Nonlinear Dynamics, Psychology, and Life Sciences, Vol. 15, No. 2, pp. 229-252.

Innovation, Imitation, and Problem-Solving in a Networked Group

Thomas N. Wisdom[1], *Indiana University, Bloomington, IN*
Robert L. Goldstone, *Indiana University, Bloomington, IN*

Abstract: *We implemented a problem-solving task in which groups of participants simultaneously played a simple innovation game in a complex problem space, with score feedback provided after each of a number of rounds. Each participant in a group was allowed to view and imitate the guesses of others during the game. The results showed the use of social learning strategies previously studied in other species, and demonstrated benefits of social learning and nonlinear effects of group size on strategy and performance. Rather than simply encouraging conformity, groups provided information to each individual about the distribution of useful innovations in the problem space. Imitation facilitated innovation rather than displacing it, because the former allowed good solutions to be propagated and preserved for further cumulative innovations in the group. Participants generally improved their solutions through the use of fairly conservative strategies, such as changing only a small portion of one's solution at a time, and tending to imitate solutions similar to one's own. Changes in these strategies over time had the effect of making solutions increasingly entrenched, both at individual and group levels. These results showed evidence of nonlinear dynamics in the decentralization of innovation, the emergence of group phenomena from complex interactions of individual efforts, stigmergy in the use of social information, and dynamic tradeoffs between exploration and exploitation of solutions. These results also support the idea that innovation and creativity can be recognized at the group level even when group members are generally cautious and imitative.*

Key Words: *social learning, innovation, imitation, problem solving, innovation diffusion*

INTRODUCTION

With the growth of online social networks and the ever-greater breadth and depth of peer-produced knowledge available on the Internet, more attention is being paid to the effects of socially mediated (as opposed to self-generated or mass-broadcast) information on human decision-making. The direct and flexible

[1]Correspondence address: Thomas N. Wisdom, Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405. E-mail: tnwidom@indiana.edu

imitation found in humans that Bandura (1965) called “no-trial learning” seems so far to be a very rare talent among animals, and its evolution in humans is likely to have conferred important survival benefits. In fact, despite the negative connotations of “conformity” in lay usage, research has shown that “conformity bias” (the tendency to imitate the behavior of others instead of exploring solutions individually) can be adaptive (Boyd & Richerson, 1985; Kameda & Nakanishi, 2002).

Some of the negative connotations of conformity are exemplified in Asch's (1951) seminal studies of the influence of confederates on the judgments of experimental subjects, and in the popular interpretation of these studies emphasizing the negative aspects of participants' behavior (Krueger & Funder, 2004). Though recent re-examinations of Asch's results have revealed substantial independence on the part of his subjects (Friend, Raferty, & Bramel, 1990), a larger problem with this kind of experiment is in its treatment of conformity as a static influence rather than part of a dynamic, interactive process. Changes in culture and technology involve large numbers of agents interacting over long time scales, and the generation and exchange of information between people cannot realistically be treated as a series of isolated events. While ignoring socially mediated information in a laboratory setting may yield a clear study of important cognitive processes of individual problem solving and choice, in many situations it represents an unrealistic constraint, and may limit the relevance or applicability of the conclusions.

Asch, and the vast majority of the subsequent research inspired by his studies of conformity, made the methodological choice of creating judgment situations with a single human subject surrounded by accomplices of the experimenter who were scripted to give specific judgments. This method is well justified from the perspective of creating a well-controlled experimental environment for exploring factors affecting individual choices to imitate. However, the cost of constraining the judgments of all but one participant in a group is that the group dynamics of imitation cannot be revealed. The impact of individual imitation choices on the group’s performance can best be discovered by allowing all participants in a decision making task to be naturally and spontaneously influenced by one another. Understanding the group dynamics of imitation and innovation is one of the main goals of our study, and so we allow all group members the opportunity to influence, and be influenced by, each other.

Innovation and Imitation

Gabriel Tarde, one of the forefathers of social psychology, considered innovation and imitation to be "the fundamental social acts" (Tarde, 1903/1969). *Innovation* (generating locally novel solutions to problems) produces a diversity of solutions, though often at a cost in resources (time or energy) or risk (in giving up the opportunity to exploit previous reliable solutions in favor of something new or unknown). *Imitation* allows a decision maker to employ solutions discovered and passed on by others without having to develop them independently using costly trial-and-error learning. When a decision maker acquires information from others about possible solutions to a problem, the resources not expended in information gathering can be used for other aspects of problem solving, thus improving performance overall.

Of course, domination of a system by either innovation or imitation can be problematic. Excessive innovation is unhelpful because good ideas are not propagated and extended by others. Excessive imitation is maladaptive because good but suboptimal solutions are spread to the exclusion of better alternatives that are left unexplored. A common assumption is that the costs of imitation within a group are particularly detrimental – there are typically strong incentives for individuals to pursue largely imitative strategies to avoid risk, which may lead to a dearth of good new solutions that would benefit everyone in the group. So the group is better off if imitation is not universally chosen, but each individual is better off by choosing imitation. This idea has strong parallels in the large body of research on social dilemmas, in which individuals' self-interests are in conflict with the shared interests of themselves and others facing the same problem. Research on social dilemmas has consistently shown that when such interactions are repeated over time, and participants are able to recognize the dilemma they share with others, they are often able to adapt their behavior to ameliorate or avoid social dilemmas (Ostrom, 1990).

Research points to several factors that can modulate animals' imitation decisions (Laland, 2004). Many species are able to use the observed success (or lack thereof) of other individuals to guide imitation, e.g. in foraging (Templeton & Giraldeau, 1996), food preferences (Galef & Giraldeau, 2001), and mate choice (Höglund, Alatalo, Gibson, & Lundberg, 1995). This kind of performance-biased imitation is intuitively present in humans, and has been confirmed in social learning experiments (e.g. Mason, Jones, & Goldstone, 2008). In addition, a tendency to imitate solutions that are similar to one's own has been noted in studies of innovation diffusion; it is thought that this tendency allows agents to take advantage of social information while preserving the usefulness of previously-acquired information (Rogers, 2003).

Incorporating these results into our expectations for participant behavior, we can assume that there are greater incentives for imitation (in the form of immediate risk-free returns) than innovation, and that imitation will be biased primarily toward the highest-scoring solutions, and secondarily toward relatively similar solutions. An intuitive prediction is that the typical outcome is likely to show fairly small amounts of innovation, stagnant learning processes, and a lack of substantial improvements in performance over time. Such predictions are in line with the negative view of imitation as static, idle conformity, but at odds with the potential for groups to solve collective action problems given sufficient information and incentives (Ostrom, 1990).

Experiment Concept and Motivation

In order to explore the dynamics of innovation and imitation in groups, and to attempt to arbitrate between the competing predictions regarding individual incentives and group consequences, we designed an experiment in which networked groups of people explored a multidimensional problem space over a series of time-limited rounds. Participants received score feedback about their guesses after each round, and passively shared information about their guesses and scores with others. Participants could observe and copy the most recent guesses of their neighbors as they explored.

Though the task environment in this experiment did not vary over time as in previous models of social learning (e.g. Rogers, 1988; Boyd & Richerson, 1995), the problem space was designed to be more complex than in those models, in order to provide greater realism in the form of a larger number of options for exploration, as well as substantial uncertainty about optimal solutions and strategies. In addition, the difficulty of the problem was manipulated across conditions by changing the size of the problem space. A range of group sizes was tested to investigate the effects of differences in the social information environment on individual performance and strategies.

The in-game interaction between players was limited to the simple passive exchange of guess information only, so as to allow examination of imitation and innovation behavior unencumbered by more complex social interactions, such as direct communication among players. We intentionally avoided, however, the simplification of using confederates or other artificial "peers," in favor of creating real-time interactions of groups of human subjects. This methodological decision was made because we are interested in the dynamics of group behavior rather than static influence processes, and in the group as an important entity and level of analysis in its own right (Goldstone & Gureckis, 2009). If we had taken the common strategy of controlling for the group's composition and testing only one participant's behavior at a time, then we would be unable to measure group-level patterns such as the variability of solutions or the within-group distribution and dynamics of imitation and innovation. Accordingly, we purposefully sacrifice some control of an individual's group context in favor of being able to reveal group-level emergent phenomena that occur when every individual is potentially influenced by every other individual.

Our experiment on collective search connects with several core themes in nonlinear dynamics. One such theme is *decentralization*, and the notion that structured behavior may not be produced from a single authority, rule system, or architect who explicitly prescribes how the individual people or parts are to behave. Instead, structured behavior can be produced by agents that are interacting with each other locally, opportunistically, and without an eye toward macroscopic consequences. In our experiment, individuals can selectively imitate each other such that only a few solution variants are preserved for further imitation and variation, so subsequent searches are more efficiently focused on a small subset of the search space. Thus, through repeated interactions the group as a whole often converges on variants of a single solution without having been directed to do so.

At its core, our project explores the relationship between individual decisions to imitate and explore and the consequences for the group, and hence a second relevant theme of nonlinear dynamics is *emergence*. We are centrally concerned with how group behavior emerges out of individual interactions. Indeed one might go so far as to say that groups of people themselves can be interpreted as information processing systems (Goldstone & Gureckis, 2009; Theiner, Allen, & Goldstone, 2010). In our experiment, we can ask how well the group as a whole searches a highly multidimensional problem space, and the group's ability to do so does not reduce to individuals' search considered in isolation because of the pervasive interactions among individuals via imitation and differentiation.

A third theme of nonlinear dynamics that motivates our work is *interactions across levels of description*. For us, the two critical levels are individual people and groups. Whereas individual humans have probably not increased substantially in the complexity of their internal structure in recent history, the groups that they are part of have. And while social critics have argued out that people are less deeply enmeshed in their local communities (Putnam, 2001), it is undeniable that people are becoming more broadly connected, as the widespread use of internet-facilitated social networking and worldwide travel both attest. Our interest in studying the emergence of group behavior is partly motivated by an appreciation that we, as individuals, are strongly affected by the groups to which we belong. Structured cognitive behavior can be described at multiple levels, and our thoughts both depend upon and determine the social structures that contain us as elements.

A fourth, prominent theme of nonlinear dynamics is interactions among parts through *stigmergy*. Stigmergy is a form of indirect communication between agents that is achieved by agents modifying their environment and also responding to such modifications (Dorigo, Bonageau, & Theraulaz, 2000; Moussaid, Garnier, Therulaz, & Helbing, 2009). This effect has been well documented in ant swarms, in which ants lay down pheromones as they walk that attract subsequent ants (Theraulaz & Bonabeau, 1995). This allows ants to perform collective feats such as finding the shortest path between their nest and a food source using only a pheromone-facilitated bias to follow trails used by others, because shorter paths allow faster repeat journeys and thus greater reinforcement via increased use and pheromone deposition. Stigmergy is also apparent in human trail systems, with people often walking where others have walked before them, and in so doing, reinforcing, extending, and modifying the existing trail (Goldstone & Roberts, 2006). However, stigmergy is a much more general form of interaction than this, appearing for example on the web site Amazon.com when one customer's behavior in a buying books X and Y can affect a subsequent customer who has bought X and is told by Amazon that other customers like them have also bought Y. Stigmergy is also evident in music downloads, with listeners using the previous download prevalence as a cue to the merits of a song, and hence, their own likelihood of downloading the song (Salganik & Watts, 2009). This tendency for interactions among social animals to dynamically amplify a relatively small subset of individual-level behaviors is known as “symmetry-breaking” (Sulis, 2009). In our experiment, the otherwise symmetrical likelihood of options available for random exploration is broken as users create their own solution attempts and change the solutions that are available for imitation in their peers’ environment, influencing their peers to pursue similar solutions.

A fifth and final theme of nonlinear dynamics research that is relevant to our research is the tradeoff between exploration and exploitation. Adaptive agents often form decision strategies that tradeoff the benefits of exploring a problem space and exploiting the solutions found during this exploration (Hills, Todd, & Goldstone, in press; Holland, 1999). Exploration without exploitation generally leads to relatively poor outcomes because exploration is risky. Exploitation without exploration generally leads to poor outcomes because agents do not discover solutions that are appreciably better than their starting solutions. In accordance with this intuition, simulations of innovation markets by Allen (2009) showed that groups with a mixture of agents pursuing exploitative and exploratory strategies performed better than groups using just one of these strategies or a simple trial-and-error exploration strategy. In our experiment, participants make round-by-round decisions as to whether they will explore new areas of a problem space by developing novel solutions, or exploit others’ solutions via imitation, but they are also able to blend these options in their choices. In general, individuals’ search strategies partake of both imitation and exploration processes. For example, successful scientists may choose to copy the general area and approach of a well-received piece of scholarship, but once in the general vicinity of this work, they may choose to explore a somewhat new issue with somewhat new methods (Gilbert, 1977). This kind of blend between imitation and exploration could be well modeled by a reaction-diffusion process (Bar-Yam, 1997) that diffuses influential innovations to neighboring regions in science, but also reacts against exactly the same innovation being presented more than once. Similarly, the participants in our experiment can shift their solutions suddenly to the solution of a peer with a high score, but once they have shifted their solution, they then explore new solutions within a limited region around this solution.

Hypotheses and Predictions

Our central hypothesis for this experiment is that the ability of individuals to imitate each other’s tentative solutions in a search task will fundamentally change their search behavior, such that grouped participants will perform qualitatively differently and quantitatively better than those who are isolated, due to their ability to combine selective imitation and exploration to productively build on each other’s solutions.

For all participants, score performance was predicted to improve over rounds within each game, and across games within an experiment session, due to score feedback and simple learning effects. Conversely, solution turnover (the amount of change in a solution between rounds), and the diversity of solutions within a group, were expected to decrease over time as participants found more parts of the correct solution, narrowed their search to smaller areas of the board, and refined their search strategies. Grouped participants would be able to take advantage of feedback about others' solutions as well as their own, which would increase their mean performance relative to isolated participants, and produce even lower mean solution turnover and diversity as they performed relatively less of their own exploration and relied more on that of others.

Imitation of peer guesses with higher scores was expected to be prevalent, because their quality would be explicitly known and thus imitators could maximize their expected utilities by choosing the best-scoring guesses. It was assumed that larger groups would (at least initially) produce a greater diversity of solutions through random exploration, and thus provide better solutions to be imitated and propagated. Thus, an increase in score and decreases in turnover and solution diversity with increasing participant group size would be expected. However, higher-scoring solutions could be more difficult for individuals in larger groups to perceptually distinguish and select because there are more peer solutions to select from in the same limited time; therefore, the expected trends of increasing score and decreasing turnover and diversity with increasing group size could become nonlinear (leveling off or even reversing) if participants in larger groups were unable to consistently imitate the best-scoring peer solutions.

For similar reasons, participants who imitated others more often were expected to obtain better score performance (relative to those who imitated less often), conditional on whether they were able to imitate the best available solutions. As mentioned above, we also expected to observe imitation of guesses that were relatively more similar to the imitator's own previous guess; this is another strategy that reduces the risk of exploration by maximizing the use of individuals' previously-acquired knowledge of the problem space.

METHOD

One hundred forty-five participants were recruited from the Indiana University Psychology Department undergraduate subject pool, and were given course credit for taking part in the study. Participants populated each session by signing up at will for scheduled experiments with a maximum capacity of 9 persons, and were distributed across 39 sessions as shown in Table 1.

We implemented the experiment using custom software run in a web browser. The task was a round-based picture-matching puzzle game with score feedback given after each round. The goal picture that participants attempted to match was a randomly generated spline quantized to a grid of square pixels. The squares making up the spline were colored black, and the remaining squares colored white (see Fig. 1 for examples).

Table 1. Distribution of Participants Across Group Sizes.

Group size	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
# Sessions	8	6	9	3	3	5	1	2	2
# Participants	8	12	27	12	15	30	7	16	18

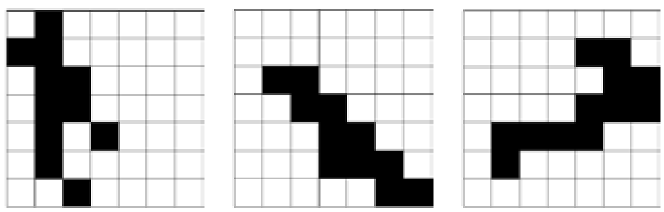


Fig. 1. Examples of randomly generated goal pictures in the experimental task.

The participant's game board was a grid of the same dimensions as the goal picture, with each square initially colored white. The color of each square on the game board could be toggled between black and white by clicking it with the mouse. Each participant's display included their own game board and most recent score (given as the number of squares, both black and white, marked correctly out of the total number of squares on the board), their neighbors' game boards and scores from the previous round, and indications of the current round in the game and the amount of time remaining in the current round (see Fig. 2). Players could copy a neighbor's most recent solution to their own at any time during the game by clicking the chosen neighbor's board with the mouse. Each game consisted of 24 rounds of 10 seconds each. After the last round in each game, participants were shown their guesses and scores for each round, along with the goal picture, and a button to click when they were ready to begin the next condition. When all participants had clicked this button, the next condition began.

Participants were instructed to maximize their scores over all rounds by matching the hidden goal picture as closely as possible. They were also informed that the picture they were supposed to match in each game was randomly generated and not representative of any particular object, shape, or symbol, and was generally not symmetrical; that the black squares were all connected to each other vertically, horizontally, or diagonally; and that the number of black squares was small relative to the size of the grid.

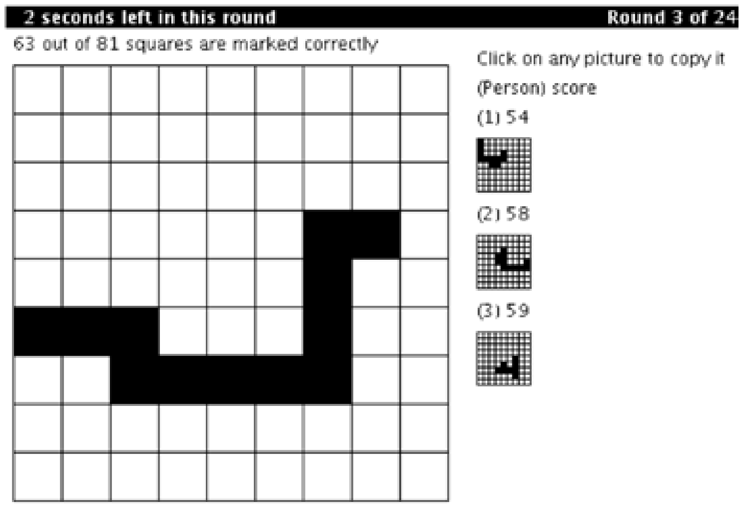


Fig. 2. Example of a participant's display**.**

A participant's *score* in each round was defined as a cell-by-cell comparison (overlap) between the participant's guess for that round and the hidden goal picture (i.e. the number of cells which the two pictures had in common), divided by the total number of squares in the goal picture, to give a percentage which could be compared between conditions of varying grid size (see Fig. 3). An *improvement* was defined as an instance of a participant obtaining a score higher than all prior scores of all players within a particular condition. Each participant's normalized *improvement share* was defined as their individually achieved proportion of the total improvements achieved by all participants in a session, multiplied by the number of participants in the session. A value of 1 indicated a "fair" share, e.g. a participant achieved one third of the improvements in a three-person session. A participant's *score rank* in a particular round was defined as the rank of their score (one being the best) among all scores in the group in that round; individuals with the same score had the same rank. *Turnover* for each round (after the first) was a measure of the amount of change between a participant's guesses over successive rounds. It was defined conversely to similarity, except that the two pictures compared were the participant's guesses from the current round and the previous round.

Imitation (a measure of whether a participant copied a neighbor's guess in a particular round) was defined as follows:

$$I_{pr} = \begin{cases} 1 : \max_i(overlap(G_{p,r}, G_{i,r-1})) > overlap(G_{p,r}, G_{p,r-1}) \\ 0 : \max_i(overlap(G_{p,r}, G_{i,r-1})) \le overlap(G_{p,r}, G_{p,r-1}) \end{cases}; p \neq i$$

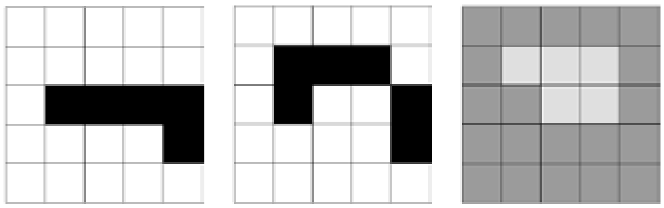


Fig. 3. Score overlap calculation: the first two pictures have 20 out of 25 squares in common (shown in dark grey on the right), so they have an overlap of 80%.

Where $G_{p,r}$ is the guess of participant p for round r, $G_{i,r-1}$ is the guess of neighbor i for the round prior to round r, *overlap* is the comparison described above for the score calculation. In other words, a participant has imitated in a particular round (I_{pr}=1) if the participant's guess is closer to the most similar neighbor's previous guess than to the participant's own previous guess. Though participants could explicitly click on a neighbor's game board to copy it as mentioned above, these click-to-copy actions were erroneously not recorded by the experiment software. Fortunately, the estimation specified in the equation above allowed us to note the occurrence of imitation whether a participant made use of this feature or performed manual copying.

Diversity (a measure of the spread of group members' guesses over the problem space within a particular round) was defined as follows:

$$D_r = 1 - \frac{\sum_s \sum_p majority(G_{spr})}{S_{tot} P_{tot}}$$

Where G_{spr} is the binary value (black or white) of square s in the guess of participant p in round r, S_{tot} is the total number of squares in the game board, P_{tot} is the total number of participants in the group, and *majority* is a binary function that conveys whether the value of G_{spr} is in agreement with the majority of participants in the group for that square in that round (0 = not in majority, 1 = in majority). Diversity as defined above is constrained to be within the 0 to 1 range, and higher values of diversity indicate more deviation of individuals' guesses from the majority guesses.

The number of squares in the game board was manipulated across two conditions: in the small board size condition, the game board was 7 squares on each side for a total of 49 squares, and in the large board size condition, the game board was 9 squares on each side for a total of 81 squares. The larger board was hypothesized to be more difficult to fully search. There were 4 repetitions of each condition, for a total of 8 games in each session. The probability distribution of scores among all possible game board states in each of the *board size* conditions described above is shown in Fig. 4. The size of the group participating in each session was treated as a covariate; *group size* ranged between 1 and 9. Another factor considered was the (randomized) position of each condition in an experiment session; this was called the *game order*.

Summary of Planned Analyses

We used the following analyses to illuminate important factors in the dynamics of activity in the collective search task. First, we performed linear regression analyses of the factors of board size, round, and game order on score, turnover, imitation, and solution diversity over the entire dataset. The analyses over board size would show the effect of changing task difficulty on participant strategies and performance, while those over round and game order would show the effects of participants' learning and adaptation to the task over time.

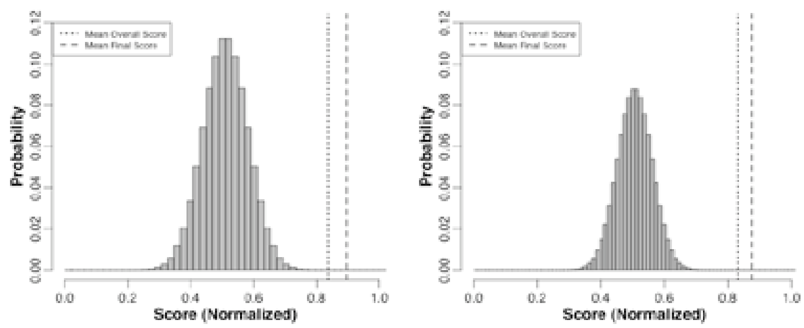
Second, we performed similar linear regressions of participant group size on score, turnover, imitation, and diversity, which would show the effect of differences in participants' social information environments on their inclinations toward exploration and exploitation, and the results of these strategy differences on performance. To examine the selectivity of imitation decisions, we measured the aggregate probabilities of participants to imitate based on relative score difference (whether the imitated guess was better or worse than their own previous guess), as well as on score rank (whether the imitated guess was the best, second-best, etc. within the group). We also examined the correlations between overall score performance and mean tendencies in turnover and imitation within individuals, as well as the aggregate measures for groups, in order to explore the relative performance of different "strategies." We analyzed distributions of individual improvement shares in order to examine the range and frequency of individual differences in performance within groups. Finally, we compared the similarity of guesses that were imitated and not imitated across all grouped participants in all rounds to test for the presence of similarity-biased imitation strategies.

RESULTS

For most analyses, dependent variables were averaged across all participants within a group to give measures for the group's aggregate activity. In this manner, the fundamental level of analysis was the group, not the individual, and dependencies between individuals within a group do not lead to elevated Type I statistical errors.

In the aggregate, participants achieved final scores of .893 (i.e. 89.3% of the way from the worst score to the best score), and average scores (over all rounds) of 0.833. Mean final scores were slightly but significantly (about 2 percentage points) lower in the larger board size condition ($t(38)=-2.88$, $p<.01$; see Fig. 4). The average guess turnover rate per round was 7.3% of the game board, and participants engaged in imitation on 25.8% of guesses. There were no significant differences in turnover or imitation rate between the two board size conditions.

The data were averaged across all participants and all conditions in each group to give dependent variable measures for each group within each round. Linear mixed-effects models were used to examine trends across rounds for each dependent variable, with a random effect of group membership on the slope over rounds. A preliminary examination of guess content confirmed experimenter observations that participants' first round guesses often (approximately 18.5% of the time) consisted of all white squares, because the resulting score would reveal how many squares were correctly marked as white, and thus how many black squares were in the solution. This was a clever and useful tactic for participants, but tended to skew trends across rounds. For this reason, the first round was excluded from analysis.



(a) Board size of 49 (b) Board size of 81

Fig. 4. Distribution of scores for all possible game board states in each board size condition. Note that the probabilities are nonzero across the entire range [0,1], and that those for scores in the upper and lower tails are too small to plot visibly due to the number of possible game board states in each condition (approximately $5.6\text{x}10^{14}$ and $2.4\text{x}10^{24}$, respectively). Thus the actual mean overall and final scores appear to be plotted outside of the distributions, but in fact are just very rare scores in the upper tails.

Analysis of score versus round showed a strongly significant positive trend ($F(1,857)=139.91$, $p<.0001$, $\beta=0.577$; see Fig. 5a). Similarly, a strongly significant negative trend was observed for turnover versus round ($F(1,857) = 169.06$, $p<.0001$, $\beta = -0.527$; see Fig. 5b). A significant negative trend was also found for imitation rate versus round: participants tended to imitate each other less often as each game progressed ($F(1,713) = 14.37$, $p < .001$, $\beta = -0.182$; see Fig. 5c). Guess diversity was subjected to a similar analysis after normalizing it for participant group size, which was accomplished by dividing all values by the mean diversity value in the second round for the appropriate group size, which was generally at or near the maximum due to the first-round blank-board phenomenon noted above. The analysis showed that the diversity of guesses in a group decreased significantly over the course of a game ($F(1,713)=33.38$, $p < .0001$, $\beta=-0.415$; see Fig. 5d).

Similar linear mixed-effects models were used to examine trends for dependent variables across game order within sessions, averaged across all participants and all rounds in each game. Once again, participant group was used as a random effect in each model. Analysis of score versus game order showed a significant positive trend ($F(1,272) = 52.69$, $p < .0001$, $\beta = 0.437$; see Fig. 6a), while a similar analysis of turnover showed a significant negative trend ($F(1,272) = 23.08$, $p < .0001$, $\beta = -0.305$; see Fig. 6b). Imitation increased significantly across game order ($F(1,246)=11.86$, $p = .0007$, $\beta = 0.214$; see Fig. 6c), while guess diversity decreased significantly across game order ($F(1,111) = 7.27$, $p < .01$, $\beta = -0.282$; see Fig. 6d).

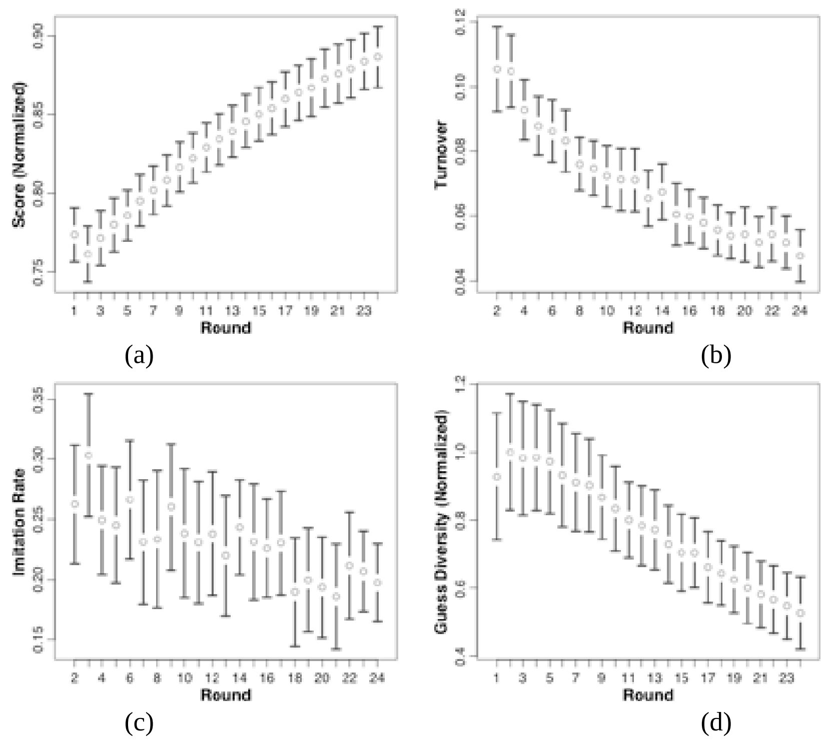
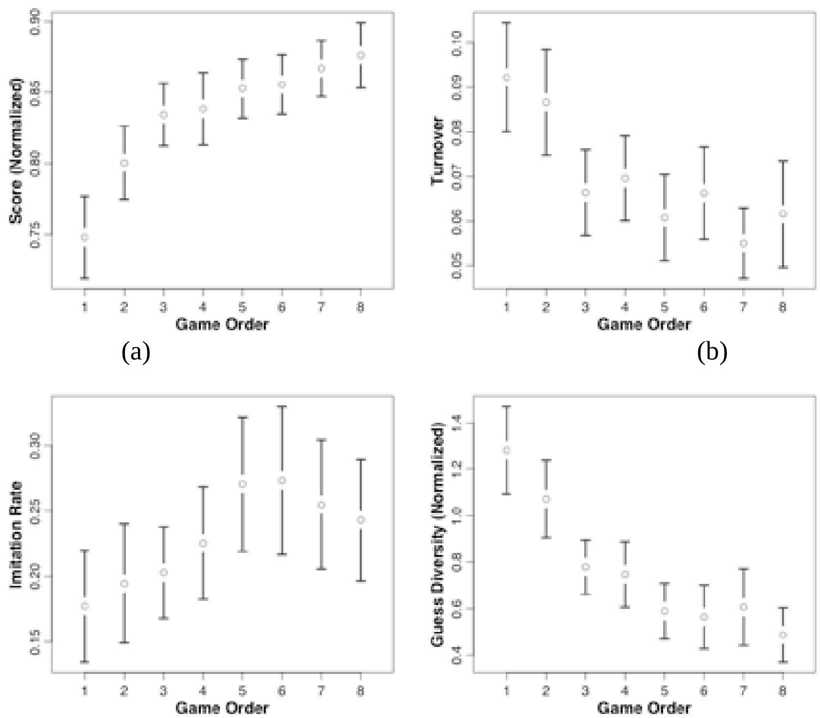


Fig. 5. (a) Mean score increased, while (b) turnover, (c) imitation rate, and (d) guess diversity decreased as more rounds were played within a game.

A mild upward linear trend was observed for score versus participant group size, as well as a marginal quadratic trend which peaked at a group size of 4 ($F(2,36) = 4.33$, $p < .05$; see Fig. 7a). Similar, stronger upward linear and quadratic trends were also found for imitation rate versus group size ($F(2,28) = 16.94$, $p < .001$; see Fig. 7c). No significant trends were found for turnover or guess diversity across group size, although both displayed substantial variance across group sizes, and both seemed to be generally inversely associated with score (see Fig. 7b & 7d). Analyses of the targets of imitation showed that nearly all instances of imitation were of those with higher scores than the imitator's (see Fig. 8a), implying that imitation behavior was generally purposeful and not random. However, there was a strong bias for imitating the top-scoring solution in smaller groups that weakened substantially (suggesting that it was more difficult) in larger groups (see Fig. 8b).



(a) (b) (c) (d)

Fig. 6. (a) Mean score and (c) imitation rate increased, while (b) turnover and (d) guess diversity decreased as more games were played within an experimental session.

To further investigate the relationship between strategy and performance, we performed regression analyses of score versus mean rates of imitation and turnover for individuals and groups. A linear regression of mean individual score versus mean individual imitation rate showed a significant positive relationship for those in group sizes of 4 or less ($F(1,49) = 7.69$, $p = .008$, $\beta = 0.368$), but none in groups of 5 or larger (see Fig. 9a). Likewise, a significant positive relationship was found between mean group score and mean group imitation rate in groups of 4 or smaller ($F(1,16) = 9.92$, $p = .006$, $\beta = 0.619$) but none in groups of 5 or larger (see Fig. 9b). Across all group sizes, there was a significant positive relationship between an individual's score and

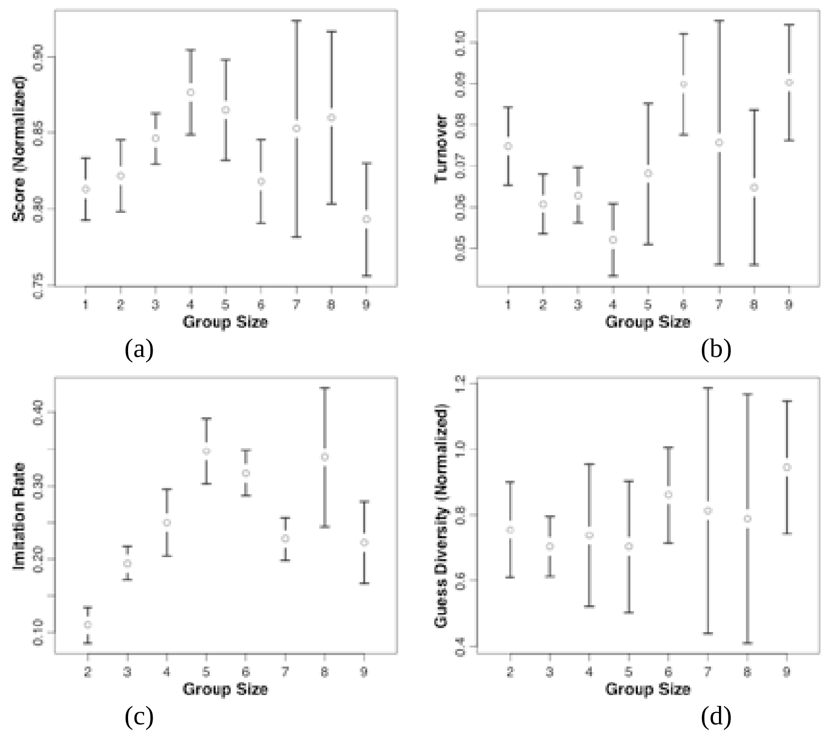


Fig. 7. (a) Mean score and (c) imitation rate showed significant quadratic trends across participant group sizes, while (b) turnover (d) guess diversity showed no significant trends.

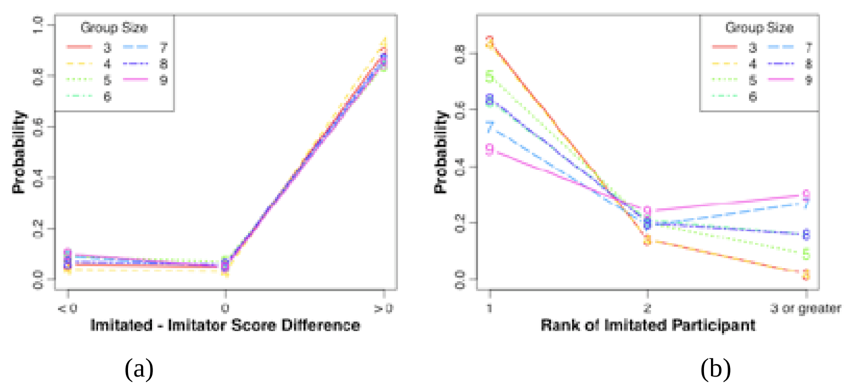


Fig. 8. (a) Nearly all imitation was of guesses with higher scores than the imitator's, and (b) there was a strong bias toward imitating top-scoring participants, which weakened in larger groups.

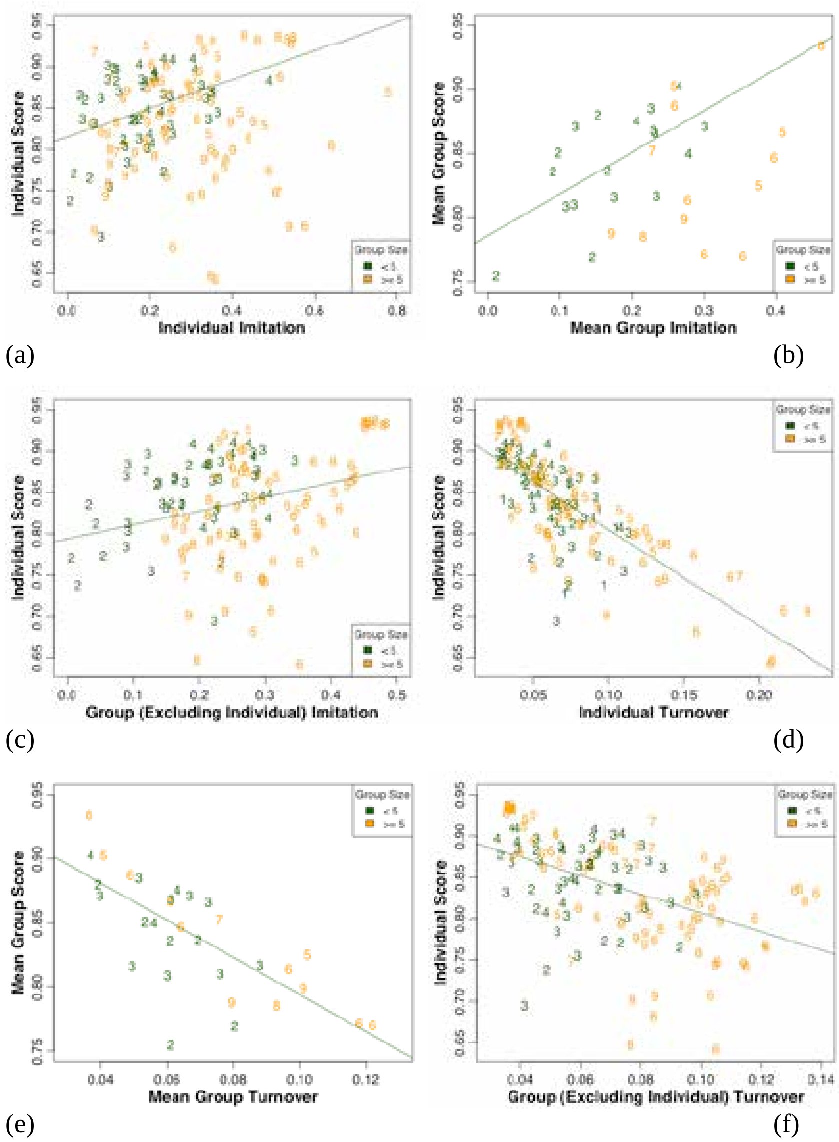


Fig. 9. (a-b) For smaller groups (< 5 participants), higher scores were associated with higher imitation rates; however, these relationships did not hold for larger groups. (c) For all group sizes, regardless of a particular individual's imitation rate, the individual's score tended to increase as the imitation rate of others in the group increased. (d-f) Higher scores were associated with lower turnover rates at all three levels of analysis noted above.

the mean imitation rate of all other group members, excluding the individual ($F(1,135) = 11.68$, $p < .001$, $\beta = 0.282$; see Fig. 9c); that is, regardless of what an individual did, she/he was likely to have a higher score if the others in her/his group imitated more often. Similar analyses of score versus mean turnover showed strong negative relationships for all group sizes, at all levels: for individual score versus individual turnover ($F(1,143) = 198.9$, $p < .0001$, $\beta = -0.763$; see Fig. 9d), group mean score versus group mean turnover ($F(1,29) = 34.59$, $p < .0001$, $\beta = -0.738$; see Fig. 9e), and individual score versus group others' mean turnover ($F(1,135) = 40.0$, $p < .0001$, $\beta = -0.478$; see Fig. 9f). In addition, for imitative guesses, we estimated a value for "innovation" by calculating the proportion of an imitative guess that was different from both the imitator's previous guess and the guess that was imitated. The correlation of scores with this value were nearly identical to those found for turnover above. The mean turnover rate for guesses that resulted in improvements was significantly smaller than that of non-improvements (0.055 for improvements vs. 0.074 for non-improvements; $t(2040) = 12.11$, $p < .0001$). No significant difference was found for mean imitation rate in improvements versus non-improvements.

An examination of participants' normalized improvement share showed a distribution with an unequal skew; approximately 57.7% of all participants achieved less than a "fair" improvement share of 1, while a small minority achieved much higher shares (see Fig. 10). We wished to differentiate the distributions of improvement sums from outcomes generated by a static random process, in order to demonstrate that such a process (e.g. all players having a uniform random probability of discovering improvements) could not easily explain these results. Toward this end, we constructed a Poisson distribution of improvement sums for each participant group with the lambda (mean value) parameter equal to the mean improvement sum for that group. and found the range of values which contained 50% of the density in this artificial Poisson distribution. In over 80% of groups with more than one participant, the 50% density range from the associated Poisson distribution contained less than 50% of the density of the actual individual improvement sums, indicating that they had a greater skew than would be expected by chance.

A comparison between the similarity of imitators' most recent guesses to those which they imitated, and to those which they did not imitate, revealed that there was significantly greater similarity to imitated guesses than to non-imitated guesses (77.7% for imitated vs. 72.3% for non-imitated; $t(4914) = -18.23$, $p < .0001$; see Fig. 11a). In other words, imitation tended to be biased toward guesses that were more similar to the imitator's own prior guess. This difference held over all rounds within a game (see Fig. 11b), even though mean guess diversity decreased such that solutions generally converged (see Fig. 5d). No significant trends were observed in linear regressions of similarity versus imitated score rank, or the score difference between imitator and imitated participants.

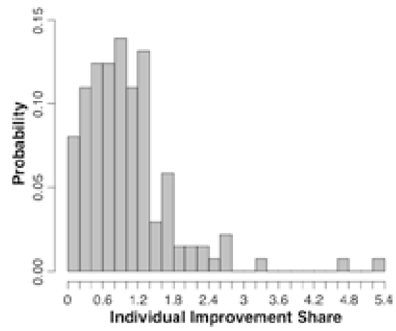


Fig. 10. Histogram showing the unequal distribution of improvements across the participants within groups. (A value of 1 indicates an even share, e.g. an individual achieved one-third of the total improvements in a three-person group.)

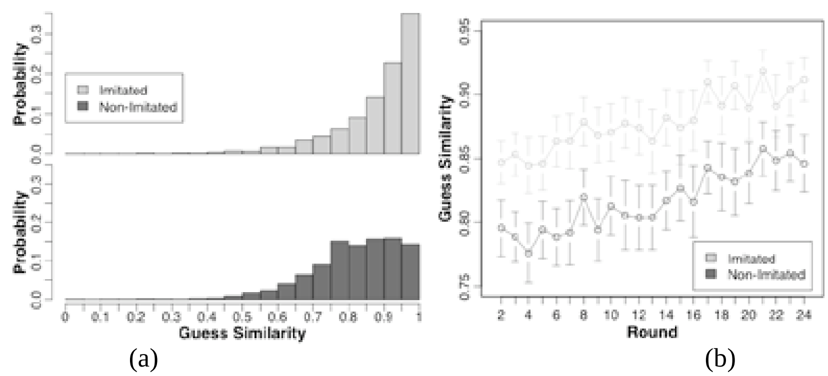


Fig. 11. Similarity bias for imitation. (a) Imitators' previous guesses showed greater similarity to the guesses they imitated than to those they did not imitate. (b) The bias toward imitating more similar guesses was consistent across all rounds in a game.

DISCUSSION

Dynamics and Strategies

The larger board size had a significant negative effect on final scores, which confirmed its use as a proxy for problem difficulty, but this change in difficulty had no significant effect on the other dependent variables. However, we observed revealing patterns in participants' behavior that gave some clues about their strategies.

Increasing mean scores across rounds and game order showed that participants in groups learned the task and that their drawings converged toward the hidden goal picture over rounds in each game and over the course of the entire session. Participants accomplished their improvements through the use of fairly conservative strategies, as evidenced by the low mean turnover rate. Furthermore, the dynamics of these strategies caused solutions to become increaseingly entrenched over the course of the game. This happened in two ways (which may have been mutually reinforcing): participants' rates of imitation and general turnover decreased across rounds, and the imitation that did occur was biased toward more similar solutions. This entrenchment carried over to the group level as well, shown by the decreasing group solution diversity across rounds. Of course, this result is likely partially due to participants converging toward the goal picture, but the average final score of approximately 89% of the maximum suggests that groups often converged before finding the optimal solution.

The problem space used in this task is quite large (on the order of 5×10^{14} possible solutions for the smaller board size), and any change of more than one pixel can easily result in a situation where score-decreasing changes cancel out score-increasing changes, which makes score feedback difficult to interpret. Thus it makes sense that high scores were consistently associated with low turnover, and that mean turnover was significantly lower for guesses that resulted in improvements. Rather than large, revolutionary changes, participants made small, incremental improvements by changing only a few cells, typically just one. These small changes allowed participants to make accurate comparative inferences about their effects on score.

The unequally skewed distribution of improvements within each group showed that not all participants were skilled at finding good new solutions, but imitation allowed some participants to take advantage of other participants innovations and maintain high mean scores. The fact that average turnover was higher for non-improvement guesses shows, however, that non-improvers were not just idly waiting to imitate others' improvements.

Benefits of Imitation

Imitation was biased toward higher-scoring and more-similar guesses, as expected. The latter allowed participants a way to take advantage of others' good solutions while maintaining low turnover and higher continuity with their own previous guesses, preserving the value of their existing knowledge of the problem space.

The association of higher scores with greater imitation rates at both the individual and group levels (at least for participants in smaller groups, which were better able to distinguish the top-scoring peer solution when imitating) shows that imitation is not necessarily harmful to innovation and performance improvements. Though it might have been reasonable to expect improvements to be associated with a lower imitation rate (because those who only imitate others cannot do better than those they imitate), we found that the rate of imitation was about the same among improvements and non-improvements, which means that improvements were often achieved by imitating a relatively successful participant's solution and then slightly tweaking this solution. Once tweaked, the improved solution was then available to other participants, including the individual who was originally imitated. The association of high individual scores with high imitation rates by others in the group (regardless of the individual's behavior) reinforces the idea of a systemic benefit for imitation rather than a view of imitation as a purely self-benefiting act. It may be that, regardless of the intentions of individuals, imitation benefits the group by acting as a filter for propagating and preserving the better solutions available in a group at a given time, as was found in a recent competition of social learning strategies in a simulated environment (Rendell et al., 2010). A related result was observed in the innovation market simulations of Allen (2009), which showed that a mixture of agent "firms" with imitative and individual learning strategies resulted in better average group performance than either of the above approaches or simple "Darwinian" trial-and-error learning alone as universal strategies for all agents.

The combination of exploratory variation and selective retention processes has been explored previously as a basis for creativity (Simonton, 1988; Campbell, 1960). Building on this work, Guastello (2002) found correspondences in the use of useful initial prototype ideas as "attractors" for further innovations in both individual and group creativity contexts. We found parallel results in this experiment: successful individuals tended toward low-turnover, similarity-biased imitation strategies, and successful groups displayed diversity-reducing, convergent filtering of individual variations.

Imitation Information Overload?

There was an unexpectedly lower benefit for imitation in larger groups. Larger groups could be expected to provide more information about the distribution of solutions to their members, because there are more models for each group member to observe. If individuals can effectively process this information, each can make more informed decisions about whom to imitate and what changes to make to their guesses. However, it may have been that for larger group sizes, the larger amount of information provided was more difficult to search and compare, which led to somewhat more random imitation decisions (as indicated by the weaker bias toward imitating the top-scoring guess) and thus poorer convergence on good solutions. It is unlikely that this was a statistical artifact of purely random imitation choices among more options, because there was a universal tendency (across all group sizes) to imitate better-scoring peers than oneself, a relatively easier thing to accomplish but decidedly non-random. So, though score information was readily available, it may have been subject to cognitive load effects (Sweller, 1988). Larger groups would generally be expected to show a greater upper variance in solution quality due to chance, but an increasing inability to properly distinguish good solutions would cancel out this benefit.

CONCLUSIONS

Though our assumptions about participants' imitation strategies (favoring higher-scoring and similar guesses) were empirically supported, the related predictions of poor performance were not borne out. There was a consistent benefit for individuals to be in high-imitation groups regardless of their own behavior, and imitation was also associated with better individual and group performance when it could be done selectively and accurately.

In short, the theoretical imitation-related social dilemma did not cause a tragic outcome, which is consistent with the benefit for "conformity bias" found in previous models and experiments of social learning (Boyd & Richerson, 1985; Kameda & Nakanishi, 2002), and with the literature on social dilemmas with repeated interactions and recognition of a collective action problems (Ostrom, 1990). In fact, the degraded performance of larger groups was likely due to an impediment to imitation.

The group benefits that we found for imitation have implications for our conceptualization of human creativity. Creativity is often times contrasted to imitation and conformity. Imitation is even viewed as antithetical to creativity, as we tend to portray creative thinkers as independent spirits, as mavericks that buck norms and conventions. In contrast to this portrayal, an interpretation of our results is that useful creativity derives from imitation. Successful innovations that offer tangible improvements for a group come from individuals that were imitating previous improvements. “Art begins in imitation and ends in innovation” as Mason Cooley said, and quantifiable improvements to search problems apparently follow a similar trajectory.

More generally, creativity is typically viewed as a capacity of individuals. One contribution of embedding a search and innovation problem in a collective behavior context is to highlight that attributions of creativity can be applied at the group rather than individual level. The need for positing group-level creativity is indicated by the otherwise paradoxical effect that creativity increases as imitation increases. The paradox is resolved because the creativity is at the group level while the imitation is at the individual level. The group as whole is able to discover novel and successful solutions to a complex search problem that would have been missed if the members of the group had individually been genuinely independent and solitary pioneers. The ability of interactive imitation to bias the exploratory innovations of group members toward better solutions discovered previously by others is a form of the “symmetry breaking” discussed by Sulis (2009). The kind of creativity exhibited by our groups is certainly less striking and visionary than the creativity shown by artists and scientists. However, we believe that our paradigm offers a replicable method for investigating some of the component processes of human creativity, including exploration, filtering, and refinement. Furthermore, the collective nature of our paradigm highlights an underemphasized aspect of creativity – its larger social context and the dynamics of innovation exchange among individuals collectively finding solutions within a problem space that is so large that it cannot be feasibly covered by a single individual (Csikszentmihalyi, 1988). When individuals are constrained to pursue only small, controlled innovations in order to make reliable inferences in a large problem space, the choice of where to begin (that is, whom to imitate) becomes much more important to the process of discovery.

In this study we found that collective search behavior differs markedly from predictions made based on a static view of social learning, and that differences in the size of a group can have significant nonlinear effects on their behavior and performance. The results present several intriguing areas for further study, incorporating many areas of nonlinear dynamics including decentralization, emergence, interactions across levels of description, stigmergy, and tradeoffs between exploration and exploitation in search processes. Overall, there are strong implications in these data for testing the predictions of past work in the area of group creativity, and potential applications to many real-world problems of knowledge sharing and the progress of technology and culture.

ACKNOWLEDGMENT

The authors wish to express thanks to Seth Frey, Winter Mason, and Michael Roberts for helpful suggestions on this work, and to Frances Kidwell for help with running the experiment. This work was funded by National Science Foundation REESE grant 0910218 and National Science Foundation IGERT training grant 0903495. More information about the laboratory can be found at http://cognitrn.psych.indiana.edu

REFERENCES

Allen, P. (2009). Complexity, evolution, and organizational behavior. In S. J. Guastello, M. Koopmans, & D. Pincus (Eds.), *Chaos and complexity in psychology: The theory of nonlinear dynamical systems* (pp. 41-72). New York: Cambridge University Press.

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.) *Groups, leadership, and men* (pp. 177-190). Pittsburgh: Carnegie Press.

Bandura, A. (1965). Vicarious processes: A case of no-trial learning. In L. Berkowitz (Ed.) *Advances in experimental social psychology, Vol. II* (pp. 1-55). New York: Academic Press.

Bar-Yam, Y. (1997). *Dynamics of complex systems*. Reading, MA, Addison-Wesley.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. J. (1995). Why does culture increase human adaptability? *Ethology and Sociobiology, 16*, 125-143.

Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review, 67,* 380-400.

Csikszentmihalyi, M. (1988). Society, culture, and person: a systems view of creativity. In R. J. Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives* (pp. 325-339). Cambridge, MA: MIT Press.

Dorigo, M., Bonabeau, E., & Theraulaz, G. (2000). Ant algorithms and stigmergy. *Future Generation Computer Systems*, *16*, 851-871.

Friend, R., Rafferty, Y., & Bramel, D. (1990). A puzzling misinterpretation of the Asch 'conformity' study. *European Journal of Social Psychology, 20*, 29-44.

Galef, B. G. Jr., & Giraldeau, L. A. (2001). Social influences on foraging in vertebrates: causal mechanisms and adaptive functions. *Animal Behaviour, 61*, 3-15.

Gilbert, G. N. (1977). Competition, differentiation and careers in science. *Social Science Information*, *16*, 103-123.

Goldstone, R. L., & Gureckis, T. M. (2009). Collective Behavior. *Topics in Cognitive Science, 1*, 412-438.

Goldstone, R. L., & Roberts, M. E. (2006). Self-organized trail systems in groups of humans. *Complexity*, *11*, 43-50.

Guastello, S. J. (2002). *Managing emergent phenomena: Nonlinear dynamics in work organizations.* Mahwah, NJ: Lawrence Erlbaum Associates.

Hills, T., Todd, P. M., & Goldstone, R. L. (in press). The central executive as a search process: Exploration and exploitation in generalized cognitive search processes. *Journal of Experimental Psychology: General*.

Höglund, J., Alatalo, R. V., Gibson, R. M., & Lundberg, A. (1995). Mate-choice copying in black grouse. *Animal Behaviour*, *49*, 1627-1633.

Holland, J. H. (1999). *Emergence: From chaos to order*. New York: Basic Books.

Kameda, T., & Nakanishi, D. (2002). Cost-benefit analysis of social/cultural learning in a non-stationary uncertain environment: An evolutionary simulation and an experiment with human subjects. *Evolution and Human Behavior*, *23*, 373-393.

Krueger, J. I., & Funder, D. C. (2004). Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences*, *27*, 313-376.

Laland, K. N. (2004). Social learning strategies. *Learning & Behavior*, *32*, 4-14.

Mason, W. A., Jones, A., & Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General*, *137*, 422-433.

Moussaid, M., Garnier, S., Theraulaz, G., & Helbing, D. (2009). Collective information processing in swarms, flocks and crowds. *Topics in Cognitive Science*, *1*, 469-497.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action.* New York: Cambridge University Press.

Putnam, R. D. (2001). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., Fogarty, L., Ghirlanda, S., Lillicrap, T., & Laland, K. N. (2010). Why copy others? Insights from the social learning strategies tournament. *Science*, *328*(5975), 208-13.

Rogers, A. R. (1988). Does biology constrain culture? *American Anthropologist, 90*, 819-831.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.

Salganik, M. J., & Watts, D. J. (2009). An experimental approach to the study of collective behavior in cultural markets. *Topics in Cognitive Science, 1*, 439-468.

Simonton, D. K. (1988). Creativity, leadership, and chance. In R. J. Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives* (pp. 386-426). Cambridge, MA: MIT Press.

Sulis, W. (2009). Collective intelligence: Observations and models. In S. J. Guastello, M. Koopmans, & D. Pincus (Eds.), *Chaos and complexity in psychology: The theory of nonlinear dynamical systems* (pp. 41-72). New York: Cambridge University Press.

Sweller, J. (1988) Cognitive load during problem solving. *Cognitive Science, 12*, 257–285.

Tarde, G. (1969). *The laws of imitation.* (Elsie Clews Parsons, Trans.). Chicago: University of Chicago Press. (Original work published 1903.)

Theiner, G, Allen, C., & Goldstone, R. L. (2010). Recognizing group cognition. *Cognitive Systems Research, 11*, 378-395.

Theraulaz, G., & Bonabeau, E. (1995). Modeling the collective building of complex architectures in social insects with lattice swarms. *Journal of Theoretical Biology, 177*, 381-387.

Templeton, J. J., & Giraldeau, L.-A. (1996). Vicarious sampling: The use of personal and public information by starlings foraging in a simple patchy environment. *Behavioral Ecology and Sociobiology, 38*, 105-113.