Cognitive Psychology 61 (2010) 201–227

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Learning to cooperate without awareness in multiplayer minimal social situations

Andrew M. Colman [a,*], Briony D. Pulford [a], David Omtzigt [a], Ali al-Nowaihi [b]

[a] School of Psychology, University of Leicester, Leicester LE1 7RH, United Kingdom
[b] Department of Economics, University of Leicester, Leicester LE1 7RH, United Kingdom

ARTICLE INFO

Article history:
Accepted 20 May 2010

Keywords:
Adaptive learning
Cooperation
Decisions from experience
Explorative sampler
Game theory
Minimal social situation
Monte Carlo simulation
Reinforcement learning
Value-assessment model

ABSTRACT

Experimental and Monte Carlo methods were used to test theoretical predictions about adaptive learning of cooperative responses without awareness in minimal social situations—games in which the payoffs to players depend not on their own actions but exclusively on the actions of other group members. In Experiment 1, learning occurred slowly over 200 rounds in a dyadic minimal social situation but not in multiplayer groups. In Experiments 2–4, learning occurred rarely in multiplayer groups, even when players were informed that they were interacting strategically and were allowed to communicate with one another but were not aware of the game's payoff structure. Monte Carlo simulation suggested that players approach minimal social situations using a noisy version of the win–stay, lose–shift decision rule, deviating from the deterministic rule less frequently after rewarding than unrewarding rounds.

1. Introduction

In a certain class of repeated dyadic games, pairs of players can learn to cooperate without any deliberate intention to do so and, more remarkably still, without awareness of their strategic interdependence or even of each other's existence. Sidowski discovered this phenomenon during his doctoral research into the effects of reward and punishment in games (Sidowski, 1957; Sidowski, Wyckoff, & Tabory, 1956). Since then, a small number of researchers have continued to study adaptive learning under conditions of profound ignorance that Sidowski et al. named the *minimal social situation* or

* Corresponding author. Address: School of Psychology, University of Leicester, Leicester LE1 7RH, United Kingdom. Fax: +44 (0) 116 229 7196.
E-mail address: amc@le.ac.uk (A.M. Colman).

0010-0285/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.cogpsych.2010.05.003

		II: *C*	II: *D*
I	*C*	+, +	–, +
I	*D*	+, –	–, –

Fig. 1. Mutual Fate Control game.

MSS. A slender thread of MSS research has stretched from the 1950s into the second millennium (Colman, 2005; Delepoulle, Preux, & Darcheville, 2000, 2001; Mitropoulos, 2001, 2003).

The type of game most commonly used to study the MSS, named *Mutual Fate Control* by Thibaut and Kelley (1959, chap. 7), has the payoff matrix displayed in Fig. 1. According to standard game-theoretic notation, Player I chooses a row labeled *C* (cooperate) or *D* (defect), Player II independently chooses a column *C* or *D*, and the pair of symbols in each cell are the payoffs to Player I and Player II in that order for the corresponding outcome of the game. It is unnecessary at this stage to assign numerical utilities to the payoffs; we assume merely that every possible outcome yields either a positive (+) or a negative (−) payoff to each player.

If both players simultaneously choose the cooperative strategy *C*, then the outcome is shown in the upper-left cell, where it is evident that both players are rewarded with positive payoffs. If both simultaneously choose *D*, then both suffer negative payoffs. If one chooses *C* while the other chooses *D*, then the player who cooperates receives a negative payoff and the one who defects receives a positive payoff. In the MSS, the game is played under severely *restricted information* insofar as both players know their own strategy set $\{C, D\}$ but are ignorant of the co-player's strategy set and of both players' payoff functions—the payoffs that result from every possible pair of strategy choices, represented by the pairs of symbols in the cells of the payoff matrix. A distinction needs to be made between the *strict* MSS, in which both players know that they are involved in a decision-making task but are oblivious of the fact that they are playing a game, and consequently that their payoffs are affected by the decisions of another player, and the *informed* MSS, in which they know that they are playing an interactive game, although they do not know the payoff functions. In experimental MSS research, the game is usually repeated (iterated) many times to allow adaptive learning to occur.

The Mutual Fate Control game can be interpreted as a degenerate Prisoner's Dilemma game, with the level of measurement in the payoff functions reduced from an interval to a binary scale (Sozański, 1992). This is immediately obvious from inspection of Fig. 2. Panel (a) shows a standard Prisoner's Dilemma game. In (b), the original payoffs are replaced by ordinal numbers 0, 1, 2, and 3, with 0 replacing the smallest number, 1 replacing the second-smallest, and so on. In (c), the two smallest numbers are replaced by 0 and the two largest numbers by 1. The change in measurement scale makes a big difference: the strategic properties of the Mutual Fate Control game differ from those of the Prisoner's Dilemma game in the following ways.

The Prisoner's Dilemma game (Fig. 2a) has a unique *Nash equilibrium* at (D, D), representing joint defection. A Nash equilibrium is an outcome in which each player's strategy is a *best reply* to that of the co-player, inasmuch as it yields the highest available payoff to the player choosing it, given the co-player's chosen strategy. Against a co-player who chooses *D*, the *D* strategy yields the highest

(a)

		II: *C*	II: *D*
I	*C*	3, 3	0, 5
I	*D*	5, 0	1, 1

(b)

		II: *C*	II: *D*
I	*C*	2, 2	0, 3
I	*D*	3, 0	1, 1

(c)

		II: *C*	II: *D*
I	*C*	1, 1	0, 1
I	*D*	1, 0	0, 0

Fig. 2. The Mutual Fate Control game as a degenerate Prisoner's Dilemma game. (a) Standard Prisoner's Dilemma game. (b) Prisoner's Dilemma game with ordinal-scale payoffs. (c) Mutual Fate Control game corresponding to Prisoner's Dilemma game with binary-scale payoffs.

payoff (it yields 1, whereas a *C* choice yields 0), hence at (*D*, *D*) the players' strategies are in Nash equilibrium, and this property is lacking at each of the other three possible outcomes of the game. Furthermore, the *D* strategies are *strictly dominant* for both players, inasmuch as *D* yields a strictly higher payoff than *C*, irrespective of the strategy chosen by the co-player—a payoff of 5 rather than 3 if the co-player chooses *C*, and of 1 rather than 0 if the co-player chooses *D*. It follows that *D* is the unconditionally best, uniquely rational strategy in the Prisoner's Dilemma game. In the Mutual Fate Control game (Fig. 2c), on the other hand, all four outcomes (*C*, *C*), (*C*, *D*), (*D*, *C*), (*D*, *D*) are weak Nash equilibria, each player's strategy yielding a payoff no better but no worse than any other, given the co-player's strategy, and neither strategy is strictly dominant. The defining property of the Mutual Fate Control game, as its name implies, is that each player's payoff is determined exclusively by the co-player's choice—a strategy choice has no effect on the payoff of the player choosing it. This property is not shared by the Prisoner's Dilemma game, in which payoffs are determined jointly by the choices of both players, as is normally the case in strategic games.

This analysis relies on key concepts of game theory that originated from assumptions about human rationality. However, when games are repeated many times, primitive adaptive processes can mimic rational choice, and the key game-theoretic concepts turn out to have relevance even if the players are unaware that they are playing a game. This was first pointed out by Nash (1950): "It is unnecessary to assume that the participants have full knowledge of the total structure of the game, or the ability and inclination to go through any complex reasoning process" (p. 21). Young (1998) later explained: "Evolutionary forces often *substitute* for high (and implausible) rationality when the adaptive process has enough time to unfold" (p. 5, italics in original).

Mutual Fate Control can occur naturally in everyday strategic interactions. Kelley and Thibaut (1978, pp. 5–13) discussed at some length an example taken from Tolstoy's novella, *Family Happiness*, and Colman (1995, pp. 40–50) discussed a lifelike example of a strict MSS involving data analysts logging into a central computer from different locations and potentially disrupting each other's work. Here is a more homely example. Every morning, Alf chooses whether to give his son raisins or cheese sticks to snack on during the day. Similarly, Beth chooses between popcorn or peanuts for her daughter's snack. The children are friends and always share their snacks with each other at school, although their parents know nothing about this. Alf's son is allergic to peanuts and gets ill if he eats any of his friend's peanuts, and Beth's daughter is allergic to cheese and gets ill if she eats any of her friend's cheese sticks. The upshot is that although each parent's snack choice has no effect whatsoever on his or her own child's wellbeing, in each case one option leaves the other parent's child well and its parent happy, whereas the alternative option makes the other child ill and upsets its parent. The choices of Alf and Beth govern each other's fates, and if we consider each day as a single round of the game, then this is evidently a strict MSS, repeated daily.

Although games played without knowledge of the relevant payoff functions are common in everyday life, it may seem unlikely that players could learn to cooperate in any repeated game in which they were ignorant not only of the nature but also of the fact of their interdependence. However, experimental evidence suggests that adaptive learning, resulting in increasing cooperation, can occur even in the strict MSS, through a process that Kelley (1968) called *interpersonal accommodation*, without any conscious intention to cooperate or awareness of the opportunity for cooperation.

The earliest MSS experiments (Sidowski, 1957; Sidowski et al., 1956) used a methodology and incentive scheme that would be difficult to navigate through a present-day research ethics committee or institutional review board. Pairs of players were seated in separate rooms, unaware of each other's existence, and electrodes capable of delivering painful shocks were attached to their non-preferred hands. Each player was provided with a pair of buttons for choosing strategies and a digital display showing the cumulative number of points scored. The game was repeated over many rounds, with both players pursuing the goal of maximizing rewards (points) and minimizing punishments (shocks). Rewards and punishments were arranged according to the Mutual Fate Control payoff matrix shown in Fig. 1, but the players were not shown the matrix and were ignorant of the dyadic linkage and general setup of the experiment. For half the players, pressing the left-hand button (labeled *C* in Fig. 1) caused the co-player to be rewarded with points, and pressing the right-hand button (*D*) caused the co-player to be punished with an electric shock, and for the rest of the players, the functions of the left and right-hand buttons were reversed.

Sidowski (1957) ran some of his participants in an informed MSS treatment condition in which they were told that "there is another *S* in another room who controls the number of shocks and the number of scores that you will receive. You in turn control the number of shocks and scores which the other *S* receives" (p. 320). In this condition, although the participants were not informed about the precise nature of their interdependence (the payoff functions), they knew that they were making interactive decisions. Interestingly, this additional information produced no increase in the relative frequency of cooperative choices (p. 324), although in a later experiment (Kelley, Thibaut, Radloff, & Mundy, 1962), it did produce a small increase.

Immediately following Sidowski's earliest experiments, researchers dispensed with electric shocks. Most subsequent experiments involved incentive schemes based merely on positive and negative points, and the participants were instructed to try to maximize their cumulative (net) point totals (e.g., Crawford & Sidowski, 1964; Kelley et al., 1962; Rabinowitz, Kelley, & Rosenblatt, 1966; Sidowski & Smith, 1961). The points were not converted into money or any other tangible rewards. As in the earlier experiments, pairs of players generally learned over repeated rounds to coordinate on the efficient (C, C) outcome, even in strict MSS conditions in which they were unaware of their strategic interdependence. A typical finding was that the proportion of *C* choices reached approximately 75% after 200 rounds, with a small number of pairs settling down at some stage to choosing *C* on every round. In other words, some players behaved as if they were learning to cooperate, although the situation was, from their point of view, non-interactive, and they did not know—and did not guess, as early experimenters confirmed through postexperimental interviews—that they were linked to co-players in interactive decisions that afforded opportunities for cooperation. It seems reasonable, therefore, to describe this behavior as adaptive learning to cooperate without awareness of strategic interdependence. This is not inconsistent with evidence that awareness of stimulus contingencies is a necessary condition for human conditioning to occur (Lovibond & Shanks, 2002), because a decision maker can be aware of the stimulus contingencies without becoming aware of the involvement of a strategically interdependent co-player.

Following these early experiments, further investigations of the MSS were published, using human and occasionally animal players, and the findings were broadly similar (Arickx & Van Avermaet, 1981; Bertilson & Lien, 1982; Bertilson, Wonderlich, & Blum, 1983, 1984; Boren, 1966; Delepoulle et al., 2000, 2001; Mitropoulos, 2001, 2003; Molm, 1981). For example, in an informed MSS experiment reported by Mitropoulos (2003), the relative frequency of cooperative choices reached 66% after only 100 rounds.

1.1. MSS theory

How can adaptive learning without awareness be explained, especially in the strict MSS? The most persuasive suggestion was put forward by Thibaut and Kelley (1959), who hypothesized that MSS players tend to adopt a myopic *win–stay, lose–shift* (WSLS) decision rule. This is a decision rule, applicable to any repeated binary decision task, according to which a decision maker chooses arbitrarily on the first round and then repeats on every subsequent Round t any choice that was followed by a positive payoff on Round $t-1$ and switches to the alternative option on Round t after receiving a negative payoff on Round $t-1$. Rapoport and Chammah (1965, pp. 73–74) rediscovered WSLS in a study of the repeated Prisoner's Dilemma game and rather unkindly called it *Simpleton*. The mathematicians Kraines and Kraines (1989) rediscovered it yet again in their research into evolutionary games and called it *Pavlov* because of its reflex-like character, and this term has persisted in the burgeoning literature of evolutionary games. To a psychologist, WSLS is essentially a formalization of Thorndike's (1898, 1911, 1927) *law of effect*:

> Of several responses made to the same situation, those which are accompanied or closely followed by satisfaction to the animal will, other things being equal, be more firmly connected with the situation, so that, when it recurs, they will be more likely to recur; those which are accompanied or closely followed by discomfort to the animal will, other things being equal, have their connections with that situation weakened, so that, when it recurs, they will be less likely to occur. The greater

the satisfaction or discomfort, the greater the strengthening or weakening of the bond. (Thorndike, 1911, p. 244).

This is a fundamental behavioral law (Bonem & Crossman, 1988; Dennett, 1975; Ferster & Skinner, 1957; Skinner, 1969; Teigen, 2002), in spite of certain controversies that have attended it (Bitterman & Schoel, 1970; Herrnstein, 1970), and learning processes approximating it occur pervasively in nature (Bitterman, 1975; Nowak & Sigmund, 1993).

Camerer (2003, chap. 6) reviewed a wide range of approaches to learning in games, including reinforcement learning, belief-learning models, experience-weighted attraction (EWA) models, anticipatory learning and sophistication, imitation learning, learning direction theory, and rule learning theories. Apart from reinforcement learning, these approaches are inapplicable to the MSS because of the impoverished information that is available to MSS players. Simple reinforcement learning theories (e.g., Barron & Erev, 2003; Erev & Roth, 1998; Feltovich, 2000; Roth & Erev, 1995), on the other hand, are closely related to WSLS, and the most relevant forms of reinforcement learning will be revisited in Section 6.4.

We formalize the WSLS theory as follows. We represent the outcomes of successive rounds of the MSS by a sequence of ordered pairs, the elements of each pair corresponding to the strategy choices of Players I and II respectively. We assume that the game is repeated an indefinite number of times, that the players' initial choices are arbitrary, and that on all subsequent rounds they implement WSLS deterministically.

What happens if the initial, arbitrary choices of both players are C (cooperate)? Each player rewards the co-player, and because both payoffs are positive, neither player ever switches to the alternative strategy:

$$(C, C) \to (C, C) \to (C, C) \to \cdots.$$

What happens if both players initially choose D (defect)? In this case, both receive negative payoffs on the first round, causing them to shift to C on the second round, after which (as shown above) they repeat these C choices on all subsequent rounds:

$$(D, D) \to (C, C) \to (C, C) \to \cdots.$$

Last, what happens if one player initially chooses C and the other D? In that case, the C-chooser receives a negative payoff on the first round and therefore shifts to D on the second round, and the D-chooser receives a positive payoff on the first round and therefore stays with D. On the second round, both players accordingly choose D, followed (as shown above) by C on all subsequent rounds:

$$(C, D) \to (D, D) \to (C, C) \to (C, C) \to \cdots,$$
$$(D, C) \to (D, D) \to (C, C) \to (C, C) \to \cdots.$$

It is clear from this analysis that players who use WSLS learn to cooperate—to choose mutually rewarding strategies—by the third round at the latest, and continue cooperating indefinitely after that. However, this is not what actually happens in MSS experiments, in which joint cooperation develops only very slowly over many repetitions and never attains 100%, with all pairs choosing C on every round. This implies that players do not adhere strictly to WSLS or the law of effect. For example, in informed MSS experiments, Mitropoulos (2001, 2003) found little direct evidence for strict WSLS and considerable inertia in choice behavior, even in a treatment condition in which the players were told that the payoff scheme was such that each player's choice completely determined the payoff of the co-player (Mitropoulos, 2003, p. 29). Nevertheless, there are at least three reasons for assuming that players probably use a nondeterministic, noisy WSLS-like rule in the MSS. First, the slow adaptive learning that has been observed by many independent experimenters is difficult to explain without assuming that decision makers use some such decision rule, especially under strict MSS information conditions, because it is hard to imagine how else they could solve the problem at the behavioral level, as many of them do, at least to a degree. Second, empirical evidence has been found for reinforcement learning in an almost unlimited range of learning environments, and there is no obvious reason to suppose that the MSS is an exception. Third, Edgeworth's (1881) famous dictum, "The first principle of Economics is that every agent is actuated only by self-interest" (p. 16), is equally applicable to the

MSS as to any other decision task and, in the MSS, it seems to mandate WSLS or some functionally equivalent decision rule, because it is the natural way for decision makers to realize their aspirations (Posch, Pichler, & Sigmund, 1999).

1.2. Multiplayer MSS theory

The *multiplayer minimal social situation* is a formal generalization of the MSS to an arbitrary number of players, modeled as a Galois field of integers modulo 2 (Coleman, Colman, & Thomas, 1990). The set of $n \geqslant 2$ players, each with a uniquely designated predecessor and successor, is represented by a cyclic graph of valency 2. The easiest way to visualize this structure is to imagine n players, labeled $1, 2, \ldots, n$, seated round a table, making decisions sequentially, in anticlockwise order. Each player is then flanked by a predecessor on her left and a successor on her right. Player l's predecessor is Player n and Player n's successor is Player 1. Each player has a choice of two strategies, C and D, hence the choices of the n players on each round can be represented by an n-vector of Cs and Ds called a *configuration*. Whenever a player chooses C, that player's successor receives a positive payoff, and whenever a player chooses D, that player's successor receives a negative payoff.

According to WSLS, any player who receives a rewarding payoff will repeat the same strategy on the following round, and one who receives an unrewarding payoff will shift strategies on the following round. Under WSLS, every configuration is therefore followed by a uniquely specified configuration on the following round. It is obvious that a jointly cooperative configuration consisting exclusively of C choices will be repeated on all subsequent rounds. The two-player MSS is a special (dyadic) case of the multiplayer MSS, and the analysis presented earlier shows that all initial MSS configurations lead to joint cooperation by the third round at the latest. Coleman et al. (1990) proved that this result generalizes to groups of size 2, 4, 8, 16, and so on—groups of size 2^n with $n = 1, 2, 3, \ldots$. They also proved that in odd-sized groups, no adaptive learning takes place under deterministic WSLS, so that joint cooperation occurs only when all players happen to make the same arbitrary choice (all C or all D) on the first round, the probability of this being $(1/2)^{n-1}$ and decreasing rapidly with the number n of players: the probability is 1/4 in a three-player group, 1/16 in a five-player group, 1/64 in a seven-player group, and so on. In groups that are even-sized but not powers of 2 (group sizes 6, 10, and so on), adaptive learning occurs from some initial configurations but not others. If n is even, and if k is the highest power of 2 that divides n evenly, then the configurations that lead to joint cooperation are $(x_1, \ldots, x_n)$ such that $x_i = x_{i-k}$ for all i (mod n). For example, in a six-player group, it is easy to verify that the configuration (D, C, D, C, D, C) leads quickly to joint cooperation under WSLS, because it is followed by (D, D, D, D, D, D), and then by (C, C, C, C, C, C), repeated indefinitely. But the six-player configuration (D, D, C, C, D, D) generates the following sequence: $(D, D, C, C, D, D) \rightarrow (C, C, D, C, D, C) \rightarrow (C, C, D, D, D, D) \rightarrow (D, C, D, C, C, C) \rightarrow (D, D, D, D, C, C) \rightarrow (D, C, C, C, D, C) \rightarrow (D, D, C, C, D, D)$, returning to the starting point after six rounds and therefore cycling forever without ever reaching (C, C, C, C, C, C). Of the 64 possible configurations in a six-player MSS, just four satisfy the condition $x_i = x_{i-k}$ for all i (mod n), hence 1/16 of them, or 6.25%, progress to joint cooperation. For formal proofs, see Coleman et al. (1990).

Although Kelley et al. (1962) were unaware of these admittedly nonintuitive theoretical properties of the generalized MSS, they included 28 three-player and 14 four-player strict MSS groups in one of their experiments. The results were C choices close to 50%—no better than chance—for both group sizes, and no significant increases over 150 rounds. Smith and Murdoch (1970) studied strict and informed three-player and four-player MSS games, with 16 groups in each combination of group size and information condition. They found no differences between strict and informed groups and small increases in C choices over 150 rounds in four-player groups only. Findings from the second of these studies appear broadly consistent with the multiplayer MSS theory of Coleman et al. (1990), assuming that players use a noisy form of WSLS, although these earlier researchers lacked a theoretical basis for formulating hypotheses about adaptive learning in their multiplayer groups.

WSLS is a plausible method of making decisions in the dyadic MSS and more generally in the multiplayer MSS. Given the limited information supplied to the players, especially in strict MSS conditions, it is difficult to imagine how the adaptive learning reported by experimenters could be explained without assuming that some players, at least, used a form of WSLS with misimplementation noise or something functionally equivalent to it. But WSLS theory has not been rigorously tested in the multi-player MSS, where it yields predictions that are far from obvious but are nevertheless empirically testable.

1.3. Rationale for current research

Our experiments were designed to test some key predictions about adaptive learning in MSS groups of different sizes. Our experimental methodology corrected some of the deficiencies of early experiments in this area. One obvious weakness of most previous experiments is that procedures were not fully automated. In many cases, players made their choices as and when they wished, and payoffs were delivered on each round by an experimenter operating a control unit manually after both players had made their choices, potentially undermining the strict information conditions by providing subtle clues, especially in the timing of payoffs, that outcomes depended on something apart from individual participants' own choices.

Another weakness of most MSS experiments, including the multiplayer MSS experiments of Kelley et al. (1962) and Smith and Murdoch (1970), was the lack of tangible incentives, raising questions about the degree to which the players were motivated to maximize the valueless points that were assumed to represent the relevant utilities governing their behavior. Mitropoulos (2001, 2003) is among the very few researchers to have used financial incentives in MSS research, but his MSS experiments were informed rather than strict. It is widely accepted nowadays that experimental games require tangible incentives (Camerer, 2003). Without incentives, participants' behavior tends to be relatively "weak, erratic, and … readily satiated" (Smith, 1976, p. 277). Incentives reduce the variance in participants' behavior and generally improve decision making, bringing decisions closer to predictions of rational choice theory and game theory (Camerer & Hogarth, 1999; Smith & Walker, 1993), especially in tasks of intermediate difficulty (Hertwig & Ortmann, 2001).

The experiments described below were designed to test the most fundamental predictions of WSLS theory, namely that adaptive learning should occur in two-player and four-player groups, and in a small proportion of six-player groups, but not in three-player groups. Our experiments were implemented on networked computers, enabling synchronized decision making and rigorous control, and our participants were motivated with substantial financial incentives.

2. Experiment 1: strict MSS

This experiment was designed primarily to discover whether adaptive learning in different group sizes follows the pattern specified by the theoretical multiplayer MSS analysis of Coleman et al. (1990). The theoretical predictions regarding group sizes apply only to deterministic WSLS, and there are reasons to doubt that they hold under noisy WSLS. For example, following a conjecture by al-Nowaihi (personal communication, July 1, 2003), Colman (2005, Theorem 5) proved that, in a noisy version of the WSLS strategy that he called *Optimistic Pavlov*, in which a player chooses C whenever a deterministic WSLS player would choose C, and chooses D with probability p ($0 < p < 1$) whenever a deterministic WSLS player would choose D with certainty, the probability of cooperation converges to unity as the number of repetitions goes to infinity in MSS groups of every size, including odd-sized groups. A secondary aim of Experiment 1 was to check the validity of earlier findings on the two-player, strict MSS under conditions of automated feedback and significant monetary incentives.

2.1. Participants

The participants were 85 undergraduate and postgraduate students (49 women and 36 men), mean age 21.67 (SD = 4.00) years, recruited via an online experimental participant volunteer panel on the web pages of the University of Leicester. They volunteered to take part in what was described as an experiment on decision making, and each participant earned between £3.00 ($6.00 at the conversion rate at that time) and £16.00 ($32.00), M = £11.56 ($23.12) according to their payoffs in the games. Participants were assigned randomly to ten two-player groups (n = 20), five three-player groups (n = 15), five four-player groups (n = 20), and five six-player groups (n = 30).

2.2. Apparatus and general instructions

The experiment was conducted in a laboratory furnished with six computer terminals separated by partitions. Depending on treatment condition, players sat at either two, three, four, or six terminals that were networked and programmed to deliver feedback according to MSS. Each player began by reading the following instructions from a printed sheet:

> Thank you for coming here to take part in this experiment. The experiment is very simple, and the instructions will appear on the screen in front of you. First, please fill in the consent forms, which also record your age and sex.
>
> You will be making a series of 200 decisions, by pressing J or K on the keyboard in front of you. The computer will not respond to any other key. You will be given £3 to begin with, and then, after each decision, you will be awarded either 5 pence or nothing. This means that, including the £3 that has already been awarded to you, you can earn anything up to £13 (if every decision is followed by 5 pence). This is real money, and you will be paid your £3 plus whatever you have earned on top of that at the end of the experiment.
>
> Please do not confer with anyone else doing the experiment, and do not make any attempt to see what anyone else is doing. Everyone will be performing the same task. Although I cannot tell you how the payoffs are determined—that is for you to try to figure out—I can tell you that they are not random. If you concentrate and try hard, there is a possibility of increasing your earnings.

2.3. Procedure

The experimenter initiated the first of the 200 rounds by instructing the players to "please press J or K now." As soon as a player pressed one of these keys, the following message appeared on the screen: "Thank you. Calculating..." After both or all players in the group had registered their choices, a random delay uniformly distributed between 1 and 3 s occurred, then each player received a screen message stating either "You received 0p from your last decision," or "You received 5p from your last decision," followed by "CONSIDER YOUR NEXT DECISION" and then, after a uniformly distributed random delay between 2 and 4 s, "MAKE YOUR DECISION NOW." The lengths of the delays, based on the results of a pilot study, were designed to create a natural experience of interfacing with a stand-alone computer and to eliminate obvious clues pointing to the interdependence of the group members.

After 200 rounds, the participants responded to a brief questionnaire about how they believed their payoffs were determined. They were then paid in cash the amounts that they had earned and were promised a full debriefing by e-mail as soon as the experiment was completed.

2.4. Results and discussion

The postexperimental questionnaire responses confirmed that all participants believed that they were involved in individual decision making, and none gave any indication of having guessed that they had been involved in interactive decisions. This corroborates what other researchers have reported.

Proportions of cooperative choices over four trial blocks, each representing 50 rounds, are shown in Fig. 3. There appear to have been more cooperative choices in two-player than in larger groups, in the first trial block and throughout the 200 rounds of the experiment, but the effect of group size is nonsignificant: $F(3, 81) = 1.56$, *ns*. For two-player groups, the proportions of cooperative choices appear to have increased across trial blocks, but this effect narrowly fails to reach significance: $F(3, 57) = 2.28$, $p = .09$. For three-player groups, differences across trial blocks are nonsignificant, $F(3, 42) = 0.33$, *ns*, and similarly for four-player groups, $F(3, 57) = 0.36$, *ns* and six-player groups, $F(3, 87) = 0.59$, *ns*.

The data in this experiment and the ones that follow are actually time series, with potential serial dependence between successive outcomes over rounds, and time series analysis provides more appropriate and much more sensitive methods of data analysis than repeated-measures or mixed-design ANOVA. We therefore carried out several standard time series analysis procedures—curve estimation, exponential smoothing, and ARIMA model identification, estimation, and diagnosis—separately for two-player, three-player, four-player, and six-player groups, using the proportion of cooperative

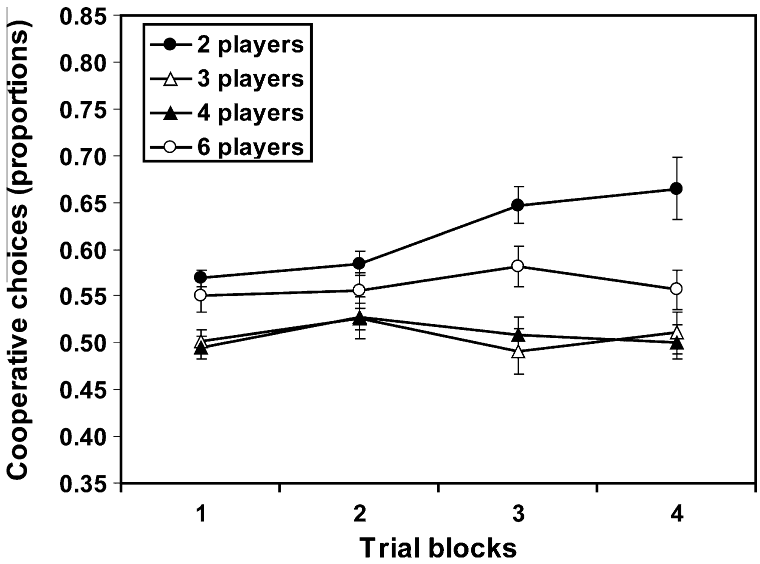

Fig. 3. Experiment 1: strict MSS. Proportions of cooperative choices over four trial blocks in groups of varying sizes. Error bars represent ±1/2 standard errors of the means.

choices per group per round as the unit of analysis. Details of the time series analyses of all experiments are given in Appendix A. For the data of Experiment 1, time series analysis confirmed a significant linear increase in cooperative choices over the 200 rounds in two-player groups only, with no evidence of any adaptive learning whatsoever in three-player, four-player, and six-player groups, where the data appear to have been random walks. The original time series and exponential smoothing model fits for all group sizes are displayed in Fig. 4.

2.4.1. Win–shift, lose–stay error analysis

Proportions of win–stay, lose–shift, win–shift, and lose–stay choices are plotted separately for two-player, three-player, four-player, and six-player groups in Fig. 5. According to the deterministic WSLS rule, win–stay and lose–shift choices are mandatory, whereas win–shift and lose–stay choices are errors. The first and most important point to notice is that win–stay choices were approximately twice as frequent as win–shift choices throughout the 200 rounds of the game in groups of all sizes. This means that whenever players cooperated, their successors repeated their previous strategy choice more than twice as often as they switched strategies, and this trend increased over trial blocks in two-player groups. The difference between lose–shift and lose–stay is less clear. It is evident from Fig. 5 that win–stay choices were generally more frequent than lose–shift choices in groups of all sizes, although both are mandatory in deterministic WSLS. In two-player groups only, the proportions of win–stay choices increased over trial blocks from .40 in Trial Block 1 to .55 in Trial Block 4. In two-player groups, lose–shift choices were much less common, and they declined slightly, from .22 in Trial Block 1 to .16 in Trial Block 4. In larger groups, win–stay choices hovered between .30 and .40, and lose–shift choices remained close to .25 and did not rise or fall markedly over trial blocks. Turning now to an examination of errors, win–shift choices remained largely static (between .10 and .18) in two-player and multiplayer groups of all sizes. Lose–stay choices were also largely static though slightly more frequent (between .16 and .26) in groups of all sizes, suggesting a degree of inertia leading to more frequent errors following unrewarding rounds.

2.4.2. Correlations within and between groups

An implication of the noisy WSLS model is that, as players within a group learn to cooperate over repeated rounds, correlations between their strategy choices should increase relative to correlations between players in different groups. To test this prediction, we computed phi coefficients (r_φ, equivalent to product-moment correlations for binary variables) between the strategy choices of all pairs of

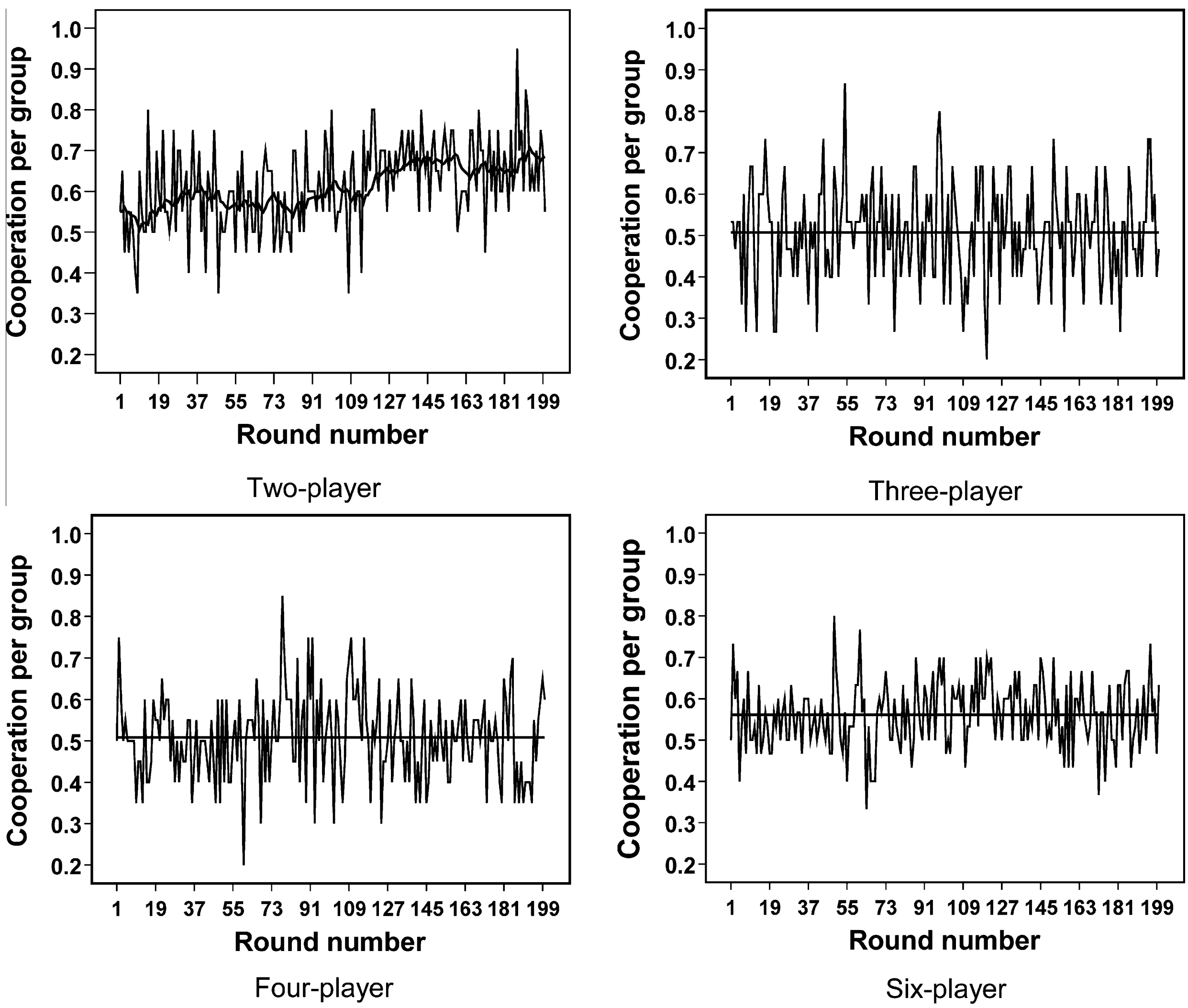

Fig. 4. Experiment 1: original time series and exponential smoothing model fits. The smoothed series are constant for group sizes three, four, and six.

players, separately for each group size and for each trial block, and we then calculated the mean values of those coefficients between pairs of players within the same group and the mean values between pairs of players in different groups. For two-player groups, the mean phi coefficients for Trial Blocks 1–4 were .12, .32, .23, and .42, respectively, between players within the same groups, and .00, .01, .02, and .01, respectively, between players in different groups. For three-player, four-player, and six-player groups, the mean phi coefficients were all close to zero, ranging from −.06 to .07 between players within the same groups and in different groups. These results are clearly in line with predictions from the WSLS model, given that adaptive learning occurred only in two-player groups.

2.4.3. Discussion

The results are broadly in line with earlier research on the MSS (Mitropoulos, 2001, 2003), with evidence for choice inertia in other areas of research (e.g., Cooper & Kagel, 2008), and with accumulated evidence that reward tends to drive behavior change more effectively than punishment (Skinner, 1969). In two-player groups, time series analysis confirms significant but slow adaptive learning over the 200 rounds of the game. This is consistent with earlier research and with a theoretical interpretation based on a form of WSLS with misimplementation noise. In larger groups, time series analysis revealed no evidence of adaptive learning over rounds. Even in four-player groups, in which cooperative choices should increase rapidly according to deterministic WSLS theory, there was no evidence of any significant adaptive learning. Joint cooperation appears to be very difficult to achieve in groups larger

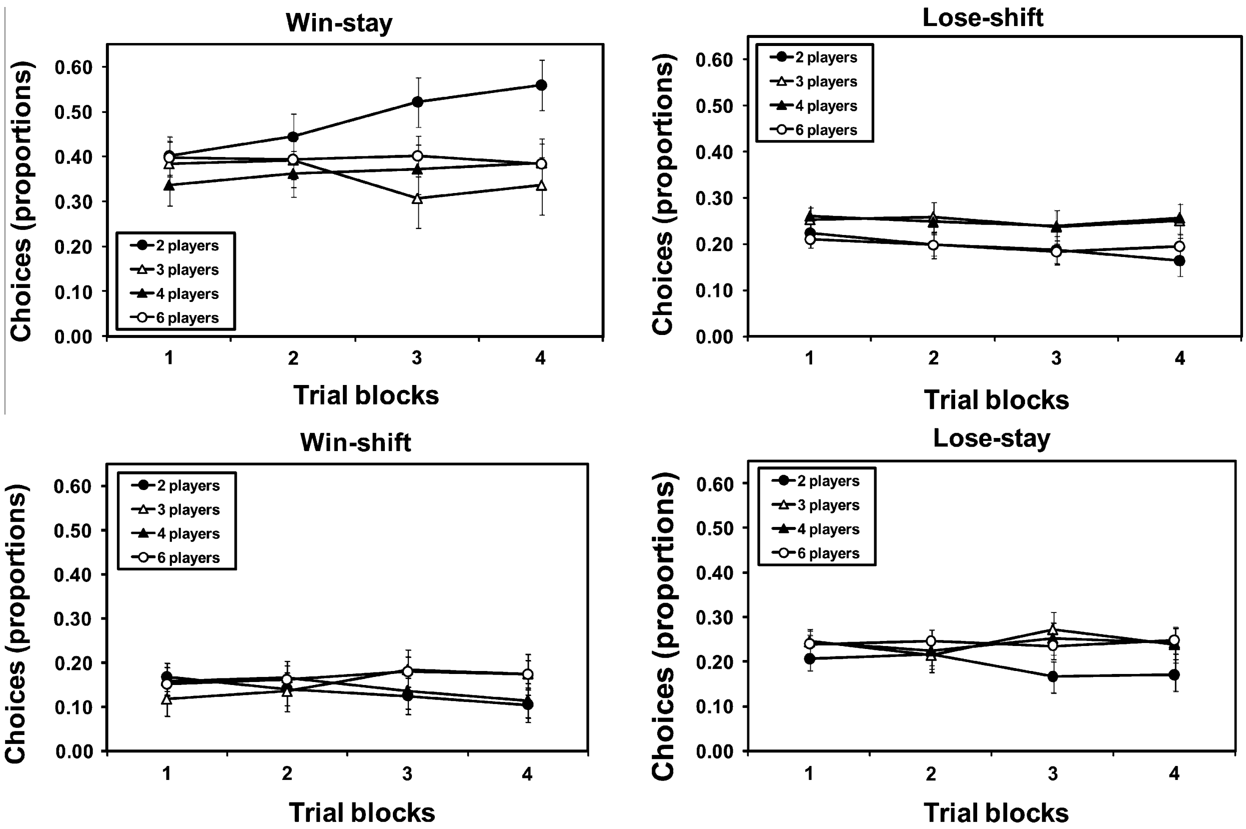

Fig. 5. Experiment 1: proportions of win–stay, lose–shift, win–shift, and lose–stay choices over trial blocks for two-player, three-player, four-player, and six-player groups.

than the dyad, presumably because it takes only one wrong move for the whole process to unravel. In a dyad, there is a reasonable chance of direct feedback reinforcing a cooperative response, but in larger groups the feedback depends on two or more intermediary players who must also choose appropriately for the first choice to be reinforced. Because players probably tend to use a form of WSLS with misimplementation noise, the noise may drown the WSLS signal in larger groups, effectively blocking efficient adaptive learning. Experiment 2 was designed to establish whether the cognitive environment of an informed MSS would reduce noise levels sufficiently to facilitate adaptive learning in groups larger than the dyad.

3. Experiment 2: informed MSS

This experiment was essentially a replication of Experiment 1 with players who were informed that they were involved in interactive decision making, although they were not informed about the payoff structure. Given the lack of adaptive learning in larger groups in Experiment 1, only two-player, three-player, and four-player group sizes were investigated.

3.1. Participants

The participants were 45 students from a sixth-form college (29 women and 16 men), mean age 17.44 (*SD* = 0.62) years, recruited as volunteers via staff members at the college. The volunteers were told that they would take part in an experiment on decision making, and each participant earned between £5.00 ($10.00) and £14.50 ($29.00), *M* = £10.25 ($20.50), according to their payoffs in the games. Participants were assigned randomly to five two-player groups (*n* = 10), five three-player groups (*n* = 15), and five four-player groups (*n* = 20).

3.2. Procedure

The instructions and general procedure were as in Experiment 1, apart from the following additional information that was read out to each group before the sequence of decisions began:

> One further piece of information I can give you is that all computers are linked together, which means that your payments may be affected by others in the group. It is for you to work out how the payments are determined, and the aim of the experiment is to maximize the amount of money that you earn. If you concentrate and try hard, there is a possibility of increasing your earnings.

This was designed to replicate the informed MSS conditions studied by previous researchers, possibly facilitating adaptive learning of mutually rewarding cooperative choices. As in Experiment 1, 200 rounds were played.

3.3. Results and discussion

Proportions of cooperative choices over trial blocks are shown in Fig. 6. Once again, cooperative choices appear to have been more frequent in two-player than in larger groups. In this case, the effect of group size is significant, $F(2, 42) = 3.94$, $p = .03$. A Tukey-HSD post hoc test confirmed that the mean number of cooperative choices is larger in the two-player than the three-player groups, and that the means in the three-player and four-player groups do not differ significantly, nor do the means in the two-player and four-player groups. In two-player groups, cooperative choices appear to have increased over the first three trial blocks and to have declined in the fourth, but the overall trial block effect is nonsignificant: $F(3, 27) = 1.24$, *ns*. For the larger groups, there is no evidence of adaptive learning, and the trial block effect is nonsignificant for three-player groups, $F(3, 42) = 0.71$, *ns*, and four-player groups, $F(3, 57) = 1.35$, *ns*. The adaptive learning that must have occurred, given the significant main effect of group size, was over by the end of Trial Block 1. Time series analysis (see Appendix A) provided no evidence for sustained adaptive learning across the entire time series in groups of any size.

3.3.1. Discussion

The results confirm the slightly surprising finding of previous researchers that adaptive learning is not noticeably better under informed than strict MSS conditions. Although our informed dyads cooperated significantly more than three-player or four-player groups, they actually performed slightly

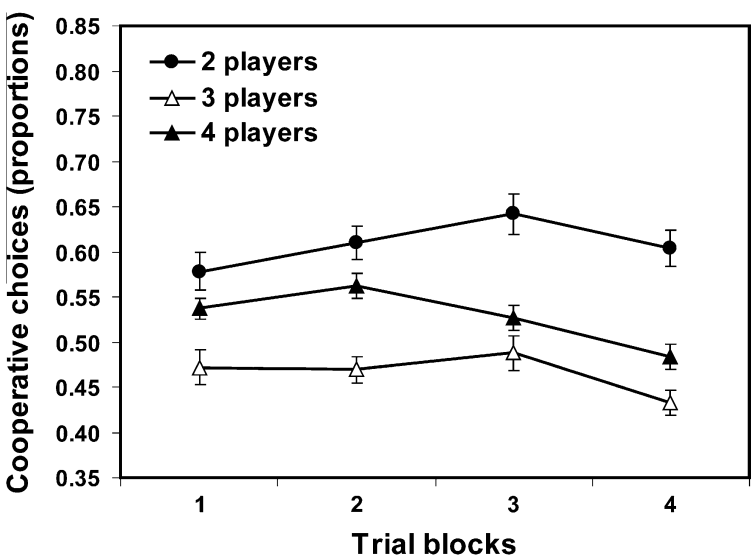

Fig. 6. Experiment 2: informed MSS. Proportions of cooperative choices over four trial blocks in groups of varying sizes. Error bars represent ±1 standard errors of the means.

worse than the uninformed dyads in the strict MSS in Experiment 1. Under the informed conditions of this experiment, dyad members, although they knew that they were strategically interdependent, were unable to coordinate their choices even as effectively as the dyads in Experiment 1, and knowledge of strategic interdependence seems only to have confused them. Perhaps this is a manifestation of Humphrey's law (Humphrey, 1951), according to which, in some circumstances, ignorance is bliss, and knowledge can interfere with efficient automatic processing. A familiar example of this is the Stroop effect, in which color names such as BLUE and RED are displayed in unrelated colors, and knowledge of what the words mean impedes one's ability to name the colors in which they are displayed (Stroop, 1935). The results from three-player and four-player groups were essentially the same as those of Experiment 1: neither time series showed any evidence of significant adaptive learning. In three-player and four-player groups, cooperative choices actually declined slightly over rounds, although ARIMA modeling suggested that the series were random walks.

Another way of looking at the results is to recognize that players must have entertained and attempted to test numerous idiosyncratic hypotheses in an effort to understand the pattern of rewards, and one reason why performance was worse in this experiment, when players knew that there were other players involved, may have been that this knowledge increased the number of false hypotheses to consider.

4. Experiment 3: discursive MSS

In this modified replication of Experiment 2, only 50 rounds were played, and the players were not informed at the outset that they were involved in interactive decision making, but the procedure was stopped after every 10 rounds to allow them to discuss the task among themselves. In addition to the period discussion breaks, after 30 rounds, players were informed that they were involved in interactive decision making, but they were not informed about the payoff structure, and after 50 rounds the game was terminated. The aim of this experiment was to determine whether these additional manipulations would prove sufficient to enable adaptive learning to occur. After the failure of mere information about being involved in interactive decision making to facilitate adaptive learning in Experiment 2, the primary purpose of Experiment 3 was to discover whether periodic discussion among the players might achieve this effect in two-player, three-player, and four-player groups.

4.1. Participants

The participants were 45 undergraduate and postgraduate students (27 women and 18 men), mean age 20.40 ($SD = 2.40$) years, recruited via an online experimental participant volunteer panel on the web pages of the University of Leicester. They volunteered to take part in what was described as an experiment on decision making, and each participant earned between £3.05 ($6.10) and £9.50 ($19.00), $M =$ £5.25 ($10.50), according to their payoffs in the games. Participants were assigned randomly to five two-player groups ($n = 10$), five three-player groups ($n = 15$), and five four-player groups ($n = 20$).

4.2. Procedure

The instructions and procedure were as in Experiment 2, apart from the restriction to 50 rounds, the breaks for discussion among the members of each group after every 10 rounds, and the withholding of information about the player' strategic interdependence until after 30 rounds. Partly because of the smaller number of rounds, incentive payments were increased from 5p (10 cents) or nothing to 50p ($1.00) or nothing on each round, with no "free" money for merely showing up. The instructions included the following:

> You will be making a series of 50 decisions, by pressing J or K on the keyboard in front of you. The computer will not respond to any other key. After each decision, you will be awarded either 50p or nothing. This means that you can earn anything up to £25 (if every decision is followed by 50p). This is real money, and you will be paid at the end of the experiment.

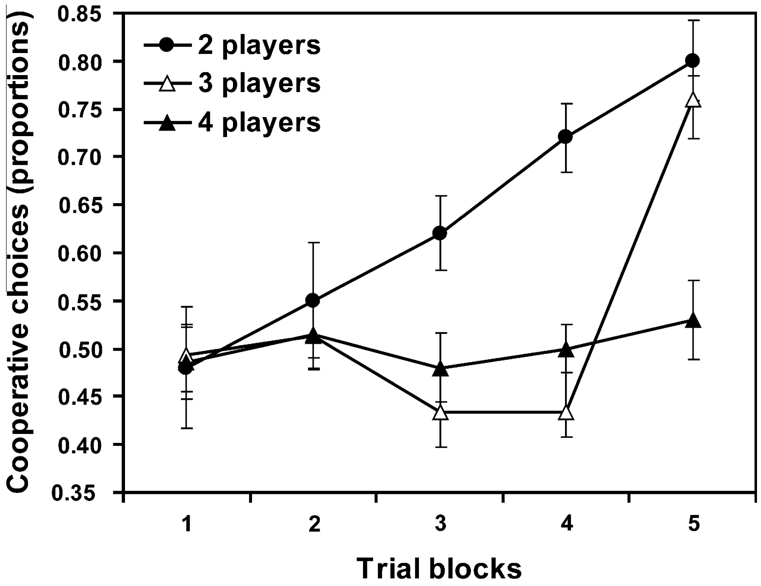

Fig. 7. Experiment 3: discursive MSS. Proportions of cooperative choices over five trial blocks in groups of varying sizes. Error bars represent ±1 standard errors of the means.

> Please do not confer with anyone else doing the experiment, and do not make any attempt to see what anyone else is doing. Everyone will be performing the same task. Although I cannot tell you how the payoffs are determined—that is for you to try to figure out—I can tell you that they are not random. If you concentrate and try hard, there is a possibility of increasing your earnings.

After each block of 10 rounds, the experimenter invited the group members to discuss the task for up to one minute. After 30 rounds, the experimenter made the following announcement: "I can tell you that your payoffs are affected by the decisions of other members of the group." After 50 rounds, the game was terminated.

4.3. Results and discussion

Proportions of cooperative choices over trial blocks are shown in Fig. 7. As in Experiments 1 and 2, the proportion was greater in two-player groups ($M = .63$) than three-player groups ($M = .53$) or four-player groups ($M = .50$), and the three means differ significantly: $F(2, 147) = 9.99$, $p < .001$. A posteriori Tukey-HSD tests confirmed that there were more cooperative choices in two-player groups than in either of the larger-sized groups but that the frequency of cooperative choices did not differ significantly between three-player and four-player groups.

Adaptive learning began earlier and occurred relatively rapidly in two-player groups, reaching .80 after 50 rounds. In three-player groups, it was only after the players were told at the end of the third trial block that they were strategically interdependent that adaptive learning occurred rapidly and efficiently. This finding should be viewed in perspective, taking into account that it relates to just 15 participants (five three-player groups). An examination of the performance of individual groups reveals that two of the five groups were cooperating perfectly in the final trial block, and two—interestingly, the groups whose cooperation rates were highest in the first two trial blocks—showed little or no increase in cooperation in the last two trial blocks. It appears that two three-player groups solved the problem through insight once they were told, after the third trial block, that they were strategically interdependent, and the other three-player groups did not. There is little evidence of any adaptive learning in the four-player groups. Time series analysis (see Appendix A) confirmed an increase in cooperative choices over rounds in two-player and three-player groups, with data from four-player groups resembling random walks.

4.3.1. Discussion

Once again, significantly more cooperation occurred in two-player than in larger groups. Allowing players to stop after every 10 rounds to discuss the task, and informing them after 30 of the 50 rounds that they were involved in interactive decision making, appear to have had some effect, though less than might have been anticipated. Adaptive learning occurred in two-player and possibly three-player groups, although in three-player groups it was evident only in two groups, after players were informed that they were making interactive decisions. Discussion and information about interactive linkage appears to have been insufficient to stimulate adaptive learning in four-player groups.

5. Experiment 4: gain–loss MSS

This experiment was a replication of Experiment 2 with a considerably enhanced incentive structure and just 50 rather than 200 rounds. Because our earlier experiments had unambiguously confirmed that adaptive learning can occur in two-player MSS groups, only three-player and four-player MSS groups were investigated. Players were given £5.00 ($10.00) at the beginning of the experiment, and over the course of 50 rounds were rewarded with £1.00 ($2.00) every time players designated as their predecessors made cooperative choices and fined 50p ($1.00) every time their predecessors made noncooperative choices. This high-incentive gain–loss incentive scheme was motivated by the finding first reported in a classic study of Tversky and Edwards (1966) that significantly more normatively optimal choices are observed when monetary gains and losses are assigned to decision makers' choices (see Smith & Walker, 1993, for a review of experimental studies corroborating and extending this finding).

5.1. Participants

The participants were 35 undergraduate and postgraduate students (16 women and 19 men), mean age 20.69 (SD = 3.99) years, recruited via an online experimental participant volunteer panel on the web pages of the University of Leicester. They volunteered to take part in what was described as an experiment on decision making, and participants ended up with earnings ranging from zero (one participant) to £50.00 ($100.00, two participants), M = £16.44 ($32.88), according to their payoffs in the games. Participants were assigned randomly to five three-player groups (n = 15) and five four-player groups (n = 20).

5.2. Procedure

The instructions and general procedure were as in Experiment 2 (informed MSS), but only 50 rounds were played for much larger financial incentives, incorporating both gains and losses. The key part of the instructions was as follows:

> You will be making a series of 50 decisions, by pressing J or K on the keyboard in front of you. The computer will not respond to any other key. You will be given £5 to begin with, and then, after each decision, you will be awarded either £1.00 if you make the correct decision or lose 50p if you make the incorrect decision. This means that you can earn anything up to £50 (if every decision is followed by £1.00). This is real money, and you will be paid at the end of the experiment.
> I cannot tell you exactly how the payments are determined, but I can tell you that they are not random. One further piece of information I can give you is that all the computers are linked together, which means that your payments may be affected by others in the group. It is for you to work out how the payments are determined, and the aim of the experiment is to maximize the amount of money that you earn. If you concentrate and try hard, there is a possibility of increasing your earnings.
> Please do not confer with anyone else doing the experiment, and do not make any attempt to see what anyone else is doing. Everyone will be performing the same task.

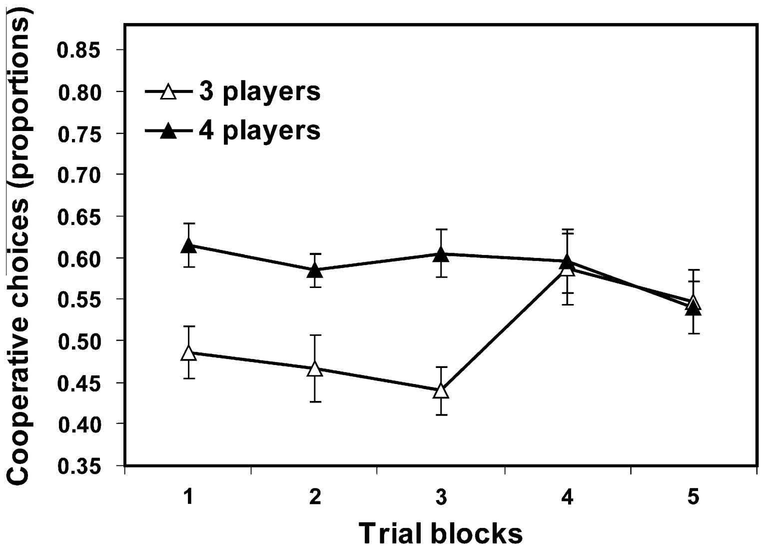

Fig. 8. Experiment 4: gain–loss MSS. Proportions of cooperative choices over five trial blocks in groups of varying sizes. Error bars represent ±1 standard errors of the means.

5.3. Results and discussion

Proportions of cooperative choices over trial blocks are shown in Fig. 8. Overall, the proportion was slightly larger in four-player groups ($M = .59$) than three-player groups ($M = .51$), but this difference is nonsignificant: $F(1, 8) = 0.53$, *ns*. The effects of trial blocks, $F(1, 8) = 0.11$, and the Group Size × Trial Blocks interaction, $F(1, 8) = 0.80$, are both nonsignificant. Time series analysis (see Appendix A) suggested no clear trend in three-player or four-player groups.

5.3.1. Discussion

Very large financial incentives, involving both gains and losses, did not help to activate adaptive learning in three-player or four-player informed MSS groups, at least not rapidly enough to become significant within the 50 rounds of the experiment, even under sensitive time series analysis. Taken together with the findings of the experiments described above, this establishes without any remaining doubt that adaptive learning is extremely difficult in MSS groups larger than the dyad.

6. Monte Carlo simulation: parameters for noisy MSS

It is clear from the four experiments described above that adaptive learning occurs slowly in two-player MSS groups but is extremely difficult to elicit in three-player or larger MSS groups. According to deterministic WSLS theory, rapid adaptive learning should occur in two-player and four-player groups, and also in a small proportion of six-player groups (depending on initial choices). However, the standard WSLS decision rule incorporates an unrealistic assumption of deterministic choice. Our hypothesis is that players implement a degraded form of WSLS with misimplementation noise. This seems the simplest and most natural decision rule suitable for the informationally impoverished conditions of a strict MSS.

The most important challenge is to explain the results of strict MSS experiments, both two-player and multiplayer. If, as we hypothesize, the slow adaptive learning observed in two-player MSS groups is explained by players implementing a degraded form WSLS with misimplementation noise, then why does this not also occur in four-player groups? After all, from the point of view of the individual player, the task appears identical—the player does not even know that the group exists, still less what size it is.

One obvious explanation for this group size effect is that noisy decision making, characteristic of human behavior, undermines the efficacy of the WSLS decision rule in groups larger than a dyad.

For example, in a two-player MSS with noise, a joint cooperative outcome or sequence of outcomes might be disrupted by a player misimplementing a *D* instead of a *C* choice on Round *t*, in spite of having received a rewarding payoff on Round $t - 1$. In the absence of any further errors, joint cooperation is re-established within two further rounds. But the picture is very different in a multiplayer MSS with misimplementation noise. Following a joint cooperative outcome, if a single player misimplements *D* instead of *C* immediately after receiving a rewarding payoff, the outcome is followed, in the absence of any further errors, by joint cooperation only after several further rounds—up to four rounds in a four-player MSS, and later still in larger groups. With high noise levels, it is unlikely that these corrective rounds will be completed without further misimplementation errors, and any error that does occur is more likely to lengthen than to shorten the path back to joint cooperation. Although all outcome configurations in group sizes that are powers of 2, and some configurations in other even-sized groups, lead back to joint cooperation under deterministic WSLS, misimplementation noise is likely to disable the WSLS corrective mechanism in groups larger than the dyad.

Another way of looking at repetitive decisions under restricted information is to interpret initial exploratory decisions as noise, and to assume, in line with research on *k*-armed bandit problems (Gittins, 1989; Holland, 1975; Schmalensee, 1975), that exploration tends to decline over trials. Whether or not there exist declining noise levels that might generate results similar to those observed in our experiments is an empirical question that can be investigated through Monte Carlo methods. Accordingly, we programmed a noisy WSLS computational algorithm, using the assumption of exponentially decaying misimplementation noise over successive rounds.[1] Exponential noise decay is mathematically simple inasmuch as it assumes that the rate of decrease is a constant proportion of the absolute noise level, and it is also in line with Holland's proof that, for optimal learning in two-armed bandit problems, exploratory trials should be allocated to the better arm at an exponentially increasing rate and to the worse arm at an exponentially decreasing rate. Our algorithm accepts as input parameters the initial noise level (the probability of a choice being misimplemented according to deterministic WSLS) and the half-life for exponential noise decay, and these parameters can be set independently for choices following rewarding and unrewarding rounds, to take account of the possibility of closer adherence to win–stay than to lose–shift. Using this computational algorithm, we performed grid searches for approximate values that minimized the sum of absolute differences between observed and predicted trial block means for each of our experiments. All simulation results were averaged over 10,000 replications.

6.1. Experiment 1 simulation

For Experiment 1, the grid search revealed an asymmetric set of parameters that provided a good fit to most of the data. These parameters included initial noise levels of 20% after rewarding rounds and 50% after unrewarding rounds, in general agreement with our finding, and that of Mitropoulos (2001, 2003), of closer adherence to win–stay than to lose–shift. With the half-life of the exponential noise decay set to 200 rounds, the simulation results are compared with the empirical results in Table 1. A striking feature of the simulation is that our noisy WSLS model predicts substantial cooperative learning only in two-player groups, and not in four-player groups, in line with our experimental findings, although deterministic WSLS would yields rapid cooperative learning in both two-player and four-player groups. The fit is reasonably close for most group sizes, though noticeably less so for six-player groups. This was partly due to one of the six-player experimental groups cooperating much more than the others. When this unusually successful group is omitted, the simulated values are closer to the empirical proportions.

6.2. Comparison with empirical error data

The win–shift, lose–stay error analysis from Experiment 1, shown in Section 2.4, provided empirical data that can be usefully compared with the simulation results. Dividing the empirically observed

[1] A ready-to-use version of the simulation program, with simple explanatory notes, is available at http://hdl.handle.net/2381/7816 and the source code, written in C++, may be requested from D_Omtzigt@msn.com.

Table 1
Proportions of cooperative choices in Experiment 1 compared to Monte Carlo results.

Group size	Rounds			
	1–50	51–100	101–150	151–200
2				
Experiment	.57	.59	.65	.67
Monte Carlo	.54	.58	.62	.67
3				
Experiment	.50	.53	.49	.51
Monte Carlo	.50	.50	.51	.52
4				
Experiment	.50	.53	.51	.50
Monte Carlo	.50	.50	.50	.51
6				
Experiment[a]	.55	.56	.58	.56
	(.51)	(.53)	(.57)	(.52)
Monte Carlo	.50	.50	.50	.50

Note: Initial noise levels: 20% after rewarding rounds; 50% after unrewarding rounds. Number of rounds to 50% exponential noise reduction: 200. Proportions shown are means from 10,000 replications.

[a] Figures shown in parentheses are experimental results for six-player groups omitting an outlier group.

Table 2
Proportions of cooperative choices in Experiment 2 compared to Monte Carlo results.

Group size	Rounds			
	1–50	51–100	101–150	151–200
2				
Experiment	.58	.61	.64	.60
Monte Carlo	.54	.58	.62	.67
3				
Experiment	.47	.47	.49	.43
Monte Carlo	.50	.50	.51	.52
4				
Experiment	.54	.56	.53	.48
Monte Carlo	.50	.50	.50	.51

Note: Initial noise levels: 20% after rewarding rounds; 50% after unrewarding rounds. Number of rounds to 50% exponential noise reduction: 200. Proportions shown are means from 10,000 replications.

win–shift choices by the sum of the win–stay and win–shift choices yields the proportion of "erroneous" choices following rewarding rounds; and dividing the lose–stay choices by the sum of the lose–shift and lose–stay choices yields the proportion of "erroneous" choices following unrewarding rounds. In two-player groups, the proportions of errors following rewarding rounds decline steadily from .29 in Trial Block 1 to .16 in Trial Block 4, and in larger groups, the proportions lie between .37 and .23 for all trial blocks, without any clearly discernable pattern. The proportions of errors following unrewarding rounds lie between .56 and .45 in groups of all sizes, again with no obvious pattern.

The empirical frequency of win–shift choices—between .16 and .32 for 14 of the 16 trial block means across all four group sizes—is roughly consistent with the Monte Carlo estimate of initial 20% misimplementation noise following reward. Following unrewarding rounds, the empirical proportions of erroneous lose–stay choices were between .45 and .56 for all 16 trial block means, and this is roughly consistent with the Monte Carlo estimate of 50% initial noise following unrewarding rounds. These comparisons of empirically observed deviations from deterministic WSLS and Monte

Table 3
Proportions of cooperative choices in Experiment 3 compared to Monte Carlo results.

Group size	Rounds				
	1–10	11–20	21–30	31–40	41–50
2					
Experiment	.48	.55	.62	.72	.80
Monte Carlo					
(20, 50; 200)[a]	.52	.53	.54	.55	.55
(20, 50; 25)[a]	.53	.61	.68	.75	.81
3					
Experiment	.49	.51	.43	.43	.76
Monte Carlo					
(20, 50; 200)[a]	.50	.50	.50	.50	.50
(20, 50; 25)[a]	.50	.51	.52	.53	.55
4					
Experiment	.49	.52	.48	.50	.53
Monte Carlo					
(20, 50; 200)[a]	.50	.50	.50	.50	.50
(20, 50; 25)[a]	.49	.50	.51	.54	.58

[a] (x, y; z) denotes initial noise levels: x% after rewarding rounds; y% after unrewarding rounds; z rounds to 50% exponential noise reduction. Proportions shown are means from 10,000 replications.

Carlo noise estimates reinforce our confidence in our interpretation of how the experimental data may have been generated.

6.3. Experiments 2–4 simulations

Simulation and empirical results for Experiment 2 are compared in Table 2. In this experiment, the players were informed that their decisions were interactively linked with those of other group members, but our proposed model, with initial noise levels of 20% after rewarding rounds and 50% after unrewarding rounds, decaying exponentially with a half-life of 200 rounds, remains a moderately good fit to the data. The most striking discrepancies are in the final block of rounds, especially in the two-player and three-player groups, where the experimental data show slight declines in cooperation not predicted by the model (compare Fig. 6).

In Experiment 3, players discussed the task among themselves after every block of 10 rounds, and after 30 rounds they were informed that they were involved in interactive decision making. The simulation results are compared to the empirical results in Table 3. In this case, a better fit was obtained,

Table 4
Proportions of cooperative choices in Experiment 4 compared to Monte Carlo results.

Group size	Rounds				
	1–10	11–20	21–30	31–40	41–50
3					
Experiment	.49	.47	.44	.59	.55
Monte Carlo					
(20, 50; 200)[a]	.50	.50	.50	.50	.50
(20, 50; 25)[a]	.50	.51	.52	.53	.55
4					
Experiment	.62	.59	.61	.60	.54
Monte Carlo					
(20, 50; 200)[a]	.50	.50	.50	.50	.50
(20, 50; 25)[a]	.49	.50	.51	.54	.58

[a] (x, y, z) denotes initial noise levels: x% after rewarding rounds; y% after unrewarding rounds; z rounds to 50% exponential noise reduction. Proportions shown are means from 10,000 replications.

most notably so for two-player groups, by assuming more rapid exponential noise decay, with a half-life of 25 rounds. This is easily understood as a consequence of the reduction in the total number of rounds from 200 to 50 and the discussion among the players under the conditions of this experiment. The sudden increase in cooperation in three-player groups in the last trial block was not predicted by our model. It was caused by the behavior of two groups whose members apparently gained sudden insight into the problem after being told that the group members were linked. Our WSLS model predicts adaptive learning in the absence of insight, and it is obvious that the learning process can be overridden or short-circuited if collective insight is achieved in a group. This occurred in just two groups whose members were permitted to discuss the problem regularly among themselves and were also given a big clue after 30 rounds. Behavior in these unusual circumstances falls outside the scope of our WSLS model, but it is a real and important phenomenon.

Finally, in Experiment 4, players knew that their decisions were interactive, as in Experiment 2, and large financial gains and losses were assigned to the game. Simulation and empirical data are compared in Table 4. Once again a slightly better fit was obtained by assuming rapid noise decay, with a half-life of 25 rounds. In this case, it seems that the increased rate of noise decay was a consequence of the reduction in the total number of rounds from 200 to 50 and of raising the stakes and thereby presumably encouraging players to invest more cognitive effort in the task.

6.4. Models of decisions from experience

A number of models have been developed to predict "decisions from experience" involving reinforcement learning in tasks that, from the decision maker's point of view, are similar to the MSS, but that include a wider range of binary decision tasks (Erev & Haruvy, in press). Some of these models are capable of predicting the results of our experiments.

Barron and Erev (2003) developed a *value-assessment model* that they showed to be successful at predicting repetitive individual decisions when information available to decision makers is limited to feedback about outcomes on previous trials. In the value-assessment model, the value of a decision j is initially the subjective expected value from random choice; then the adjusted value A_j on trial $t + 1$ is

$$A_j(t+1) = (1 - w_t)A_j(t) + (w_t)\,v(x_t),$$

where $v(x_t)$ is the subjective value of the payoff x_t weighted by $0 < w_t < 1$. Decision makers are assumed to use exploration (random choice) on early trials. The probability of exploration is 1 on the first trial and declines with experience. To model greater reliance on recent outcomes than on outcomes in the more distant past (myopic decision making), $w_t = \alpha$ if t is an exploration trial, the probability of this being the case decreasing with experience, and $w_t = \beta$ otherwise, with $0 \leqslant \beta \leqslant \alpha < 1$. To model loss aversion (greater sensitivity of decision makers to losses than to gains of equivalent absolute size), $v(x_i) = x_i$ if $x_i \geqslant 0$ and $v(x_i) = \lambda x_i$ if $x_i < 0$, with $\lambda > 1$. Except on exploration trials, a decision maker is assumed to choose the alternative with the highest subjective value. Barron and Erev found that this simple model can predict several nontrivial decision-making phenomena. With suitable parameter settings it can emulate (generate the same general pattern of results as) noisy WSLS in minimal social situations (I. Erev, personal communication, April 12, 2010).

The *explorative sampler* model (Erev, Ert, & Yechiam, 2008) was the most successful baseline model for repeated decisions from experience in a choice prediction competition organized by Erev et al. (2010). In this model, decision makers use two cognitive strategies, namely exploration (random choice) and exploitation. The probability of exploration is 1 on the first trial and declines toward an asymptote with experience, the rate of decline depending on the total number of trials that the decision maker anticipates. On an exploitation trial t, the decision maker draws from memory, randomly with replacement, a sample of m_t past experiences—outcomes from previous trials for each alternative. The recalled subjective value of an outcome x is affected by regression to the mean of all experiences with the same alternative, and with diminishing sensitivity (decreasing sensitivity to changes in absolute payoff with increasing distance from zero) and loss aversion. The regressed value is $R_x = (1 - w)x + (w)A_j(t)$, where $0 < w < 1$ is a free parameter and $A_j(t)$ is the average payoff from alternative j on trial t. Diminishing sensitivity and loss aversion are modeled with the value function

$$sv(x) = \begin{cases} R_x^{\alpha(t)} & \text{if } R_x \geqslant 0, \\ -(-R_x)^{\alpha(t)} & \text{if } R_x < 0, \end{cases}$$

where $\alpha(t) = (1 + V_t)^{(-\rho)}$, $\rho \geqslant 0$, is a free parameter capturing the effect of diminishing sensitivity, and V_t is a measure of payoff variability. On every exploitation trial, a decision maker is assumed to choose the alternative with the higher subjective value.

According to Ido Erev (personal communication, April 12, 2010), the explorative sampler model emulates noisy WSLS in minimal social situations, and good emulation occurs even with the parameters that best fit the individual choice task studied by Erev et al. (2010). In spite of this, for interpreting results from minimal social situations, the WSLS decision rule has two major advantages over the explorative sampler, value-assessment model, and other leading models of decisions from experience. The first arises from Ockham's razor, the principle of economy of explanation according to which assumptions should not be multiplied beyond necessity (*entia non sunt multiplicanda praeter necessitatem*), and hence simple explanations should generally be preferred to more complex ones. Popper (1959, Section 43), in a discussion of "simplicity and degree of falsifiability," argued that simpler rules or laws are generally preferable because they are more refutable. Noisy WSLS is obviously far simpler than the value-assessment model, the explorative sampler, and other models reviewed by Erev et al. (2010) and Erev and Haruvy (in press). Noisy WSLS should therefore be preferred as an explanation of behavior in minimal social situations, even if some of the more complex models can emulate it. Second, it encapsulates the most primitive form of the law of effect, for which there is over a century of empirical support across a very wide range of behavioral domains. Noisy WSLS takes no specific account of well established decision-making phenomena such as loss aversion and inaccurate weighting of rare events, and it is therefore all the more impressive that it provides a compelling interpretation and remarkably good predictions of behavior in minimal social situations. However, we should not expect it to be notably successful at predicting results from binary decision tasks unrelated to minimal social situations.

7. General discussion

Experiments reported here provide clear evidence that adaptive learning of cooperative behavior can occur without awareness in the two-player MSS, even under strict information conditions, with properly controlled experimental conditions implemented through networked computers. This confirms the findings of previous experiments, invariably (in the case of strict MSS) performed without adequate incentives for the players and in many cases lacking adequate experimental control. However, in spite of our significant financial incentives and automated control techniques, we found no evidence of adaptive learning in three-player, four-player, or six-player MSS groups.

It seems likely that at least some human decision makers respond to the strict MSS with a noisy form of WSLS. Because WSLS is merely a version of the law of effect applicable in binary choice situations with severely restricted information, and bearing in mind that the law of effect has been found to have very wide applicability, it would be surprising if it did not govern behavior in MSS situations. Furthermore, without assuming that decision makers use noisy WSLS, or some functionally equivalent decision rule of decisions from experience, it seems impossible to explain the adaptive learning in the two-player MSS that has been reported by previous experimenters and that we have replicated in our own research, reported above.

In Experiments 2–4, we examined the effects of various manipulations designed to facilitate adaptive learning, but they proved largely unsuccessful. Even when very large incentive payments were on offer and group members were allowed to discuss the task among themselves at frequent intervals, cooperative choices generally failed to increase significantly over rounds in multiplayer groups. It was only in three-player groups in Experiment 3, when very large incentive payments were offered, discussion among group members was allowed, and players were informed after 30 rounds that they were strategically interdependent, that significant increases in *C* choices began abruptly in two groups in the final trial block, evidently as a consequence of sudden collective insight in two groups—an important phenomenon that falls outside the scope of our WSLS adaptive learning model. The results from our four-player groups present a particularly challenging problem of interpretation, because deterministic WSLS predicts a rapid increase in cooperative choices in four-player groups. However, the results of our experiments showed no significant increases in cooperative choices in four-player groups.

We interpret the results from all four experiments as evidence that WSLS with misimplementation noise is generally ineffective at eliciting cooperation in groups larger than the dyad, presumably because an aberrant choice—a *D* choice that should have been a *C* choice according to deterministic WSLS, or vice versa—takes several rounds to work its way through a multiplayer group, and with a high noise level this process is unlikely to be completed before another aberrant choice is made by one of the players.

This interpretation is confirmed by our Monte Carlo simulations. Assuming that players follow the spirit of WSLS but begin with errors on 20% of rounds immediately following rewarding payoffs and on 50% of rounds immediately following unrewarding payoffs, and that these noise levels decay exponentially, we were able to generate patterns of results approximating those of all four experiments. This enables us to suggest a coherent explanation of our findings, and in particular of our failure to observe significant adaptive learning in four-player MSS groups.

The deterministic WSLS decision rule works as reliably in the four-player MSS as it does in the two-player MSS, but its built-in error-correcting property breaks down in the four-player MSS under sufficiently high levels of misimplementation noise. The following examples, using the noise levels derived from our Monte Carlo simulations, should help to clarify this. Suppose that the configuration (D, C, C, C) occurs in a four-player MSS. Under deterministic WSLS, this would be followed in four steps by joint cooperation: (D, D, C, C), (D, C, D, C), (D, D, D, D), (C, C, C, C). Assuming (reasonably) that errors are uncorrelated, the probability of all four of these steps occurring correctly with 20% noise following a rewarding round and 50% noise following an unrewarding round is $(4/5)^7 \times (1/2)^9 = .0004$. In a two-player MSS, by way of contrast, the configuration (D, C) is followed in two steps by joint cooperation: (D, D), (C, C), and the probability of this occurring correctly with 20% noise following a rewarding round and 50% noise following an unrewarding round is $(4/5) \times (1/2)^3 = .10$. Hence, in the two-player MSS, there is a fighting chance of success, enabling cooperation to develop slowly, but in the four-player MSS there is virtually none, and as group size increases, the probability diminishes even further. These indicative calculations suggest that adaptive learning can occur, though with some difficulty, in two-player MSS groups but not in larger MSS groups. In a multiplayer MSS, a group that strays from the WSLS path faces a route back that is too long to navigate before further wrong turns are taken, and the errant group is therefore likely to get lost unless it stumbles on joint cooperation by chance.

Experimentation with our WSLS simulator tends to corroborate this interpretation in four-player groups. With initial noise up to about 20%, decaying exponentially with a half-life of 200 rounds, adaptive learning is rapid in four-player groups; but with initial noise above about 30%, negligible improvement is evident even after 200 rounds.

We therefore suggest that human decision makers approach MSS decision tasks using a WSLS decision rule, starting with something like 20% misimplementation noise after rewarding rounds and 50% after unrewarding rounds. What we call "noise" may, of course, include myriad idiosyncratic hypotheses, of which we are ignorant, that players formulate and try to test through exploratory decisions, the number of false hypotheses increasing when the players know that they are involved in interactive games. The result is very different from what would occur under the decision rule that we conjectured before performing our experiments, namely Optimistic Pavlov, in which misimplementation noise follows unrewarding rounds only. Under that rule, adaptive learning should theoretically occur remarkably quickly and easily in MSS groups of all sizes, even with very noisy responses following unrewarding rounds.

The implications of our findings for everyday social, economic, and political life are clear. Whereas dyads, such as married couples or business partners, may be able to rely to a certain extent on coordination without explicit communication, larger groups simply cannot. Reverting to our dyadic example in the introduction of Alf and Beth choosing snacks for their children, if they followed the noisy WSLS rule that our experimental participants evidently followed, assuming vaguely that their snack choices might be making their own children ill from time to time, then each would eventually learn to reward the other by avoiding the snack that, in reality, makes the other child ill. However, this remarkable social mechanism would not work in larger groups of snack-sharing children although, if the parents discussed their children's allergies, they might gain insight into the game. More generally, tacit coordination can evolve in dyads, but larger groups can hope to coordinate their actions only by explicit communication and planning. If successful cooperation is to be achieved in larger groups or organizations, then mechanisms need to be put in place to facilitate the coordination of members' actions and expectations of one another's actions.

In conjunction with our experimental results, our Monte Carlo simulations solve the problem of providing a coherent explanation of how human decision makers behave in MSS games. As in almost all domains, behavior tends to follow the law of effect. However, in the MSS, players begin by implementing the WSLS decision rule, mandated by the law of effect, rather noisily, frequently deviating from it after rewarding rounds and more frequently after unrewarding rounds, but gradually implementing it more rigorously over successive rounds. This would explain not only our own results but also those of previous investigators.

Acknowledgments

This project was supported by Grant No. LRG-37265 awarded by the British Academy Larger Research Grants Scheme to Andrew M. Colman and Ali al-Nowaihi. The authors are grateful to Clare Davies for assistance in collecting the data, to Tony Andrews for developing software for running the experiments, to Robert Goldstone and Ido Erev for several helpful suggestions, and to Natalie Gold and other participants in "Teamwork: A Multi-Professional Workshop" at St. Catherine's College, Oxford, April 9–10, 2003, for a discussion of the theoretical issues underlying this research before the experiments were designed. David Omtzigt is now at Radboud University Nijmegen, The Netherlands.

Appendix A

A.1. Time series analysis

To provide more detailed insight into changes in relative frequencies of cooperative choices over rounds, we performed time series analyses by least squares curve estimation, exponential smoothing, and autoregressive integrated moving average (ARIMA) model fitting, using the numbers of cooperative choice per group on each round as the data series. Exponential smoothing is a method of estimating the maximum likelihood value of a general parameter α, ranging from $\alpha = 0$ when previous values are weighted as heavily as recent ones in fitting the model to $\alpha = 1$ when the single most recent value is used exclusively; a parameter γ indicating the degree of trend, not necessarily linear; a seasonality parameter δ; and a damping parameter φ, when a series shows trend that tends to die out. ARIMA (p, d, q) modeling estimates the maximum likelihood values of the p and q parameters, analogous to beta coefficients in multiple regression, and the degree of differencing involved in fitting the model to the time series. The value of p indicates the number of autoregression parameters, d the degree of differencing, indicative of linear or nonlinear trend, and q the number of moving average parameters.

A.2. Experiment 1

A.2.1. Curve estimation

For two-player groups, linear and quadratic curve estimation produced a least-squares linear model with $R^2 = .18$ and standardized regression coefficient $\beta = .42$, $F(1, 198) = 42.30$, $p < 001$, indicating a significant linear increase in cooperative choices over the 200 rounds of the experiment. The least-squares quadratic model yielded $R^2 = .18$, with standardized regression coefficients β (linear) = .40 and β (quadratic) = .02, $F(2, 197) = 21.05$, $p < .001$. The amount of incremental variance explained by the quadratic component over and above the linear component is zero, so the curvilinear trend can be disregarded. For three-player groups, R^2 (linear) = .00 and R^2 (quadratic) = .00, confirming the absence of any significant trend in cooperative choices over rounds. For four-player groups, R^2 (linear) = .00, R^2 (quadratic) = .01, and the incremental variance explained by the quadratic trend is nonsignificant, $F(2, 197) = 1.98$, *ns*, hence failing to provide evidence of any significant trend. For six-player groups, R^2 (linear) = .01, R^2 (quadratic) = .02, and the incremental variance explained by the quadratic trend is nonsignificant, $F(2, 197) = 1.60$, *ns*, failing once again to detect any significant trend. The results of curve estimation therefore confirm the impression conveyed by Fig. 3 that it was only in the two-player groups that cooperation increased significantly over rounds.

A.2.2. Exponential smoothing

For two-player groups, a grid search showed that the smoothing parameters having the smallest sum of squared errors were $\alpha = .1$ and $\gamma = 0$. The autocorrelation function (ACF) and partial autocorrelation function (PACF) of the residuals from the smoothed series—after smoothing with these parameters—are within the upper and lower 95% confidence limits for all lags up to 16, confirming that the residuals are white noise, as required for a satisfactory model fit. For three-player, four-player, and six-player groups, exponential smoothing yielded stationary smoothed series, suggesting that the original time series lacked any statistically significant structure. The original time series and exponential smoothing model fits for all group sizes are displayed in Fig. 4.

A.2.3. ARIMA model identification, estimation, and diagnosis

The Box–Jenkins procedure was used to identify and estimate an autoregressive integrated moving average or ARIMA (p, d, q) model for each group size. For two-player groups, analysis of the ACF and PACF plots of the original series, in conjunction with the associated Box–Ljung statistics, suggested an ARIMA (0, 1, 1) model, with differencing at lag 1 to produce stationarity. The moving average parameter q was marginally significant ($t = 1.79$, $p = .07$) suggesting that the two-player data were generated by a process in which the proportion of cooperative choices increased gradually over the 200 rounds of the game, and that the score on each round was determined marginally by a random disturbance on the immediately preceding round. For three-player, four-player, and six-player groups, ARIMA (0, 1, 0) models were derived, suggesting that the time series from these groups were random walks, without significant autoregressive or moving average parameters. For diagnosis, the ACF and PACF plots of the residual (error) series were analyzed, confirming absence of significant structure in the residuals after model fitting for all group sizes.

A.3. Experiment 2

A.3.1. Curve estimation

For two-player groups, curve estimation yielded a nonsignificant overall effect: $F(3, 27) = 1.24$, *ns*; similarly for three-player groups, $F(3, 42) = 0.71$, *ns* and four-player groups, $F(3, 57) = 1.45$, *ns*. On the basis of curve estimation, there is no evidence of significant trend for any of the group sizes.

A.3.2. Exponential smoothing

For two-player groups, a grid search yielded smoothing parameters of $\alpha = 0$ and $\gamma = 0$ having the smallest sum of squared errors, with ACF and PACF plots of residuals within the upper and lower 95% confidence limits for all but one of the lags up to 16, and an inspection of the sequence plot revealed no hint of structure. Essentially the same result emerged for three-player and four-player groups.

A.3.3. ARIMA model identification, estimation, and diagnosis

ACF and PACF plots of the original series, in conjunction with the associated Box–Ljung statistics, suggested that all three time series were best represented by an ARIMA (0, 1, 0) random walk, and this was confirmed by inspection of the ACF and PACF plots of the residuals after fitting these models.

A.4. Experiment 3

A.4.1. Curve estimation

A sequence graph of the two-player data showed evidence of trend, hence linear and quadratic curve estimation were performed. The least-squares linear model yielded $R^2 = .32$, standardized regression coefficient $\beta = .56$, $F(1, 48) = 22.28$, $p < 001$, indicating a significant linear increase in cooperative choices over the 50 rounds of the experiment. The least-squares quadratic model yielded $R^2 = .18$, with little incremental variance explained by the quadratic component over and above the linear component. For three-player groups, R^2 (linear) = .17 and R^2 (quadratic) = .39, $F(2, 47) = 14.75$, $p < .001$. For four-player groups, R^2 (linear) = .02 and R^2 (quadratic) = .03, with standardized regression coefficients $\beta = 0.00$ (linear) and $\beta = .16$ (quadratic), $F(2, 47) = 0.50$, *ns*, and there is no evidence of any significant trend in cooperative choices over rounds. Curve estimation suggests, therefore, that it was only in the two-player and three-player groups that cooperation increased significantly over rounds.

A.4.2. Exponential smoothing

For two-player groups, a grid search revealed the parameters minimizing the sum of squared errors to be $\alpha = .1$ and $\gamma = 0$, with ACF and PACF plots of the residuals from the smoothed series lying within the upper and lower 95% confidence limits for almost all lags up to 16. For three-player groups, the smoothing parameters were $\alpha = .3$ and $\gamma = 0$, and for four-player groups the smoothed series was stationary.

A.4.3. ARIMA model identification, estimation, and diagnosis

Analysis of the ACF and PACF plots and Box–Ljung statistics suggested, for two-player and three-player groups, ARIMA (1, 1, 0) models, with differencing at lag 1 to produce stationarity. For both group sizes, the autoregressive parameters were significant ($t = 2.18$, $p = .03$ for two-player groups; $t = 3.57$, $p < .01$ for three-player groups). These results suggest that, in both two-player and three-player groups, the proportions of cooperative choices increased over the 50 rounds of the experiment, and that the outcome on each round was influenced by the outcome on the immediately preceding round, although in three-player groups, the significant increase in cooperative choices appears to have occurred mainly from the fourth to the fifth and final trial block. For four-player groups, no significant autoregressive or moving average parameters were found, hence the time series is best described as an ARIMA (0, 1, 0) random walk.

A.5. Experiment 4

A.5.1. Exponential smoothing

For three-player groups, a sequence graph suggested the possibility of a weak trend over the 50 rounds of the experiment. However, an exponential smoothing grid search yielded a best-fitting model with only one parameter, $\alpha = .2$, and ACF and PACF plots of the residuals from the smoothed series were within the upper and lower 95% confidence limits for all lags up to 16. This suggests that outcomes in three-player groups were slightly influenced by outcomes on previous rounds but that no significant trend was present in the time series.

For four-player groups, a sequence graph showed no suggestion of any pattern in the data. Exponential smoothing yielded a best-fitting model with $\alpha = 0$ and ACF and PACF plots of the residuals from the smoothed series within the upper and lower 95% confidence limits for all lags up to 16, confirming an absence of any meaningful pattern in the time series.

A.5.2. ARIMA model identification, estimation, and diagnosis

For three-player and four-player groups, analysis of ACF and PACF plots and Box–Ljung statistics of the original time series yielded nothing suggestive of any significant effects, in spite of the slight increase in cooperation in three-player groups in the second half of the series. Both series are best modeled by ARIMA (0, 1, 0) random walks.

References

Arickx, M., & Van Avermaet, E. (1981). Interdependent learning in a minimal social situation. *Behavioral Science, 26*, 229–242.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making, 16*, 215–233.

Bertilson, H. S., & Lien, S. K. (1982). Comparison of reaction time and interpersonal communication tasks to test effectiveness of a matching strategy in reducing attack-instigated aggression. *Perceptual and Motor Skills, 55*, 659–665.

Bertilson, H. S., Wonderlich, S. A., & Blum, M. W. (1983). Withdrawal, matching, withdrawal-matching, and variable-matching strategies in reducing attack-instigated aggression. *Aggressive Behavior, 9*, 1–11.

Bertilson, H. S., Wonderlich, S. A., & Blum, M. W. (1984). Withdrawal and matching strategies in reducing attack-instigated aggression. *Psychological Reports, 55*, 823–828.

Bitterman, M. E. (1975). The comparative analysis of learning. *Science, 188*(4189), 699–709.

Bitterman, M. E., & Schoel, W. M. (1970). Instrumental learning in animals: Parameters of reinforcement. *Annual Review of Psychology, 21*, 367–436.

Bonem, M., & Crossman, E. K. (1988). Elucidating the effects of reinforcement magnitude. *Psychological Bulletin, 104*, 348–362.

Boren, J. J. (1966). An experimental social relation between two monkeys. *Journal of the Experimental Analysis of Behavior, 9*, 691–700.

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. New York: Russell Sage Foundation.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in economics experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty, 18*, 7–42.

Coleman, A. A., Colman, A. M., & Thomas, R. M. (1990). Cooperation without awareness: A multiperson generalization of the minimal social situation. *Behavioral Science, 35*, 115–121.

Colman, A. M. (1995). *Game theory and its applications in the social and biological sciences* (2nd ed.). London: Routledge.

Colman, A. M. (2005). Evolution of cooperation without awareness in minimal social situations. In N. Gold (Ed.), *Teamwork: Multi-disciplinary perspectives* (pp. 216–235). Basingstoke: Palgrave Macmillan.

Cooper, D., & Kagel, J. (2008). Learning and transfer in signaling games. *Economic Theory, 34*, 415–439.

Crawford, T., & Sidowski, J. B. (1964). Monetary incentive and cooperation/competition instructions in a minimal social situation. *Psychological Reports, 15*, 233–234.

Delepoulle, S., Preux, P., & Darcheville, J.-C. (2000). Evolution of cooperation within a behavior-based perspective: Confronting nature and animats. *Artificial Evolution Lecture Notes in Computer Science, 18291*, 204–216.

Delepoulle, S., Preux, P., & Darcheville, J.-C. (2001). Selection of behavior in social situations: Application to the development of coordinated movements. *Applications of Evolutionary Computing, Proceedings, 2037*, 384–393.

Dennett, D. C. (1975). Why the law of effect will not go away. *Journal for the Theory of Social Behaviour, 5*, 169–187.

Edgeworth, F. Y. (1881). *Mathematical psychics: An essay on the application of mathematics to moral sciences*. London: Kegan Paul.

Erev, I., & Haruvy, E. (in press). Learning and the economics of small decisions. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (Vol. 2). Princeton, NJ: Princeton University Press.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., et al (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making, 23*, 15–47.

Erev, I., Ert, E., & Yechiam, E. (2008). Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions. *Journal of Behavioral Decision Making, 21*, 575–597.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review, 88*, 848–881.

Feltovich, N. (2000). Reinforcement-based vs. beliefs-based learning in experimental asymmetric-information games. *Econometrica, 68*, 605–641.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.

Gittins, J. C. (1989). *Multi-armed bandit allocation indices*. New York: Wiley.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13*, 243–266.

Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists? *The Behavioral and Brain Science, 24*, 383–403.

Holland, J. H. (1975). *Adaptation in natural and artificial systems* (2nd ed.). Ann Arbor, MI: University of Michigan Press.

Humphrey, G. (1951). *Thinking: An introduction to its experimental psychology*. London: Methuen.

Kelley, H. H. (1968). Interpersonal accommodation. *American Psychologist, 23*, 399–410.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence*. New York: Wiley.

Kelley, H. H., Thibaut, J. W., Radloff, R., & Mundy, D. (1962). The development of cooperation in the "minimal social situation". *Psychological Monographs, 76*(19, Whole No. 538).

Kraines, D., & Kraines, V. (1989). Pavlov and the Prisoner's Dilemma. *Theory and Decision, 26*, 47–79.

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes, 28*, 3–26.

Mitropoulos, A. (2001). Learning under minimal information: An experiment on mutual fate control. *Journal of Economic Psychology, 22*, 523–557.

Mitropoulos, A. (2003). An experiment on the value of structural information in a 2×2 repeated game. *Economics Letters, 78*, 27–32.

Molm, L. D. (1981). A contingency change analysis of the disruption and recovery of social exchange and cooperation. *Social Forces, 59*, 729–751.

Nash, J. (1950). Non-cooperative games. Doctoral dissertation, Princeton University. (Reprinted in *The essential John Nash*. H. W. Kuhn & S. Nasar (Eds.) (2002). Princeton, NJ: Princeton University Press, pp. 53–4.).

Nowak, M. A., & Sigmund, K. (1993). A strategy of win–stay, lose–shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature, 364*, 56–58.

Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson (Original German work published 1934).

Posch, M., Pichler, A., & Sigmund, K. (1999). The efficiency of adapting aspiration levels. *Proceedings of the Royal Society of London B, 266*, 1427–1435.

Rabinowitz, L., Kelley, H. H., & Rosenblatt, R. M. (1966). Effects of different types of interdependence and response conditions in the minimal social situation. *Journal of Experimental Social Psychology, 2*, 169–197.

Rapoport, A., & Chammah, A. M. (1965). *Prisoner's Dilemma*. Ann Arbor, MI: University of Michigan Press. p. 34.

Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior, 8*, 164–212.

Schmalensee, R. (1975). Alternative models of bandit selection. *Journal of Economic Theory, 10*, 333–342.

Sidowski, J. B. (1957). Reward and punishment in the minimal social situation. *Journal of Experimental Psychology, 54*, 318–326.

Sidowski, J. B., & Smith, M. (1961). Sex and game instruction variables in a minimal social situation. *Psychological Reports, 8*, 393–397.

Sidowski, J. B., Wyckoff, L. B., & Tabory, L. (1956). The influence of reinforcement and punishment in a minimal social situation. *Journal of Abnormal and Social Psychology, 52*, 115–119.

Skinner, B. F. (1969). *Contingencies of reinforcement*. New York: Appleton-Century-Crofts.

Smith, V. L. (1976). Experimental economics: Induced value theory. *American Economic Review, 66*, 274–279.

Smith, G. F., & Murdoch, P. (1970). Performance of informed versus noninformed triads and quartets in the "minimal social situation". *Journal of Personality and Social Psychology, 15*, 391–396.

Smith, V. L., & Walker, J. M. (1993). Monetary rewards and decision costs in experimental economics. *Economic Inquiry, 31*, 245–261.

Sozański, T. (1992). A combinatorial theory of minimal social situations. *Journal of Mathematical Sociology, 17*, 105–125.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Teigen, K. H. (2002). One hundred years of laws in psychology. *American Journal of Psychology, 115*, 103–118.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review Monograph Supplement, 2*(4, Whole No. 8), 1–109.

Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. New York: Macmillan.

Thorndike, E. L. (1927). The law of effect. *American Journal of Psychology, 39*, 212–222.

Tversky, A., & Edwards, W. (1966). Information versus reward in binary choice. *Journal of Experimental Psychology, 71*, 680–683.

Young, P. (1998). *Individual strategy and social structure: An evolutionary theory of institutions*. Princeton, NJ: Princeton University Press.