CBESS Discussion Paper 11-07

Salience as an emergent property

by Federica Alberti[1], Robert Sugden[2] and Kei Tsutsui[3]

[1] School of Economics and Centre for Reasoning, University of Kent, Canterbury CT2 7NP, United Kingdom (e-mail: f.alberti@kent.ac.uk)
[2] School of Economics and Centre for Behavioural and Experimental Social Science, University of East Anglia, Norwich NR4 7TJ, United Kingdom (e-mail: r.sugden@uea.ac.uk)
[3] School of Economics and Centre for Behavioural and Experimental Social Science, University of East Anglia, Norwich NR4 7TJ, United Kingdom (e-mail: kei.tsutsui@uea.ac.uk)

Abstract
We offer an evolutionary model of the emergence of concepts of salience through similarity-based learning. When an individual faces a new decision problem, she chooses an action that she perceives as similar to actions that, when chosen in similar previous problems, led to favourable outcomes. If some similarities are more reliably perceived than others, this process will favour the emergence of conventions that are defined in terms of reliably-perceived similarities. We present experimental evidence of such learning in recurrent play of similar but not identical pure coordination games.

JEL classification codes
C72, C73, C91

Keywords
Focal point; emergent property; similarity; coordination game

Acknowledgements
The experiment discussed in the later part of this paper was joint work with Shaun Hargreaves Heap.

University of East Anglia

CBESS

Centre for
Behavioural and
Experimental
Social
Science

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

In the classical theory of noncooperative games, the formal representation of a game in normal or strategic form takes no account of how players and strategies are *labelled*. Arbitrary player labels (such as 'row player' and 'column player') and strategy labels (such as 'up' and 'down') may be used by the theorist as an aid to analysis, but these are not intended to represent how the players describe the game to themselves, and are not treated as part of the specification of the game; solution concepts are defined so that they are independent of such labels. However, it is now widely recognised that in real-world games, players often *do* take account of the labels that feature in their own descriptions of those games, and that these labels can play an important role in equilibrium selection. As first hypothesised by Thomas Schelling (1960), players recognise (and expect their co-players to recognise) that, by virtue of differences in labelling, some equilibria are more 'prominent' or 'salient' than others; there is a systematic tendency for the most salient equilibrium (the 'focal point') to be selected. There is now a large body of experimental evidence which confirms this hypothesis, at least in relation to pure coordination games (e.g. Mehta et al., 1994a, 1994b; Bacharach and Bernasconi, 1997; Crawford et al., 2008; Bardsley et al., 2010); and there have been a number of attempts to incorporate labels into formal game theory (e.g. Lewis, 1969; Gauthier, 1975; Sugden, 1995; Bacharach and Stahl, 2000; Casajus, 2001; Janssen, 2001; Bacharach, 2006). But the features that make some labels more salient than others are still not well understood.

Salience has usually been analysed in the context of one-shot games, where it has been seen as an equilibrium selection mechanism. It is sometimes suggested that the concept of salience is redundant when games are played recurrently in a population. In this case, it is said, equilibria are reached by dynamic processes of experiential learning, such as replicator dynamics or fictitious play; players have no need to search for the simultaneous 'meeting of minds' that features in Schelling's (1960, pp. 83, 96, 106, 163, 298) account of focal points. For example, Brian Skyrms (1996, pp. 83–94) argues that David Lewis's (1969) analysis of the role of salience in the reproduction of conventions fails to explain how perceptions of salience emerge. According to Skyrms, evolutionary game theory avoids this problem by locating the origin of conventions in symmetry-breaking perturbations: which convention emerges is 'a matter of chance, not salience' (p. 93).

It is certainly true that most evolutionary game theory has not taken account of labelling, implicitly presupposing that in a game with multiple equilibria, which equilibrium emerges is independent of how the players describe that game to themselves. But is that presupposition justified? Robin Cubitt and Robert Sugden have argued that it is not (Cubitt and Sugden, 2003; Sugden, 2004, 2010). Their argument starts from a problem in standard evolutionary game-theoretic accounts of learning.[1] Experiential learning requires inductive inferences which project perceived regularities in a person's experience of previous games. Since no two games are exactly alike, these projections must rely on perceptions of similarity between non-identical games. Such perceptions are subjective, and so are likely to be sensitive to labelling. In the recurrent real-world interactions that evolutionary game theory is attempting to represent, a person's past observations can typically be fitted to a vast number of different logically possible patterns of similarity, with different projections onto new games. Inductive inference works because only a small number of these patterns are recognised and perceived as 'natural' or 'obvious' or 'salient'. The only regularities that have the potential to reproduce themselves as conventions are those that fit pre-existing perceptions of salience that are shared by members of the relevant population.

If this conclusion is correct, it raises the question of how shared conceptions of salience originate. One might hope that an evolutionary theory of the emergence of conventions would be able to explain the emergence of those conceptions of salience on which other forms of experiential learning depend. Conversely, Skymrs's (1996, pp. 83–84) reservations about Lewis's theory seem to reflect scepticism about whether, in a purely evolutionary theory, shared conceptions of salience can pre-exist conventions. (Skyrms's discussion of the emergence of conventions, which is counterposed to Lewis's, is entitled 'Birds do it'. The suggestion is that Skyrms's symmetry-breaking theory is compatible with the natural selection of animal behaviour, while Lewis's theory is not.) In this paper, we offer an evolutionary explanation of the emergence of salience, and present some supporting experimental evidence.

Our explanation combines two ideas. The first is the hypothesis that one of the mechanisms by which experiential learning works is a tendency for an individual, when facing a new decision problem, to recall previous problems that are perceived as similar to it and to choose an action that is perceived as similar to actions that, when chosen in those problems, were followed by favourable outcomes. This mechanism, applied in the context of games against nature, has been modelled by Itzhak Gilboa and David Schmeidler (1995) as

[1] This problem seems to have been first noticed by Lewis (1969). See also Goyal and Janssen (1996).

‘case-based decision theory’. The underlying idea is common to many theories of inductive learning; indeed, Gilboa and Schmeidler cite David Hume’s (1739–40/ 1987) theory of induction as an inspiration. The second idea is that, in the context of recurrent games, perceptions of similarity can be based on labelling, and that some kinds of labelling similarity are more reliably perceived than others. Thus, a putative convention is more likely to emerge and reproduce itself, the more capable it is of being described in terms of reliably-perceived similarities. The intuition that some features of the labelling of strategies are more readily perceived than others, and so are more likely to seed conventions, has been expressed before (e.g. Schlicht, 1988; Sugden, 2004), but we believe that our analysis in terms of similarity relations is new.

We begin, in Section 1, with an intuitive account of how the emergence of conventions may be the product of similarity-based learning. In Section 2, we present a formal model of this learning mechanism in a very simple environment, in which pure coordination games of the kind studied by Schelling are played recurrently. Deliberately, we model a version of the mechanism that makes minimal demands on individuals’ cognition and memory. In acting in accordance with this mechanism, individuals do not engage in any kind of strategic reasoning; they simply attempt to repeat actions that have been successful in the immediate past. We show that, even in a population of such low-rationality agents, there is a tendency for salience-based conventions to emerge, ‘salience’ being defined in terms of reliable perceptions of similarity. In Sections 3 and 4, we present evidence from an experimental investigation of recurrent play of coordination games. This evidence suggests that pairs of co-players learn to use similarity-based rules which increase the success with which they coordinate, and that the process of learning is based on the replication of previously successful actions.

1. The Right Turn Problem

At most British road junctions, driving behaviour is governed by priority rules that are designated by the highway authorities. Normally, one road through a junction is assigned priority; ‘Give way’ signs and road markings on the other routes signify that vehicles on those routes must yield priority to the traffic on the ‘major road’. This system works well at T-junctions, but has proved to be dangerous at crossroads – so much so that, outside towns, crossroads have progressively been replaced by staggered junctions (that is, pairs of slightly offset T-junctions). Where crossroads survive, drivers continually confront what we will call the Right Turn Problem. Consider a crossroads at which the east-west road has priority. Suppose this road is clear. One vehicle is approaching the junction on the minor road from the north, indicating the intention to go straight ahead; another is approaching on the minor road from the south, signalling to turn right. Since Britain drives on the left, their paths will cross. Which vehicle should give way to which? Failure to resolve this coordination problem can result in two stationary vehicles in the middle of the junction – a very dangerous outcome, since major-road drivers (who do not expect to have to give way to anyone) may be approaching the junction at speed.

Surprisingly, the Highway Code (the official codification of the rules of the road in Britain) provides no guidance on this question. If one thinks of the minor road as a continuous route across the junction, it may seem natural to give priority to the vehicle going straight ahead; but that perception is weakened by the significance that has been assigned to the major road. An opposing thought, encouraged by experience of staggered junctions, is that the right-turning vehicle is joining the major road, and so inherits the priority of major-road vehicles after it has turned. In practice, the Right Turn Problem is usually resolved by giving priority to whichever of the minor-road vehicles reached the junction first; if there are queues, a vehicle is deemed to have reached the junction only when it gets to the front of its queue.

How has this (imperfectly recognised) convention emerged, and how does it reproduce itself? The most convincing answer, we suggest, is that each driver learns from his experience. Having observed what he perceives as regularities in the behaviour of other drivers across a set of similar situations, he projects those regularities onto the new driving problems he confronts. As an explanation of how an *already existing* convention reproduces itself, this is perhaps unproblematic: if almost all drivers are already following a particular priority rule, learning that rule from one's observations may not be too difficult. But if one is trying to explain how a convention emerges in the first place, and particularly if one is trying to explain *which* convention is most likely to emerge, one has to consider how people learn from experiences which do *not* exhibit very obvious patterns.

In this context, subjective perceptions of similarity may play a crucial role. Suppose you are facing a new instance of the Right Turn Problem. You were the second driver to arrive at the junction and you are turning right. You recall just one previous instance of the problem. In that case, you were the first arrival and you were turning right, and the other driver clearly gave way to you. Does that recollection induce some expectation that the new driver will give you priority (since you are turning right) or that she will assume priority herself (since she was the first arrival)? The answer seems to depend on which similarity relationships are more salient to you. If most people perceive ‘first arrival’ to be a more salient dimension of similarity than ‘turning right’, then one might expect a ‘first arrival’ convention to be more likely to emerge. A related issue is the interpersonal reliability of judgements with respect to given dimensions of similarity. Suppose, for example, that when two vehicles arrive at a junction at approximately the same time, judgements about which was the first arrival are subject to noise. Then, even if both drivers follow the rule of giving priority to the perceived first arrival, they will not necessarily coordinate. One might expect that effect to tend to work against the emergence of the ‘first arrival’ convention.

We conclude that an adequate analysis of the emergence of conventions needs to take account of the relative salience and reliability of different conceptions of similarity that can be applied to the labelling of games. This kind of analysis may seem to involve a major departure from mainstream game theory, but we will try to persuade the reader that the problems that have to be solved are theoretically tractable and amenable to experimental investigation.

2. A model

In this section, we develop a model of learning behaviour in a family of *recurrently similar* games – that is, games that are similar but not identical to one another and that are played recurrently in a population. We assume a population that is sufficiently large to legitimate the use of the law of large numbers. In each *period* $t = 1, 2, \ldots,$ pairs of individuals are drawn at random from this population to interact as co-players. Each individual participates in one such interaction in each period.

Each interaction is a pure coordination game, defined by a set of *n labels*, where $n \geq 2$. In this section we assume $n = 2$, but we use a notation that allows the analysis to be generalised. Independently and without communication, each player sees the two labels and chooses one of them. Each receives a payoff of 1 if their choices *match* – that is, if both choose the same label – and a payoff of zero otherwise. To allow a simple representation of similarity between games, we assume that labels are of two types, A and B. In each game, one label is of type A and the other is of type B. Although our modelling strategy attaches special significance to the distinction between these two types of label, we do *not* assume that players conceptualise the problem in terms of the distinction between ‘A’ and ‘B’.

For example, suppose that in each game, the two labels are alternative descriptions of the behaviour of two vehicles facing the Right Turn Problem at some crossroads. The only difference between the two descriptions in any given game is which vehicle gives way to which. Since the players’ common objective can be interpreted as that of coordinating on an assignment of priority to one of the two vehicles, this example can be interpreted as a stylised model of a real-world Right Turn Problem.[2] Suppose that in each game, there is one description (type A) in which the second arrival gives way to the first, and one (type B) in which the first arrival gives way to the second. However, the fact that every game has this feature is not made explicit to the players. Across games, other features of the descriptions – such as whether the first arrival is going straight ahead or turning right, the types of vehicle involved, the flow of traffic on the main road, the weather, and so on – may vary.

As a model of individual behaviour, we postulate the following *replication heuristic*. The heuristic has two ‘settings’. In any period in which a given individual i is using the *default setting*, that individual chooses between labels in some way that is independent of her experience of previous games. Her choice might be random, or it might be influenced by properties of the relevant pair of labels. We simply assume that, in any randomly selected game, the *default probabilities* with which a randomly selected player chooses A and B are q_A and q_B, where $0 < q_A, q_B < 1$ and $q_A + q_B = 1$. In period 1, each individual acts on the default setting.

In each period $t > 1$, each individual i uses the default setting if and only if she failed to match with her co-player in period $t - 1$; otherwise the *similarity setting* is operative. In this case, she tries to replicate the previous match by choosing whichever label in the period t game she perceives to be more similar to the label that she chose in the previous game. We model this process by defining measures r_A, r_B of the *intrinsic replicability* of the two types

[2] One feature of the real-world problem that has been abstracted from the model is the conflict of interest between real drivers about *whose* vehicle has priority. In the game, players are not identified with particular drivers. Thus, our model is a pure coordination game rather than a Battle of the Sexes game.

of label, where $0 \leq r_A, r_B < 1$. For a randomly selected player in a randomly selected game, the probability with which A (respectively B) is chosen, conditional on that player having chosen A (respectively B) in the previous round and having matched with her co-player, is denoted by s_A (respectively s_B); these are measures of *gross replicability*. We model the process of replication by:

$$s_A = r_A + (1 - r_A)q_A, \quad \text{and} \tag{1}$$

$$s_B = r_B + (1 - r_B)q_B. \tag{2}$$

This specification ensures that s_A and s_B are strictly positive and that default probabilities have some influence on choice even when a player is trying to replicate a previous match; and it imposes the natural restriction that a player is at least as likely to choose a given type of label when trying to replicate another label of that type as when acting on default probabilities.[3]

The replication heuristic can be interpreted as a particularly simple and cognitively undemanding form of similarity-based inductive learning, in the same spirit as Gilboa and Schmeidler's (1995) case-based decision theory. Since it responds only to the success or failure of the individual's own actions, it does not involve any theory of mind or strategic reasoning. It requires only a one-period memory, and does not keep track of the relative success of alternative actions. Recalling Skyrms's scepticism about salience, it does not seem implausible to suggest that even birds might be capable of this kind of mental operation.

For any period t, let π_t be the relative frequency with which, in the whole population, A is chosen in that period. We define a function f such that, for all t, $\pi_{t+1} = f(\pi_t)$. Each game played in period t must have one of three outcomes – a match on A, a match on B, or no match. These outcomes occur with the respective probabilities π_t^2, $(1 - \pi_t)^2$, and $2\pi_t(1 - \pi_t)$. It follows immediately from the specification of the replication heuristic that

$$\text{for all } t\text{: } f(\pi_t) = \pi_t^2 s_A + 2\pi_t(1 - \pi_t)q_A + (1 - \pi_t)^2 (1 - s_B) \tag{3a}$$

$$= \pi_t^2[1 + s_A - s_B - 2q_A] + \pi_t[2q_A + 2s_B - 2] + [1 - s_B]. \tag{3b}$$

[3] This formulation allows the model to be extended to cases where $n > 2$. The general model has label types $j = 1, ..., n$, default probabilities $q_j \in (0, 1)$ and intrinsic replicability measures $r_j \in [0, 1)$. Following a match on j, j is chosen with probability $r_j + (1 - r_j)q_j$, while each $k \neq j$ is chosen with probability $(1 - r_j)q_k$.

A stationary state or *equilibrium* of the model is defined by a relative frequency $\pi^* \in [0, 1]$ such that $f(\pi^*) = \pi^*$.

To investigate the properties of this equilibrium, notice that f is a quadratic function with $f(0) = 1 - s_B$ and $f(1) = s_A$; f is everywhere convex (respectively: concave) if $1 + s_A - s_B - 2q_A$ is positive (respectively: negative). Equivalently, f is everywhere convex (concave) if $(s_A - s_B) - (q_A - q_B)$ is positive (negative). Given these properties of f, the following result is an immediate implication of the assumption that s_A and s_B are strictly positive:

> *Result 1*: There is exactly one equilibrium π^*. This equilibrium satisfies $0 < \pi^* < 1$ and is globally stable.

(Proofs are presented in Appendix 1.) Figure 1 illustrates equilibrium as the intersection of the graph of $y = f(\pi)$ with the line $y = \pi$. (In the case illustrated, $q_A = q_B = 0.5$, $r_A = 0.6$ and $r_B = 0.2$, implying $\pi^* = 0.59$.)

[Figure 1 near here]

It follows from (1), (2) and (3a) that, for all $\pi \in (0, 1)$, $f(\pi)$ is increasing in r_A and q_A and decreasing in r_B, implying the following comparative static properties of equilibrium:

> *Result 2*: Other things being equal, π^* increases as r_A and q_A increase, and as r_B decreases.

In other words, the replication heuristic tends to favour label types that have higher default probabilities and greater intrinsic replicability.

One can get some feel for the trade-off between decreases in default probability and increases in replicability by considering the conditions under which the equilibrium frequency of A is greater than (respectively: equal to, less than) 0.5. It is evident from Figure 1 that π^* is greater than (equal to, less than) 0.5 if and only if $f(0.5)$ is greater than (respectively: equal to, less than) 0.5. Using (3a) and the fact that $q_A + q_B = 1$, it is straightforward to derive:

> *Result 3*: π^* is greater than (equal to, less than) 0.5 if and only if $s_A - s_B$ is greater than (equal to, less than) $q_B - q_A$.

Our next result concerns the effects of varying the intrinsic replicability of A and B *together* while maintaining equality between r_A and r_B. Because of the symmetry between A and B, there is no loss of generality in considering only the case in which $q_A \geq 0.5$:

Result 4: Assume $r_A = r_B = r$ and $q_A \geq 0.5$. Then:

(i) if $r = 0$, $\pi^* = q_A$;

(ii) if $q_A = 0.5$, then $\pi^* = 0.5$ for all $r \in [0, 1)$;

(iii) if $q_A > 0.5$, then as $r \rightarrow 1$, $\pi^* \rightarrow 1$;

(iv) if $q_A > 0.5$, then $\pi^* > q_A$ for all $r \in (0, 1)$.

This result shows that the overall effect of the replication heuristic is to magnify dispersion in the distribution of default probabilities, and hence to increase the frequency of coordination relative to the default benchmark. Obviously (given the symmetry between the two label types, and given the assumption that they have equal intrinsic replicability r), if A and B have equal default probabilities, then they are chosen with equal frequency in equilibrium. But part (iv) of Result 4 establishes that if A has strictly greater default probability and if $r > 0$, the frequency with which A is chosen is not only greater than 0.5; it is strictly greater than the default probability q_A. Intuitively, this is because replication is activated by matching. If players choose according to default probabilities, the ratio between *choices* of A and *choices* of B is $q_A{:}q_B$, but the ratio of *matches* on A to *matches* on B is $q_A^2{:}q_B^2$. If $q_A > 0.5$, the latter ratio is greater than the former, and so replication works disproportionally in favour of A. The higher the value of r, the greater the effect of this disproportion on the equilibrium; as r tends to one, the equilibrium frequency of A choices tends to one also.

Finally, we consider whether each individual benefits by using the replication heuristic, rather than by acting on default probabilities. Consider any individual i. For the purposes of this analysis, we use q_A, q_B, r_A and r_B to denote the default probabilities and intrinsic replicabilities that are relevant in determining i's behaviour; we continue to assume $0 < q_A, q_B < 1$, $q_A + q_B = 1$, and $0 \leq r_A, r_B < 1$. Let $\pi \in [0, 0]$ be the (constant) relative frequency with which A is played in the population, and hence (given the assumption that the population is large) the probability that A is chosen by i's co-player in each game. We make no assumptions about the determinants of π or about the relationship between π and q_A. In particular, we do *not* assume that i's co-players have the same default probabilities as i does, nor that they use the replication heuristic. We assume only that their behaviour has some consistent pattern, described by π. Is *i*'s expected payoff per round greater if her decision rule is the replication heuristic ('rule *R*') than if it is the use of default probabilities ('rule *D*')?

In general, the answer to this question depends on q_A, r_A, r_B and π. (For example, if $\pi > 0.5$ and $r_A < r_B$, replication works in the 'wrong' direction; if this effect is sufficiently strong, *i* might get a higher expected payoff by using *D* rather than *R*.) But a sharp answer is possible for all cases in which A and B have the same intrinsic replicability:

> *Result 5*: Consider any individual *i* for whom the default probabilities and intrinsic replicabilities of A and B are q_A, q_B and r_A, r_B, where $r_A = r_B = r > 0$. Suppose that, in every period, *i*'s co-player chooses A with some constant probability $\pi \in [0, 0]$. Let $v(D, t)$ be *i*'s expected payoff in period *t*, conditional on her using rule *D* in all periods. Let $v(R, t)$ be *i*'s expected payoff in period *t*, conditional on her using rule *R* in all periods. Then $v(R, t) \geq v(D, t)$ for all *t*. If $\pi \neq 0.5$, $v(R, t) > v(D, t)$ for all $t > 1$.

At first sight, it might seem surprising that, given only the assumptions $r_A = r_B > 0$ and $\pi \neq 0.5$, rule *R* can be shown to be unambiguously superior to rule *D*. The key to the proof is that, when individual *i* uses rule *R*, the effect of replication is always to increase the probability with which she chooses the label type that her co-players choose more frequently. This is the case because the replication heuristic does not try to replicate whatever co-players do; it tries only to replicate *matches*. For example, consider the case $\pi = 0.6$, $q_A = 0.9$, $r = 0.3$. If *i* tried to replicate her co-players' behaviour in general, the probability with which she chose A would tend to fall from its default level. But because the replication heuristic is activated only by matches, this probability will tend to increase. Suppose, for example, that in period 1, *i* matches with her co-player. The posterior probability that this match was on A is $(0.9)(0.6)/\ [(0.9)(0.6) + (0.1)(0.4)] = 0.931$. Thus, the probability that *i* chooses A in period 2 is $(0.931)[0.3 + (0.7)(0.9)] + (0.069)(0.7)(0.9) = 0.909$, which is greater than q_A.

To sum up, we have described a very simple heuristic based on the principle of choosing actions that are similar to actions that have proved successful in the immediate past. When this heuristic is used by populations of players of recurrently similar pure coordination games, there is a tendency for the emergence of conventions that are based on shared perceptions of similarity. Other things being equal, this process tends to favour those putative conventions that have higher default probabilities and higher intrinsic replicability.

We suggest that this process can be interpreted as the emergence of conceptions of salience, and that the conceptions that are favoured by this process have at least something in common with those features of labels that distinguish focal points in one-shot coordination games. In the theoretical and experimental literature, one recurring idea is that focal points are grounded in *primary salience* – that is, in individuals' predispositions to choose some labels rather than others, in the absence of any strategic or payoff-related reasons to do so (Lewis, 1969; Mehta et al., 1994a; Bardsley et al., 2010). This idea can be developed by using *level*-k or *cognitive hierarchy* theories, in which pre-strategic dispositions are attributed to players who reason at 'level 0' (Stahl and Wilson, 1995; Camerer et al., 2004; Crawford et al., 2008). The default probabilities of our model capture something of the same idea. Another recurring idea is that focal points are distinguished by properties of uniqueness in their labelling. Thus, in games with a finite number of labels, labels that are perceived as 'odd-ones-out' tend to be chosen, even if they are not primarily salient (Schelling, 1960; Bacharach and Stahl, 2000; Casajus, 2001; Janssen, 2001; Bacharach, 2006; Bardsley et al., 2010). It is plausible to suppose that, in families of games with clearly-defined odd-ones-out, the odd-one-out type of label will generally have high intrinsic replicability.

3. An experiment

In this and the following section, we discuss some evidence of learning behaviour in an experimental implementation of recurrently similar pure coordination games.

This experiment is described in more detail in Alberti et al. (2010). Here we merely summarise its main features. Subjects were paired randomly and anonymously; pairings were maintained for the duration of the experiment. The experiment used forty pure coordination games. Each pair of subjects played all of these games, but not in the same order. Each game was defined by a set of four labels or *images*. Players were instructed to try to match with their co-players, and were rewarded in money in proportion to the number of matches they achieved. After each game, players were told which image their co-players had chosen, thus allowing opportunities for them to learn from one another's behaviour. Games were presented in 'blocks' of five similar games; each pair of subjects played all the games in one block before moving on to another block. There were four blocks of *culture-laden* games and four of *abstract* games. Each pair of subjects either played all the culture-laden games before playing the abstract ones, or vice versa; which type of game was played first (as ‘part 1’ of the experiment) was counterbalanced.

In each game in each culture-laden block, each of the four images was an example of a different *style*; the same four styles appeared in each game. However, this feature of the design was deliberately not made explicit to subjects, who were told only that every game was made up of four ‘pictures’. Figure 2 shows two games from a block of culture-laden games in which the images are fabric designs and styles represent particular periods in the history of fashion in western society. In the figure, images with the same style appear above one another. In the experiment, however, the positions of the images were randomised independently for each co-player, so that position could not be used as a coordinating device. In two blocks the images were fabric designs; in the other two they were paintings. In the latter case, each style corresponded with a specific painter; within a game, the subject matter of the four paintings was (as far as possible) similar. Each block used a distinct set of four styles (fashion periods or painters). The images for all culture-laden games are shown in Appendix 3 [intended for on-line publication only].

[Figure 2 near here]

Figure 3 shows two games from a block of abstract games. In these games, there were no pre-determined styles. Although the twenty abstract games were presented to subjects in blocks of five, all games were constructed on the same principles. Each image is chequered pattern of coloured squares, using three distinct colours in a common pattern. In any given game, two *fixed colours* and their respective positions are held constant across all four images. The third *variable colour* is different in each of the four images. This design feature induces a general resemblance between the images in any given game. Subject to the constraints we have described (and the additional constraint that no two images in a given game should be identical) colours were selected at random from a pre-determined set of forty-eight colours, which had been constructed so that no two colours were difficult to distinguish from one another; the location of the variable colour was also selected at random. The images for all abstract games are shown in Appendix 4 [intended for online publication only].

[Figure 3 near here]

A further feature of the experiment is relevant for our analysis. Before playing the twenty abstract (respectively: culture-laden) games, each subject completed a questionnaire in which, in turn, she was shown eight sets of four abstract (culture-laden) images. For each set, she was asked to record which of the four images she 'liked most' and which she 'thought the person whom she was paired with liked most'. In fact (although this was not revealed to the subject at the time), these were the sets of images that would appear in the first and last games that the subject played in each of the four blocks of abstract (culture-laden) games.[4] The order of games within blocks was only partially randomised, so that every subject's questionnaire referred to the same games.

One difference between the experimental set-up and the model is that the experimental games have four labels rather than two. However, as explained in Section 2, the replication heuristic can be generalised to apply to any number of labels. A second difference is that, in the experiment, each subject interacted with the same co-player in all games, while in the model players are randomly re-matched between games. Thus, the experimental set-up allowed different conventions to emerge for different pairs of subjects, while in the model conventions must be properties of the whole population. However, the replication heuristic itself does not differentiate between the two cases: it is simply a mechanism by which each individual player seeks to replicate previously successful choices. In Appendix 2 we consider the implications of the replication heuristic when player pairings are fixed.

Before analysing the data that are directly relevant to replication, we briefly review some of the general findings of the experiment.

For any given game, we define the *matching frequency* as the relative frequency with which *co-players* chose the same image. Following Mehta et al (1994a), we define the *coordination index* for a game as the probability that two distinct players, *selected at random from the whole population of players,* chose the same image. This latter statistic uses individuals' actual choices, but (in contrast to the matching frequency) does not take account of who was paired with whom. As a benchmark, notice that (given that we are dealing with a four-label game) if all players choose at random, then the expected values of both the matching frequency and the coordination index are 0.25. The extent to which the coordination index exceeds 0.25 is a measure of positive correlation between choices *among*

[4] Each subject also completed the same questionnaire after playing the twenty relevant games, but these 'after play' responses are less relevant for our current purposes.

subjects in general. The extent to which the matching frequency exceeds the coordination index measures any additional *pair-specific* correlation.

Table 1 presents some summary statistics about matching frequencies and coordination indices for different types of game and at different stages in the experiment. Recall that 'part 1' and 'part 2' refer respectively to the first and second sets of twenty games played by pairs of subjects. For some pairs, part 1 contained all the culture-laden games and part 2 contained all the abstract games; for others, this order was reversed. We use the term 'round' to refer to the order in which games were played. Thus, for each set of twenty games of the same type (culture-laden or abstract), 'rounds 1–10' refers to the first ten games of that type faced by any pair of subjects, while 'rounds 11–20' refers to the remaining games of that type. Because the order in which games were faced by different pairs was randomised, comparisons between data for 'part 1' and 'part 2', and between data for 'rounds 1–10' and 'rounds 11–20', pick up the effects of experience.

[Table 1 near here]

Notice that coordination indices, although consistently greater than the 0.25 benchmark, are typically quite close to 0.25, even in the later stages of the experiment. In contrast, matching frequencies are always markedly greater than coordination indices, and (particularly in the case of abstract games) increase over the course of the experiment. The marked increase in the matching frequency for abstract games between rounds 1–10 (0.361) and rounds 11–20 (0.423) suggests that, by using their experience of playing recurrently similar games, co-players were able to increase their success in matching. It is perhaps not surprising that a similar effect was not observed for culture-laden games, for which each five-game block had its own characteristics and styles. However, it is interesting that the matching frequency for abstract games also increased between part 1 (0.345) and part 2 (0.446), suggesting that experience of playing culture-laden games facilitated matching in abstract games.

It seems clear that some kind of similarity-based learning occurred. Since co-players were matched at random, the only possible explanation of the excess of matching frequencies over coordination indices is that subjects learned something from their co-players' behaviour, that this learning facilitated matching, and that different pairs followed different learning paths. In a general sense, any learning that facilitates matching in recurrently similar games *must* exploit similarity relations between games: if players did not perceive similarities between the games they played, they would have no way of using what they learned in one game to guide their decisions in another. But the fact that learning was predominantly pair-specific provides a further clue about what was being learned. During the course of the experiment, the only feedback that each player received was information about her co-player's choices. So this information (and this information alone) must have been the input to pair-specific learning. That is, players must have adapted their own choices in response to their co-players' earlier choices. Since the overall effect of this adaptation was to increase matching, and since different pairs followed different learning paths, the obvious inference is that in some way, players adapted their decision rules to favour choices that were similar to their co-players' previous choices.

Another significant finding concerns the role of subjects' 'likings', as reported in the questionnaire. The graphs in Figure 4 plot the proportion of cases in which the image chosen by a player in a game was the same as the one that she had reported she most liked.[5] Notice that points on the graphs aggregate over games (which may have had different sets of labels) which were played at the same position in the sequence of twenty culture-laden or abstract games. One very obvious feature of these graphs is that, when playing their first culture-laden or first abstract game, subjects were very likely to choose the label that they most liked (the proportions are 0.686 for culture-laden games and 0.593 for abstract). The frequency with which most-liked images are chosen tends to decline over the twenty relevant games, but remains well above 0.25 throughout; and there is a spike at the start of each new block of games. This pattern suggests that subjects used their own likings to determine their default choices, but that as they played more games of a given type, they made less use of that default rule. That the use of subject-specific default rules declined as more games were played is consistent with the conclusion that, after observing their co-players' choices, subjects adapted their own decision rules to favour choices similar to those that had been made by their co-players.

[Figure 4 near here]

[5] Which image a subject most liked was a better predictor of her choices than which image she thought her co-player most liked.

4. Similarity and replication in experimental choices

We have shown that, in general terms, behaviour in the experiment revealed the effects of similarity-based learning. So far, however, we have not presented evidence that, as would be the case if the replication heuristic were being used, players specifically replicated previous *matches* (rather than their co-players' previous choices in general). And we have not asked *which* features of the images players replicated, or *how successfully* different features were replicated. One of the most significant implications of the model presented in Section 2 is that the evolution of conventions is influenced by the *relative* replicability of the labelling characteristics of alternative putative conventions. It is therefore particularly interesting to investigate whether, in the experiment, different labelling features had different degrees of replicability.

For each block of culture-laden games, the four styles provide a pre-defined set of similarity relationships between games, corresponding with the two label types in the model. We will investigate how subjects used these relationships. We must stress, however, that these are not the only concepts of similarity that players can use. (Indeed, the concept of 'liking' is itself a concept of similarity: from the viewpoint of a given player, an image in one game can be similar to an image in another by virtue of its being the 'most liked'.) For our purposes, the significance of styles is merely that they allow us to identify *one* set of well-defined similarity relationships. We can then investigate whether there was any tendency for these particular similarity relationships to be used by players to replicate previous matches.

In contrast, the abstract games do not have pre-determined styles. However, we suggest that *one* salient way of distinguishing between the four images in a game is in terms of the variable colour. Thus, for example, in abstract game A2 (shown in the top row of Figure 3), the variable colours (from left to right) might be perceived as 'pale turquoise', 'purple', 'green' and 'dark blue'. Suppose that, in one round of the experiment, a pair of players face this game and match on the image on the left. In the next round, they face abstract game A3 (shown in the bottom row of Figure 4). If a player tries to replicate the previous match, which image in game 3 is most similar to the image chosen in game 2? One possible answer is that since the image chosen in game 2 was the palest of the four, the palest image in game A3 (presumably the third from the left) should be chosen.

By virtue of the way in which the experiment was computerised, each of the forty-eight colours is described by a unique triple of parameters (x_R, x_G, x_B) where $x_j \in [0, 1]$ is the intensity of colour *j*, where *j* is one of the three primary colours for light (R = red, G = green, B = blue). In Appendix 4, the three parameters which identify the variable colour in each image are shown (as percentages) below the image. These parameters allow us, as analysts, to define colour-based similarity rules in an objective way. (Of course, these definitions of colours will not always coincide with the subjective perceptions of individual players.) For the purposes of our analysis, we define eight different *similarity rules*. One rule is 'choose the most red', which we operationalise as 'maximise $x_R/(x_R + x_G + x_B)$'. (Because colour mixes are defined in terms of light rather than pigment, a high value of $(x_R + x_G + x_B)$ represents a 'pale' or 'unsaturated' colour, that is, a colour close to white. Our definition of 'most red' uses a relative rather than absolute concept of 'redness' so that, for example, a pale pink is treated as more red than a dark purple.) The rules 'choose the most green' and 'choose the most blue' are defined analogously. Similarly, the rule 'choose the least red' is operationalised as 'maximise $(1 - x_R)/[(1 - x_R) + (1 - x_G) + (1 - x_B)]$'. The rules 'choose the least green' and 'choose the least blue' are defined analogously. Colours at the 'least blue' extreme appear yellow; 'least green' colours appear purple, and 'least red' colours appear turquoise. Our final two rules use the unsaturated/saturated dimension. The rules 'choose the most pale' and 'choose the least pale' are operationalised as 'maximise $x_R + x_G + x_B$' and 'minimise $x_R + x_G + x_B$' respectively.

Typically, each of these rules identifies a unique image in each game. (In approximately five per cent of cases, the maximisation or minimisation criterion of a similarity rule is satisfied by two or more images.[6]) Since there are eight similarity rules but only four images in each game, a given image is often picked by more than one rule. (For example, in game A2, image 3 is 'least pale', 'most green', and 'least blue'.) Thus, even on the assumption (which we do not make) that the only similarity relationships perceived by players are those specified by our eight rules, successful coordination after a match requires more than that each co-player replicates *some* similarity property: the similarity properties that they replicate must pick out the same image in the following game. Nevertheless, we can investigate the extent to which players use these various similarity rules in trying to replicate previous matches.

[6] For example, in game A3, with x_R, x_G, x_B expressed as percentages, the variable colours of images 1, ... 4, are (100, 0, 100), (100, 63, 48), (100, 100, 100) and (0, 100, 100) respectively. Images 1 and 4 jointly satisfy the 'most blue' criterion.

Table 2 summarises the relevant data. For each block of culture-laden games, we consider four similarity rules, each corresponding with a pre-determined style. For the abstract games, we consider the eight colour-based similarity rules explained above. To simplify the exposition, we will refer to the corresponding colour properties (such as ‘most pale’ and ‘least red’) as ‘styles’. For each similarity rule *j*, the table reports the following data.

[Table 2 near here]

Column 1 reports the number of games in which rule *j* was defined and (after the slash) the total number of relevant games. For example, the ‘19/20’ entry for ‘most pale’ indicates that the similarity rule ‘choose the most pale image’ was potentially applicable to twenty games (i.e. all twenty abstract games) and was in fact uniquely defined for nineteen of these. In the culture-laden games, each similarity rule is uniquely defined for all games in the relevant block, and so these rules have the entry ‘5/5’. For those games in which rule *j* was uniquely defined, column 2 reports (as a percentage) the relative frequency with which subjects’ choices were consistent with that rule. We will call this the *choice frequency* for style *j*. Notice that if subjects’ choices were random, the expected choice frequency would be 0.25 for every style.

Columns 3–8 refer to cases (that is, combinations of a subject *i* and a round *t*) in which (i) rule *j* was uniquely defined for the game played by *i* in round *t* – 1, (ii) *i*’s choice in round *t* – 1 was consistent with rule *j*, and (iii) rule *j* was uniquely defined for the game played by *i* in round *t*. Notice that these are the cases in which it is meaningful to ask whether *i*’s choice in round *t* replicated an immediately preceding choice that was consistent with rule *j*. These cases can be subdivided into two categories: those in which *i*’s choice in *t* – 1 was matched by that of her co-player, and those in which it was not. In the context of an investigation of learning mechanisms, it is particularly relevant to compare the relative frequencies of *j* choices in these two categories. Notice that if subjects took no account of the behaviour of their co-players, there would be no systematic difference between the two frequencies. However, a subject who used the replication heuristic would respond differently in the two cases. If she had matched in *t* – 1, her choice in *t* would be an attempt to replicate her previous choice. If she had not matched in *t* – 1, her choice in *t* would be determined by her default rule. Thus, to the extent that style *j* is reliably perceived by subjects, we should expect *j* choices to be more frequent following a match.

Column 3 reports the number of relevant cases in which there was a match in round *t* – 1. Column 4 reports the number of these cases in which *i*'s choice in *t* was consistent with rule *j*. Column 5 shows this as a percentage of possible cases (i.e. as a percentage of the entry in column 3); we will call this the *post-matching replication rate* for style *j*. Similarly, column 6 reports the number of relevant cases in which there was *not* a match in round *t* – 1. Column 7 reports the number of these cases in which *i*'s choice in *t* was consistent with rule *j*. Column 8 shows this as a percentage of possible cases; we will call this the *baseline replication rate* for style *j*.

The entries in columns 3, 4, 6 and 7 can be used to construct a 2×2 contingency table in which the rows are 'matched in *t* – 1' and 'failed to match in *t* – 1' and the columns are '*j* chosen in *t*' and '*j* not chosen in *t*'. (For example, for block 1, style 1, the numbers in the cells of the first row are 5 and 19; the second-row numbers are 12 and 53.) Column 8 reports the chi-squared statistic for the relevant contingency table. Recognising that the observations in these tables are not independent, we do not treat the entries in this column as the results of formal statistical tests; we offer them merely as convenient summary statistics.

The data in Table 2 reveal a clear tendency for subjects to use similarity rules to replicate previous matches. This tendency is most obvious when one aggregates over rules. Averaging over all rules for culture-laden games, the baseline replication rate is 31.3 per cent.[7] The corresponding post-matching replication rate is 51.3 per cent. Choices in abstract games reveal a similar but less pronounced pattern; here the baseline rate is 31.3 per cent and the post-matching rate is 39.1 per cent. That the pattern is less pronounced for abstract games should not be surprising, given that we have specified eight styles for games with only four labels.[8]

We now consider whether some similarity rules were used more, or with more success, than others. For a given similarity rule, we can assess the success with which it was used by comparing its post-matching and baseline replication rates. If the former is markedly

[7] That this is greater than 25 per cent may reflect between-subject variation in likings for different styles. Such variation would induce autocorrelation in subjects' default choices. An alternative explanation is that some subjects persisted in attempting to replicate previous matches even after one such attempt had failed.

[8] When there is a one-to-one relationship between similarity rules and labels, as in the culture-laden games, it is in principle possible for every similarity rule to have a post-matching replication rate of 100 per cent. Clearly, this is not possible when different similarity rules have conflicting implications about how a given choice should be replicated.

greater than the latter, and if the chi-squared statistic is relatively large,[9] it is reasonable to infer that subjects used the rule's concept of similarity in attempting to replicate matches made in preceding rounds – and that these attempts were to some extent successful. Interpreting such a finding in terms of our model, we will say that it is evidence of the *intrinsic replicability* of the corresponding style.

Notice that, because of the interaction between default choices and replication, a style can have high intrinsic replicability even if its choice frequency is relatively low, and conversely. Intuitively, a style has intrinsic replicability by virtue of its *distinctiveness*, while (as we have shown) default choices reflect the *likeability* of styles. A style that is distinctive but not generally liked may high intrinsic replicability but a low choice frequency. Conversely, a style may express some common property that subjects tend to like, but of which they are not consciously aware. Such a style may have a high choice frequency but low intrinsic replicability.

Perhaps the most striking feature of the data in Table 2 is the extent to which intrinsic replicability varies between styles. Among the sixteen styles used in the culture-laden games, there is one very obvious outlier: style 2 in block 2, for which the post-matching and baseline replication rates are respectively 92.9 and 45.6 per cent. Our intuition is that the similarity relationship between different instances of this style is particularly strong, making the action of choosing it particularly easy to replicate from one game to another. *In this sense*, the style is very distinctive. Notice that this is not the same thing as saying that, *as seen in any one game*, the style is highly salient (and hence an obvious focal point in Schelling's sense); our impression is that, in this latter sense, it does not stand out particularly strongly. (Readers are invited to look at Appendix 3 and compare their intuitions with ours.) But even if one finds it unsurprising that subjects recognised this style, it is still remarkable that, in over ninety per cent of cases, subjects who had matched on this style in one game chose the same style in the following game.

Four other culture-laden styles stand out as having high degrees of intrinsic replicability: style 3 in block 2, style 4 in block 3, and styles 1 and 3 in block 4. (For each of these styles, the post-matching replication rate is more than twenty-five percentage points larger than the baseline rate and the chi-squared statistic is far higher than the conventional

[9] As a benchmark: for a 2×2 contingency table, the critical value of the chi-squared statistic at the 95 per cent confidence level is 3.84.

benchmark of 3.84.) Our intuition is that each of these styles is distinctive in the sense explained in the previous paragraph – although not to the same degree as the style we were then referring to.

It is perhaps worth noticing a culture-laden style that is an outlier in a different sense: style 3 in block 1. This has by far the highest choice frequency of any style (44.7 per cent, compared with the next highest frequency of 33.9 per cent), but does not stand out in terms of intrinsic replicability. We suggest that this is an example of a style that subjects generally liked, while failing to recognise *as a style*.[10]

When similarity relations are defined stylistically and applied to fabric designs or paintings, any discussion of those relations inevitably draws on subjective judgements and intuitions. We see no need to apologise for this, sharing Schelling's (1960: 98) view that normative game theory must be prepared to analyse whatever features of coordination games players are able to use in their mutual interest.[11] However, the abstract games allow us to analyse styles which, although applied to images with aesthetic properties, can be defined more objectively.

Comparing the eight colour-based styles, we again find a high degree of variation in intrinsic replicability. Two styles stand out from the others as highly replicable: 'most pale' and 'most blue'. For each of these styles, the post-matching replication rate is more than thirteen percentage points greater than the benchmark rate, and the chi-squared statistic is far higher than the conventional benchmark. Interestingly, these styles have very different choice frequencies: the frequency of 'most blue' (35.5 per cent) is the highest of any colour-based style, while that of 'most pale' (24.8 per cent) is almost exactly equal to the random-choice benchmark. We conjecture that subjects were more inclined to choose blue images than pale images when making default choices, but found both styles distinctive in the sense of being able to recognise and replicate them.

Summing up, the data in Table 2 provide direct evidence of a tendency for subjects to replicate choices that, in immediately preceding rounds, had been matched by their co-

[10] Compare the proverbial response to the pretentions of high culture: I may not be an expert, but I know what I like.

[11] Sugden and Zamarrón (2006) examine the fundamental differences between Schelling's pragmatic understanding of game-theoretic rationality and the concept of rationality used in classical game theory.

players. They show that, in replicating such choices, subjects used specific similarity relations between the images by which strategies were labelled, and that some such relations were used much more, or much more effectively, than others. These findings support the argument, presented in Section 2 through the medium of our model, that the relative replicability of different labelling features can have significant effects on the evolution of conventions.

It is natural to ask whether styles with greater intrinsic replicability were chosen more frequently in later rounds than in earlier ones. (Recall that a corresponding effect is an implication of the model.) Our prior expectation was the twenty-game sequence of abstract games would allow such an effect to reveal itself. Intuitively, one might expect that the choice frequencies of the most intrinsically replicable abstract styles would show upward trends. In fact, there is no such effect. Table 3 reports the choice frequencies of the eight abstract styles for successive five-round blocks. There seem to be no obvious trends; contrary to expectation, the choice frequency for ‘most blue’ is slightly lower in later rounds than in earlier ones. We offer an explanation of this non-result in Appendix 2. In brief, the explanation is that, in a game with four labels, the degrees of intrinsic replicability exhibited by even the most replicable abstract styles in the experiment are too low to induce quantitatively significant upward trends in choice frequencies. Although this finding suggests the need for caution in extrapolating the results of the model, one must bear in mind that the abstract games *are abstract*: they are almost completely devoid of cues that might relate to subjects’ prior experiences. Further problems for players of these games arise because the colour-based similarity rules that can be applied to them are sometimes non-defined and sometimes conflict with one another. One should not assume that these similarity concepts are comparable in replicability with those that can be used in everyday life.

[Table 3 near here]

5. Conclusion

Theoretically and experimentally, we have shown how a simple form of experiential learning can lead to the emergence of conventions that are defined in terms of similarity relations between the ‘labels’ that individuals use to describe games to themselves, and that evolutionary game-theoretic analysis usually treats as irrelevant. We have also shown how the evolutionary selection of conventions can be influenced by the comparative replicability properties of different similarity relations. We do not claim that our stripped-down representation of experiential learning can explain all the ingredients of 'salience', understood as the defining characteristic of focal points in one-shot coordination games. But we suggest that it provides some clues about how conceptions of salience might begin to emerge even among low-rationality agents. In doing so, it answers some sceptical objections to the use of salience as an explanatory device in evolutionary game theory.

Table 1: Matching frequencies and coordination indices

	average matching frequency (MF) and coordination index (CI) for:			
	abstract games		culture-laden games	
	MF	CI	MF	CI
all games	0.392	0.284	0.369	0.280
culture-laden block 1 (fabrics)			0.356	0.316
culture-laden block 2 (fabrics)			0.407	0.263
culture-laden block 3 (paintings)			0.346	0.280
culture-laden block 4 (paintings)			0.366	0.263
games played in rounds 1–10	0.361	0.298	0.367	0.282
games played in rounds 11–20	0.423	0.274	0.370	0.281
games played in part 1	0.345	0.274	0.350	0.276
games played in part 2	0.446	0.297	0.384	0.290

Table 2: The use of similarity rules

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

				j chosen in *t*–1;
				match in *t* – 1;
				j defined in *t*
							j chosen in *t*–1;
							no match in *t* – 1;
							j defined in *t*			

similarity rule (*j*)		games with *j* defined	choice frequency of *j* (%)	total	*j* chosen	%	total	*j* chosen	%	χ^2
Culture-laden games										
Block 1:	style 1	5/5	19.7	24	5	20.8	65	12	18.5	0.06
	style 2	5/5	21.5	36	13	36.1	76	22	28.9	0.58
	style 3	5/5	44.7	106	61	57.5	103	50	48.5	1.70
	style 4	5/5	14.1	12	1	8.3	50	13	26.0	1.73
Block 2:	style 1	5/5	19.8	24	9	37.5	75	17	22.7	2.07
	style 2	5/5	33.9	84	78	92.9	68	31	45.6	41.39
	style 3	5/5	27.5	50	32	64.0	84	30	35.7	10.09
	style 4	5/5	18.8	24	12	50.0	63	18	28.6	3.53
Block 3:	style 1	5/5	20.3	38	10	26.3	65	19	29.2	0.10
	style 2	5/5	25.8	48	14	29.2	75	18	24.0	0.41
	style 3	5/5	21.7	22	7	31.8	71	18	25.3	0.36
	style 4	5/5	32.2	58	34	58.6	95	28	29.5	12.69
Block 4:	style 1	5/5	25.9	50	32	64.0	72	25	34.7	10.16
	style 2	5/5	27.3	38	10	26.3	87	23	26.4	0.00
	style 3	5/5	23.1	38	23	60.5	69	23	33.3	7.39
	style 4	5/5	23.7	30	9	30.0	88	31	35.2	0.27
total				682	350	51.3	1206	378	31.3	
Abstract games										
most pale		19/20	24.8	201	85	42.3	296	84	28.4	10.32
least pale		19/20	27.6	230	76	33.0	326	94	28.8	1.13
most red		19/20	23.5	162	54	33.3	303	86	28.4	1.23
most green		19/20	22.8	159	55	34.6	302	84	27.8	2.27
most blue		19/20	35.5	348	190	54.6	359	144	40.1	14.88
least red		18/20	32.0	260	99	38.1	313	95	30.4	3.79
least green		20/20	28.3	270	90	33.3	363	127	35.0	0.19
least blue		19/20	21.3	147	46	31.3	285	83	29.1	0.22
total				1777	695	39.1	2547	797	31.3	

Table 3: Evolution of choice frequencies in abstract games

	choice frequency (%) in rounds:				
style	1–5	6–10	11–15	16–20	all
most pale	25.1	25.2	23.5	25.2	24.8
least pale	25.8	28.5	28.7	28.1	27.6
most red	23.1	24.4	23.0	23.4	23.5
most green	22.3	22.4	22.3	25.0	22.8
most blue	37.6	37.3	33.2	33.2	35.5
least red	32.5	35.7	31.0	28.8	32.0
least green	27.5	30.2	27.1	28.0	28.3
least blue	20.5	22.3	22.5	20.2	21.3

Figure 1: Equilibrium

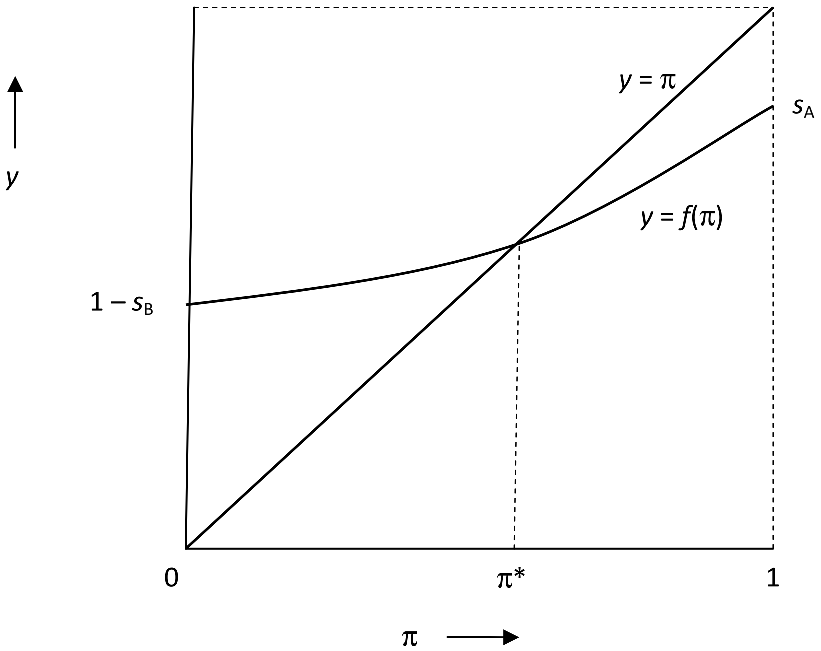


Figure 2: Two culture-laden games

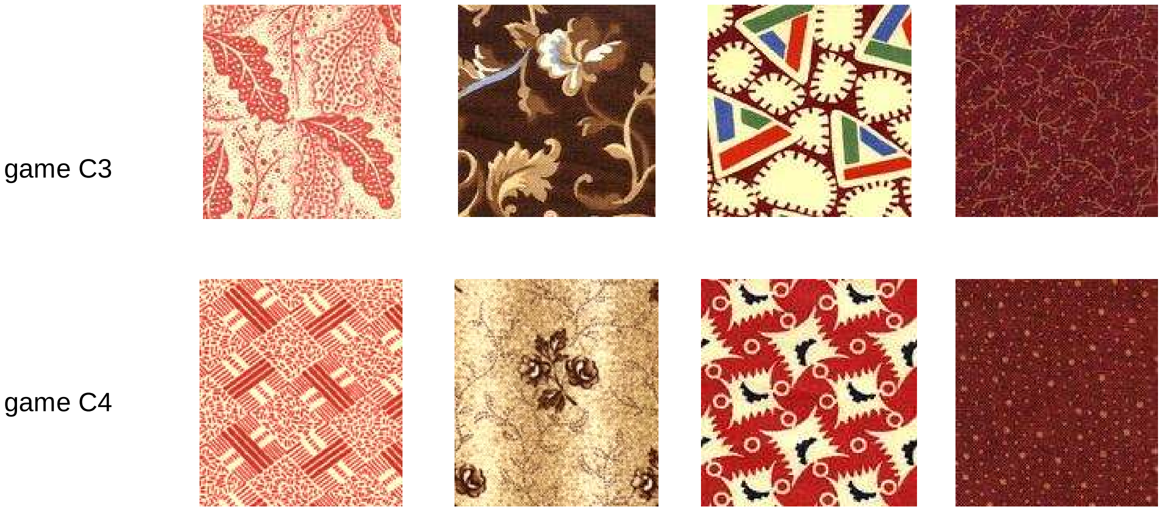


Figure 3: Two abstract games

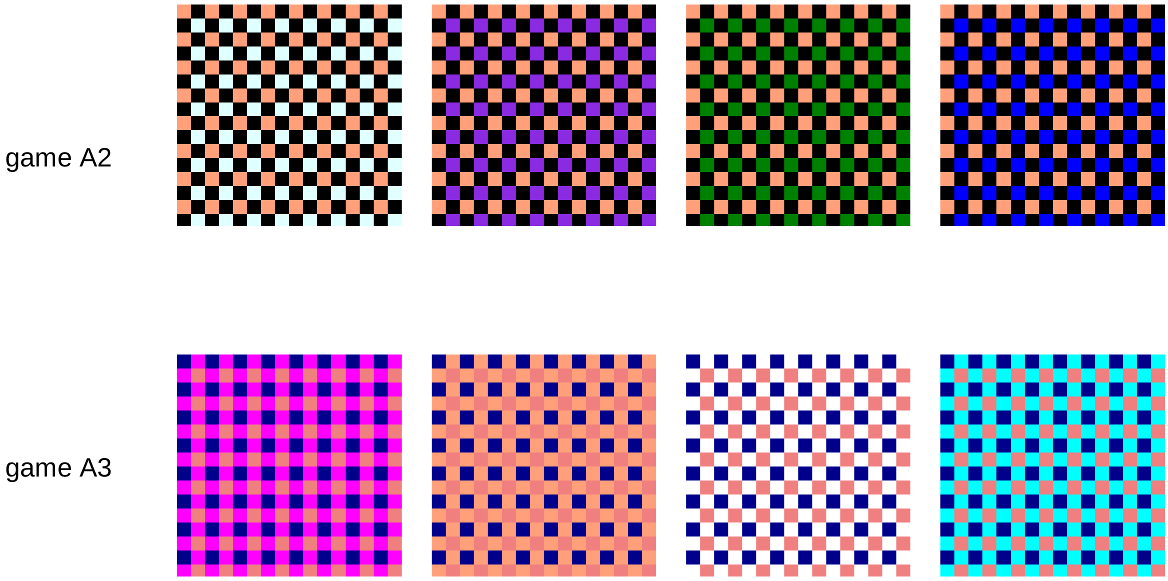
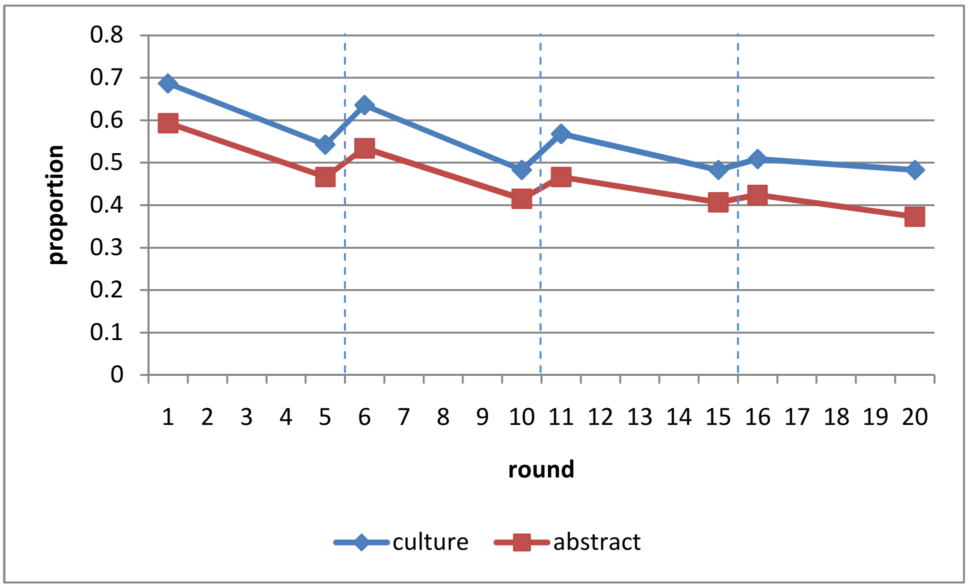


Figure 4: Frequency of choice of ‘most liked’ image



References

Alberti, Federica, Shaun Hargreaves Heap and Robert Sugden (2010). The emergence of salience: an experimental investigation. CBESS Discussion Paper 11-01, Centre for Behavioural and Experimental Social Science, University of East Anglia.

Bacharach, Michael (2006). *Beyond Individual Choice: Teams and Frames in Game Theory* (Natalie Gold and Robert Sugden, eds). Princeton, NJ: 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): 1–45.

Bacharach, Michael, and Dale O. Stahl (2000). Variable-frame level-*n* theory. *Games and Economic Behavior* 33(2): 220–46.

Bardsley, Nicholas, Judith Mehta, Chris Starmer, and Robert Sugden (2010). Explaining focal points: cognitive hierarchy theory *versus* team reasoning. *Economic Journal* 120 (March): 40–79.

Camerer, Colin F., Teck Ho and Kuan Chong (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics* 119(3): 861–98.

Casajus, André. (2001). *Focal Points in Framed Games: Breaking the Symmetry*. Berlin: Springer-Verlag.

Crawford, Vincent P., 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.

Cubitt, Robin P, and Robert Sugden (2003). Common knowledge, salience and convention: a reconstruction of David Lewis's game theory. *Economics and Philosophy* 19(2): 175–210.

Gauthier, David (1975). Coordination. *Dialogue* 14: 195–221.

Gilboa, Itzhak, and David Schmeidler (1995). Case-based decision theory. *Quarterly Journal of Economics* 110 (3): 605–639.

Goyal, Sanjeev and Maarten C.W. Janssen (1996). Can we rationally learn to coordinate? *Theory and Decision* 40: 29-49.

Hume, David (1739–40/ 1987). *A Treatise of Human Nature*. Oxford: Clarendon Press.

Janssen, Maarten C.W. (2001). Rationalising focal points. *Theory and Decision* 50(2): 119–48.

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

Mehta, Judith, Chris Starmer, and Robert Sugden (1994a). The nature of salience: an experimental investigation of pure coordination games. *American Economic Review* 84(3): 658–73.

Mehta, Judith, Chris Starmer, and Robert Sugden (1994b). Focal points in pure coordination games: an experimental investigation. *Theory and Decision* 36(2): 163–85.

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

Schlicht, Ekkehart (1988). *On Custom in the Economy*. Oxford: Oxford University Press.

Skyrms, Brian (1996). *Evolution of the Social Contract*. Cambridge: Cambridge University Press.

Stahl, Dale O. and Wilson, Paul W. (1995). On players’ models of other players. *Games and Economic Behavior* 10(1): 218–54.

Sugden, Robert (2004). *The Economics of Rights, Co-operation and Welfare* (second edition). Basingstoke: Palgrave Macmillan. First edition 1986.

Sugden, Robert (1995). A theory of focal points. *Economic Journal* 105 (May): 533–550.

Sugden, Robert (2010). Salience, inductive reasoning and the emergence of conventions. *Journal of Economic Behavior and Organization*, forthcoming.

Sugden, Robert and Ignacio Zamarrón (2006). Finding the key: The riddle of focal points. Journal of Economic Psychology 27: 609–621.

Appendix 1: Proofs of results

Proof of Result 4

Assume $r_A = r_B = r$ and $q_A \geq 0.5$.

To prove part (i), assume $r = 0$. Then $s_A = q_A$ and $s_B = 1 - q_A$. Thus, using (3a), $f(\pi) = q_A$ for all $\pi \in (0, 1)$. Since equilibrium is defined by $f(\pi^*) = \pi^*$, $\pi^* = q_A$.

To prove part (ii), assume $q_A = 0.5$. Then $s_A - s_B = q_B - q_A = 0$. Thus, by Result 3, $\pi^* = 0.5$.

To prove part (iii), assume $q_A > 0.5$. As $r \to 1$, s_A, $s_B \to 1$. Thus, using (3a):

$$\text{for all } \pi \in (0, 1)\text{: as } r \to 1,\ f(\pi) - \pi \to \pi^2 + 2\pi(1 - \pi)q_A - \pi. \qquad \text{(A1)}$$

But

$$\pi^2 + 2\pi(1 - \pi)q_A - \pi = \pi(2q_A - 1)(1 - \pi), \qquad \text{(A2)}$$

which (given $q_A > 0.5$) is positive for all $\pi \in (0, 1)$. Thus, in the limit as $r \to 1$, $f(\pi) > \pi$ at all $\pi < 1$, implying $\pi^* \to 1$.

To prove part (iv), it is sufficient to prove that if $r > 0$,then $q_A > 0.5$ implies $f(q_A) - q_A > 0$. From (3b):

$$f(q_A) - q_A = q_A^2[1 + s_A - s_B - 2q_A] + q_A[2q_A + 2s_B - 3] + [1 - s_B]. \qquad \text{(A4)}$$

Using (1) and (2) and rearranging:

$$f(q_A) - q_A = q_A\, r\,(2q_A - 1)(1 - q_A), \qquad \text{(A5)}$$

which is strictly positive if $q_A > 0.5$ and $r > 0$.

Proof of Result 5

Notice that if $\pi = 0.5$, then $v(D, t) = v(R, t) = 0.5$ for all t. Notice also that, because the two decision rules prescribe the same behaviour in period 1, $v(D, 1) = v(R, 1)$. Thus, given the symmetry between A and B, it is sufficient for a proof of Result 5 to show that if $r > 0$ and $\pi > 0.5$, then $v(R, t) > v(D, t)$ for all $t > 1$.

Assume $r > 0$ and $\pi > 0.5$. Consider any period $t > 1$. First suppose that *i did not* match with her co-player in period $t - 1$. Then, irrespective of whether she is using rule *D* or rule R, she chooses A in period t with probability q_A, and so her expected payoff in t is the same for both rules.

Now suppose instead that i *did* match with her co-player in $t-1$. If she is using rule D, she chooses A in t with probability q_A. But suppose she is using rule R. For periods $T = 1, ..., t$, let ρ_T denote the probability with which i chose (or chooses) A in T. First, we show (*Lemma* 1) that $\rho_{t-1} \geq q_A$ implies $\rho_t > q_A$.

Assume $\rho_{t-1} \geq q_A$. Conditional on i having matched in $t-1$, the probability that this match was on A is given by:

$$m_{t-1} \equiv \rho_{t-1}\,\pi / [\rho_{t-1}\,\pi + (1-\rho_{t-1})(1-\pi)]. \qquad \text{(A6)}$$

Thus:

$$m_{t-1} > q_A \iff \rho_{t-1}\,\pi / [\rho_{t-1}\,\pi + (1-\rho_{t-1})(1-\pi)] > q_A. \qquad \text{(A7)}$$

The right-hand side of (A6) is increasing in ρ_{t-1}. Thus, since $\rho_{t-1} \geq q_A$,

$$q_A\,\pi / [q_A\,\pi + (1-q_A)(1-\pi)] > q_A \;\Rightarrow\; m_{t-1} > q_A\,. \qquad \text{(A8)}$$

But $\pi > 0.5$ implies that the antecedent of (A8) holds, proving that $m_{t-1} > q_A$.

From the assumption that i uses rule R:

$$\rho_t = m_{t-1}\,[r + (1-r)q_A] + (1-m_{t-1})(1-r)q_A$$

$$= q_A + r(m_{t-1} - q_A). \qquad \text{(A9)}$$

Since $m_{t-1} > q_A$ and $r > 0$, (A9) implies $\rho_t > q_A$, proving Lemma 1.

Maintaining the assumptions that $t > 1$, that i matched in $t-1$, and that i is using rule R, exactly one of the following cases must hold:

Case 1: Either (i) $t = 2$, or (ii) $t > 2$ and i failed to match in $t-2$.

Case 2: Either (i) $t = 3$, or (ii) $t > 3$ and i matched in $t-2$, but failed to match in $t-3$.

Case 3: Either (i) $t = 4$, or (ii) $t > 4$ and i matched in $t-2$ and $t-3$, but failed to match in $t-4$.

Case 4 ...

In Case 1, $\rho_{t-1} = q_A$, and so $\rho_t > q_A$ by Lemma 1. In Case 2, $\rho_{t-2} = q_A$, and so $\rho_{t-1} > q_A$ by Lemma 1. Applying Lemma 1 again, $\rho_t > q_A$. And so on for Cases 3, 4, ... , establishing $\rho_t > q_A$ in all cases.

If $\pi > 0.5$, i's expected payoff in any given period is higher, the higher the probability with which she chooses A. If i uses rule D, she chooses A with probability q_A in all periods. We have established that if $\pi > 0.5$ and if i uses rule R, the probability with which she chooses A in any period $t > 1$ is strictly greater than q_A if she matched in $t - 1$, and is equal to q_A otherwise. Since $q_A, r \in (0, 0)$, the probability of matching when using rule R is non-zero in all periods. Thus $\pi > 0.5$ implies $v(R, t) > v(D, t)$ for all $t > 1$, completing the proof.

Appendix 2: The replication heuristic used by fixed pairs of co-players

In this appendix, we adapt the model presented Section 2 so that it applies to settings like that of the experiment, where a recurrently similar coordination game is played by fixed (rather than randomly re-matched) pairs of co-players.

Consider a large population of individuals, randomly assigned to pairs which then remain fixed. In each period $t = 1, 2, ...$, each pair plays a recurrently similar coordination game. Each game is defined by a set of n labels, where $n \geq 2$. In each game, there is one label of each of the types or *styles* $j = 1, ..., n$. For simplicity, we assume that all individuals are identical. For each style j, there is a default probability $q_j \in (0, 1)$ satisfying $\sum_j q_j = 1$ and an intrinsic replicability measure $r_j \in [0, 1)$. In period 1, choices are determined by default probabilities. In each period $t > 1$, players who have failed to match with their co-payers in period $t - 1$ act on default probabilities. If, in period $t - 1$, a player achieved a match on some style j, then in period t that style is chosen with probability $s_j \equiv r_j + (1 - r_j)q_j$; each style $k \neq j$ is chosen with probability $(1 - r_j)q_k$.

As an illustration, we consider a game where (as in the culture-laden games of the experiment) $n = 4$. To allow a simple analysis, we assume that only one of the styles, style 1, is recognised by the players. Thus, for each $k = 2, 3, 4$, we set $q_k = (1 - q_1)/3$ and $r_k = 0$, implying $s_k = q_k$. To simplify notation, we suppress the style subscript and use q, r and s to denote q_1, r_1 and s_1 respectively. Given the assumptions we have made, the only case in which default probabilities are *not* used is where, in some period $t > 1$, a pair of co-players matched *on style 1* in $t - 1$. For a given pair of co-players and a given period t, we can define two mutually exclusive and exhaustive *states*: either there is a match on style 1 (state M) or there is not (state N). The replication heuristic induces a Markov process in which the transition probabilities from M to M and from N to N are s^2 and $1 - q^2$ respectively.

Considering the whole population, let ρ_t be the probability that a randomly-selected pair of co-players is in state M in period t. (If the population is sufficiently large, ρ_t can also be interpreted as the proportion of pairs in state M in period t.) Then, for all t:

$$\rho_{t+1} = \rho_t s^2 + (1 - \rho_t)q^2. \tag{A10}$$

The stationary-state solution is defined by the condition $\rho_{t+1} = \rho_t = \rho^*$. Combining this condition with (A10):

$$\rho^* = q^2/(1 + q^2 - s^2). \tag{A11}$$

Let π_t be the probability with which a randomly-selected member of the population chooses style 1 in period t. (Equivalently, if the population is sufficiently large, π_t is the relative frequency with which style 1 is chosen in period t.) Consider a randomly-selected individual i in any period $t + 1$. By definition, the probability that i matched on style 1 in period t is ρ_t. Thus, the specification of the replication heuristic implies:

$$\pi_{t+1} = \rho_t s + (1 - \rho_t)q. \tag{A12}$$

Let π^* be the value of π_t in the stationary-state solution. Then, from (A12):

$$\pi^* = \rho^* s + (1 - \rho^*)q. \tag{A13}$$

Notice that if $r = 0$ (implying $s = q$), default probabilities are used at all times, with the trivial implication that, in every period t, $\pi_t = q$ and $\rho_t = q^2$. If $r_1 > 0$ (implying $s > q$), $\pi^* > q$ and $\rho^* > q^2$. If style 1 has non-zero intrinsic replicability, the frequency with which that style is chosen increases over time from the default value q, converging to the higher stationary-state value π^*. In relation to the findings of the experiment, it is illuminating to consider the implications of specific values of r for the size of this effect and the speed with which convergence takes place.

For illustrative purposes, we set $q = 0.35$, implying that individuals have a small default bias in favour of style 1, and consider how the choice frequency π_t evolves in the cases $r = 0.1$, $r = 0.5$, $r = 0.7$ and $r = 0.9$. For each of these cases, Table A1 shows the values of $\pi_1, ..., \pi_6$, and the stationary-state value π^*. When $r = 0.1$, the stationary-state choice frequency is only slightly higher than the default probability ($\pi^* = 0.358$), and convergence is extremely quick (to three significant figures, convergence is completed by period 2). Even with $r = 0.5$, the stationary-state frequency is fairly close to the default probability ($\pi^* = 0.410$), and convergence is almost complete after five periods. In contrast, with $r = 0.9$, the stationary-state choice frequency is well above the default probability ($\pi^* = 0.639$) and convergence occurs more slowly. (Even in this case, more than half of the gap between the default probability and the stationary-state frequency has been closed after four periods.)

[Table A1 near here]

One implication of these calculations is that, when the relevant coordination game has four or more labels (and corresponding styles), the intrinsic replicability of a style will have a significant effect on choice frequencies only if the degree of replicability is quite high. Intuitively, when there are many styles, the default probabilities for individual styles will typically be low. Thus, after any period in which there has not been a match, the probability of achieving a match on any given style will be low. Because the probability that default choices match on a given style is *the square of* the relevant default probability, increases in the number of labels have a disproportionately negative effect on transitions from non-matching to matching. Further: other things being equal, a given level of intrinsic replicability (r) is associated with lower gross replicability (s) when there are more styles, with the result that matching on a given style, if achieved, is less likely to be sustained.

If we are to relate these theoretical conclusions to the experimental results, we must consider what ranges of values of intrinsic replicability are plausible, given the data. Within our model, the gross replicability s_j of any style j is the same thing as that style's post-matching replication rate. Observed values of these rates are shown in column 5 of Table 2. For culture-laden games, the highest such rate is 0.929 for style 2 of block 2; six other styles have rates in the range from 0.500 to 0.640. For abstract games, the highest rates are 0.546 (most blue), 0.423 (most pale) and 0.381 (least red). By definition, $r_j \equiv (s_j - q_j)/(1 - q_j)$. On the neutral assumption that for each style $q_j = 0.25$, s_j values of 0.929, 0.640, 0.544 and 0.425

imply r_j values of 0.905, 0.520, 0.392 and 0.233 respectively.[12] These estimates suggest that intrinsic replicability might have significant effects on the evolution of choice frequencies for at least one, and possibly for several, culture-laden styles, but in the case of abstract styles, any effects would be relatively small and (because of the speed of convergence) would not be observed as persistent trends in choice frequencies.

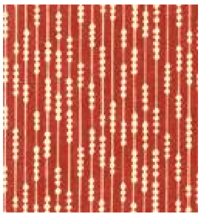
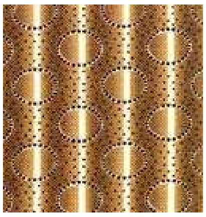
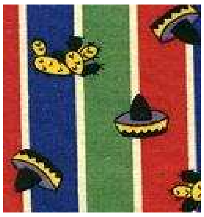
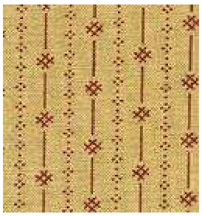
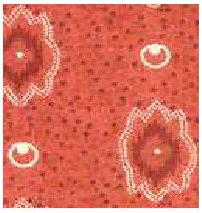
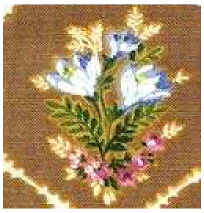
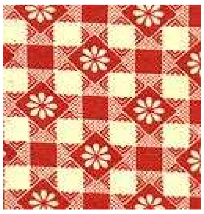
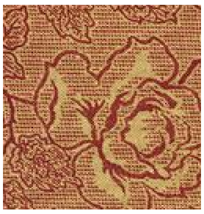
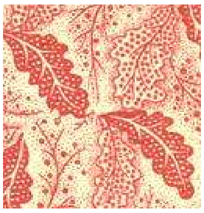
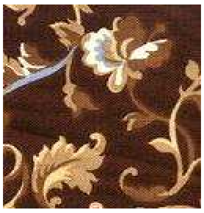
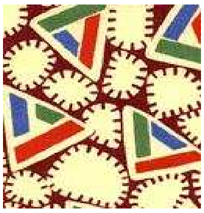
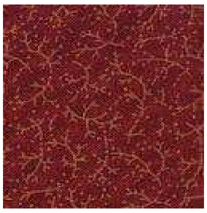
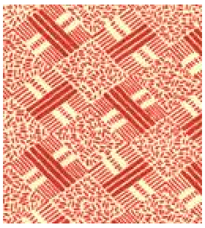
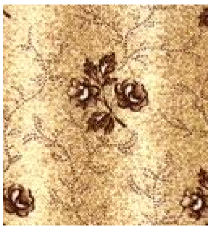
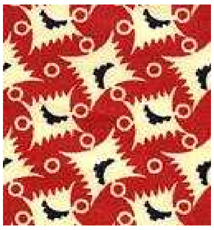
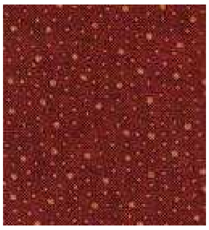
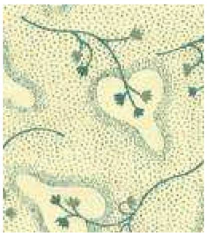
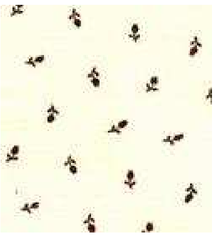
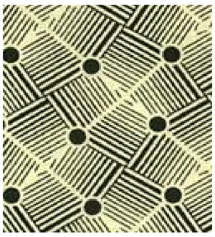
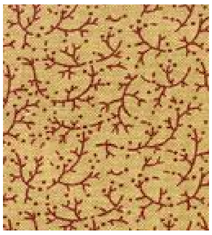
Table A1: Evolution of choice frequencies at different values of *r* (with *n* = 4 and *q* = 0.35)

	$r = 0.1$	$r = 0.5$	$r = 0.7$	$r = 0.9$
	(s = 0.415)	(s = 0.675)	(s = 0.805)	(s = 0.935)
π_1	0.350	0.350	0.350	0.350
π_2	0.358	0.390	0.406	0.422
π_3	0.358	0.403	0.435	0.476
π_4	0.358	0.407	0.450	0.516
π_5	0.358	0.409	0.459	0.546
π_6	0.358	0.409	0.463	0.569
π^*	0.358	0.410	0.467	0.639

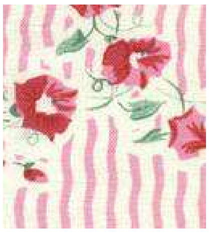
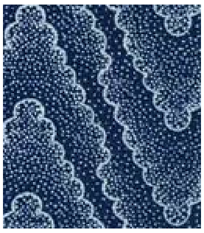
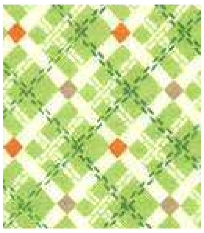
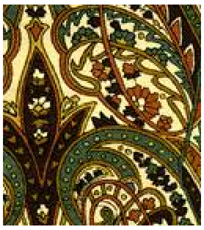
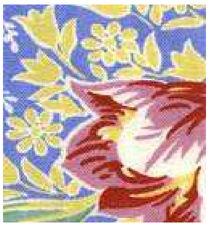
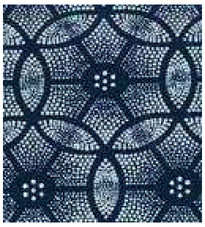
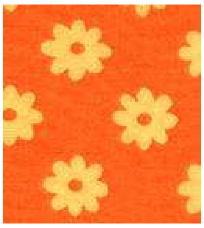
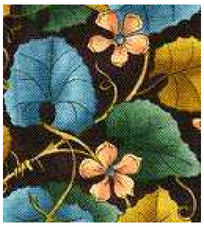
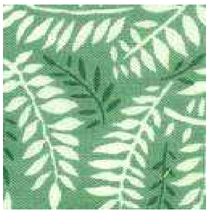
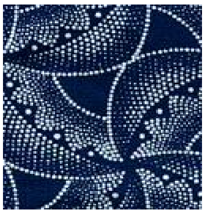
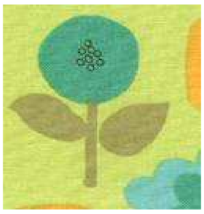
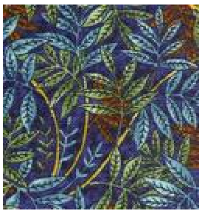
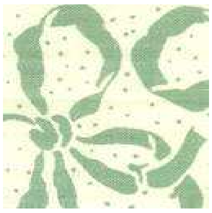
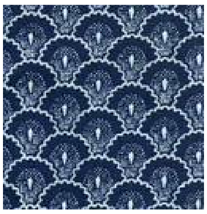
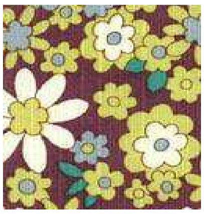
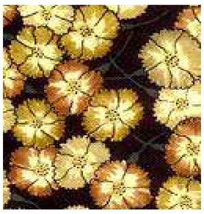
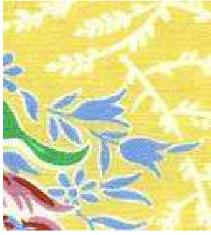
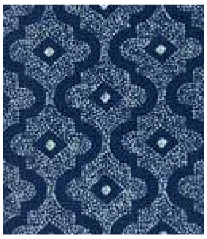
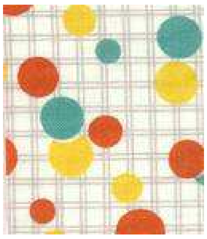
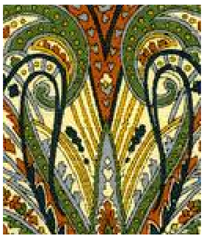
[12] Notice that higher assumed values of q_j imply lower estimates of r_j. To the extent that (as intuition perhaps suggests) intrinsic replicability tends to be positively associated with default probability, these values of r_i are over-estimates.

Appendix 3: Images used in culture-laden games [for online publication only]

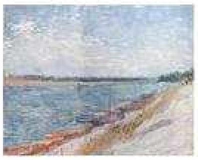
Block 1 (fabrics)

	Style 1	Style 2	Style 3	Style 4
Game C1				
Game C2				
Game C3				
Game C4				
Game C5				

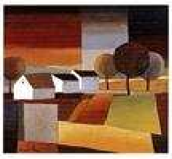
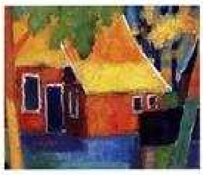
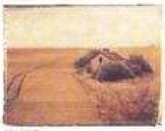
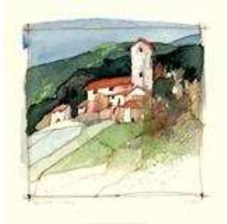
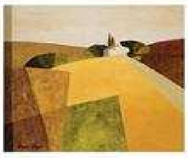
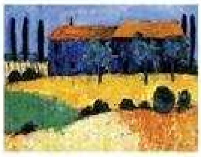
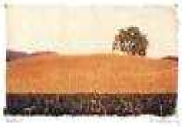
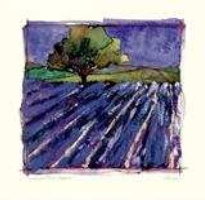
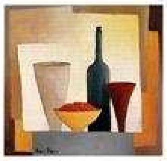
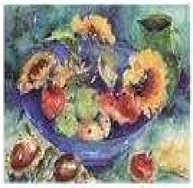
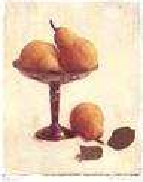
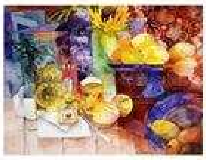
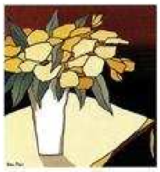
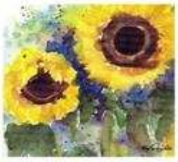
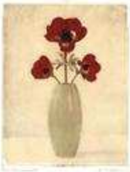
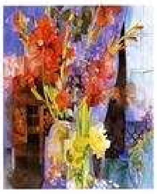
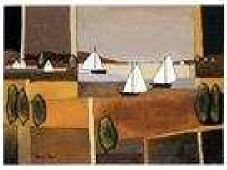
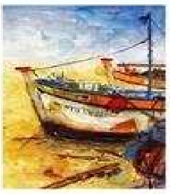
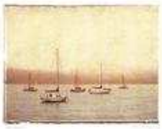
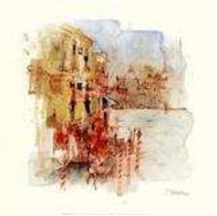
Block 2 (fabrics)

	Style 1	Style 2	Style 3	Style 4
Game C6				
Game C7				
Game C8				
Game C9				
Game C10				

Block 3 (paintings)

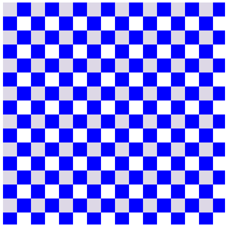
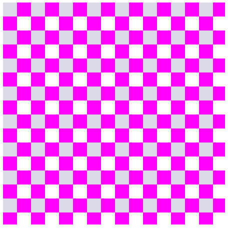
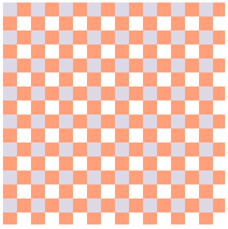
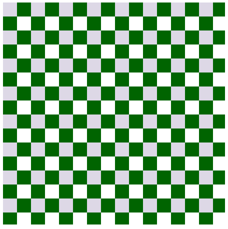
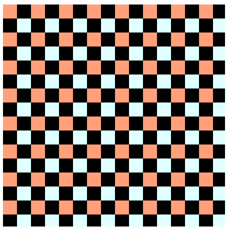
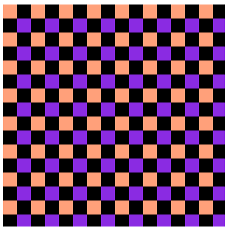
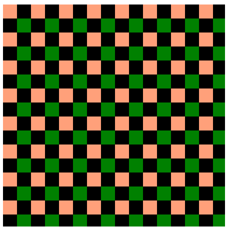
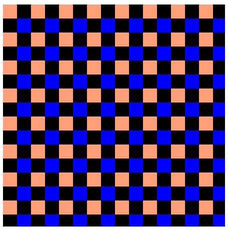
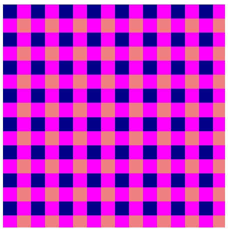
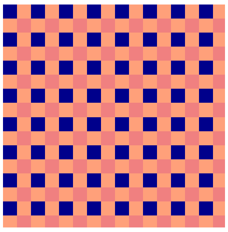
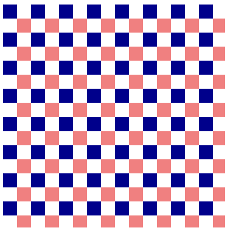
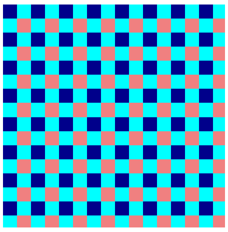
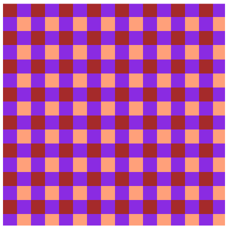
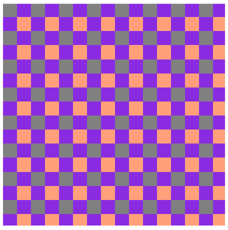
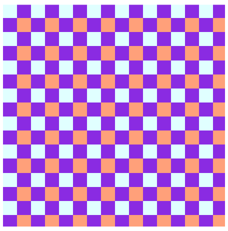
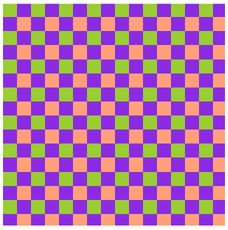
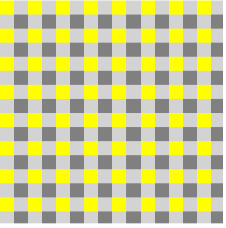
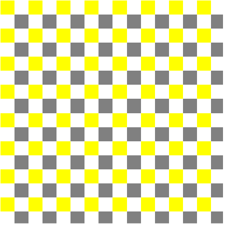
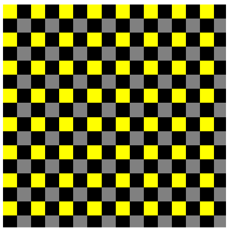
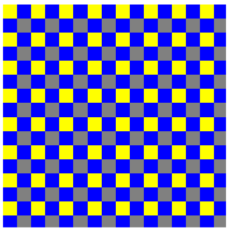
	Style 1	Style 2	Style 3	Style 4
Game C11				
Game C12				
Game C13				
Game C14				
Game C15				

Block 4 (paintings)

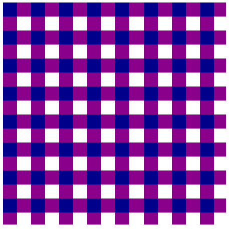
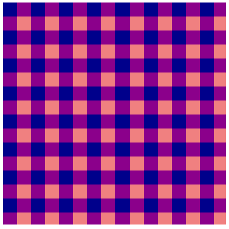
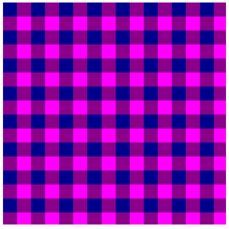
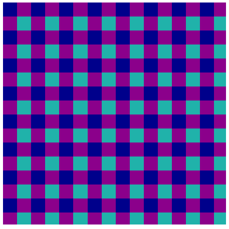
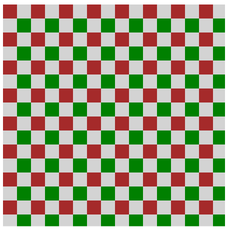
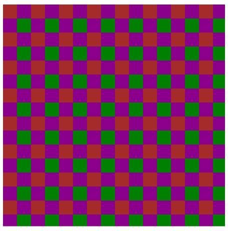
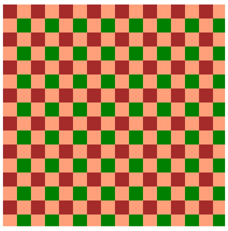
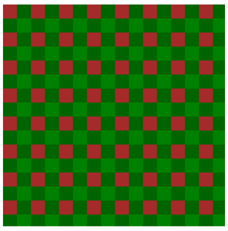
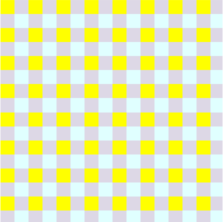
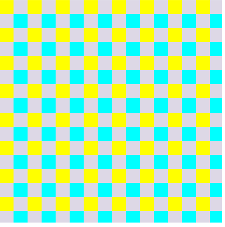
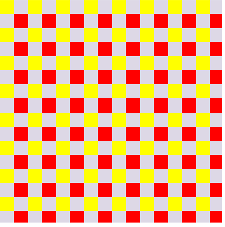
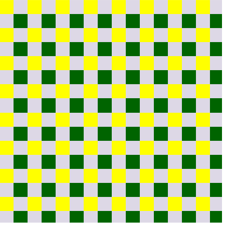
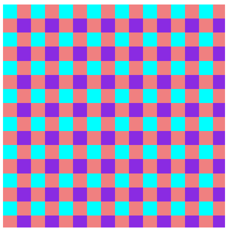
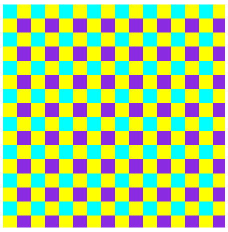
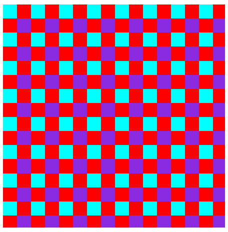
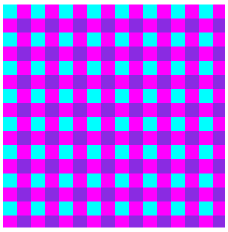
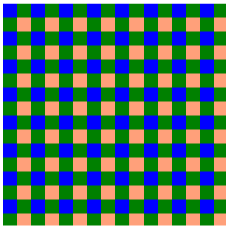
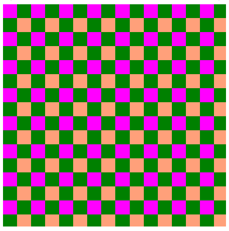
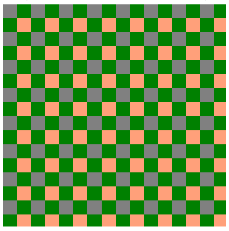
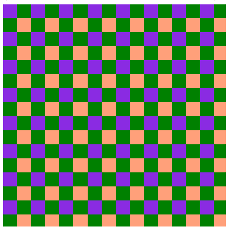
	Style 1	Style 2	Style 3	Style 4
Game C16				
Game C17				
Game C18				
Game C19				
Game C20				

Appendix 4: Images used in abstract games [for online publication only]

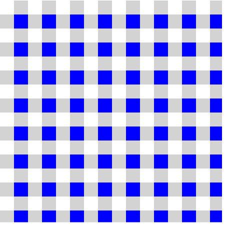
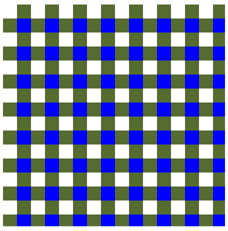
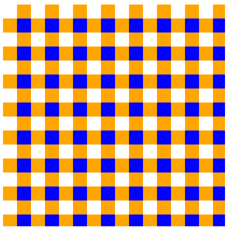
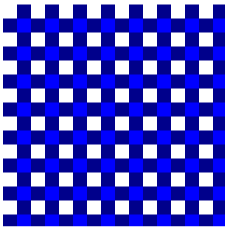
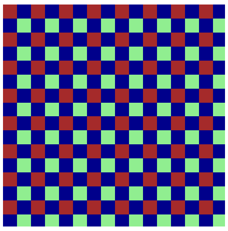
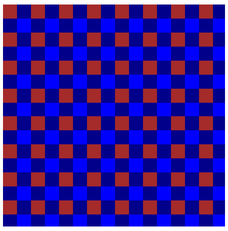
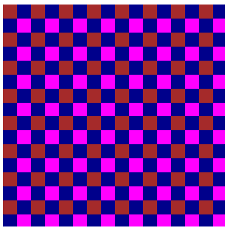
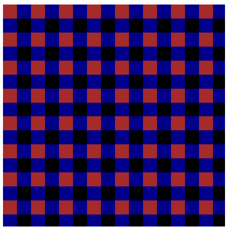
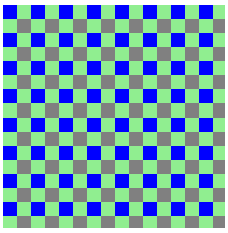
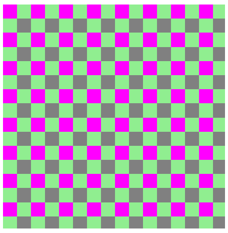
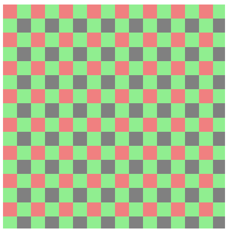
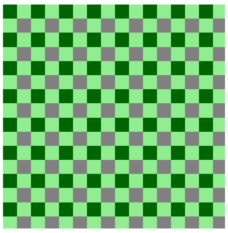
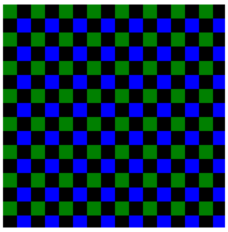
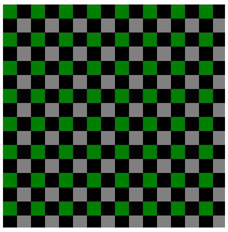
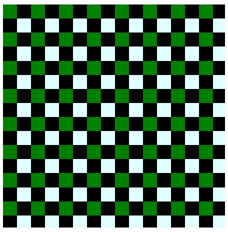
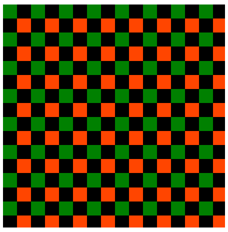
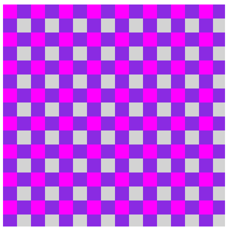
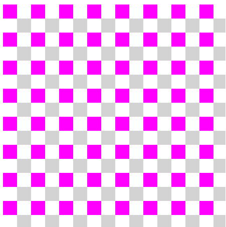
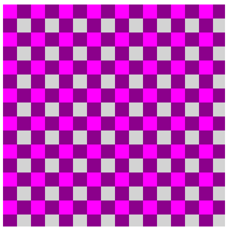
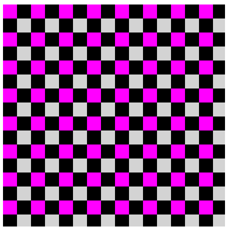
Block 1

	Image 1	Image 2	Image 3	Image 4
Game A1	 (0%, 0%, 100%)	 (100%, 0%, 100%)	 (100%, 63%, 48%)	 (0%, 39%, 0%)
Game A2	 (88%, 100%, 100%)	 (54%, 17%, 89%)	 (0%, 50%, 0%)	 (0%, 0%, 100%)
Game A3	 (100%, 0%, 100%)	 (100%, 63%, 48%)	 (100%, 100%, 100%)	 (0%, 100%, 100%)
Game A4	 (65%, 16%, 16%)	 (50%, 50%, 50%)	 (88%, 100%, 100%)	 (60%, 80%, 20%)
Game A5	 (83%, 83%, 83%)	 (100%, 100%, 100%)	 (0%, 0%, 0%)	 (0%, 0%, 100%)

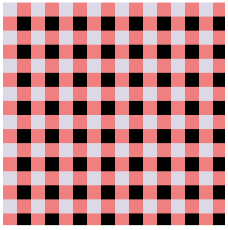
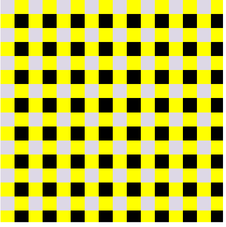
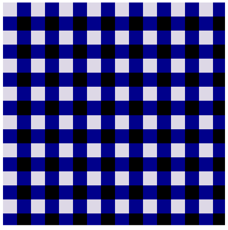
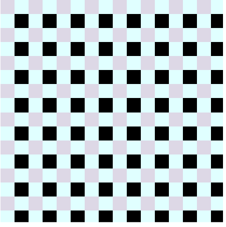
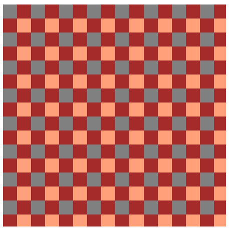
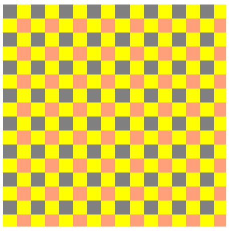
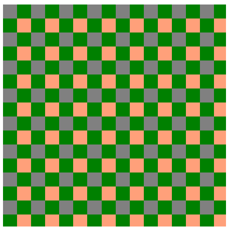
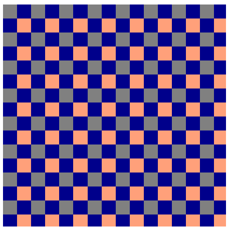
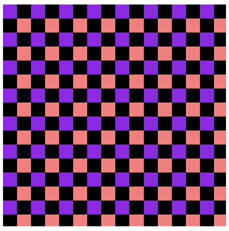
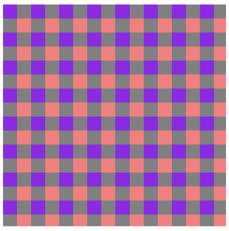
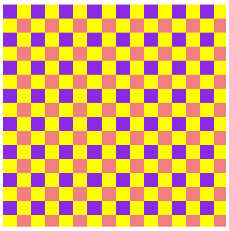
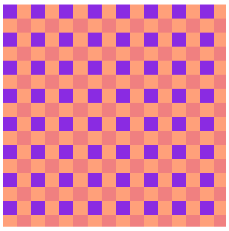
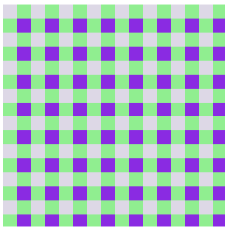
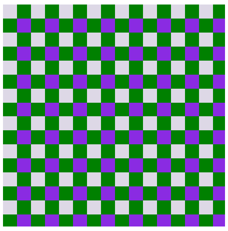
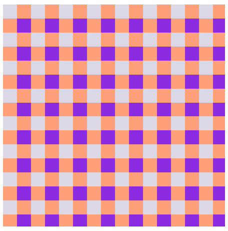
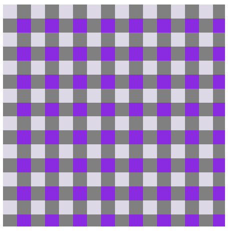
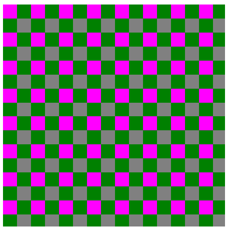
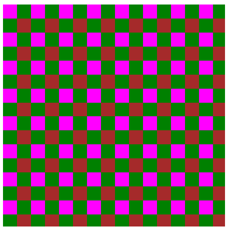
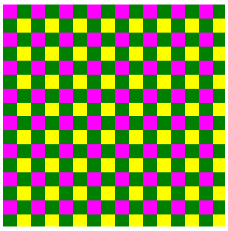
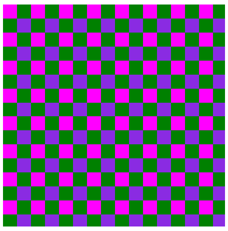
Block 2

	Image 1	Image 2	Image 3	Image 4
Game A6	 (100%, 100%, 100%)	 (94%, 50%, 50%)	 (100%, 0%, 100%)	 (13%, 70%, 67%)
Game A7	 (83%, 83%, 83%)	 (55%, 0%, 55%)	 (100%, 63%, 48%)	 (0%, 39%, 0%)
Game A8	 (88%, 100%, 100%)	 (0%, 100%, 100%)	 (100%, 0%, 0%)	 (0%, 39%, 0%)
Game A9	 (94%, 50%, 50%)	 (100%, 100%, 0%)	 (100%, 0%, 0%)	 (100%, 0%, 100%)
Game A10	 (0%, 0%, 100%)	 (100%, 0%, 100%)	 (50%, 50%, 50%)	 (54%, 17%, 89%)

Block 3

	Image 1	Image 2	Image 3	Image 4
Game A11	 (83%, 83%, 83%)	 (33%, 42%, 18%)	 (100%, 65%, 0%)	 (0%, 0%, 55%)
Game A12	 (56%, 93%, 56%)	 (0%, 0%, 100%)	 (100%, 0%, 100%)	 (0%, 0%, 0%)
Game A13	 (0%, 0%, 100%)	 (100%, 0%, 100%)	 (94%, 50%, 50%)	 (0%, 39%, 0%)
Game A14	 (0%, 0%, 100%)	 (50%, 50%, 50%)	 (88%, 100%, 100%)	 (100%, 27%, 0%)
Game A15	 (54%, 17%, 89%)	 (100%, 100%, 100%)	 (55%, 0%, 55%)	 (0%, 0%, 0%)

Block 4

	Image 1	Image 2	Image 3	Image 4
Game A16	 (94%, 50%, 50%)	 (100%, 100%, 0%)	 (0%, 0%, 55%)	 (88%, 100%, 100%)
Game A17	 (65%, 16%, 16%)	 (100%, 100%, 0%)	 (0%, 50%, 0%)	 (0%, 0%, 55%)
Game A18	 (0%, 0%, 0%)	 (50%, 50%, 50%)	 (100%, 100%, 0%)	 (100%, 63%, 48%)
Game A19	 (56%, 93%, 56%)	 (0%, 50%, 0%)	 (100%, 63%, 48%)	 (50%, 50%, 50%)
Game A20	 (50%, 50%, 50%)	 (65%, 16%, 16%)	 (100%, 100%, 0%)	 (54%, 17%, 89%)