 2006), most mentioned item (Sugden 1995), odd-one-out, standing out and favouriteness (Bardsley et al. 2010). Many of these concepts have been represented in formal models and/or operationalised in experiments. However, very little has been done, either theoretically or empirically, to disentangle them. Section II will present a list of concepts that may act as general principles and discusses each concept. These concepts must be well-defined and relevant regarding how people perceive choices in a coordination situation. For these reasons, concepts proposed in the cognitive psychological literature studying how people conceptualise the world are included: similarity (Tversky 1977), prototypicality and typicality (Rosch 1977).

The second dimension is the level of reasoning. Are people operating in picking mode (e.g. "Pick your favourite") or guessing mode? (e.g. "Guess the other person's favourite"). Bardsley et al. (2010) elicit the level of reasoning about choices (e.g. "Pick an animal" and "Guess the animal that the other person picked"). This paper is interested in the level of reasoning of concepts though.

The third dimension is the possible sensitivity to whether the choice set is open (Mehta et al. 1994) or closed (Bardsley et al. 2010). In open sets, subjects are given the name of the set (category) but not the elements of the set (e.g. "Choose an animal"). In closed sets, subjects are given the name of the category and the elements of a subset of the category (e.g. "Choose one of the following animals: dog, cat, lion, tiger, monkey").[4] Some informal evidence exists for difference in the application of the concept. The concept of odd-one-out can work well in closed sets, but it could seem difficult to operationalise in open sets. Bardsley et al.'s (2010)

[4] This is a difference with respect to Bardsley et al. (2010). In their paper, subjects are not told the category of the elements. This may cue different categories to different subject (e.g. wild animals, pets, felines, etc). In order to keep things as comparable as possible between open and closed, the name of the category is present in both open and closed.

game with {frog, lion, tiger, leopard} has an obvious odd-one-out. However, which animal in the open set is the odd-one-out? It is not easy to determine. It could be duck-billed platypus, Komodo dragon, etc. Bardsley et al. (2010) also point out the different results between their experiment and Mehta et al.'s (1994) experiment. An example of this difference would be that in the "Choose a car" game, in open sets, subjects tend to choose Ford while in closed sets, which include Ford and Ferrari, choose Ferrari. This phenomenon can be due to subjects using different general principles between open and closed sets. "Choose a car" prompts thoughts about everyday cars, and Ford is the archetypical everyday car. However, when both cars are in a closed set, the simple fact that Ferrari is in the set may cue thoughts about racing cars, dream cars, etc. If there is no such difference between open and closed sets, then there are reasons to expect individuals to be more successful in closed sets than in open sets. The obvious reason is that the number of possible items to coordinate in closed sets is delimited. Furthermore, individuals have all information available in the closed sets, whereas there is uncertainty about which items of the category the other player can recall in open sets. On the contrary, if such difference exists, it is not clear which set would be more successful. Although there is already data on this matter (Mehta et al. 1994 for open sets; Bardsley et al. 2010 for closed sets), comparing both experiments cannot provide a controlled test due to two reasons. First, subjects in each experiment are from different populations, from different years and facing different problems. Second, the differences stated in Bardsley et al. (2010) between open and closed sets might be due to the fact that the closed sets were *constructed* by the experimenter. This design removes the experimenter's selection of the closed set by taking the 5 items most chosen in the open set to be the closed set. This procedure implies that the items chosen in the open sets are available in closed sets and, therefore, produce a more controlled comparison between open and closed sets.

More concretely, this methodology consists of two experiments. In experiment I, choices according to concepts reviewed in section II were elicited at a personal level (picking mode - e.g. "Choose your favourite"), and trying to match another person's choices (guessing mode - e.g. "Choose the other person's favourite"). Thus, by paying subjects for a correct guess in the guessing treatment, experiment I is an incentivised experiment. These different concepts in both modes of reasoning were run for open and closed sets. Then, experiment II elicits choices in PCG for the same problems using a different sample from the same population. Thus, there is no contamination between concepts and PCG choices. This methodology provides the opportunity to investigate the content of focality in a systematic manner.

The rest of the paper is organized as follows. Section II reviews and disentangles different concepts proposed in the scientific literature. Section III introduces the details of experiments I and II. Section IV takes a look at experiments I and II independently, then maps choices from the former to the latter. The results show that there are evidences of general principles in open sets (i.e. prototypes, typical and prominent items in open sets) which are similar to one another. There are no general principles in closed sets, despite the fact that subjects are very successful at coordinating. In general, concepts in guessing mode can explain coordination better than in picking mode. Furthermore, evidences for sensitivity between open and closed sets are found. Not only are choices different between open and closed sets, but also subjects are surprisingly more successful with open sets than closed sets. Section V finishes with a discussion of the proposed methodology and results.

II. Concepts

This section produces a list of concepts that will be included in experiment I. On one hand, these concepts must be searched within the relevant scientific literature. For this, the literature on focal points and cognitive psychology was reviewed, because their parallel research programs have certain similarities. The focal point literature has repeatedly strived to analyse how people perceive the labels of the strategies in PCG, both experimentally and in reality. Cognitive research has developed general principles on how people perceive and organize information from the environment. On the other hand, the list must be operational. So, the concepts used must be small in number and well-defined. The latter implies that concepts are sharp and can be expressed in a few words in order to be operational inside the laboratory. In each concept, there is also a discussion of the implications of such concept in open and closed sets.

Schelling (1960) and Lewis (1969) are two early classics in the literature of focal points. Schelling does not commit himself to a formal definition of focal points. In fact, all the following concepts can be read throughout his book: prominence, conspicuousness, uniqueness, simplicity, precedent; some rationale that makes one strategy qualitatively different from the continuum of possible alternatives. Most of these features of focal points appear as interchangeable. Although Schelling gives some special status to prominence, he immediately reduces that status by a context-dependent claim:

“A prime characteristic of most of these ‘solutions’ to the problems, that is, of the clues or coordinators or focal points, is some kind of prominence or conspicuousness. But it is a

prominence that depends on time and place and who the people are… Equally essential is some kind of uniqueness; the man and his wife cannot meet at the 'lost and found' if the store has several… Partly this may reflect only that uniqueness conveys prominence; but it may be more important that uniqueness avoids ambiguousness" (pp.57-58).

Lewis (1969) uses the concept of salience in two different ways in his discussion:
"*They try for a coordination equilibrium that is somehow salient: one that stands out from the rest by its uniqueness in some conspicuous respect… It merely has to be unique in some way the subjects will notice, expect each other to notice and so on. If different coordination equilibria are unique in different conspicuous ways, the subjects will need to be alike in the relative importance they attach to different respects of comparison; but often they are enough alike to solve the problem. How can we explain coordination by salience? The subjects might all tend to pick the salient as a last resort, when they have no stronger ground for choice. Or they might expect each other to have that tendency, and act accordingly; or they might expect each other to expect each other to have that tendency and act accordingly, and act accordingly; and so on*" (pp.35-36).

First, salience is defined as conspicuously unique, which is in the same line as Schelling. Second, according to Lewis, people make some choices as a last resort when they have no reason to choose otherwise. This second definition is called in the literature primary salience (Mehta et al. 1994). Lewis links the two meanings by proposing the empirical hypothesis that primarily salient choices favour conspicuously unique options. Primary salience serves as a non-rational anchorage for rational choices. As people know that there is primary salience, they have reasons to choose items they expect to have primary salience (secondary salient). According to Lewis, this hierarchy can reach higher levels, but most people use only a few levels. This idea has been modelled by level-k reasoning (specifically for coordination games, see Bacharach and Stahl 2000; Crawford et al. 2008; Bardsley et al. 2010).

The terms presented above are not very clear because there are several ways in which a choice can be salient (e.g. salient by virtue of being the favourite, by virtue of being unique, by virtue of being the most typical, etc.). Besides, salient, prominent, unique and all the other concepts are often used interchangeably or their definitions and implications are not clearly distinguishable, as can be seen in the above quotation by Schelling.

Favouriteness

Bardsley et al. (2010) investigated whether choices in PCG are favourites. Favouriteness is a key and clear concept. It links with the idea of taste and preference for certain environmental signals (as opposed to taste and preference for choices associated with payoffs). For example, in Bardsley et al. (2010), Florida was chosen from the set {Colorado, Florida, Louisiana, Nevada, Ontario} or Paris from the set {Calais, Berlin, Paris, Prague, Rome} even though there is a clear odd-one-out in both sets: Ontario and Calais. Their post-experimental questionnaire shows that these answers are related with favouriteness. In experiment I, favouriteness will be elicited as "Choose your favourite". Like commonly-used preferences, if it is assumed that there is a complete ranking of items by favouriteness and this is the concept used in both, open and closed; then, open and closed sets must have the same order.

Prominence

There are different ways in which prominence can be understood. The concept of standing out is not clear, because there are at least two possible meanings. Standing out can be understood as the odd-one-out. This concept has a long tradition in focal point literature and will be presented later. More generally, standing out can be also interpreted as a synonym of salience, conspicuousness and so on. This concept links with the Aristotelian idea of Telos. The natural ranking of some category, for instance, category C is in terms of what makes a "good" C. What makes a mountain a good mountain? Height. What makes a footballer a good footballer? Skill at playing football. So, Everest is the top of the most natural ranking of mountains (highest), Maradona or Pelé is the top of the most natural ranking of footballers (most skilled) and so on. For example, Grand Central Station seems likely to be the top of the most natural ranking of meeting places in New York (at Schelling's time). In experiment I, it will be elicited as "Choose the top of the most natural ranking".[5] If the top of the most natural ranking is used in both sets, there should be no differences between open and closed because natural ranking produces a complete ranking as well.

Typicality

This concept is present in cognitive psychology and Sugden's (1995) normative theory of focal points. An element that appears more in the world requires diminishing costs to store and retrieve it. In Sugden (1995), subjects should choose the more mentioned item because it increases the probabilities of coordination. There are two differences between the two ideas.

[5] This paper does not enter into the discussion of this concept being scientifically or philosophically legitimate. The important point here is whether ordinary people can make sense of the concept and reach the same conclusions.

First, there is a difference between the psychological concept of mental accessibility and the rational-choice concept of maximising the probability of coordination, when both players use the same rule: i.e. team-reasoning. Second, the reality of individuals is not always public, accessible to any individual, but local and conveyed by words or thoughts not perceived straightaway by other individuals. For example, if you ask people to tell you the best known national teams in the World Cup, they will probably go for Brazil, Germany, Italy, Argentina, etc. However, if you ask people to tell you the most frequently mentioned in their conversations, they will say England (given that we are in England). So, I call best known "public typicality" and the most frequently mentioned in conversations "local typicality". In experiment I, public typicality will be elicited as "Choose the best known" and local typicality as "Choose the most frequently mentioned in your conversations". Typicality should produce the same choices in open and closed sets, because again there should be a complete ranking of answers that should not change between open and closed.

Prototypicality

Rosch (1977) introduces the idea of prototypes as cognitive process. The two premises behind the idea of prototypes is that the world is highly organized (e.g. creatures with feathers are more likely to also have wings than creatures with fur) and the process for perceiving and storing that information is also highly structured. In this setup, prototypes are the most characteristic members of the set. In fact, typicality would be one of its determinants. In addition, prototype is the element which takes lower cognitive effort to be retrieved. For example, when you ask children to draw a car or flower, they draw something similar to a Ford or daisy. Binmore and Samuelson (2006) show evolutionary evidence for imperfect monitoring (lower cognitive effort); however, their premises are completely different from Rosch. Mehta et al. (1994) show that John is the most chosen boy's name, Ford as car manufacturer. In experiment I, subjects will be asked to "Choose an example". This concept only has sense when there are no other elements in comparison but only the category, as in open sets.

Odd-one-out

Some papers invoke a concept of uniqueness (Schelling 1960), standing out, odd-one-out (Bardsley et al. 2010) and rarity preference (Bacharach 1993; 2006). Bacharach and Bernasconi (1997) find certain support for such choices. Uniqueness and stand out are not sharper concepts, because it is not clear why an item may be unique or stand out. In fact, stand out might be related to prominence as seen above. Similarly, Everest may be unique by virtue of being the top of the most natural ranking, but Everest may also be unique by virtue of being the odd-one-out. Odd-one-out is related to Bacharach's idea of rarity preference. Imagine a situation where you have 5 cubes, 4 red and 1 blue. The problem can be presented as pick a thing (0.2 expected probability of success if both players follow this strategy), pick a red cube (0.25 expected probability of success if both players follow this strategy) or choose the blue cube (1 expected probability of success if both players follow this strategy). For example, in Bardsley et al. (2010), Mannheim was chosen in the set {Mannheim, Berlin, Brussels, Lissabon, Madrid} and glass in the set {glass, diamond, emerald, sapphire}. This concept has a lot of power in closed sets, because all the items of the list are known and easy to compare. In open sets, this is not the case. Actually, the more an item satisfies the concept of odd-one-out, the more difficult it should be to recall (e.g. the duck-billed platypus). Odd-one-out and rarity preference are in this sense paradoxical concepts for open sets. In experiment I, this concept is elicited as "Choose the least similar".

Similarity

Similarity is a relevant theory in cognition and it is proposed as a principle organizing the classification of objects. Tversky (1977) presents similarity as a function of three arguments: those features that are common to both objects, those features of one element that do not belong to the other, and those features of the other element that do not belong to the first one. This function increases with the first argument and decreases with the other two. As mentioned before, odd-one-out can be read as the least similar item. Similarity can also provide another definition of prototypicality as the element maximizing similarity of itself with respect to the category. Besides, there is a psychological literature (reviewed in Tversky 1977) where similarity can be seen as a general principle driving subjects' choices. So, similarity becomes relevant for an investigation of focality. In experiment I, the concept is elicited as "Choose the most similar". Similarity needs to compare different items. Like odd-one-out, this is problematic in open sets, but may have some power in closed sets.

In short, the review above has provided the following sharp concepts: prototypicality (example), similarity (most similar), odd-one-out (least similar), public typicality (best known), local typicality (most frequently mentioned in conversations), favouriteness and top of the most natural ranking.

III. Experimental Details

The experimental design is divided into two experiments. Experiment I is an incentivised questionnaire and elicits the concepts of section II. Experiment II is an experiment[6] eliciting focal points in PCG. The categories and the sets used are the same in both experiments, so choices according to concepts can be mapped into choices of the PCG. One sample of subjects without replacement from the same population is taken to participate in each experiment. Thus, the mapping is meaningful because subjects are from the same population but, at the same time, there is independence between choices in the PCG and the concepts.

The following procedure to include categories into the experiment is applied. The categories must require that people have a basic knowledge. Some categories have already appeared in previous experiments on focal points. For instance, the famous examples of positive numbers from Schelling (1960), colours from Mehta et al. (1994), etc. Then, a short list of categories is drawn-up. A pre-experiment is run in which categories are randomly allocated to different treatments. The questions chosen for the experiments are the 16 which appeared in the pre-experiment as open set questions.[7] This method provides the possibility to construct the closed sets in a different manner to Bardsley et al. (2010). In this case, the 5 most named items in the coordination[8] treatment of the pre-experimental open sets form the closed sets. The items in closed sets are presented in random order across players. The chosen items are shown in Table 1.

[6] There are more treatments ('picking' and 'guessing' as in Bardsley et al. 2010), but they are independent of the ones presented here. A description of the other treatments and results can be requested to the author.

[7] However, category "firms" used in the pre-experiment is *a priori* problematic in terms of interaction with 3 previous categories (cars, supermarkets and fast food chains). Therefore, it was removed and "metals" was randomly selected among the other categories.

[8] If there is a tie, this is broken by looking at the most named items in open guessing and if there is still a tie, by looking at open picking.

Table 1: Labels for strategies in closed sets

Category	Strategy 1	Strategy 2	Strategy 3	Strategy 4	Strategy 5
City	NEW YORK	BIRMINGHAM	LONDON	TOKYO	NORWICH
Fruit	ORANGE	MANGO	APPLE	PEAR	BANANA
Sport	TENNIS	CRICKET	FOOTBALL	SWIMMING	RUGBY
Footballer	ROONEY	SEAMAN	BECKHAM	MARADONA	RONALDO
Positive number	1	2	5	7	10
Supermarket	MORRISONS	SAINSBURY´S	TESCO	WAITROSE	ASDA
Flower	TULIP	DAFFODIL	ROSE	SUNFLOWER	DAISY
Furniture	DESK	SOFA/COUCH	TABLE	BED	CHAIR
Car	HONDA	MERCEDES	FORD	FERRARI	BMW
Fastfood	KFC	BURGER KING	MCDONALDS	PIZZA HUT	SUBWAY
Animal	LION	MONKEY	DOG	TIGER	CAT
Colour	GREEN	BLACK	RED	YELLOW	BLUE
Metal	ALUMINIUM	GOLD	IRON	SILVER	STEEL
Transport	TRAIN	AEROPLANE	CAR	BIKE	BUS
Drink	WATER	JUICE	BEER	TEA	COKE
Superhero	BATMAN	HULK	SUPERMAN	IRONMAN	SPIDERMAN

Experiment I is a 2x2 design (Table 2). There are open (O) and closed sets (C), and concepts are in picking (P) and guessing (G) mode. In P, subjects must answer an item according to one of the aforementioned concepts. For example, “Choose your favourite” or “Choose the top of the most natural ranking”. In G, subjects are paired with another subject from P and they must guess the item their match has picked following that concept. For example, “Choose the other person’s favourite” or “Choose the other person’s top of the most natural ranking”. These instructions are given at the beginning of each task (see Annex I). Feedback is not provided until the end of the experiment.

Table 2: 2x2 Design of experiment I

	Open	Closed
Picking	OP	CP
Guessing	OG	CG

Subjects can only take part in open or closed sets. Each subject faces P and then G. The G questions are only announced once subjects have answered P. Subjects answer to all the concepts for 8 categories in P and then, in G, they answer all the concepts for 8 different categories (Table 3). This feature of the experiment resembles Bardsley et al. (2010).

Table 3: Concepts elicited in experiment I

Sets	Prototypicality	Similarity	Odd-one-out	Typicality (public)	Typicality (local)	Favourite	Prominence
Open	example	most similar	least similar	best known	most frequently mentioned	favourite	top of the most natural ranking
Closed	N/A	most similar	least similar	best known	most frequently mentioned	favourite	top of the most natural ranking

A similar incentive mechanism to Bardsley et al. (2010) was used to avoid collusion between subjects. In each session, there were two groups. Both groups had an identical 'pot' that consisted of the number of people times £5. Subjects were randomly matched with a person from the other group. At the end of the experiment, a random lottery mechanism picked a question that the other person answered in picking and compared that answer with the one given by the subject in guessing. The pot was divided equally among those members of the group that guessed correctly. All subjects were given a participation fee of £2.

Experiment I was paper based. As perfect randomization was impossible, there were four different models of booklets with a random order of concepts, categories and, in closed questions, a random order of items in the set.

In experiment II, PCG were presented to the subjects. In each session, half of the subjects started with open and then closed questions; and the other half with closed and then open.[9] This way provided a counterbalance to control for order effects. Subjects faced 16 questions following the aforementioned experimental procedure. Experimental instructions were given (Annex II). The first part of the instructions instructed subjects how open and closed question screens looked like, and how to proceed in each screen to record data (Annex III). Then, the actual instructions are given. Finally, subjects faced a questionnaire to see whether they had understood the instructions and mechanism. In the actual experiment, each subject was matched with another subject randomly in the room throughout the session, although they did

[9] The reason for this difference between experiments was that, in experiment I, there were more questions to answer. So, it was easier for subjects to face only one kind of question.

not know who this subject was. As in the previous experiment, subjects did not receive feedback until the end of the experiment.

The incentive mechanism was similar to experiment I. In each session, there was an amount of money (£5 for each person in that session) comprising the 'pot'. The pot was divided equally among those subjects who gave the same answer as their co-player to a randomly selected question. However, the experimental software, Ztree (Fischbacher 2007), cannot compare answers to open questions. Ztree records written words in a box, but cannot compare whether the word written by a given subject is the same as the word written by another subject.[10] On top of this, retrieving a specific written word in a later period is impossible. In order to overcome these difficulties, the conditional information lottery (Bardsley 2000) was used to incentivise the whole experiment. The main idea is that subjects were told that they would be paid for a real task. There are other tasks that they will not be paid for (because they cannot be incentivised). However, as they would only know which one is for real at the end of the experiment, it was in their best interest to play as if they were all real. Apart from that incentive, subjects were given a participation fee of £2. The relevant information was conveyed to the subjects by the general instructions (Annex II).

The experiments were run between March and December 2009. 400 subjects took part in the experiments (198 in experiment I and 202 in experiment II). Subjects were recruited through an email via the distribution list of the University of East Anglia (Norwich, UK). There were 28 sessions (13 for experiment I and 15 for experiment II). Each treatment in both parts had 50 subjects, except one treatment in experiment I which had 48 subjects and one treatment in experiment II which had 51.[11] The average payment was £7 for an average of 45 minutes work.

[10] The data from open sets have two features. First, subjects are requested to write down *one* word (e.g. juice instead of orange juice, strawberry juice, apple juice and so on). Second, there are words with identical meaning (e.g. soccer and football). In those cases, all those different words are grouped under the same label: in sports, soccer and football; in furniture, sofa and couch; in transport, underground and tube; in drinks, coke, coca-cola and cola. The same procedure is applied in both experiments.

[11] One subject was excluded because he had taken part in the pilot study.

IV. Results

Design results

There are no important experimental artefacts explaining the results. In experiment I, the 4 different models of booklets were assigned chronologically; model 1 was assigned to the first sessions, then model 2 and so on. In addition, open sessions were run before closed sessions. Thus, if there was any learning or ordering effect, the distribution of answers between models should be quite different. This was significantly different for only 31 questions out of 416, being 11 questions in OP (112), 7 questions in OG (112), 9 questions in CP (96) and 4 questions in CG (96). This represents less than 7.5% of the total questions. In experiment II there was only one question out of 64 presenting order effects. Furthermore, there was no position effect in closed sets or learning between sessions.[12] Notice that no position effects in closed sets means that subjects do not use the rule "Choose the first item in the row" or "Choose the one in the middle".

Experiment I

Section II disentangles the different concepts proposed in the literature from a theoretical point of view. Experiment I provides empirical data on this matter. The test of two concepts producing the same distribution of choices was performed for each different category. As choices according to different concepts come from the same subjects, a Stuart-Maxwell test was performed with the null hypothesis of marginal homogeneity between concepts. In Table 4, the number in each cell is the number of times that the null hypothesis of the aforementioned test was not rejected across 16 categories. For example, the hypothesis of homogenous concepts between favourite in picking mode and most similar in picking mode was not rejected in 14 cases out of 16 in open sets, while in closed sets the number reduced to 5 cases. Similarly, the hypothesis of homogeneous concepts between favourite in guessing and most similar in guessing mode was not rejected in 9 cases in open and 2 cases in closed sets. Finally, the hypothesis of homogeneous concepts[13] between favourite in picking mode and most similar in guessing mode was not rejected in 3 cases, whereas the same hypothesis between most similar in picking mode and favourite in guessing mode was not rejected in 5 cases.

[12] Answers from first and last sessions were compared.

[13] For comparison between picking and guessing, a Fisher's exact test was performed because the comparison was between subjects.

As can be seen in Table 4, there are three main results. First, there is a difference between open and closed sets in picking/picking and guessing/guessing. Concepts in open sets are more correlated between themselves than concepts in closed sets. Second, the correlation in open sets decreases from picking/picking to guessing/guessing, except for a group of concepts (i.e. example, best known, most frequently mentioned and top of the natural ranking). The higher level of reasoning (guessing) makes subjects start to see differences between concepts and narrow them down. Third, concepts in picking are related to the same concepts in guessing (look main diagonals). Yet there are other high numbers off the diagonal showing that some concepts (e.g. prototype) are related to others (e.g. typicality), as explained in section II.

Table 4: Relationship between concepts

Pick/Pick	Open							Closed					
	example	most similar	least similar	best known	frequently mentioned	favourite	top ranking	most similar	least similar	best known	frequently mentioned	favourite	top ranking
example		**15**	**11**	**14**	**16**	**13**	**14**	n/a	n/a	n/a	n/a	n/a	n/a
most similar			**13**	**9**	**14**	**14**	**15**		**0**	**2**	**5**	**5**	**3**
least similar				**10**	**13**	**14**	**14**			**0**	**2**	**0**	**2**
best known					**16**	**9**	**13**				**5**	**1**	**7**
frequently mentioned						**14**	**15**					**5**	**5**
favourite							**15**						**3**
Guess/Guess	Open							Closed					
	example	most similar	least similar	best known	frequently mentioned	favourite	top ranking	most similar	least similar	best known	frequently mentioned	favourite	top ranking
example		**11**	**5**	**15**	**15**	**9**	**15**	n/a	n/a	n/a	n/a	n/a	n/a
most similar			**9**	**8**	**9**	**9**	**11**		**0**	**1**	**5**	**2**	**2**
least similar				**1**	**5**	**6**	**6**			**0**	**1**	**1**	**1**
best known					**14**	**9**	**14**				**8**	**1**	**6**
frequently mentioned						**10**	**14**					**2**	**6**
favourite							**13**						**6**

Pick (r) / Guess (c)	Open							Closed					
	example	most similar	least similar	best known	frequently mentioned	favourite	top ranking	most similar	least similar	best known	frequently mentioned	favourite	top ranking
example	**7**	**10**	**1**	**5**	**5**	**7**	**6**	n/a	n/a	n/a	n/a	n/a	n/a
most similar	**4**	**15**	**3**	**3**	**3**	**5**	**3**	**16**	**0**	**2**	**5**	**1**	**1**
least similar	**1**	**1**	**15**	**0**	**1**	**1**	**0**	**0**	**14**	**0**	**1**	**0**	**1**
best known	**11**	**7**	**1**	**12**	**10**	**4**	**10**	**2**	**0**	**10**	**6**	**3**	**7**
frequently mentioned	**10**	**10**	**2**	**7**	**11**	**10**	**8**	**6**	**1**	**5**	**9**	**5**	**5**
favourite	**2**	**3**	**3**	**1**	**1**	**6**	**3**	**3**	**0**	**2**	**0**	**9**	**3**
top ranking	**7**	**8**	**1**	**7**	**8**	**6**	**10**	**4**	**1**	**6**	**3**	**6**	**15**

Experiment II

The first issue to be addressed was the extent to which coordination was achieved (see Table 5). In closed sets, every category had a significantly higher coordination index than the mixed-strategy Nash equilibrium: random choices.[14] For open sets, this test could not be performed because random choices were ill-defined. The coordination rate in open sets was, however, intuitively higher than random choices. For example, the coordination rate for drinks was 0.1735. For this rate to be induced by random choices, there should be only 5-6 different drinks, although there were 11 different drinks that people used in the PCG.

Furthermore, as closed sets were constructed from the most chosen items of open sets, the distribution of answers over the 5 items in closed sets and the distribution of answers over the same items in open sets could be compared. Given that some expected frequencies were less than 5, they needed to be grouped. All the items were grouped around 3 items: focal point in open set, focal point in closed set and the remainder. If the focal points in open and closed sets were the same; then, there were only 2 groups. As can be seen in Table 5, the hypothesis that open and closed distributions are generally the same was disconfirmed. Therefore, there was sensitivity between open and closed sets with respect to choices.

Table 5: Open and closed c index and focal points, and p-value equal distribution

	Open c index	Closed c index	p-value	Comparison between open and closed distributions
City	London 0.422	Norwich 0.4571	0	different
Fruit	Apple 0.6229	Apple 0.4139	0.054	not different
Sport	Football 0.8457	Football 0.6465	0.041	different
Footballer	Beckham 0.4106	Beckham 0.5633	0.701	not different
Pos. num.	1 0.1796	10 0.2653	0.005	different
Supermarket	Tesco 0.6345	Tesco 0.4196	0.092	not different
Flower	Rose 0.5265	Rose 0.3527	0.048	different

[14] One-tail bootstrap method as in Bardsley et al. (2010) is used.

Furniture	Chair 0.2645	Bed 0.3224	0	different
Car	Ford 0.4237	Ferrari 0.3314	0	different
Fastfood	McDonald's 0.7078	McDonald's 0.6147	0.214	not different
Animal	Dog 0.3608	Dog 0.3412	0.277	not different
Colour	Blue 0.338	Red 0.2531	0.019	different
Metal	Iron 0.2416	Gold 0.5747	0	different
Transport	Bus 0.4171	Car 0.3363	0.002	different
Drink	Coke 0.2481	Beer 0.2314	0.006	different
Superhero	Superman 0.5078	Superman 0.418	0.092	not different

In order to compare the coordination rates between open and closed sets, the bootstrap method (Efron 1979) was used. The null hypothesis was that choices in the closed set are drawn from a distribution with the same relative frequencies as the actual responses in the open set. A two-tail test was used to see whether the rejections of the null hypothesis of equal coordination index are in the direction of either open or closed sets. The critical values of the coordination index were generated by repeated sampling with the relative frequencies of open sets. The estimates of these critical values were constructed with 20,000 simulated samples. When the null hypothesis was rejected at the 5% (or 1%) significance level, it was indicated by * (or **). In Table 6, the hypothesis that subjects are generally equally successful in both sets was disconfirmed. It was also disconfirmed, surprisingly, that coordination rates are in general greater in closed sets than in open sets.

Table 6: Coordination indexes and bootstrap results (open vs closed)

	Open c index	Closed c index	Results	Open or Closed more successful?
City	0.422	0.4571	ns	**No difference**
Fruit	0.6229	0.4139	**	**Open**
Sport	0.8457	0.6465	**	**Open**
Footballer	0.4106	0.5633	ns	**No difference**
Number	0.1796	0.2653	ns	**No difference**
Supermarket	0.6345	0.4196	*	**Open**
Flower	0.5265	0.3527	*	**Open**
Furniture	0.2645	0.3224	ns	**No difference**
Car	0.4237	0.3314	ns	**No difference**
Fastfood	0.7078	0.6147	ns	**No difference**
Animal	0.3608	0.3412	ns	**No difference**
Colour	0.338	0.2531	*	**Open**
Metal	0.2416	0.5747	**	**Closed**
Transport	0.4171	0.3363	*	**Open**
Drink	0.2481	0.2314	ns	**No difference**
Superhero	0.5078	0.418	ns	**No difference**

Relationship between experiment I and experiment II

The key point of the proposed methodology is to provide a picture of which concepts elicited in experiment I are consistent with experiment II, PCG, and therefore provide answers to the positive question of what a focal point is and some insights about the reasons for the sensitivity between open and closed sets. The data can be organized in two different ways.

First, data can be arranged by concepts and how many categories are predicted by those concepts. Thus, a chi-square test of equal distribution between the distribution implied by a certain concept if everyone is going to use this concept and the distribution in PCGs is run for each category. The number of times the null hypothesis of same distribution is not rejected appears in Table 7. For example, the number of cases that the distribution over choices of people following the concept “favouriteness” in picking mode, “Choose your favourite”, is

consistent from the distribution over choices in the open coordination treatment is 0 out of 16. Similarly, the number of cases that the distribution over choices of people following the concept "best known" in guessing mode, "Choose the other person's best known", is consistent from the distribution over choices in the closed coordination treatment is 7 out of 16.

Table 7: Comparing rule distribution in the "questionnaire" and choice distribution

Times null not rejected	Picking		Guessing	
	Open/16	Closed/16	Open/16	Closed/16
example	5	n/a	15	n/a
most similar	3	4	8	5
least similar	1	2	1	1
best known	12	8	13	7
frequently mentioned	9	7	14	8
favourite	0	4	6	8
top ranking	5	8	11	6

The fact that 6-7 concepts were introduced in this experiment could explain that choices in every coordination problem were idiosyncratically mimicked by some concepts.[15] However, a general principle requires that the same concept can explain coordination across tasks. The results in Table 7 suggest that guessing appears to be a better predictor in open sets and not different from picking in closed sets. In terms of concepts, "example" (prototypicality), "best known" (public typicality), "frequently mentioned" (local typicality) and "top of the natural ranking" (prominence) can explain across tasks in open sets and, therefore, act as general principles. Furthermore, these concepts are related among themselves (see Table 4) and the common feature may be that they all require the lowest effort to be stored or retrieved. This last point is also suggested by the fact that open and closed sets have similar response times before subjects pass to the next question. Despite the fact that subjects in open sets have to type and may have to correct spelling errors, subjects go, on average, to the next screen in 18 seconds (3 seconds later than closed sets).

Second, data can be arranged by categories and how many concepts are explaining each category. This arrangement also provides further insight about the results in Table 7, where there seems to be no general principle for closed sets. This could indicate that either (i) some

[15] The only category where no concept could map coordination choices was closed "cities".

categories are explained by one concept, some by another; or (ii) some categories are explained by many concepts, some categories are explained by hardly any. In order to shed some light on this issue, Table 8 presents the number of concepts consistent with each category. Only concepts in guessing mode (e.g. "Choose the other person's favourite") are presented.[16] There are several reasons for this. First, concepts in guessing mode are more successful from a normative point of view.[17] In addition, concepts in picking mode have no more predictive power for behaviour than guessing (Table 7). The fact that most categories are explained by many rules in open sets has been already established. In fact, this result also appears in Table 8, where open sets have many concepts explaining each category. However, most categories are explained by a lower number of concepts in closed sets than in open sets. Therefore, closed sets cannot present the same indication as open sets, this is, some categories are explained by many concepts, some categories by hardly any. Otherwise they should be similar numbers. So, the answer must be then that some games are explained by one concept, some by another by logical impossibility.

Table 8: Density of concepts 'explaining' each category

Categories	Open	Closed
cities	1	0
fruits	4	4
sports	5	3
footballers	6	2
numbers	5	2
supermarkets	5	2
flowers	6	1
furniture	2	0
car	2	2
fastfood	5	4
animals	6	0
colours	5	4
metals	5	3
transport	2	1
drinks	4	2
supeheroes	5	5

[16] The category "cities" in closed sets is not explained by any concept in this experiment. Besides, "furniture" and "animals" in closed sets are explained by P mode concepts.

[17] Some analysis was run in the normative domain, but it is not presented here because this paper is concerned with the positive domain. The normative question answers: if all subjects had followed the same concept, which concepts should I have chosen? Which mode of reasoning should be more successful? Which set would be more successful? The results show that (i) the concept "best known" is the most successful in both open and closed sets, (ii) every concept is more successful in G than in P and (iii) the same concept is also more successful in closed than in open sets.

V. Discussion

This paper is concerned with identifying the content of the focal points occurring in one-shot coordination games; and in particular with whether the content arises from the use of some general principle of focality or is idiosyncratic to the particular coordination problem. Some concepts, for example "prototype" (Rosch 1977), "typicality" as best known and frequently mentioned (Rosch 1977; Sugden 1995) and "top of the natural ranking", act as general principles behind focality in open sets. These concepts are correlated between themselves, and the underlying feature of these different concepts is that the choices are easy to recall. On the other hand, no single concept can apparently explain or predict across tasks in closed sets. This result might appear to contradict Bardsley et al. (2010), who *do* find evidence supporting the existence of general principles in closed sets constructed to point out the use of a particular concept. However, evidence that subjects are very successful at coordinating in different concepts for different closed sets is found here. Bardsley et al.'s (2010) evidence and this success inevitably imply a relationship between the construction of a closed set and the choice of a concept to coordinate that is yet to be discovered.

This difference between open and closed sets also shows that (i) the distribution of answers between open and closed sets is generally not the same and (ii) subjects are surprisingly more successful at open than closed coordination games. One reason for this is that concepts have the same implications in open sets but are different in closed sets. Therefore, coordination is more difficult.

What enables these conclusions to be drawn is the methodology used here to investigate the content of focal points. This methodology uses elicited choices according to different concepts to see which of these concepts are consistent with choices in PCG. The methodology has several interesting features. First, by having guessing, experiment I is an incentivised method. Second, there is no contamination between the use of concepts and choices in PCG, because different samples from the same population are used. Third, this method is validated, because although the tasks faced by subjects in experiment I do not look like economic questions, they predict behaviour in problems that are important for economics. Finally, this method of asking people is extremely direct and natural. In fact, the methodology can be applied to a wide range of interesting questions in this field.

References

Aumann, Robert. 2008. “Rule Rationality vs. Act Rationality.” Discussion Paper #497. The Center for the Study of Rationality. The Hebrew University.

Bacharach, Michael. 1993. “Variable Universe Games.” In *Frontiers of Game Theory*, ed, Ken Binmore, Alan Kirman and Piero Tani, 255–276. Cambridge, MA: MIT Press.

Bacharach, Michael. 1999. “Interactive Team Reasoning: A Contribution to the Theory of Cooperation.” *Research in Economics*, 53: 117–147.

Bacharach, Michael. 2006. *Beyond Individual Choice: Teams and Frames in Game Theory*. Princeton University Press.

Bacharach, Michael and Michele Bernasconi. 1997. “The Variable Frame Theory of Focal Points: An Experimental Study.” *Games and Economic Behavior*, 19: 1-45.

Bacharach, Michael and Dale Stahl. 2000. “Variable-Frame Level-*n* Theory.” *Games and Economic Behavior*, 33: 220-246.

Bardsley, Nicholas. 2000. "Control Without Deception: Individual Behaviour in Free-Riding Experiments Revisited." *Experimental Economics*, 3(3), 215-240.

Bardsley, Nicholas, Judith Mehta, Chris Starmer and Robert Sugden. 2010. “Explaining Focal Points: Cognitive Hierarchy Theory versus Team Reasoning.” *Economic Journal*, 120(543), 40-79.

Binmore, Kenneth and Larry Samuelson. 2006. “The Evolution of Focal Points,” *Games and Economic Behavior*, 55, 21–42.

Crawford, Vincent, Uri Gneezy and Yuval Rottenstreich. 2008. “The Power of Focal Points is Limited: Even Minute Payoff Asymmetry May Yield Large Coordination Failures.” *American Economic Review*, 98 (4), 1443-1458.

Efron, Bradley. 1979. “Bootstrap Methods: Another Look at the Jacknife.” *Annals of Statistics*, 7: 1–26.

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

Lewis, David. 1969. *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.

Mehta, Judith, Chris Starmer and Robert Sugden. 1994. “The Nature of Salience: An Experimental Investigation.” *American Economic Review*, 84: 658–673.

Rosch, Eleanor. 1977. "Classification of Real-World Objects: Origins and Representations in Cognition." in *Thinking: Readings in Cognitive Science*, ed, Johnson-Laird, P.N. & Wason, P.C., pp.212-222. Cambridge: Cambridge University Press.

Schelling, Thomas. 1960. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Sugden, Robert. 1993. “Thinking as a Team: Towards an Explanation of Non-Selfish Behavior.” *Social Philosophy and Policy*, 10: 69–89.

Sugden, Robert. 1995. “A Theory of Focal Points.” *Economic Journal*, 105: 533–550.

Tversky, Amos. 1977. “Features of Similarity.” *Psychological Review*, 84: 327–352.

ANNEX I (instructions for experiment I)

OPEN PICKING

In each question, you will be asked to think about a "category". You will be asked to pick one item from this category according to some criterion. For example, the category might be "political leaders" and you might be asked "Which is the most charismatic?". You answer this question by writing down one item belonging to that category. For example, you might write down "Obama". *You must write down one (and only one) item*. Even if you do not feel strongly about your answer, please follow this instruction.

You will also be asked to indicate how strongly you feel about your answer. In the example you would be asked to show how strongly you feel that the political leader you have picked is the most charismatic. You show this by marking a point on a scale from 0 ("not at all") to 5 ("very strongly").

OPEN GUESSING

In part 1 you were asked a series of questions. Each question asked you to write down an item from some category according to some criterion (for example, your "favourite"). Different people in this room answered different questions. Before the experiment started, we divided you all into two groups. Each group had its own set of questions. You are in one group and we have matched you with someone from the other group. You will never be told who that other person was, and that person will never be told who you are.

The only difference between the questions you answered and the questions the other person answered was that they used different categories. For example, you answered eight questions, each asking you to write down the "best known" item in a category. The other person also answered eight questions, each asking him/her to write down the "best known" item in a category, but the categories in his/her questions were different from the categories in yours.

You will now see the categories that appeared in the other person's questions. What we want you to do is to say which item you think the other person wrote down in each case.

When you have finished, we will see if you guessed correctly what the other person wrote down.

How do you win money?

You will be paid according to how well you guessed what the other person wrote down on one question. You will not know which question this is until the end of the experiment. So, you should guess in each case as if you were going to be paid for guessing correctly.

In addition to the participation fee of £2, we have £xxx for each group in this session. If you guessed correctly what the other person picked, you will win some MONEY. If not, you will win NOTHING (but you will be paid the participation fee anyway).

How much money will you win?

It will depend on the number of people in your group who have given a correct answer. The £ will be divided equally between the people who guessed correctly. If everyone guesses correctly, you will earn £5 (plus the participation fee). If only you guess correctly, you will earn £ (plus the participation fee).

You answer this question by circling one item in the list. *You must circle one (and only one) item from the list.* Even if you do not feel strongly about your answer, please follow this instruction.

You will also be asked to indicate how confident you feel about your guess. You show this by marking a point on a scale from 0 ("not at all") to 5 ("very strongly").

CLOSED PICKING

In each question, you will be shown a list of five items in random order, all of which belong to the same category. For example, the category might be political leaders and the list might be:

Obama Merkel Sarkozy Berlusconi Hu

You will be asked to pick one of these items according to some criterion. For example, you might be asked "Which of the items in the list is the most charismatic?".

You answer this question by circling one item in the list. *You must circle one (and only one) item from the list.* Even if you do not feel strongly about your answer, please follow this instruction.

You will also be asked to indicate how strongly you feel about your answer. In the example, you would be asked to show how strongly you feel that the political leader you have picked is the most charismatic of the five leaders in the list. You show this by marking a point on a scale from 0 ("not at all") to 5 ("very strongly").

CLOSED GUESSING

In part 1 you were asked a series of questions. Each question showed you a list of five items, and asked you to pick one of them according to some criterion (for example, your "favourite"). Different people in this room answered different questions. Before the experiment started, we divided you all into two groups. Each group had its own set of questions. You are in one group and we have matched you with someone from the other group. You will never be told who that other person was, and that person will never be told who you are.

The only difference between the questions you answered and the questions the other person answered was that they used a different list of items. For example, you answered eight questions, each asking you to pick "the best known" from a list of five items. The other person also answered eight questions, each asking him/her to pick the "best known" from a list of five items, but the lists in his/her questions were different from the lists in yours.

You will now see the lists of items that appeared in the other person's questions. However, the items will not necessarily be listed in the same order as they were for the other person. What we want you to do is to say which item you think the other person picked in each case.

When you have finished, we will see if you guessed correctly what the other person picked.

How do you win money?

You will be paid according to how well you guessed what the other person picked on one question. You will not know which question this is until the end of the experiment. So, you should guess in each case as if you were going to be paid for guessing correctly.

In addition to the participation fee of £2, we have £xxx for each group in this session. If you guessed correctly what the other person picked, you will win some MONEY. If not, you will win NOTHING (but you will be paid the participation fee anyway).

How much money will you win?

It will depend on the number of people in your group who have given a correct answer. The £xxx will be divided equally between the people who guessed correctly. If everyone guesses correctly, you will earn £5 (plus the participation fee). If only you guess correctly, you will earn £xxx (plus the participation fee).

You answer this question by circling one item in the list. *You must circle one (and only one) item from the list.* Even if you do not feel strongly about your answer, please follow this instruction.

You will also be asked to indicate how confident you feel about your guess. You show this by marking a point on a scale from 0 (“not at all”) to 5 (“very strongly”).

ANNEX II (instructions for experiment II)

You have been paired with another person in this room. These pairings have been made at random, and you will never be told who you have been paired with. The other person is being given exactly the same instructions as you. He or she will answer exactly the same questions as you will do, except that, where there are lists of items, the order in which YOU see the items may be different from the order in which the other person sees them. You will be presented with 16 questions.
Your aim is to give the same answer as the other.

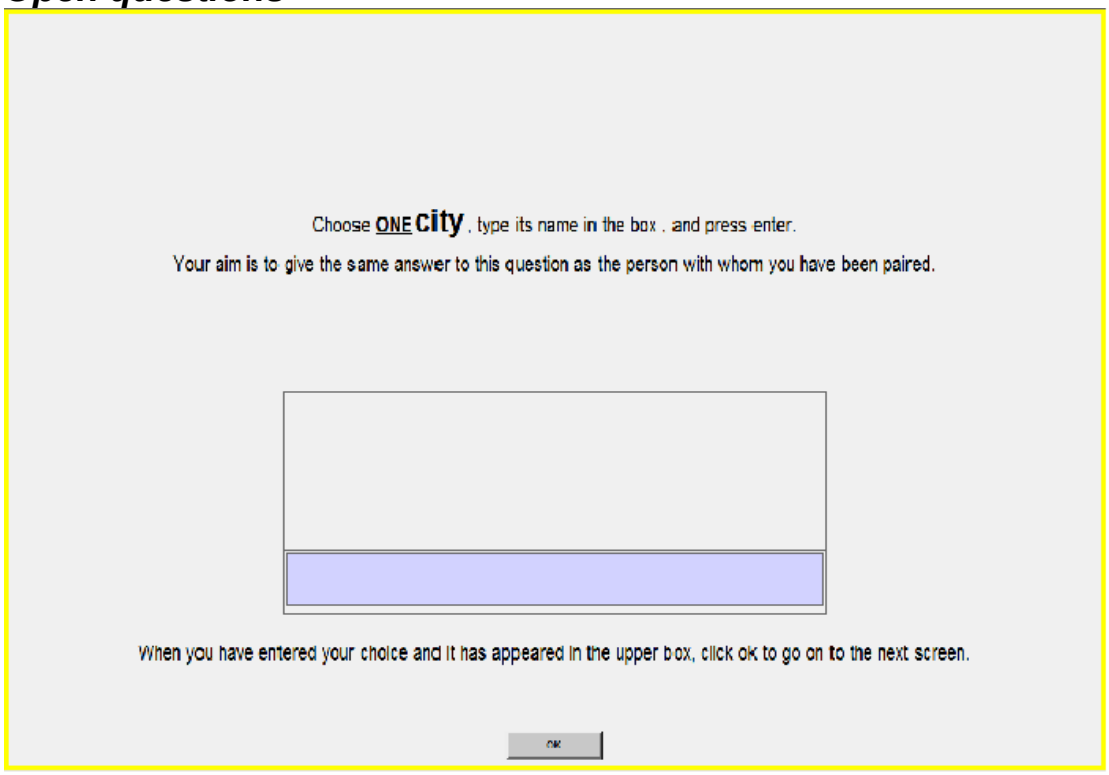
How do you win money?
You will be paid according to your answer to one of the questions. This question will be the same for you and the person you are matched with. You will not know which this is until the end of the experiment. So, you should answer each question as if you were going to be paid for it.
In addition to the participation fee of £2, we have £xxx[18] for this session. If you give the same answer as the other person, you will win some MONEY. If not, you will win NOTHING (but you will be paid the participation fee anyway).

How much money will you win?
It will depend on the number of people in this session who have given the same answers as the people they have been paired with. The £xxx will be divided equally between the people who do this. If everyone does this, you will earn £5 (plus the participation fee). If only two people do this, each will earn £xxx/2 (plus the participation fee).

[18] This number depends on the number of people participating. It was £5 times the number of subjects in the session.

ANNEX III

Open questions



Closed questions