

Coordination when there are restricted and unrestricted options

Shaun P. Hargreaves Heap¹  ·
David Rojo Arjona² · Robert Sugden³ 

© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract One might expect that, in pure coordination games, coordination would become less frequent as the number of options increases. Contrary to this expectation, we report an experiment which found more frequent coordination when the option set was unrestricted than when it was restricted. To try to explain this result, we develop a method for eliciting the general rules that subjects use to identify salient options in restricted and unrestricted sets. We find that each such rule, if used by all subjects, would generate greater coordination in restricted sets. However, subjects tend to apply different rules to restricted and unrestricted sets.

Keywords Coordination · Salience · Focal points · Odd-one-out · Typicality · Prototypicality

JEL Classification C72 (noncooperative games) · C78 (bargaining theory, matching theory) · C91 (design of experiments: laboratory, group behaviour)

We thank participants in many conferences and seminars for comments on earlier versions of this paper.

Electronic supplementary material The online version of this article (doi:[10.1007/s11238-017-9589-9](https://doi.org/10.1007/s11238-017-9589-9)) contains supplementary material, which is available to authorized users.

✉ Shaun P. Hargreaves Heap
s.hargreavesheap@kcl.ac.uk

David Rojo Arjona
dra9@le.ac.uk

Robert Sugden
r.sugden@uea.ac.uk

¹ Department of Political Economy, King's College, The Strand, London WC2 R2, UK

² Department of Economics, Leicester University, University Road, Leicester LE1 7RH, UK

³ School of Economics, University of East Anglia, Norwich NR4 7TJ, UK

1 Introduction

Subjects manage to coordinate their decisions with a surprisingly high frequency in laboratory experiments on pure coordination games (e.g. see Schelling 1960; Mehta et al. 1994; Bardsley et al. 2010). This is both impressive and important. It is impressive because the games are symmetric in pay-offs. Strategies have different labels and the player's objective is to match the choice of the other, defined in terms of its label. The pay-off from matching is always the same, whichever label it occurs around, and it is zero otherwise. There is nothing in terms of pay-offs to distinguish the strategy labels and so it seems that subjects are able to exploit some asymmetry in the labels themselves when they manage to match to a surprising degree.¹ It is important because many real-world interactions appear to have multiple equilibria. The selection of one involves solving a coordination game and it seems we have a surprising capacity to do this in the laboratory, even in the difficult cases of pure coordination. This paper is concerned with this apparent capacity to solve pure coordination games.

First, we consider whether the capacity to coordinate in the laboratory depends on restricting the number of strategy options. This is potentially important because laboratory experiments often restrict the strategy options, whereas choices outside the laboratory are more typically unconstrained. As a result, the observed capacity to coordinate in the laboratory could to some degree be an artifact of the laboratory setting and its constraint on choice. Towards this end, our first experiment tests for differences in coordination when otherwise identical coordination problems have either restricted or unrestricted options. Second, we examine with the aid of another experiment how the observed levels of coordination in the first experiment might be explained by appealing to the idea that people follow rules that are based on the labels of the strategies in such games.

The difference between restricted and unrestricted options seems likely to be important because the scope for dis-coordination appears to grow as the possible number of objects in the choice set increases. For instance, when there is one option, the only outcome is a coordinated one; whereas, when there are two options, only half of the possible outcomes are coordinated, when there are three options, the number of possible outcomes that are coordinated drops to one third, and so on. Our first Hypothesis follows from this reasoning.

H1: Coordination is higher when the choice set is restricted.

We test H1 in our first experiment by asking some subjects to play 16 pure coordination games with a restricted number of strategy options and others to play 16 corresponding games with no restrictions. In each coordination game, the options belong to a particular category (i.e. a class of objects). For example, one category is car manufacturers. In the restricted version for this category, we ask subjects to choose

¹ Crawford et al. (2008) supply some evidence from games where there are asymmetric pay-offs which suggests that slight asymmetries in pay-offs can interfere with whatever reasoning players use to coordinate in pure coordination games. See also Isoni et al. (2013).

one of the following options: {HONDA, MERCEDES, FORD, FERRARI, BMW}. The identified car manufacturers are the strategy labels. The analogous unrestricted version of this coordination game asks subjects to choose a car manufacturer. In both cases, the subjects are told that they will be randomly matched with another subject and ‘if you give the same answer as the other person you will win some MONEY. If not, you will win NOTHING’.

In fact, we reject H1. Contrary to expectation, coordination is typically *higher* when options are unrestricted than when they are restricted. It seems *in this respect* that laboratory experiments with restricted sets may understate the likely coordination when there are unrestricted options outside the laboratory. The result sets a problem that we take up in our second experiment: why is coordination higher when options are unrestricted?

We address this problem from a particular perspective. We follow Schelling (1960) and Sugden (1995) in assuming that individuals use a selection rule based on the labels attached to the strategies in such games. In particular, subjects are randomly matched and we elicit their application of seven rules in a ‘naïve’ and in a ‘strategic’ way. Thus, suppose the rule is ‘favourites’ and is applied to the category of car manufacturers. In the ‘naïve’ elicitation, we ask the subject to choose their favourite: ‘Circle/write down your favourite car manufacturer’. When applied ‘strategically’, we ask the subject to choose the other person’s favourite: ‘Circle/write down the other person’s favourite car manufacturer’. The seven rules are drawn from the psychology literature.

We use these elicitations to assess whether any of three conceptually distinct reasons, within this framework, explains why one observes more frequent coordination in unrestricted than restricted sets of the first experiment. One possible reason is that individuals are using the *same* rule and coordination improves because the individual interpretations of that rule become closer to each other when there are unrestricted options. For example, with an ‘odd one out’ rule it may help such convergence in interpretation when selecting a number from the unrestricted set of all positive integers as compared with a restricted set like {1, 2, 3, 7, 9}. This is because, in the restricted set, both ‘1’ and ‘9’ might stand out as, respectively, the lowest and highest while ‘2’ could do the same as the only even number. However, in the unrestricted set there is no highest number and there are lots of even numbers, so only the lowest number, ‘1’ stands out. This is the basis of our second Hypothesis. We define the average own rule concordance rate for rule i as the average coordination rate that would be achieved across all categories when all subjects use rule i in the manner suggested by the elicitations of rule i in the second experiment.

H2: The average own rule concordance rate is higher with unrestricted than with restricted options.

Another possible reason is that each individual could be using a constant rule across the categories but the constant rule is not the same for all individuals and greater coordination is achieved because the application of these different rules yields greater concordance on the unrestricted than the restricted options. For example, people who select the object with the most features associated with that category of objects (i.e. the

‘most similar’ rule) could plausibly choose ‘1’ on both the unrestricted and restricted positive integer sets because in both cases every number is divisible by ‘1’ (but not by any other). As a result, if, as suggested above, people using the ‘odd one out’ rule are more likely to select ‘1’ on the unrestricted than the restricted set, then they will achieve more coordination with those using the ‘most similar’ rule on the unrestricted than on the restricted set. This is our third Hypothesis. We define the average cross-rule concordance rate between rule i and j as the average coordination rate that would be achieved across the categories when one individual uses rule i and the other individual uses rule j (in the manner suggested by the rule elicitation in the second experiment). H3 is distinct from H2, but they are *not* mutually exclusive.

H3: The average cross-rule concordance rate is higher with unrestricted rather than restricted options.

The final possible explanation builds on the second. Individual subjects use different rules *and* the frequency of their use of each rule changes between the unrestricted and restricted options sets. For example, the ‘odd one out rule’ might seem obviously applicable in the restricted set of car manufacturers like {HONDA, MERCEDES, FORD, FERRARI, BMW} when subjects come from the US because FORD is the only US manufacturer while all others originate from foreign countries. But, this rule may not be as attractive to use on an unrestricted domain of car manufacturers because being a domestic manufacturer no longer singles out one option and there are no obvious alternative bases for an ‘odd one out rule’ that would work on car labels. H4 concerns whether there is evidence of different frequency of rule use on restricted and unrestricted options.

H4: Rules are used with different frequencies by individuals when options are unrestricted as compared with when they are restricted.

We find no evidence for H2. Indeed for each of our rules when used by all, the concordance rate is higher with the restricted options than the unrestricted ones. Nor do we find evidence in favour of H3: the cross-rule concordance between different rules in our experiment is very similar and almost always slightly higher when options are restricted rather than unrestricted. Finally in relation to H4, we find evidence that subjects use rules with different frequencies on the restricted as compared with the unrestricted sets. Using the different estimated frequencies of use, we are able to generate the predicted coordination rates when subjects are given restricted and unrestricted options. These predictions are below, but close to what was observed in the first experiment and the predicted coordination rate is higher, albeit only marginally so, when the options are unrestricted. In short, there is evidence that different rules are triggered when options are restricted as compared with when they are unrestricted and this contributes to the explanation of why we observe higher coordination when options are unrestricted.

The first experiment on restricted and unrestricted coordination games and its results are described in Sects. 2 and 3. Sections 4 and 5 do the same for the second experiment on the possible explanation of the differences in coordination observed in the first experiment. Section 6 discusses these results and concludes.

Table 1 Labels for strategies in restricted sets

Cities: {NORWICH, LONDON, NEW YORK, TOKYO, BIRMINGHAM}
Fruits: {APPLE, ORANGE, BANANA, MANGO, PEAR}
Sports: {FOOTBALL, SWIMMING, CRICKET, TENNIS, RUGBY}
Football players: {BECKHAM, ROONEY, RONALDO, MARADONA, SEAMAN}
Positive Numbers (greater than 0): {10, 1, 7, 5, 2}
Supermarket chains: {TESCO, MORRISONS, ASDA, SAINSBURYS, WAITROSE}
Flowers: {ROSE, DAFFODIL, SUNFLOWER, DAISY, TULIP}
Items of furniture: {BED, SOFA/COUCH, TABLE, CHAIR, DESK}
Car manufacturers: {FERRARI, FORD, MERCEDES, BMW, HONDA}
Fast food chains: {MCDONALDS, SUBWAY, BURGER KING, PIZZA HUT, KFC}
Animals: {DOG, CAT, LION, TIGER, MONKEY}
Colours: {RED, BLUE, BLACK, GREEN, YELLOW}
Metals: {GOLD, STEEL, ALUMINIUM, SILVER, IRON}
Means of transport: {CAR, BUS, AEROPLANE, BIKE, TRAIN}
Drinks: {BEER, TEA, WATER, COKE, JUICE}
Superheroes: {SUPERMAN, BATMAN, SPIDERMAN, HULK, IRONMAN}

2 Experiment I

Individuals play 16 pure coordination games. Each game is distinguished by the category of labels used to identify possible strategy choices. The categories were selected because each has more than five possible labels and subjects are likely to have sufficient knowledge of the category to give access to at least this number of label options. The categories are set out in Table 1. Flowers, car manufacturers, positive numbers and colours have been used in previous experiments on focal points.

Subjects play a block of eight games in an unrestricted version where there is no restriction on the labels (e.g. “Choose the same colour as the other person”) and the remaining block of eight games in a restricted version where there are five labels (e.g. “Choose the same color as the other person: {green, black, red, yellow, blue}”). The blocks and versions are counterbalanced within each session. Further details and screenshots appear in the Supplementary Material. Subjects see the restricted version labels in a row and the order of the labels is independently randomized for each subject. They are randomly and anonymously matched to another subject throughout the experiment and feedback is provided only at the end of the experiment. Choices are incentivized through the creation of a ‘pot’ (£5 for each person in that session). This is equally divided at the end of the experiment between those pairs of players who choose the same label for a randomly chosen game. This incentive mechanism leads to a separation of coordination from other-regarding motivations. In practice, the unrestricted version could create disagreements between experimenter and subjects about outcomes. For example, ‘soccer’ and ‘football’ typically refer to the same sport for British

citizens but different ones for American citizens.² As a result, we did not actually use the unrestricted version for payment. Nevertheless, *both* versions are incentivized using Bardsley's (2000) conditional information lottery by *truthfully* telling subjects that, as only one category is for payment and it will be announced at the end of the experiment, it is in their best interest to play as if all categories were for payment.

The five labels were those that were chosen most frequently in the pilot of the unrestricted version of each game. This selection procedure prevents any discretion by the experimenter, avoids possible confounds and increases the statistical power of tests for difference. For instance, while a random selection of labels would avoid experimenter discretion, it creates a possible confound because only some of the labels in the restricted set might be recognized by subjects. Likewise, if the design included a random selection of labels for the restricted version, there might be no common labels for comparison or the frequency of choices of these common labels would be too small to perform a non-parametric test.³ Table 1 also gives the five labels that were available in the restricted version of each game. They are listed in the order of frequency from left (most chosen) to right (least chosen).

The experiment was conducted in seven sessions at the University of East Anglia in March 2009. 100 subjects from the general student population were recruited randomly through an email via the distribution list of the university. Subjects participated anonymously at computer workstations. Instructions were read aloud (Appendix A). Every subject received £2 for participating and could expect to get £5 from the 'pot' for an average of 45 minutes work.

3 Results: Experiment I

Table 2 presents a summary of the results for each category. Column (1) presents the number of labels chosen by at least one subject in the unrestricted version. The only category where less than 5 labels were mentioned is Sports; the median number across all categories is 6.5 (average is 6.75). Columns (2) and (3) show the modal label of the distribution; and columns (4) and (5) the degree of coordination (c) for the unrestricted and restricted version, respectively, defined as:

$$c = \sum_g m_g (m_g - 1) / [N(N - 1)] \quad (1)$$

where m_g is the number of subjects choosing strategy labelled g for a given category and version, and N is the total number of subjects facing that decision problem. It

² The actual cases in which different words have the same meaning 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. In addition, subjects in the unrestricted version are requested to write down *one* word (e.g. juice instead of orange juice, strawberry juice, apple juice and so on).

³ This method to create the five labels in the restricted version could be undermined if the results from the pilot and from the unrestricted version in Experiment I would differ significantly. From the 80 labels in Experiment I, 16 labels did not appear in the pilot. However, from the alternative 16 labels mentioned in the pilot, only three labels have a frequency of 5% (Paris, Rabbit and White). All other 13 labels are mentioned residually.

Table 2 Summary results: Experiment I

	# Labels in unrestricted version (1)	Modal response in unrestricted version (2)	Modal response in restricted version (3)	c-index in unrestricted version (4)	c- index in restricted version (5)	Significance (6)
Cities	5	London	Norwich	0.422	0.4571	ns
Fruits	5	Apple	Apple	0.6229	0.4139	##
Sports	4	Football	Football	0.8457	0.6465	##
Football players	9	Beckham	Beckham	0.4106	0.5633	ns
Positive numbers	10	1	10	0.1796	0.2653	ns
Supermarket chains	5	Tesco	Tesco	0.6345	0.4196	#
Flowers	7	Rose	Rose	0.5265	0.3527	#
Items of furniture	5	Chair	Bed	0.2645	0.3224	ns
Car manufacturers	11	Ford	Ferrari	0.4237	0.3314	ns
Fast food chains	6	McDonald's	McDonald's	0.7078	0.6147	ns
Animals	8	Dog	Dog	0.3608	0.3412	ns
Colours	6	Blue	Red	0.338	0.2531	#
Metals	7	Iron	Gold	0.2416	0.5747	**
Means of transport	5	Bus	Car	0.4171	0.3363	#
Drinks	8	Coke	Beer	0.2481	0.2314	ns
Superheroes	7	Superman	Superman	0.5078	0.418	ns

*, ** and *** denote significance at the 10, 5 and 1% significance levels if the two-tail test that there is no difference is rejected in the correct direction for H1 (i.e. coordination is higher in the restricted version). #, ## and ### denote significance at the 10, 5 and 1% significance level in the opposite direction to H1

measures, given the actual responses of subjects, the probability of two randomly picked individuals choosing the same label.

To test H1, we use a bootstrap method with 20,000 simulated samples, to compare, for every category, whether the *c*-index in the restricted version can be obtained from the distribution of choices in the unrestricted version. Column (6) in Table 2 shows the levels of significance in a two-tail test and the direction in which the null hypothesis is rejected. There is only one category where coordination is higher in restricted than unrestricted at 95% or higher confidence levels (Metal). The remaining 15 categories tell against the hypothesis: the coordination level is significantly higher in unrestricted than restricted sets for 6 and there is no difference at 95% confidence levels in 9 categories. So, we reject the Hypothesis.

4 Experiment II

4.1 The rules

We consider seven possible rules and locate each briefly in the experimental and broader psychology literatures. They, in different ways, cash out why something might ‘stand out’. Their precise operationalization for the naïve elicitation of Experiment II is given in parenthesis. Of course, the list is selective but we shall provide a cross check on their plausibility by considering whether the recommendation of each rule would be good advice for someone playing against the population from Experiment I.

Favourite (“Choose your favourite”) has a long history in the experimental literature on coordination games (see [Mehta et al. 1994](#)). [Bardsley et al. \(2010\)](#), for instance, provide evidence of a correlation between choices in coordination games and a post-experiment questionnaire on favourites.

Odd-one-out (“Choose the least similar”) is central to Schelling’s (1960) original work and it has been elaborated by [Bacharach \(1993, 2006\)](#) as a rarity preference. [Bacharach and Bernasconi \(1997\)](#) and [Bardsley et al. \(2010\)](#) find some support for such choices: e.g. Mannheim was chosen in the set {Mannheim, Berlin, Brussels, Lisbon, Madrid} and glass in the set {glass, diamond, emerald, sapphire}.

Prominence (“Choose the top of the most natural ranking”) can be understood as what is pre-eminent in a natural ranking for a category. 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, Maradona or Pelé is the top of the most natural ranking of footballers, and so on. This might explain why Schelling found that Grand Central Station was the most salient place for people to meet in New York: it was the place where most people pass through on a single day.

Typicality I and II (“Choose the best known” and “Choose the most frequently mentioned”) are two versions of an availability heuristic in cognitive psychology and may be used because something that appears frequently in the world comes more readily to mind ([Tversky and Kahneman 1973](#)). For example, Tesco has the largest number of supermarket stores in the UK and so may come most easily to mind. Likewise, there are more Ford cars than the other makes and so on. This rule is also suggested by Sugden’s (1995) normative theory of focal points where subjects should

choose more mentioned items because this increases the probability of coordination under team reasoning assumptions.

Prototypicality (“Choose an example”) is another availability heuristic (see Rosch 1977). The idea 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 accordingly highly structured. Prototypes are the most characteristic members of the set and so they require least cognitive effort to retrieve. This may explain why, when children are asked to draw a flower, they draw something similar to a daisy and why, in Mehta et al. (1994), John is the most chosen boy’s name and Ford is the most chosen car manufacturer.

Similarity (“Choose the most similar”). Objects typically share an overlapping set of features (see Tversky 1977) and the one that contains most of these elements may stand out for this reason. For example, flowers usually have scent, they are colourful, often come in different varieties of the genus which are given identifying names by horticulturalists who enter them in competitions, are displayed in vases, are often given as gifts on special days, and so on. Thus, roses might be chosen with this rule because they exemplify all these characteristics, whereas the other flowers in one respect or another fall short.

4.2 Naïve (‘pick’) and strategic (‘guess’) choices

In each session, the 16 categories are divided into two blocks of 8. The subjects are divided into two and each group is presented in the first stage with one of the blocks. They are asked in the naïve version to pick using each of the seven rules: for example, with the *favourite* rule on the colour category, they are asked ‘Choose your favourite colour’. Once this stage has been completed, they receive the instructions for the second, strategic stage. In this stage, the blocks are swapped and subjects are asked to guess the label that a randomly selected subject (from the other group) chose in the first stage: e.g. ‘Choose the other person’s favourite colour’. The choices in the second stage are incentivized in a similar same way as in Experiment I. At the end of the experiment, a question is randomly selected and the subjects correctly choosing their randomly matched partner’s choice share a ‘pot’. The only difference is that each group has their own pot.

The sessions differ only with respect to whether the labels in each category are an unrestricted or restricted set. When the labels are restricted, they are the same five labels for each category as in Experiment I.

The experiment was conducted at the University of East Anglia in December 2009. 198 subjects were recruited through an email via the distribution list of the University of East Anglia. Subjects were recruited from the general student population and participated anonymously at a workstation. Subjects were assigned one out of the four models of booklets with a randomly determined order of psychological rules, categories and, in the restricted version, a random order of labels (see supplementary material). Instructions were read aloud (Appendix A). Every subject received £2 for participating and £5 in expectation from the pot; the experiment lasted on average for 45 min.

5 Results: Experiment II

With four booklets, order or learning effects would produce different distributions of answers across the booklets. However, the distribution was only significantly different for 31 of the 416 questions (i.e. 7.5%). Thus, we discount the influence of order and learning effects.

To check on the plausibility of our chosen rules, we consider whether each rule would constitute good advice to someone who was playing against the population from Experiment I. Our criteria of good advice are that someone using a rule would do significantly better in terms of coordination than would be achieved if they followed conventional game theory by selecting their strategy randomly. Tables 3 and 4 give the levels of coordination that would be achieved with each rule on each category. In addition, the coordination rate for each rule is bootstrapped following the protocol described earlier to check whether the coordination rate for each rule is significantly different from the level that would be achieved when someone uses a randomizing strategy. The success rate for randomizing is not well defined for unrestricted sets, but we also use 0.2 in the test for significance to keep comparability. The test, therefore, errs on the side of toughness in unrestricted sets. Almost all rules are plausible in the sense that a person would almost always do better by using the rule. The one exception is the ‘Odd-one-out’ rule: on unrestricted sets, it is never significantly different; and to a lesser extent on restricted sets where it is only significantly better than random in 6 of the 16 categories.

One can also compare the level of coordination that is achieved when using a rule in play with the population from Experiment II with the coordination level that was actually achieved in Experiment I. Figures 1 and 2 give this in the form of scatter diagram for the unrestricted and restricted version, respectively, where each observation is a category. When they fall around a straight line, this suggests that the use of the rule tracks well the differences in coordination across categories in Experiment I and when they fall around a 45° ray from the origin, the use of the rule is also close to the actual level of coordination achieved in Experiment I. Again, it is apparent that the ‘Odd-one-out’ rule does least well in these respects.

We test H2 by comparing the level of coordination that would be achieved by each rule on the restricted and unrestricted versions of labels, had it been used by all players in either the ‘naïve’ or ‘strategic’ form. Table 5 gives the average own concordance rate that would have been achieved by each rule. For each rule, the average rate is higher in the restricted version than the unrestricted version. This is true for both naïve and strategic elicitation. Looking at each rule in each category, we find that the number of instances in which the restricted version produces a higher *c*-index is 159 out of 192 (87 in naïve choices and 72 in strategic choices). This would not be generated randomly (binomial test *p* value <0.01). We also note that the own concordance rate for each rule is always higher in the ‘strategic’ than in the ‘naïve’ elicitation.⁴

We test H3 by generating a prediction for the coordination rate that would be observed between subjects using each possible pair of rules, based on what Experiment

⁴ A similar result is found in Bardsley et al. (2010). This suggests that the heterogeneity in tastes or interpretations of a rule is greater than is the heterogeneity in the beliefs about the modal taste or interpretation of the rule.

Table 3 Coordination rate between Exp II rule user and players from Exp I for each rule (unrestricted)

Category	Prototypicality	Similarity	Odd-one-out	Typ I best	Typ II frequent	Favourite	Prominence
Cities	0.4878***	0.3478***	0.1469ns	0.398***	0.4592***	0.3196***	0.3747***
Fruits	0.6094***	0.4661***	0.058ns	0.6155***	0.5286***	0.2473ns	0.5486***
Sports	0.8265***	0.4457***	0.0588ns	0.8641***	0.8833***	0.6996***	0.7363***
Football players	0.3441***	0.3257***	0.1551ns	0.4588***	0.4159***	0.3065***	0.3302***
Positive numbers	0.1135ns	0.1412ns	0.1054ns	0.2057ns	0.1576ns	0.0702ns	0.1388ns
Supermarket chains	0.6604***	0.5502***	0.1008ns	0.6731***	0.5845***	0.3812***	0.5237***
Flowers	0.5682***	0.389***	0.1608ns	0.6669***	0.6249***	0.6216***	0.5967***
Items of furniture	0.2673***	0.2114ns	0.1041ns	0.2608***	0.2318ns	0.1816ns	0.2506***
Car manufacturers	0.3061***	0.2824**	0.062ns	0.3057***	0.2653*	0.129ns	0.2049ns
Fast food chains	0.6445***	0.4608***	0.0502ns	0.6959***	0.6449***	0.4967***	0.5596***
Animals	0.391***	0.3253***	0.0457ns	0.4637***	0.46***	0.409***	0.3192***
Colours	0.2963***	0.2351ns	0.0922ns	0.3012***	0.2457*	0.3261***	0.2649**
Metals	0.2823***	0.2292ns	0.1335ns	0.2717***	0.2572**	0.2364ns	0.2734***
Means of transport	0.3939***	0.3249***	0.0335ns	0.3441***	0.3829***	0.2804**	0.2265ns
Drinks	0.2527**	0.1943ns	0.0832ns	0.272***	0.2045ns	0.206ns	0.2108ns
Superheroes	0.438***	0.3898***	0.0727ns	0.4706***	0.3767***	0.262*	0.5061***

ns, *, ** and *** indicate no significance, 10, 5 and 1% significance, respectively

Table 4 Coordination rate between Exp II rule user and players from Exp I for each rule (restricted)

Category	Similarity	Odd-one-out	Typ I best	Typ II frequent	Favourite	Prominence
Cities	0.240ns	0.3336***	0.2648**	0.5784***	0.2288ns	0.2304ns
Fruits	0.3588***	0.1084ns	0.5384***	0.4312***	0.3316***	0.4152***
Sports	0.3908***	0.0388ns	0.7252***	0.7404***	0.4596***	0.5432***
Football players	0.3888***	0.018ns	0.5688***	0.4964***	0.488***	0.3524***
Positive numbers	0.1656ns	0.3056***	0.2528**	0.1808ns	0.2464**	0.2936***
Supermarket chains	0.4288***	0.0308ns	0.5808***	0.5816***	0.3192***	0.3148***
Flowers	0.2264ns	0.1676ns	0.5408***	0.5072***	0.3956***	0.488***
Items of furniture	0.1624ns	0.264**	0.2844**	0.3468***	0.4128***	0.3292***
Car manufacturers	0.1492ns	0.34***	0.26**	0.216ns	0.3856***	0.3596***
Fast food chains	0.4892***	0.066ns	0.6652***	0.5088***	0.3092**	0.4804***
Animals	0.2276ns	0.0484###	0.4704***	0.43**	0.2896**	0.2268ns
Colours	0.2212ns	0.172ns	0.2652***	0.2468**	0.2616**	0.2416**
Metals	0.1015ns	0.2823**	0.6688***	0.5181***	0.5585***	0.6496***
Means of transport	0.3244***	0.1516ns	0.4192***	0.2976***	0.3324***	0.2328ns
Drinks	0.1742ns	0.3225***	0.2233ns	0.2604***	0.2258ns	0.2321ns
Supheroes	0.4216***	0.0716ns	0.4948***	0.4108***	0.3432***	0.4804***

ns, *, ** and *** indicate no significance, 10, 5 and 1% significance, respectively, in the expected direction. #, ## and ### indicate 10, 5 and 1% significance in the opposite direction

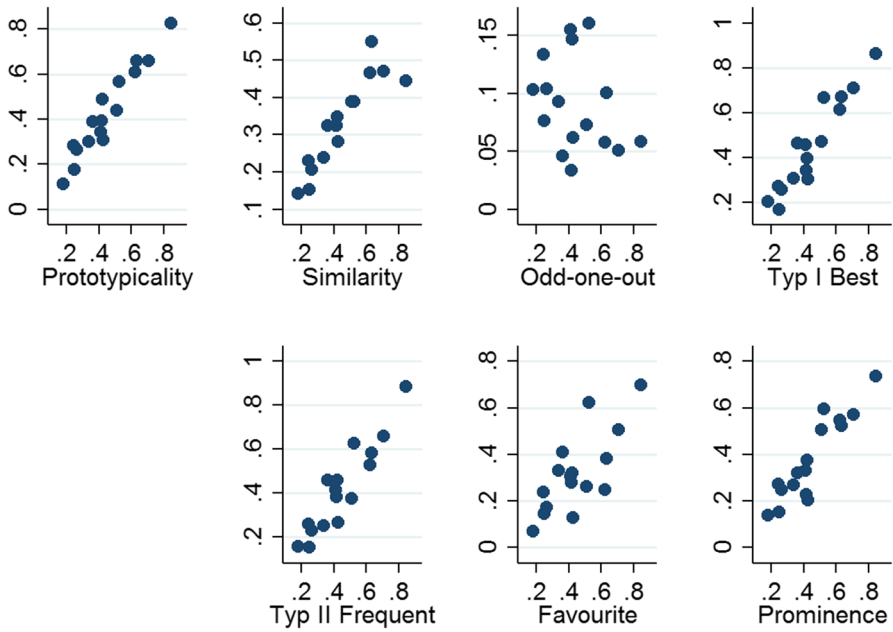


Fig. 1 Own rule concordance rate for each rule in Exp II and actual coordination in Exp I (unrestricted)

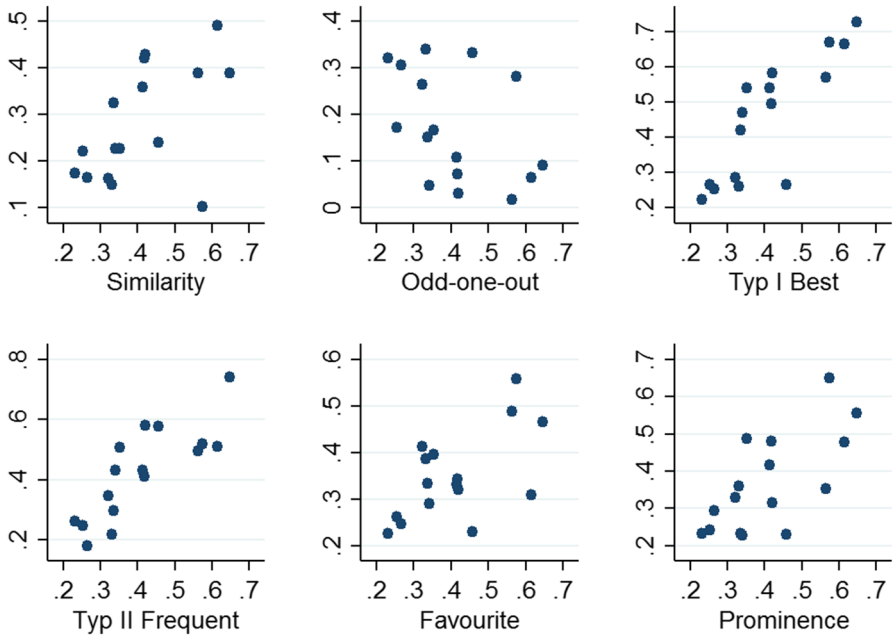


Fig. 2 Own rule concordance rate for each rule in Exp II and actual coordination in Exp I (restricted)

Table 5 Average probability of coordination (c -index) across categories for a given rule

	Prototypicality	Similarity	Odd-one-out	Typ I best	Typ II frequent	Favourite	Prominence
Naïve							
Unrestricted	0.2097	0.1717	0.0550	0.4229	0.2293	0.1249	0.2474
Restricted		0.3044	0.4580	0.5490	0.3553	0.2754	0.3508
Strategic							
Unrestricted	0.4262	0.2709	0.0953	0.4747	0.4132	0.3166	0.3459
Restricted		0.3551	0.5074	0.5788	0.5057	0.3689	0.4210

Table 6 Average concordance between rules (c-index) across categories (strategic elicitation)

	Unrestricted version					Restricted version							
	Prototypicality	Similarity	Odd-one-out	Typ I best frequent	Typ II frequent	Favourite	Prominence	Similarity	Odd-one-out	Typ I best frequent	Typ II frequent	Favourite	Prominence
Prototypicality	0.42	0.30	0.08	0.42	0.38	0.31	0.35	n/a	n/a	n/a	n/a	n/a	n/a
Similarity		0.26	0.08	0.31	0.29	0.23	0.27	0.37	0.09	0.35	0.31	0.24	0.26
Odd-one-out			0.10	0.08	0.08	0.08	0.08		0.52	0.13	0.16	0.20	0.22
Typ I best				0.44	0.40	0.31	0.37			0.59	0.50	0.39	0.44
Typ II frequent					0.38	0.31	0.34				0.52	0.34	0.38
Favourite						0.29	0.27					0.38	0.34
Prominence							0.33						0.43

II suggests subjects using those rules will choose. Table 6 gives these average cross-rule concordance rates for the 16 categories when the options are unrestricted and restricted. It is apparent that the cross concordance rates for ‘strategic’ elicitation are almost always higher on the restricted version. There is only one pair, ‘Typicality I (Best known)’ and ‘Prominence’ where the concordance rate is actually higher on the unrestricted options set, and then only very marginally. This counts against H3.⁵

We test H4 by first estimating what weights, when attached to each rule for each category, can best explain the choice of labels in Experiment I under the assumption that each rule yields the distribution of choices revealed by Experiment II (see Crawford and Iriberry 2007, for a similar approach). Using maximum likelihood, we estimate these for each category individually as well as in the aggregate by pooling all the decision problems. The comparison of the aggregate with the average estimate for each category gives some indication of how stable are the estimates of weights across categories. Tables 7 and 8 give these weights for the unrestricted and restricted versions of the coordination problems. They suggest that while both versions of ‘Typicality’ (‘Best known’ and ‘Frequent’) rules are used with similar frequency under both conditions, ‘Odd-one-out’, ‘Favourite’ and ‘Prominence’ rules are used much more frequently when the options are restricted. The corresponding rules that are much more popular in the unrestricted condition are ‘Prototypicality’ (of course, by construction) and ‘Similarity’. This is evidence in favour of subjects using the rules with different frequencies when options are restricted/unrestricted.

We examine whether these differences in rule use can account for the higher coordination when options are unrestricted using these weights to generate predicted coordination rates under these assumptions. We do this both with and without errors in the application of a rule. The results are given in Table 9. The predicted coordination rate on average is higher when the options are unrestricted, albeit only marginally so.

6 Discussion and conclusion

Our first experiment produced a surprising result: when the options in a pure coordination game are restricted to five, subjects coordinate less well than when their choice set is unrestricted. This is important because it suggests that the artifact of reducing the choice options in the laboratory will not bias upwards the estimate from these experiments of coordination when there are no such restrictions (e.g. outside the laboratory). Nevertheless, it is surprising because the scope for dis-coordination seems likely to grow as the number of options available for choice increases.

Our second experiment addresses this puzzle from a particular perspective. It develops a new method for investigating the principles that underlie salience. Under the assumption that people use one of our seven rules and that they use the same rule, we cannot explain our ‘surprising’ result that subjects appear to coordinate better when the options are unrestricted than when they are restricted by appealing to the greater

⁵ It is perhaps also worth noting that the best rule if used by everyone is ‘Best known’. However, this rule does not share the highest level of concordance with other rules as compared with several other rules (e.g. like ‘Prominence’). This means that there would be weak evolutionary pressure in a multi-rule environment for the best (if used exclusively) rule to emerge.

Table 7 Estimated frequency of each rule use on ‘unrestricted’ in the aggregate and by category

Unrestricted	Prototypicality	Similarity	Odd-one-out	Typ I best	Typ II frequent	Favourite	Prominence	ϵ	MSE
Aggregate	0.3362	0.1529	0	0.3314	0.1795	0	0	0.0342	0.001082
Cities	0.0214	0	0	0	0.789	0.1896	0	0.0234	0.0000076
Fruits	0.9646	0	0	0.0354	0	0	0	0	0.0000074
Sports	0	0	0	0	0.7201	0.2799	0	0	0.0000027
Football players	0	0.2871	0	0.0372	0.6757	0	0	0.0568	0.000334
Positive numbers	0.2117	0.381	0	0.4072	0	0	0	0	0.000384
Supermarket chains	0.3995	0.6005	0	0	0	0	0	0.0287	0.000358
Flowers	0	0	0.1608	0.421	0.4182	0	0	0	0.000338
Items of furniture	0	0	0	1	0	0	0	0	0.000946
Car manufacturers	0.5892	0.1182	0	0.2927	0	0	0	0.086	0.002421
Fast food chains	0	0.0562	0	0.9438	0	0	0	0.0499	0.000459
Animals	0.3697	0.1876	0	0	0.4427	0	0	0.0577	0.00018
Colours	0.3169	0.1839	0	0.061	0	0.4382	0	0	0.001134
Metals	0.402	0.082	0	0.516	0	0	0	0	0.002828
Means of transport	0.9186	0.0814	0	0	0	0	0	0.0388	0.00105
Drinks	0	0	0.0342	0.7107	0	0.0024	0.2527	0	0.000092
Superheroes	0.4767	0	0	0	0	0	0.5233	0.0879	0.000534
Average	0.2919	0.1236	0.0122	0.2766	0.1904	0.0569	0.0485	0.0268	0.000702

Table 8 Estimated frequency of each rule use on 'restricted' in the aggregate and by category

Restricted	Similarity	Odd-one-out	Typ I best	Typ II frequent	Favourite	Prominence	ϵ	MSE
Aggregate	0	0	0.1195	0.4084	0.3426	0.1295	0.083	0.008067
Cities	0	0.115	0.2633	0.6217	0	0	0	0.000057
Fruits	0.311	0	0.3629	0	0	0.3261	0.0964	0.000108
Sports	0	0	0	0.9042	0.0958	0	0.1281	0.000727
Football players	0	0	0.6133	0.3867	0	0	0	0.000745
Positive numbers	0.2032	0.6796	0	0	0.0335	0.0837	0	0.000125
Supermarket chains	0.5265	0	0.2163	0	0.2572	0	0	0.000141
Flowers	0	0.1974	0	0.2709	0	0.5317	0.2891	0.000331
Items of furniture	0.2737	0	0	0	0.7263	0	0.2054	0.010032
Car manufacturers	0	0	0	0.3901	0.5061	0.1038	0	0.001726
Fast food chains	0	0.0801	0.9199	0	0	0	0	0.000036
Animals	0.0792	0	0.6787	0	0	0.242	0.1997	0.000248
Colours	0	0	0.5165	0	0.3768	0.1067	0.0491	0.001045
Metals	0	0.0774	0.5047	0.3969	0.021	0	0	0.000498
Means of transport	0	0	0.4923	0.1753	0	0.3324	0	0.000707
Drinks	0	0.6282	0	0.3718	0	0	0.52	0.000446
Superheroes	0	0.031	0	0.5086	0	0.4604	0	0.000092
Average	0.0871	0.1130	0.2855	0.2516	0.1260	0.1367	0.0930	0.001067

Table 9 Estimated average coordination rate across categories using the estimated frequencies of rule use

	Average coordinate rate	
	Unrestricted	Restricted
Actual data	0.4469	0.4088
Estimated without error	0.4064	0.4029
Estimated with errors	0.3820	0.3706

precision of a rule on the unrestricted set of options. This is because own rule concordance is actually lower for all rules in the unrestricted version. Likewise, under the assumption that people use our seven rules with the same frequency in both the restricted and unrestricted versions but that this frequency differs across individuals, we cannot explain the surprising result because the cross-rule concordance is almost always higher on the restricted version. In short, our first two possible explanations of the surprising result fail. Instead, however, we do have evidence, again if we assume that subjects use our seven rules, that they use these rules with different frequencies under the unrestricted and restricted conditions. Given the own and cross-rule concordances that we generate, this difference in use could explain, albeit only in part, the otherwise ‘surprising’ result from our first experiment.

In particular, while the use of some rules like the two versions of *is used on* both ‘Typicality’ (‘Best known’ and ‘Frequent’) versions, an ‘Odd-one-out’ and a ‘Favourite’ rule are only used when options are restricted. This is an interesting difference. It also makes sense. In the restricted version of our decision problems, everyone has the same options to choose from, while in the unrestricted case, the options that come to subject minds need not be the same. This does not matter for a rule like ‘Typicality I and II’ (‘Best known’ or ‘Frequent’) where there is, in principle, one object that is most frequently mentioned or best known among all possibilities or among the restricted set of possibilities, but it does for a rule like the ‘Odd-one-out’.

The ‘odd one out’ rule depends critically on the full set of options that come to a subject’s mind because one cannot identify the ‘odd one out’ independently of this set. Hence if the options differ across individuals, as they can on the unrestricted domain, so is their likely choice of an ‘Odd-one-out’. For this reason, it is not as well suited to the restricted as the unrestricted version of the coordination problems (because, through construction, everyone has the same options in mind in the restricted case). So the rule could function in the restricted version of the experiment. Indeed this is what Table 6 reveals. It is a good rule for coordination in restricted sets when used against itself but not in unrestricted ones.⁶ The same argument could apply to the ‘Favourite’ rule in the unrestricted case because some people may think, for instance, of their favourite footballer from among those currently playing while others may think of footballers from any time (that they have known).

If this explains the use of these rules on restricted but not unrestricted sets, it also helps explain why our subjects did better on the unrestricted coordination problems.

⁶ A similar line of reasoning might explain the differential use of the ‘Favourite’ rule. If there are a variety of tastes among the population, the scope for disagreement over what is preferred most is constrained when the set of options is restricted as compared with the options that are unrestricted.

The ‘Odd-one-out’ rule, in particular, does well against itself but it does much worse than other rules when playing against other rules (see Table 6). Thus, in so far as there is a diversity of rule use (as our estimates suggest), coordination is impaired by the use of the ‘Odd-one-out’ rule in restricted coordination games. This loss of coordination does not occur in the unrestricted versions of the problem because, again, if our estimates of rule use are correct, the ‘Odd-one-out rule’ is not used.

That we might be able to explain the differences in this way in the use of rules across the restricted and unrestricted versions of these coordination problems is reassuring in the sense that if people follow rules, then the selection of the rule should be intelligible as part of the process of rule following. On one account, rules might be generated through an evolutionary process where rules that coordinate well survive. On another, they might be selected by team reasoners reflecting on which rule would coordinate best. Both make success in coordination crucial for the adoption of a rule and this is how we have attempted to make sense of the difference in the use of these rules across the restricted and unrestricted coordination problems.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

7 Appendix A

7.1 Experiment I instructions

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 £xxx⁷ 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?

⁷ This number depends on the number of people participating. It was £5 times the number of subjects in the session.

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).

7.2 Experiment II instructions

7.2.1 Unrestricted version

Naïve elicitation

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.

Strategic elicitation

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.

7.2.2 Restricted version

Naïve elicitation

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.

Strategic elicitation

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.

References

- Bacharach, M. (1993). Variable Universe Games. In K. Binmore, A. Kirman, & P. Tani (Eds.), *Frontiers of Game Theory* (pp. 255–276). Cambridge, MA: MIT Press.
- Bacharach, M. (2006). *Beyond Individual Choice: Teams and Frames in Game Theory*. Princeton: Princeton University Press.
- Bacharach, M., & Bernasconi, M. (1997). The variable frame theory of focal points: An experimental study. *Games and Economic Behavior*, *19*, 1–45.
- Bardsley, N. (2000). Control without deception: Individual behaviour in free-riding experiments revisited. *Experimental Economics*, *3*(3), 215–240.
- Bardsley, N., Mehta, J., Starmer, C., & Sugden, R. (2010). Explaining focal points: Cognitive hierarchy theory versus team reasoning. *Economic Journal*, *120*(543), 40–79.
- Crawford, V., Gneezy, U., & Rottenstreich, Y. (2008). The power of focal points is limited: Even minute payoff asymmetry may yield large coordination failures. *American Economic Review*, *98*(4), 1443–1458.
- Crawford, V., & Iriberry, N. (2007). Fatal attraction: Salience, naïveté, and sophistication in experimental “Hide-and-Seek” games. *American Economic Review*, *97*(5), 1731–1750.
- Isoni, A., Poulsen, A., Sugden, R., & Tsutsui, K. (2013). Focal points in tacit bargaining games: Experimental evidence. *European Economic Review*, *59*, 167–188.
- Mehta, J., Starmer, C., & Sugden, R. (1994). The nature of salience: An experimental investigation. *American Economic Review*, *84*, 658–673.
- Rosch, E. (1977). Classification of real-world objects: Origins and representations in cognition. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking: Readings in Cognitive Science* (pp. 212–222). Cambridge: Cambridge University Press.
- Schelling, T. (1960). *The Strategy of Conflict*. Cambridge, MA: Harvard University.
- Sugden, R. (1995). A theory of focal points. *Economic Journal*, *105*, 533–550.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.