

# Emergent Collective Sensing in Human Groups

Peter M. Krafft (pkrafft@mit.edu)\*, Robert X.D. Hawkins (rxdh@stanford.edu)†, Alex “Sandy” Pentland (pentland@mit.edu)‡, Noah D. Goodman (ngoodman@stanford.edu)†, Joshua B. Tenenbaum (jbt@mit.edu)\*

\*MIT Computer Science and Artificial Intelligence Laboratory, †Stanford Department of Psychology, ‡MIT Media Lab

## Abstract

Despite its importance in human society, the nature of human collective intelligence remains largely a mystery. Recent studies have begun to probe and enumerate the descriptive features that partially explain the presence of collective intelligence in certain human groups, but the mechanisms for the emergence of this distributed information processing ability, whether on small or large scales, are poorly understood. On the other hand, there have recently been careful experiments that have exposed the mechanisms of collective intelligence in nonhuman animal species. We leverage the experimental design of one such recent study of collective sensing in groups of fish to better understand the mechanisms for the emergence of collective intelligence in human groups, and to better understand how the human ability of theory of mind contributes to those mechanisms. We find that humans in our experiments act at a high level much like the fish from the original experiment but with two key differences. Firstly, humans appear to frequently disengage from other individuals present in the environment in phases of independent exploration. Secondly, effective humans appear to engage in targeted, as opposed to indiscriminate, mimicry behavior: they infer which other agents believe they are currently performing well and selectively copy them. The specific behavioral mechanism we identify that benefits from these distinctively human activities leads to the emergence of collective sensing in our dynamic task environment at group sizes and on times scales orders of magnitudes smaller than were observed in fish.

**Keywords:** collective intelligence; distributed cognition; social cognition; social computation; online experiments

## Introduction

Some of the most well-known phenomena of collective behavior are failures of collective intelligence. Mobs, market panics, and mass hysteria draw attention because of their apparent irrationality and painful consequences. However, the successes of collective intelligence are as remarkable as the failures are devastating. The richness of human culture, the incredible pace of our technological developments, and the gradual progression of our scientific understanding of the universe stand out as both distinctively human and heavily reliant on the emergent behavior of the interactions of many individuals. Even at a less grandiose level, humans regularly agree to work together to accomplish tasks that no individual could accomplish alone via a process of dynamic cooperative team behavior that is hypothesized to be uniquely human (Tomasello, 2014). Yet little is known about the specific mechanisms underlying this synergistic process of self-organization, even for small groups.

In general, because of the difficulties in conducting and collecting data from real-time multi-subject experiments or studies, the quantitative study of collective behavior has been largely theoretical. Many models of flocking behavior, collective decision-making, and other simple collective behaviors have been proposed based on observations, intuitions,

and analogies. More recently, though researchers have begun conducting careful experiments to test and refine these models (Goldstone & Gureckis, 2009). However, many of these experiments, with some notable exceptions (Goldstone, Roberts, Mason, & Gureckis, 2008; Kearns, 2012), have been conducted using nonhuman animal subjects. The logistical difficulties of having multiple human participants simultaneously interacting in a real-time environment are still widely regarded as prohibitive. As a result we are quickly developing a better understanding of the collective behavior of ants (Pratt & Sumpter, 2006), bees (Seeley & Buhrman, 1999), cockroaches (Amé, Halloy, Rivault, Detrain, & Deneubourg, 2006), and fish (Ward, Sumpter, Couzin, Hart, & Krause, 2008), but our empirically-grounded quantitative understanding of human collective behavior remains limited.

Furthermore, given the distinct capabilities of humans as compared to other animals, there is little reason to believe that the models of collective behavior that have been developed for other animals would be appropriate for humans. In this work, we harness recent technical advances in running real-time, collective behavior experiments on the web (Hawkins, 2014) to develop and test a behavioral model of collective human behavior. Our goal is to understand in what ways human performance differs from that of nonhuman animals.

As a point of comparison, we consider a recent experiment designed to study the collective behavior of a particular species of fish that is one of the clearest illustrations of collective intelligence in a nonhuman animal group (Berdahl, Torney, Ioannou, Faria, & Couzin, 2013). In this experiment, researchers studied a species of fish that prefers to reside in darker areas of the water, presumably to avoid predators. The researchers took advantage of this natural propensity of the fish and projected time-varying spatially correlated light fields into a fish tank. The researchers then studied the effectiveness of the fish at finding the darker areas of the tank as a function of the number of fish participating in the task. The researchers found that average group performance increased significantly as a function of group size, and they identified two simple behavioral mechanisms driving this improvement: first, individual fish tended to move more slowly in darker areas and second, individual fish also tended to turn towards conspecifics. The researchers argued that the combination of these mechanisms generated an emergent collective gradient sensing ability in groups of fish that had been absent in individual fish.

This experiment provides a beautiful example of highly effective, intelligent behavior at the group level emerging from minimal intelligence at the individual level. However, while

these simple mechanisms do give rise to surprisingly effective group behavior, they only lead to substantial gains in performance for groups of 50 to 100 fish. In contrast, we expect humans in a similar task to show significant group gains with much smaller group sizes. Most notably, we expect that humans should be able to make use of robust theory of mind, the ability to draw inferences about the underlying mental states of other players, to better utilize social information. Humans are known to be highly effective at learning from others, as well as balancing this so called “social learning” with independent exploration and exploitation of existing personal knowledge (e.g. Kameda & Nakanishi, 2003; Wisdom, Song, & Goldstone, 2013). Yet while we have a good understanding of the benefits of theory of mind and social learning to individuals, it is not known what emergent collective behavior strong theory of mind affords.

To better elucidate these potential differences between humans and fish, we developed a version of the gradient-sensing task for human participants. Specifically, we recreated the environment used by Berdahl et al. (2013) as an online real-time multi-player game. In our experiment, participants controlled avatars in a virtual world. Every location in this world corresponded to a score value that changed over time, and participants were rewarded proportionally to their cumulative scores in the game. The score of a player at a particular point in time was simply determined by the location of that player in the virtual world. Our incentives for participants to achieve high scores were designed to simulate the fishes’ preferences for darker areas in their environment. The players either played alone or in groups of varying size. As parallels to the previous researchers’ experiments with fish, we study how the performance of human groups changes as group size increases, and what behavioral mechanisms our human participants use.

## Methods

**Participants** We recruited 563 unique participants from Amazon Mechanical Turk to participate in our experiment. All participants were from the U.S., and passed a comprehension test based on their understanding of our instructions. After excluding 72 participants due to inactivity or latency after the game had started, and 6 others for disconnecting in the first half of the game, we collected usable data from 437 participants in 224 groups ranging in size from one to six individuals. Since we were only able to collect one group of size six, we ignore this group for our analysis.

**Stimuli** A  $480 \times 285$  array of score values was pre-generated for each 125ms time interval using the method reported by Berdahl et al. (2013). First, a ‘spotlight’ of high value was created, which moved smoothly from one random location to another random location over time. This spotlight was then convolved with a uniform field of Gaussian noise to yield a complex landscape with many transient local maxima but a unique time-varying global maximum.

We manipulated the weighting between the noise field and the spotlight to generate different task difficulties. We

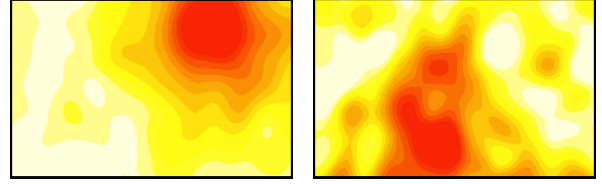


Figure 1: Example score fields in the easy (left) and medium (right) difficulties at particular points in time. Red areas in this figure indicate higher scoring areas.

used two levels of noise, corresponding to the “easy” and “medium” noise levels reported by Berdahl et al. 113 individuals (63 groups) were assigned to the easy difficulty and 324 individuals (161 groups) were assigned to the medium difficulty. To decrease variability and increase statistical power, we generated only four distinct score fields per difficulty, so multiple groups experienced the same fields. To discourage inactivity, players were not awarded any points while running into a wall, regardless of the current score field. Example of score fields are shown in Figure 1.

To model the capabilities that Berdahl et al. ascribed to fish in their environment, we severely restricted participants’ information about the score field: they were shown the score at their avatar’s location, displayed at the top of their screen, but could *not* see the scores other players were obtaining or the scores at any other locations. The positions, directions, and speeds of all other players were visible to each player. This information was updated in real-time every eighth of a second. A screenshot of the interface we used for the game is shown in Figure 2.

Players controlled their avatars using the ‘left’ and ‘right’ keys to turn (at a rate of  $40^\circ$  per second) and could hold the ‘spacebar’ to accelerate. Their avatars automatically moved forward at a constant velocity of 136 pixels per second whenever the spacebar was not depressed, and instantaneously increased to a constant velocity of 456 pixels per second whenever the spacebar was depressed. Each participant played in a single continuous game lasting for 6 minutes. We chose the speed values to match the bounds that Berdahl et al. reported for their fish, and we also matched the playing area dimensions and game length to the parameters of those experiments.

**Procedure** After agreeing to participate in our experiment, reading our instructions, and successfully completing our comprehension test, individuals were redirected to a waiting room, where they stayed for up to 5 minutes or until a pre-assigned number of other players joined the game. While in the waiting room, participants could familiarize themselves with the controls of the game. The player’s score in the waiting room was displayed as “—” unless the participant was against a wall in which case the score would change to a red “0%”. Wait times were shown to have no measurable effect on individual performance (linear regression slope  $1.993e-06$ , with 95% confidence interval  $[-1e-05, 1.4e-05]$ ). As in the ac-

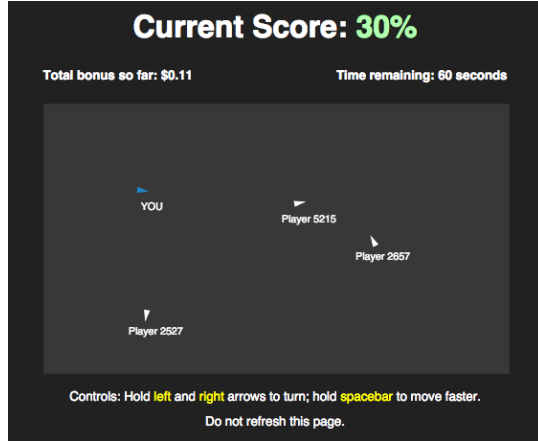


Figure 2: A screenshot of the interface that participants saw. The score displayed corresponds to the value of the score field at the location that the player is occupying.

tual game, participants in the waiting room would be removed for inactivity (if the player’s browser was active in another tab for more than 15 seconds or if the player’s avatar unmoving against a wall for 30 seconds) or for unacceptably high latencies (if the player’s ping response latency was greater than 125ms for more than 36 seconds). Participants were paid 50 cents for reading the instructions, and could receive a bonus of up to \$1.25 during six minutes of gameplay. The bonus was computed to be the player’s cumulative score divided by the total length of the game.

We implemented this experiment using the MWERT framework (Hawkins, 2014), which combines a set of recent web technologies capable of handling the challenges of real-time, multi-player web experiments, including Node.js, the Socket.io module, and HTML5 canvas. We extended the MWERT framework in several ways to handle the challenges posed by hosting larger groups of players; full source code and a demo are available online.<sup>1</sup>

## Results

We find that the performance of human groups in this game significantly increases as a function of group size in the easy task difficulty but only marginally increases as a function of group size in the medium task difficulty. Individual performance as a function of group size in each of these cases is shown in Figure 3. A linear regression on the easy difficulty points produces a significant positive slope of 0.0238 and a 95% confidence interval (95% CI) of [0.006,0.041]. A linear regression on the medium difficulty points produces a marginally significant positive slope of 0.0068, 95% CI [−0.001,0.015]. However, even this marginally significant result is driven entirely by the effect of group size in one of score fields. This particular score field (the red line in Figure 3) displays a significant effect of group size with a positive

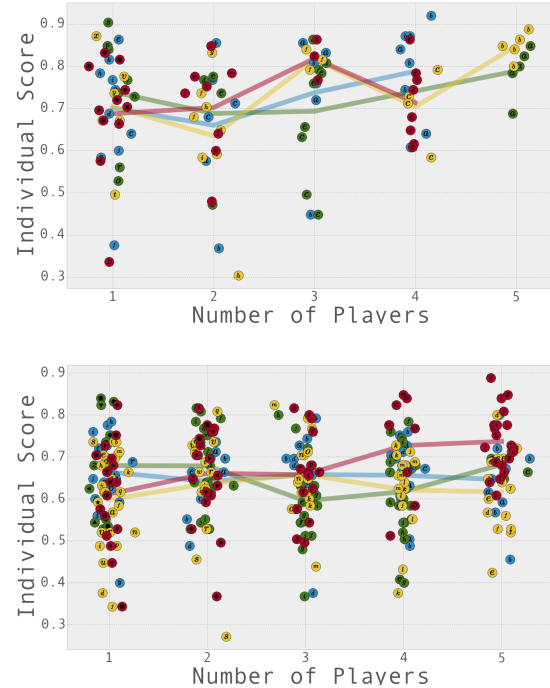


Figure 3: Individual performance as a function of group size in the easy (top) and medium (bottom) difficulties. Colors within each plot indicate which score field the player experienced. Symbols within each point identify groups of individual that played together. Lines indicate the means of each group for each score field.

slope of 0.0306, 95% CI: [0.015,0.046], while none of the others do. Qualitative inspection revealed that this particular score field seems to share spatial properties more similar to the easy score fields, which may explain the strength of the effect in that particular score field.<sup>2</sup> Overall these results indicate that larger groups do tend to perform systemically better on our task than those in smaller groups, at least in the easy game difficulty.

In order to understand the factors that may contribute to the increase in performance in the easy condition as a function of group size, we examine the behavior of the players in our games. We assume a simple state-based representation of player behavior. We then look at the factors that determine the states each player occupies at any point in time, as well as the relationship between those decisions to occupy particular states and player performance. Specifically, we assume that at any particular point in time a player is either *exploring*, *exploiting*, or *copying* (see Rendell et al., 2010, for a similar classification). Conceptually, a player is exploring if that player is looking for a good location to exploit, a player is exploiting if that player has found a location where they want to remain, and a player is copying if that player is intending

<sup>1</sup>[https://github.com/hawkrobe/couzin\\_replication](https://github.com/hawkrobe/couzin_replication)

<sup>2</sup>Indeed, a mixed-effects regression including group and score field as random effects showed a significant improvement in fit over the pure fixed effects model, and also revealed larger variability due to score field in the ‘medium’ condition than the ‘easy’ condition

to move to the location of another player.

We empirically determine the state of each player at each point in time using a set of hand-tuned filters. All of these filters depend only on information that is observable to any player in the game (i.e., the filters do not depend directly on the scores of any individuals), and hence we can use the inferred states of players as a proxy for what other players might infer as the states of those players. Moreover, the states will not be by definition related to performance, allowing for us to observe the relationship between state and performance.

Exploiting a particular location in this game is nontrivial for players since the players always move at least at a slow velocity. In order to attempt to stay put, a player can either meander around a particular location or can hold down one of the arrow keys while moving at a slow speed, which creates a tight circular motion around a particular location. We call this second activity “spinning” because of its distinctive appearance. We then define a player as exploiting if the player is spinning for 500ms or if the player moves at the slow speed for 3 seconds and has not traveled more than two thirds of the possible distance that the player could have traveled in that time. The second condition is supposed to capture the meandering behavior of individuals who have not discovered how to spin. Copying behavior is more difficult to identify, but is often characterized by fast directed movements towards other players. We thus define a player as copying if the player is moving in a straight line at the fast speed towards any particular other player consistently for 500ms. We define a player as moving towards another player if the second player is within  $60^\circ$  on either side of the first player’s straight-line trajectory. Finally, we define a player as exploring if the player is neither exploiting nor copying. Thus a player will be classified as exploring if that player is either moving slowly but not staying in the same general location, if the player is moving quickly but not towards any particular person, or if the player is moving quickly and turning.

We use these filters to analyze how players behave in our game. First, we compute the probability of a player being in a particular state conditional on the current score that the player is receiving by observing the fraction of players in each state at each score level. We find that the probability of a player occupying a particular state is closely related to that player’s score. Specifically, players in higher scoring locations are more likely to be exploiting than exploring or copying, but the probability that a player is exploring or copying increases as the player’s score decreases. These results, which are visualized in Figure 4, suggest that players are choosing their states relatively rationally. Players will tend to remain in good areas and will leave bad areas quickly either by exploring independently or by copying other individuals.

Second, we find substantial variation in the types of copying behavior that different individuals use. Some individuals appear to focus their copying behavior on other players who tend to have higher scores, whereas other individuals appear to be less discriminating in their copying behavior. Moreover,

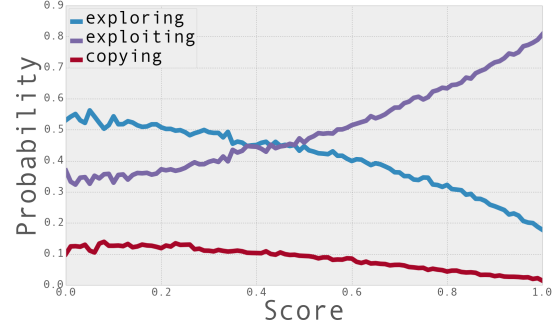


Figure 4: The probability of an individual being in a particular behavioral state as a function of the individual’s score.

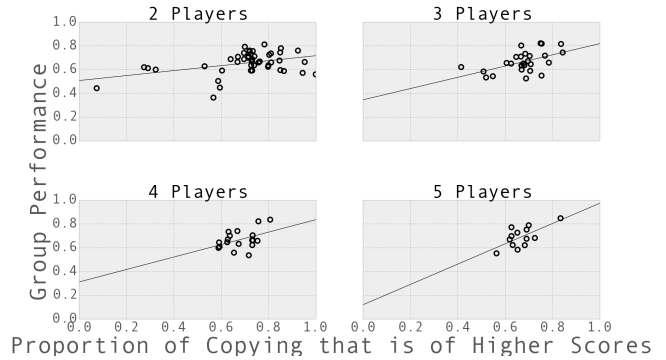


Figure 5: Average group performance as a function of the fraction of total time spent copying that consists of “intelligent copying”—copying of an individual with a higher score. Lines are individually fitted regression lines.

as shown in Figure 5, groups that contain individuals who focus their copying behavior on higher scoring individuals achieve significantly higher performance in our task (slope: 0.2639, 95% CI: [0.145, 0.383]). This result, though subject to the confounding of correlation and causation, could be explained by theory of mind assisting in individual and group performance. A player who is able to accurately infer whether another player is receiving a better score may be able to achieve higher performance on our task by leveraging these inferences to more effectively copy others.

## Behavioral Model

The trends we observe suggest a potential set of behavioral mechanisms that effective human groups may use in our task. We propose that each player in an effective group chooses a state based on the following rules:

1. If the player is in a good area, the player will remain in that area exploiting.
2. If the player is not in a good area, and the player perceives another person as possibly having a higher score, the player may choose to copy that person.
3. Otherwise the player will explore independently.

This model suggests that players can improve their individual performance by copying others who are exploiting instead of wasting time by exploring many poor quality areas. The model also has interesting emergent collective properties. When any individual finds a good area, that player will attract the other players to that location by exploiting. Then, when all the players are together in a group exploiting a particular area, one of the players will start to lose bonus as the score field shifts. This player will then either move close to the others who are still exploiting or will shift to an exploring state. If that player starts exploring but doesn't find any good locations, the player will return to the group if the group is still exploiting. If that player does find a new good area, though, the player will start exploiting that area. The rest of the group will then follow after the good spot shifts to where the exploiting player is. This mechanism creates a kind of gradual crawling that effectively tracks the moving score field. Thus by following these rules, players are not only improving their own performances directly but also that of the entire group by participating in this process of emergent collective sensing. An example of this process from our data is shown in Figure 6.

## Discussion

In our experiment, we observed that humans were able to achieve increases in performance at much smaller group sizes than fish. Fish experienced mild improvement in group performance at groups of 16 and more substantial improvements at groups of 64 and 128. However, we see significant improvements in human performance at just five players. This difference may at least be partially explained by the differences in the mechanism that humans appear to use in this task as compared to fish. Interestingly, the mechanism we identify in humans is similar to that of fish in some ways, but it is also distinct in important ways. Similar to the behavior of humans in choosing appropriate states based on current score, fish modulated their speeds based on the level of darkness that they were experiencing, moving slower in their preferred darker areas and faster in lighter areas. Similar to the copying behavior we observe, fish have a tendency for turning towards other fish. However, Berdahl et al.'s model of the behavior of the fish did not require any reference to the kind of discerning social awareness that we see in humans. Whereas fish appear to equally weight information from all nearby conspecifics, effective humans appear to modulate their copying behavior based on the inferred payoffs of other players. The strategic use of independent exploration (a form of asocial learning) was also key to the mechanism enabling human success. These key differences support recent work in social learning (Wisdom et al., 2013; McElreath et al., 2008), which find an impressive flexibility in the strategic deployment of imitation in humans.

This comparison highlights the superior capacity for distributed cognition of humans, possibly enabled by our ability for theory of mind. Perhaps an even more interesting differ-

ence that emerged between humans and fish has to do with the time scale over which the collective intelligence mechanism evolves. For fish, the ability to gain from group performance in these collective sensing tasks is likely based on innate behaviors, selected over many generations of fish facing exactly this problem over their whole lifespan. In contrast, some of our human groups, facing this particular problem for the first time, appear to have discovered reasonable collective sensing strategies in just a matter of minutes.

Our work therefore may shed light on some of the pressing puzzles of human collective intelligence and human distributed cognition. What are the mechanisms by which humans establish effective coordinated distributed information processing agents that can accomplish more than any individual alone, and how do our abilities play a role in these mechanisms? The perspective of human groups as distributed processors is already well-known in cognitive science (Hutchins, 1995), suggesting the importance of communication for collective intelligence, and theory of mind has been shown to be predictive of group performance across a battery of collective intelligence tasks in both face-to-face (Woolley, Chabris, Pentland, Hashmi, & Malone, 2010) and online groups (Engel, Woolley, Jing, Chabris, & Malone, 2014). Our work further suggests that one of the specific roles that theory of mind plays in the emergence of collective intelligence is facilitating implicit communication that allows for coordination on good collective actions. Moreover, our work also suggests that the benefit of a group's coordinating on good actions is not simply the benefit to each individual independently. By combining a natural human tendency for independent exploration with a discerning social awareness, humans are able to fluctuate between exploiting known good actions, independently exploring new options, and intelligently copying the promising choices of other individuals. A simultaneous combination of these activities by a cohesive group leads to a collective memory of recently good actions from individuals who continue to exploit, and a collective movement towards actions that promise to be good in the near future driven by independently exploring individuals who either find new good areas or return to the group. The exploiting core form the body of the group and the exploring individuals form the sticky appendages that drive the group's gradual crawl.

## Acknowledgments

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1122374 to PK and Grant No. DGE-114747 to RXDH. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This material is based upon work supported by the Center for Minds, Brains and Machines (CBMM), funded by NSF STC award CCF-1231216.

## References

- Amé, J.-M., Halloy, J., Rivault, C., Detrain, C., & Deneubourg, J. L. (2006). Collegial decision making based on social amplification leads to optimal group formation. *Proceedings of the National Academy of Sciences*, 103(15), 5835–5840.



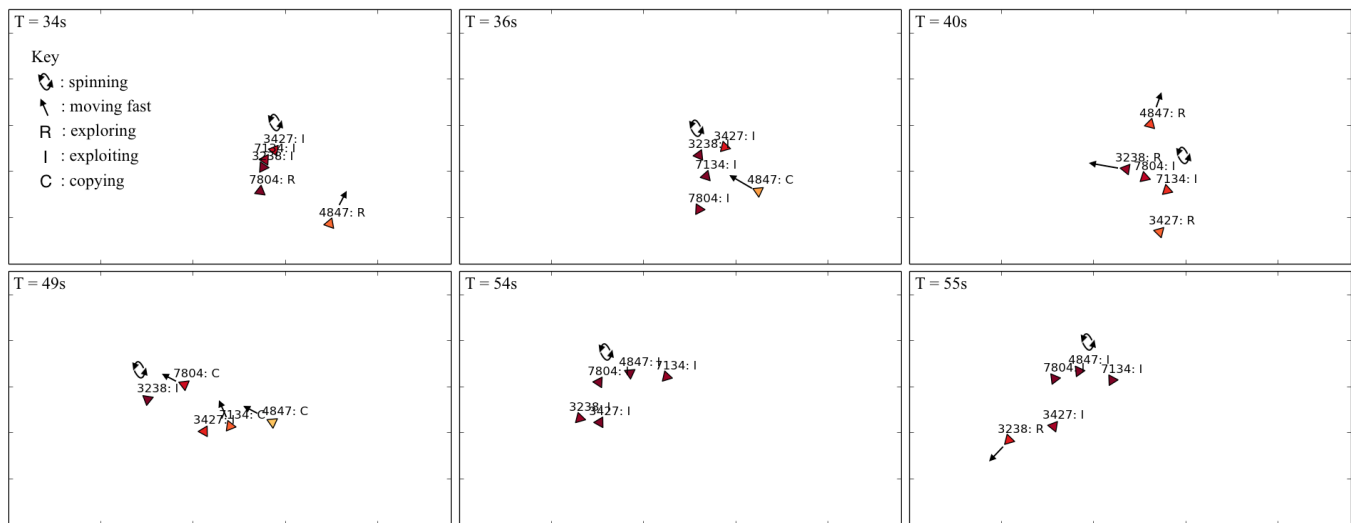


Figure 6: Reconstructions of actual gameplay in a five-person group illustrating both failed exploration leading to intelligent copying and successful exploration leading collective movement. Colors indicate the individuals' scores, with red being higher and orange/yellow being lower. The player labels indicate both player IDs and also the player states our feature extraction procedure inferred. Other annotations are provided to give a sense for the game dynamics. At 34 seconds, in the first panel, most of the group has converged on exploiting a particular area while one individual is exploring independently. At 36 seconds, the exploring individual appears to have failed to find anything good and instead copies the group. At 40 seconds, the score field has shifted and some of the group begins exploring while others continue to exploit. By 49 seconds, one of the exploring individuals found a good location, and other players have begun to move towards that individual. At 54 seconds, the entire group is exploiting the new area. Finally, at 55 seconds, the background has shifted enough again that one of the individuals begins to explore.

- Berdahl, A., Torney, C. J., Ioannou, C. C., Faria, J. J., & Couzin, I. D. (2013). Emergent sensing of complex environments by mobile animal groups. *Science*, 339(6119), 574–576.
- Engel, D., Woolley, A. W., Jing, L. X., Chabris, C. F., & Malone, T. W. (2014). Reading the mind in the eyes or reading between the lines? theory of mind predicts collective intelligence equally well online and face-to-face. *PLoS ONE*, 9(12), e115212.
- Goldstone, R. L., & Gureckis, T. M. (2009). Collective behavior. *Topics in Cognitive Science*, 1(3), 412–438.
- Goldstone, R. L., Roberts, M. E., Mason, W., & Gureckis, T. (2008). Collective search in concrete and abstract spaces. In *Decision modeling and behavior in complex and uncertain environments* (pp. 277–308). Springer.
- Hawkins, R. X. D. (2014). Conducting real-time multiplayer experiments on the web. *Behavior Research Methods*.
- Hutchins, E. (1995). *Cognition in the wild*. MIT Press.
- Kameda, T., & Nakanishi, D. (2003). Does social/cultural learning increase human adaptability?: Rogers's question revisited. *Evolution and Human Behavior*, 24(4), 242–260.
- Kearns, M. (2012). Experiments in social computation. *Communications of the ACM*, 55(10), 56–67.
- McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. (2008). Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1509), 3515–3528.
- Pratt, S. C., & Sumpter, D. J. (2006). A tunable algorithm for collective decision-making. *Proceedings of the National Academy of Sciences*, 103(43), 15906–15910.
- Reips, U.-D. (2002). Standards for internet-based experimenting. *Experimental Psychology*, 49(4), 243–256.
- Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., ... Laland, K. N. (2010). Why copy others? Insights from the social learning strategies tournament. *Science*, 328(5975), 208–213.
- Seeley, T. D., & Buhrman, S. C. (1999). Group decision making in swarms of honey bees. *Behavioral Ecology and Sociobiology*, 45(1), 19–31.
- Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press.
- Ward, A. J., Sumpter, D. J., Couzin, I. D., Hart, P. J., & Krause, J. (2008). Quorum decision-making facilitates information transfer in fish shoals. *Proceedings of the National Academy of Sciences*, 105(19), 6948–6953.
- Wisdom, T. N., Song, X., & Goldstone, R. L. (2013). Social learning strategies in networked groups. *Cognitive Science*, 37(8), 1383–1425.
- Woolley, A., Chabris, C., Pentland, A., Hashmi, N., & Malone, T. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*.