Emergent Processes in Group Behavior

Robert L. Goldstone
Michael E. Roberts
Todd M. Gureckis
Indiana University

Corresponding Address: Psychology Building
Indiana University
Bloomington, IN. 47401
rgoldsto@indiana.edu

Word Count: 2,683 for main text
Abstract

Just as networks of neurons create structured thoughts beyond the ken of any individual neuron, so people spontaneously organize themselves into groups to create emergent organizations that no individual may intend, comprehend, or even perceive. Recent technological advances have provided us with unprecedented opportunities for conducting controlled, laboratory experiments on human collective behavior. We describe two experimental paradigms where we attempt to build predictive bridges between the beliefs, goals, and cognitive capacities of individuals and group-level patterns, showing how the members of a group dynamically allocate themselves to resources, and how innovations are spread in a social network. Agent-based computational models have provided useful explanatory and predictive accounts. Together, the models and experiments point to tradeoffs between exploration and exploitation, compromises between individuals using their own innovations and innovations obtained from their peers, and the emergence of group-level organizations such as population waves, bandwagon effects, and spontaneous specialization.
Emergent Processes in Group Behavior

It is natural for psychologists to focus on the behavior of single individuals because our own introspection provides us with motivation and perspective at this level. However, in a literal sense, we are all participating in entities greater than ourselves. Self-organized collectives of people create emergent group-level patterns that are rarely the goal of any individual. A business has a style and ethos that transcends its employees. A culture has a nature, integrity, and systematicity that transcends its inhabitants while still being grounded by their interactions (Atran, Medin, & Ross, 2005). Social phenomena such as the spread of gossip, the world-wide web, the popularity of cultural icons, legal systems, and scientific establishments all take on a life of their own, complete with their own self-organized divisions of labor and specialization, dynamics, feedback loops, growth, and adaptations.

A considerable amount of early work on group behavior from social psychology focused on interpersonal relations, and the attributes that characterize good leaders or work teams. However, human collectives often organize themselves without explicit leaders, chains of command, or fixed communication networks (Ball, 2004). Examples of such spontaneous collectives include book recommendations on Amazon.com which evolve based upon similar readers’ buying habits, fans at a sport stadium, grassroots political movements, the development of a fully cross-indexed and intricately organized encyclopedia that any person can edit (www.wikipedia.org), and media that is freely accessible to both providers and consumers yet still shows striking trends of rich-get-richer popularity (www.youtube.com). In December of 2006, Time Magazine named “You” as Person of the Year in recognition of the power and sophistication of these grassroots, decentralized communities.

To understand the structure and dynamics of human collectives like these, we have developed Internet-based experimental platforms that allow groups of 20-200 people to interact with each other in real time on networked computers. The experiments implement virtual environments where participants can see the moment-to-moment actions of their peers and immediately respond to their environment. To understand the results of these experiments, we develop computational models. Several models of group behavior exist, but rarely are these models tested against detailed data sets obtained from controlled laboratory settings. Often times, there is a disconcerting mismatch between the simplicity of formal models and the complexities of the
real-world situations. Our strategy for bridging the gap between computational models and group behavior phenomena is to create relatively simple laboratory situations involving groups of people interacting in idealized environments according to easily stated “game rules.” We admittedly sacrifice some external validity in creating idealized experimental scenarios, but this loss is offset by the advantage of having the assumptions underlying our psychological experiments correspond almost exactly to the assumptions of the computational models, allowing the models to be aptly applied without sacrificing their concise explanatory value and genuine predictiveness. In what follows, we focus on two scenarios that capture ubiquitous group patterns: the competition of agents for resources, and the dissemination of innovations in social networks.

The Competitive Search for Resources
A problem faced by all mobile organisms is how to search their environment for resources. Animals forage their environment for food, web-users surf the Internet for desired data, and businesses mine the land for valuable minerals. When an organism forages in an environment that consists, in part, of other organisms that are also foraging, unique complexities arise. The resources available to each individual are affected not just by their own behavior, but also by the simultaneous actions of others.

We have been interested in experimentally exploring how human foragers allocate themselves to resources, and the time course by which that allocation is achieved. Through the Internet (available to the public at http://groups.psych.indiana.edu/), eight groups of 12-28 (average = 21) participants competed for food tokens deposited over time in a virtual environment consisting of an 80 X 80 grid of squares. In Goldstone and Ashpole (2004), two resource pools were created in the virtual environment with different rates of replenishment. The participants’ task was to obtain as many resource tokens as possible during the course of a 270 second experiment. A participant obtained a token by being the first to move on top of it. Participants moved square-by-square by pressing arrow keys on their computer keyboard.

Resources were divided between two resource pools in various ways. For example, in one condition, food tokens were split evenly (50/50) between the two pools while other conditions had a 65/35 or 80/20 split. The location of the food within the pool followed a Gaussian distribution with a mean at the center of the pool and a standard deviation of five horizontal and vertical positions. The locations of the pools were randomized under the constraint that the distance between pools was kept approximately constant. One piece of food was
delivered every $4/N$ seconds, where $N$ is the number of participants. In our visible condition, each participant could see each other and the entire food distribution. In our invisible condition, they could not see the other participants, and so they gradually acquired knowledge of the resource distributions by virtue of their histories of getting food from each location.

The dynamics of the distribution of agents to resources in each condition is shown in Figure 1, broken down by the three types of resource distribution. Although fast adaptation takes place, the asymptotic distribution of agents systematically undermatches the optimal distribution of agents to resource pools. By undermatching, we mean a distribution of agents that is less extreme than the distribution of resources. For example, in the 65/35 distribution the 65% pool only attracts an average of 60.6% of the agents. If we were efficiency consultants, we would recommend that some foragers in the less productive pool move to the more productive pool -- the resources there are being relatively underutilized despite the larger crowd there. This finding of undermatching has been obtained with other animal groups, including cichlid fish, mallard ducks, and mites (Kennedy & Gray, 1993). We have also observed this undermatching in collectives of citizens of the virtual world Second Life (http://secondlife.com/), who were invited to forage for pieces of the world’s currency that we randomly placed in two regions. This undermatching may also explain real-world human foraging behavior, such as the documented inefficiency in sperm whalers’ foraging of whales near the Galapagos islands (Whitehead & Hope, 1991).

Our results also reveal periodic fluctuations in resource use. A Fourier transformation of the time series data of participant population per pool revealed significantly greater power in the low frequency spectra for invisible than visible conditions. For all three invisible conditions, the peak power is at approximately .02 cycles/second, and was particularly high for the most uneven, 80/20 distribution. This means that in the invisible condition, agents collectively caused waves of relatively dense crowding at one pool that repeat about once every 50 seconds. There was no evidence for population cycles in the visible condition, presumably because a person who was tempted to leave their dissatisfying pool for greener pastures would be dissuaded if they saw several other people with the same idea already leaving their pool. However, in the invisible condition, agents may become dissatisfied with a pool populated with many other agents, but as they leave the pool they would not be aware that other agents are also leaving. Thus, the ironic consequence of people’s shared desire to avoid crowds is the emergence of migratory crowds!
In a second experiment (Goldstone, Ashpole, & Roberts, 2005), we ran groups of participants in conditions where food resources, but not fellow foragers, were visible, and vice versa. If people acted like buzzards, using the presence of peers as an indicator of possible food sources, then the presence of a relatively large number of participants at the richer pool would be expected to draw still more participants to the pool. In fact, when agents can see each other but not the overall pattern of food, their distribution is not extreme enough (e.g. only 62% of the participants were at the 65% pool), but when agents can see food resources but not each other, then their distribution is more extreme than the resource distribution (e.g. 73% of the participants were at the 65% pool). This suggests that people, are more like some aphid species than buzzards, avoiding sites that already have a large crowd of conspecifics. By this account, overmatching occurred because participants were attracted to the rich, productive pools, and were not dissuaded from approaching these pools by the presence of other participants (because these others were invisible).

To gain greater insight into our results, we chose to model them using an Agent-Based Model. This class of models builds social structures from the “bottom-up,” by simulating individuals as virtual agents, and allowing emergent organizations to form out of the operation of rules that govern interactions among these agents and their environment. Our EPICURE model (Roberts & Goldstone, 2006, available as an interactive simulation at http://cognitrn.psych.indiana.edu/Epicure.html) populates a world with agents that probabilistically decide from moment to moment what spatial grid location they will approach based on the locations’ values. The first factor that affects value is the distance of location. The closer a location is, the more likely it is to be selected as a target destination. Second, once a location has been selected as a target, its value is increased so that it will tend to be selected as a target at the next moment too. This is a way of incorporating consistency in target choices over time. For agents who can see all of the other agents and food, a third factor is that the value of a location increases as the density of food in its vicinity increases, and fourth, the location’s value decreases as the density of other agents increases. In the invisible condition, agents must gradually accumulate a personal history of where they have found food. Every time food is found at a location, the location’s value increases, and this increase diffuses to the nearby locations.

These simple assumptions allowed EPICURE to account for the empirically observed pattern of overmatching and undermatching for the four visibility conditions in which foragers and food could independently be visible or invisible. Why does EPICURE predict undermatching? The critical notion is
spatial “turfs.” A single agent can efficiently patrol a compact region of about 10 squares, roughly independent of the food productivity. Two pools that differ in their productivity both have the same spatial extent and variance, and so can support agents in numbers that are more similar than predicted by the pools’ productivities alone. EPICURE also predicted our observed population waves in the invisible condition, and other counterintuitive results found in animal foraging, such as that increasing the distance between two resource pools, and hence travel cost, should decrease undermatching (Baum & Kraft, 1998). One of the best ways to evaluate a model is to see whether behaviors arise that are not explicitly forced by the rules – are we getting out more than we knew we were putting in the model. By this measure, the model does a good job of explaining collective foraging behavior. Although undermatching and population waves were not explicitly stipulated by the model’s assumptions, these behaviors emerge because of agents’ tendency to avoid each other but be attracted by the same resources.

Disseminating Innovations in Social Networks

The foraging paradigm involves competitive searching for spatial resources, but we have also studied collective searching for abstract resources. Any organism that is capable of imitating its peers must decide when and how much to imitate others’ solutions versus explore one’s own solutions. To study this in an well-controlled, if somewhat artificial, setting, we had participants guess numbers between 0 and 100 using internet-connected computers (Mason, Jones, & Goldstone, 2005). Each of their computers then showed them the points earned by their guesses, based upon a hidden scoring function that had either a simple single-peaked or complex triple-peaked form, shown in Figure 2. The triple-peaked form had two local maxima, solutions that were better than their neighboring solutions but not the best possible, and one global maximum. Over 15 rounds, participants received feedback not only on their own guesses, but also on their neighbors’ guesses. Neighbors were determined by one of four network structures: a locally connected network (connections only to one’s immediate neighbors), random, fully connected (everybody connected to everybody else), and small-world (e.g. local connections plus a few long-range connections). Figure 2 shows sample networks for groups with 10 participants.

For the easy, single-peaked function, participants in the fully connected networks converged most quickly on the global maximum, with the random and locally-connected networks performing worse.
This pattern of results is readily explainable in terms of the propensity of a network to disseminate innovations quickly. Innovations disseminate most quickly in the full network because every individual is informationally connected to every other individual. For the trickier, three-peaked payout function, the small-world network performs better than the fully connected network, particularly for the first half of the trials. The truism of “the more information, the better” is not supported. Indeed, problem spaces requiring substantial exploration may benefit from networks with mostly locally connected individuals. The problem with the fully connected network is that everybody ends up knowing the same information, and they thereby become too like-minded, acting like a single explorer, rather than a federation of independent explorers. The small-world structure is an effective compromise between fully exploring a search space and also quickly disseminating good solutions once they are found (Watts & Strogatz, 1998). When the problem space is even trickier than the three-peaked function, with a single sharp needle peak for the global maximum and a broad local maximum, then the locally connected network performs best, consistent with its preservation of local and quasi-independent communities. Computational modeling work converges on these empirical results in showing that more locally connected social networks are beneficial when the problem the group has to solve is difficult (Hutchins, 1995; Lazer & Friedman, in press). Increasing connectivity among members of real-world cockpit crews has also been shown to hamper group performance by foreclosing exploration (Hutchins, 1995; see also Hinsz, Tindale, & Vollrath, 1997 for a discussion of the danger of groups’ over-reliance on shared information).

Future Prospects for Research on Collective Behavior

The previously described paradigms are united in exploring group search behavior, in both physical and abstract solution spaces. In related work, we have examined the kinds of trail systems that people create when they are motivated to take advantage of the trails left by their predecessors, and in the process of so doing, further reinforce and extend these trails (Goldstone & Roberts, 2006). The resulting trails represent a compromise between going where one wants to go, and going where others have gone before. We have begun to apply this work on imitation and exploration to modeling and predicting baby names. Names are interesting because they are roughly neutrally valued but culturally meaningful artifacts. “John” is not intrinsically a better name than “Warren” even though it occurs 35 times more frequently in the U.S. Their distribution strongly...
suggests that, as with trails and scholarly citations, the more often a name is used, the more often it will be used in the future.

Other promising areas for experimental research on collective behavior includes coalition formation and coordination, social dilemmas, group dynamics, and social specialization. The common principles that repeatedly arise in our group behavior paradigms include: 1) a tradeoff between exploration and exploitation, 2) a compromise between individuals using self- versus other-obtained information, and 3) the emergence of group-level resource usage patterns that result from individual interests but are not always favorable to these interests. Unfavorable manifestations of these principles include inefficient population waves, bandwagon effects, mismatches between agent and resource distributions, disadvantages for highly connected networks, and premature convergence of populations on suboptimal solutions. Despite these pitfalls, collective search continues to be a powerful case of distributed cognition for the simple reason that individual search often fails to provide a good solution in a limited time, and thus collaborating and sharing solutions with others can dramatically improve search efficiency.
References


Recommended Readings


Author Notes

The authors wish to express thanks to Katy Borner, Rich Shiffrin, William Timberlake, Peter Todd, Winter Mason, and Thomas Wisdom for helpful suggestions on this work. This research was funded by Department of Education, Institute of Education Sciences grant R305H050116, National Science Foundation grant 0527920, and NIH-NIMH training grant T32 MH019879-12 to the third author. On-line playable versions of experiments related to the current work can be found at http://groups.psych.indiana.edu/.

Correspondence concerning this article should be addressed to Robert Goldstone, Psychological and Brain Sciences, Indiana University, Bloomington, IN, 47405. Email: rgoldsto@indiana.edu.
Figure 1. Changes in the number of people at each of two resource pools across 270 second foraging experiments (Goldstone & Ashpole, 2004). Resources were distributed evenly (50/50), or with unequal 65/35 or 80/20 distributions. Participants either were shown the positions of other participants and resources (visible) or not (invisible). The actual distributions of resources (indicated by the straight horizontal lines) are more extreme than the distributions of participants to these resources.
Figure 2. Percentage of participants within one standard deviation of the global maximum (best solution) on each round for the simpler one-peaked and more complex triple-peaked problem spaces (Mason, Jones, & Goldstone, 2005). In the fully connected network, everybody can see each others’ guesses and outcomes. In the random network, participants only have access to a set of randomly determined neighbors. In the locally connected network, participants are informationally connected only to their close neighbors. The small world network also preserves local neighborhoods, but additionally has a few distant “short-cut” connections that bridge different local regions. For the one-peaked problem, the best group performance is initially found for the fully connected network. For the triple-peaked problem, the best performance is initially found for the small world network.