Analysis of Human Behaviors in Real-Time Swarms

Gregg Willcox
Unanimous AI
San Francisco, CA
Gregg@Unanimous.ai

Louis Rosenberg
Unanimous AI
San Francisco, CA
Louis@Unanimous.ai

Colin Domnauer
University of California, Berkeley
San Francisco, CA
Colin@Unanimous.ai

***Abstract*— Many species reach group decisions by deliberating in real-time systems. This natural process, known as Swarm Intelligence (SI), has been studied extensively in a range of social organisms, from schools of fish to swarms of bees. A new technique called Artificial Swarm Intelligence (ASI) has enabled networked human groups to reach decisions in systems modeled after natural swarms. The present research seeks to understand the behavioral dynamics of such "human swarms." Data was collected from ten human groups, each having between 21 and 25 members. The groups were tasked with answering a set of 25 ordered ranking questions on a 1-5 scale, first independently by survey and then collaboratively as a real-time swarm. We found that groups reached significantly different answers, on average, by swarm versus survey (p=0.02). Initially, the distribution of individual responses in each swarm was little different than the distribution of survey responses, but through the process of real-time deliberation, the swarm's average answer changed significantly (p<0.001). We discuss possible interpretations of this dynamic behavior. Importantly, the we find that swarm's answer is not simply the arithmetic mean of initial individual "votes" (p<0.001) as in a survey, suggesting a more complex mechanism is at play—one that relies on the time-varying behaviors of the participants in swarms. We publish a set of data that other researchers may use to investigate human swarms.**

Keywords—Artificial Swarm Intelligence, Human Swarming, Swarm Intelligence, Human Computer Interaction, Collective Intelligence, Artificial Intelligence.

I. Introduction

Extensive prior research has shown that groups of human forecasters can outperform individual forecasters by aggregating estimations across groups using statistical methods [1-3]. Often referred to as the Wisdom of Crowds (WoC) or Collective Intelligence (CI), this phenomenon was first observed over a century ago and has been applied to many fields, from predicting financial markets to forecasting geopolitical events. The most common methods involve polling a population of individuals for self-reported estimations and then aggregating the collected input statistically as a simple or weighted mean [4].

In recent years, a new method has been developed that is not based on aggregating data from isolated individuals, but instead involves groups of forecasters working together as real-time systems, their interactions moderated by AI algorithms modeled on the natural mechanisms of Swarm Intelligence.

Known as Artificial Swarm Intelligence (ASI) or simply "Human Swarming," this method has been shown in numerous studies to significantly amplify the accuracy of forecasts generated by human groups [5-11]. As one example, a study conducted at Stanford University School of Medicine asked groups of radiologists to forecast the probability that patients are positive for pneumonia based on their chest x-rays. When forecasting together as a real-time swarm, diagnostic errors were reduced by over 30% as compared to groups that voted or averaged their probabilistic diagnoses [12].

Behavioral biologists studying social organisms such as honeybees have put forth extensive theories explaining the mechanistic underpinnings and physics of Swarm Intelligence. However, no such theory has been proposed concerning the functional mechanisms of human swarms. While prior studies have shown ASI systems significantly amplify the ability of human groups across a range of tasks [7-12], from forecasting sporting events [10-12] to predicting sales volumes of new products [20], the present research focuses on describing the underlying mechanisms of such success: what influences individuals to change their responses in swarms, and what do they change their response to?

II. The Science of Swarms

The ASI system used for this study is shown in Figure 1 below, where a group of human participants, each using their own computer, is connected in real-time over a network. Each computer, which may be a desktop, tablet, or phone, runs a unique software interface that is designed to capture and stream the user's real-time input to a cloud-based processing engine. The engine runs swarming algorithms and sends back real-time output to each user, thereby creating a closed-loop system.

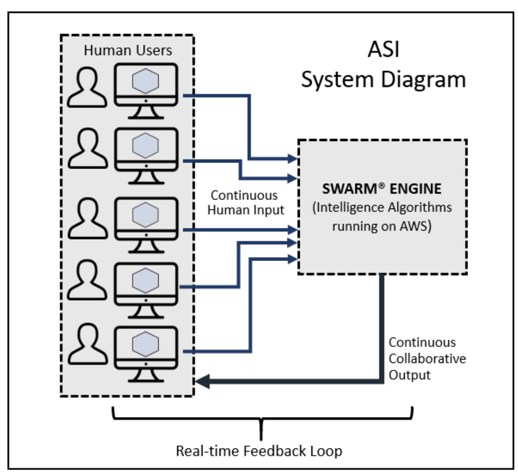

Fig.1. System Diagram for a real-time ASI System

The present study uses Swarm AI technology as the cloud-based engine. Swarm AI is modeled on the decision-making processes of honeybee swarms, which have been studied since the 1950s and share many properties with neurological brains [13,14]. Both swarms of bees and brains consist of simple decision-making units (in this case, neurons and bees) that work together in real-time to identify alternatives, weigh competing evidence, and converge on decisions. In both swarms and brains, decisions emerge from competition between subgroups of decision-making units: when a sub-group espousing a particular decision or alternative is excited beyond a threshold level, the corresponding alternative supported by that sub-group is selected as the group's collective decision. In colonial bees such as the honeybee, this enables hundreds of scout bees to work in parallel, collecting data about their environment and converging on a single decision, often selecting the optimal solution to complex multi-variable problems [15-17].

The similarities in decision-making mechanisms between "swarms" and "brains" become more evident upon a review of the mathematical models that describe these mechanisms. Primate brains are often modelled as using a mutually inhibitory leaky integration process [18], such as the Usher-McClelland model shown in Figure 2 below.

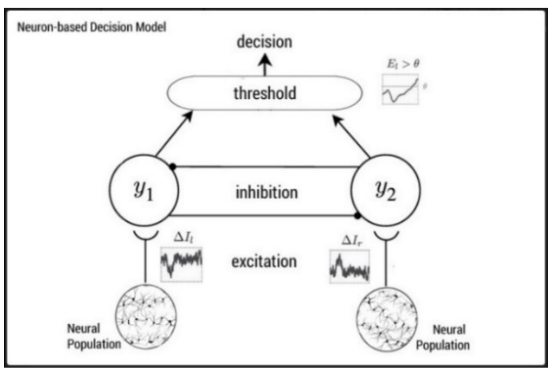

Fig. 2. The Usher-McClelland neuronal decision-making model

Swarm-based decision models are often modelled in similar ways to human brains: one mutually inhibitory model describing honeybee swarms' decision-making process that bears a striking resemblance to the Usher-McClelland model of brains is shown in figure 3. As shown below, swarm-based decisions follow a very similar process, weighing the input from sub-groups of swarm members through mutual excitation and inhibition, until a threshold is exceeded.

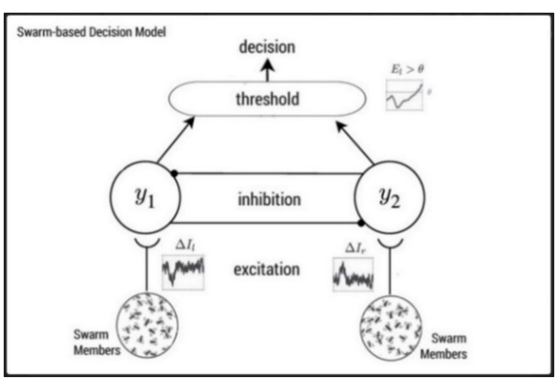

Fig. 3. A model of mutually inhibitory bee decision-making

Thus, while brains and swarms are very different forms of intelligence, both are systems that enable optimized decisions to emerge from the interactions among collections of processing units. The goal of the present study is to supplement this general model of Swarm Intelligence with a detailed quantitative examination into the mechanisms of the human swarming process, as powered by Swarm AI.

III. Swarming Software

In the natural world, swarming organisms establish real-time feedback-loops among group members. Swarming bees do this using complex body vibrations called a "waggle dance." Swarm AI technology was developed to groups of humans to swarm in a similar way. It allows groups of distributed users to form closed-loop consensus-building systems moderated by swarming algorithms [5-7]. Similarly to the decision-making processes of swarms of bees, Swarm AI enables human groups to: (a) consider evidence, (b) weigh a set of alternatives, and (c) converge on group decisions, all in real-time.

An example of this system in shown in figure 4, where a team is in the process of answering a question by collaboratively moving a 'puck' to select one answer from a set of five choices. Each participant manipulates a 'magnet' with their mouse cursor, and the magnets 'pull' on the puck in the direction of the user's cursor. The pull from each user's magnet is visible to other users, and the net force from all of the magnets controls the movement of the puck. All users adjust the direction and strength of their magnet in real-time by moving their mouse, and as a result, the swarm moves, not based on the behaviors of a single user, but the behaviors of all users. Since the behavior of one user influences the decisions and pulling behaviors of other users, however, a complex negotiation forms, where the group is empowered to negotiate among alternatives, explore the decision-space, and converge upon the most agreeable solution.

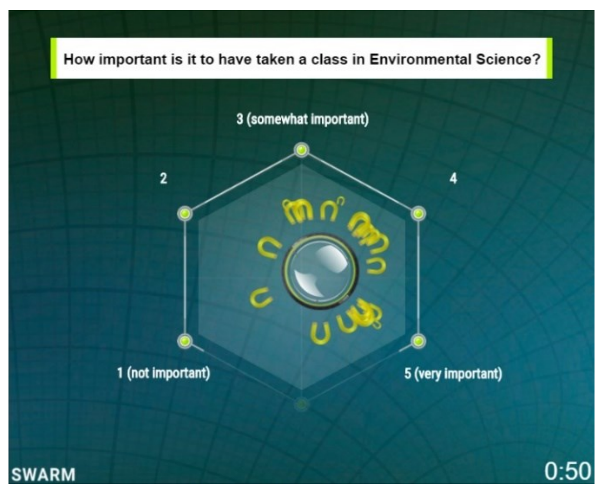

Fig. 4. A human swarm answering a question in real-time

The real-time control that participants have over the direction and magnitude of their 'pull' on the puck is important: the closer a magnet is to the edge of the puck, the more force it exerts on the puck. As the puck moves across the screen, users have to continually adjust their magnet to stay close to the puck's rim to have maximal say in the group's decision. This has the effect that participants remain fully engaged throughout the swarm's decision, continually evaluating and re-evaluating their views and adjusting their magnet's position in accordance with these views. If they stop adjusting their magnet, the distance between their magnet and the puck's rim grows (or worse, the magnet touches the rim and exerts no force) and their imparted sentiment wanes.

As a result, the participants in an artificial swarm must continuously update their preferences throughout the process or lose influence over the swarm's outcome, just like bees vibrating their bodies to express favor for a new home site in a biological swarm. At the same time, intelligence algorithms infer the conviction of each swarm participant based on their behaviors over time, and continually change the strength of each participant's magnet to reflect their conviction. Examining these behavioral characteristics of individuals in detail provides insight towards a theory of how and why human-powered swarms work.

IV. Ordered Ranking Study

A study was conducted to collect data on decision-making dynamics in human swarms. Ten groups of 23 to 25 subjects were tasked with answering a set of 25 subjective rating questions, each on a 1-5 scale. All questions were the same format, asking participants to rate the importance of an academic subject (e.g. Algebra) to a high school education: *"How important is it to have taken a class in Algebra before graduating high school?"*. Participants first provided their answers independently using a standard online survey. Upon completion, the group congregated as a real time swarm using the Swarm AI platform to answer the same set of questions.

The data collected from the swarm was analyzed in two ways: (i) as the initial responses of participants to a question, calculated as the first answer chosen by each participant before 1 second has elapsed in the swarm, and (ii) as the average contribution of each individual to the swarm, calculated as the mean response of each participant measured at quarter-second increments through the swarm.

V. Analysis and Results

All individuals in this study were verified to have completed the survey before joining the swarm, so no survey data was missing at the time of analysis. No swarm failed to select an answer in the 60 seconds allocated: all 250 questions were answered in between 9 seconds and 60 seconds, with an average time to answer of 18.8 seconds.

Comparing the survey results to the swarm behavior through time gives insight into the distinguishing characteristics of the swarming process itself. First, to ensure that users were reporting similar beliefs in both the survey and the swarm, the distribution of individual survey responses was compared to the distribution of initial individual choices (t=1 second) in the swarm. A paired two sample t-test was conducted and revealed a non-significant difference between individual survey responses and initial swarm behavior of those individuals across the full dataset ($\mu_{difference}$=0.10, p=0.47), indicating that users reported very slightly (though not significantly) higher levels of subject importance on average in the swarm than in the survey.

However, a significant difference was found between the average individual survey response and the final average individual swarm response: the final contribution of each individual in the swarm was found to be higher than the average survey response on average ($\mu_{difference}$=0.70, p<0.01), indicating that users pulled more often for answer choices that rated the subject as more positive than their initial response than for answer choices that were more negative than their initial response. In other words, by the end of the swarm users on average reported that the subject was more important than they did at the start of the swarm.

In order to compare the final swarm answers to the survey answers without the positive bias introduced through swarming, both answer methods were normalized by converting to a z-score such that the z-scores of each method had a mean value of 0 and a standard deviation of 1. These z-scores therefore enable a fair comparison between the swarm and survey ratings of subject importance, since the swarm answers are no longer biased towards rating subjects as more positive than the survey.

The z-scores of the survey answers and swarm final answers from each group were then compared by question, as shown in figure 5 below. Interestingly, even after normalization there were significant differences between the distribution of survey and swarm final responses on 44% of questions (11 of 25) on this test, indicating that the group rated 5 subjects as more important in the swarm than the survey, and 6 subjects as less important in the swarm than the survey.

These differences in the distribution of responses between the survey and swarm point to an important effect: enabling the respondents to think together as a group in real-time results in meaningfully different answers than if the group were to respond asynchronously in a survey format. It's also important to note that these differences were likely not due to the interface of swarming alone—the process of reporting a belief with a cursor rather than a survey or the layout of answers in the Swarm platform—because these significant differences did not exist when comparing the survey responses to the participant's initial responses in the swarm, before the real-time feedback of other participant answers had the time to impact the initial responses of participants. These observed differences were thus due to the interaction and deliberation among participants over time.

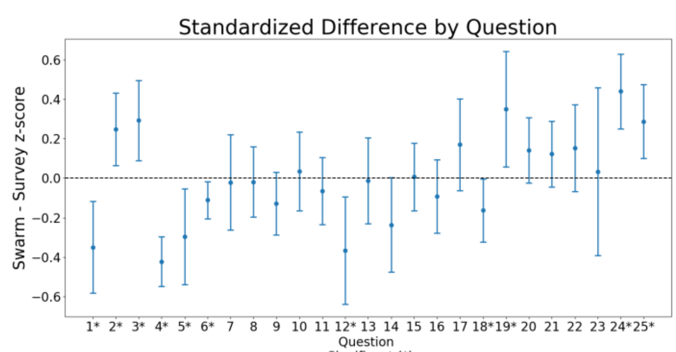

Fig. 5. 90% Confidence Interval for the difference between the Survey and Swarm Final responses on each question. * = (p<0.05)

Taking question 24 (subject: Drug Education) as a case study, we can examine in detail how each swarm's average answer changed over time when answering this question. This is depicted as a line plot in Figure 6, where each line represents the mean answer in the swarm over the course of the deliberation. The swarms' final answers were significantly different from the swarm initial mean (p=0.001), and the survey answers (p<0.001). As a result, it's reasonable to suggest that some time-based aspect of the swarming experience, such as the ability of individuals to switch answers throughout the swarm, or the influence that each participant indirectly has on every other participant's choices, was the cause of the difference between the swarm's final answers and the survey's answers.

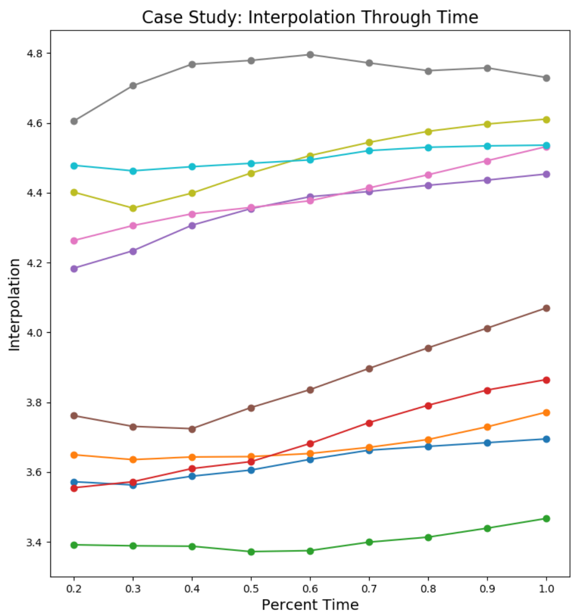

Fig. 6. Swarm Interpolations through Time for Question 23 (Drug Education)

Looking at all 238 individuals' behaviors in isolation, rather than aggregated into a swarm-level average as above, gives a better idea of how the individuals in these swarms changed their responses from over the course of the swarm. On this same question, individuals' answers in the survey were not different from their initial beliefs as expressed at the start of the swarm (p=0.77). However, individuals' average answers over the course of the swarm deliberation had a significantly higher mean ($\mu_{difference}$ = 0.14, p<0.001) and lower standard deviation ($\mu_{difference}$ = -0.157, p<0.001) than their initial swarm responses, as depicted in Figure 7.

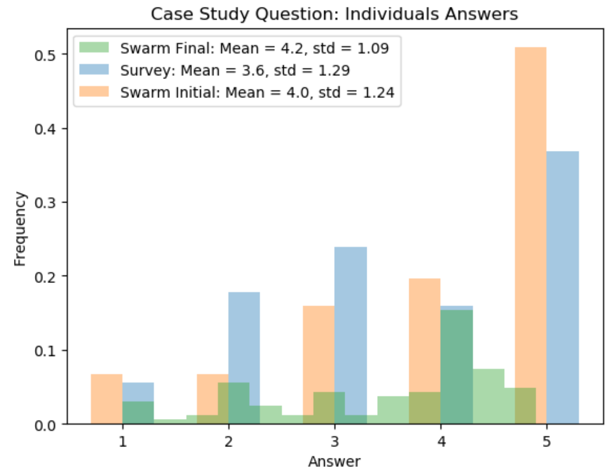

Fig. 7. Initial similarities become significant behavioral differences

These results reveal an important point: individuals are not simply changing their answer to conform to the arithmetic mean of the group. In fact, the individual's final answers in the swarm were significantly different (p<0.001) than the initial arithmetic mean of individual initial choices within the swarm. However, individuals are generally pulling closer to the arithmetic mean through the swarm than they did at the start of the swarm, leading to a lower variance of individuals' final answers than individual initial answers.

So, how and when do individuals switch their pull within a swarm, and what influences the switch? Looking again at all 25 questions, participants spent about 80% of the time pulling for their initial choice, with proportionally less time spent on answers farther from the initial choice, as shown in Figure 8. In other words, if individuals do pull for an answer that was not their original choice, they pull more often for a nearby option—one closer to their original response—than an option far away. This suggests individuals are compromising between multiple competing alternatives, weighing which answers they believe in the most against which answers other people are expressing support for, rather than simply converting to the most popular alternative or entrenching exclusively in their initial response.

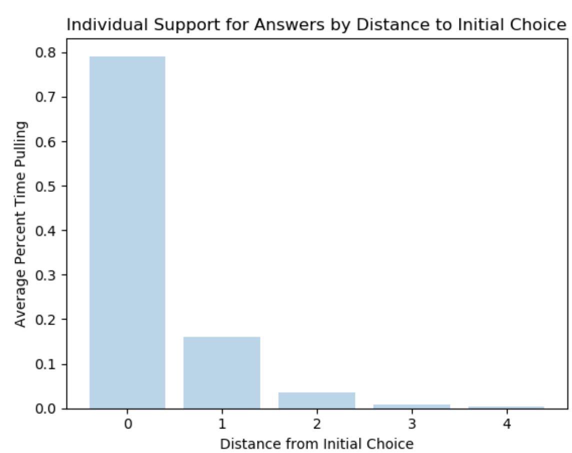

Fig. 8. Histogram of the percent of time individuals spent pulling for answer choices of varying distance from their initial choice

To visualize this compromise between competing alternatives in better resolution, a heatmap is shown in Figures 9 and 10 that compares the level of support for each of the five answers in the swarm by the initial pull of each user. 44% of all individuals switched at least once. Users preferentially pulled to answers that were (a) closer to their initial response (Figure 9), and (b) closer to the average group response (Figure 10).

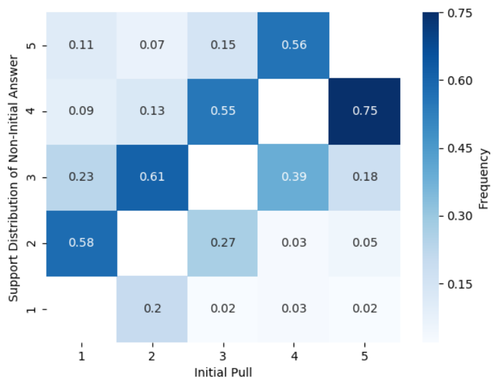

Fig. 9. Proportion of support for each answer other than each user's initial pull.

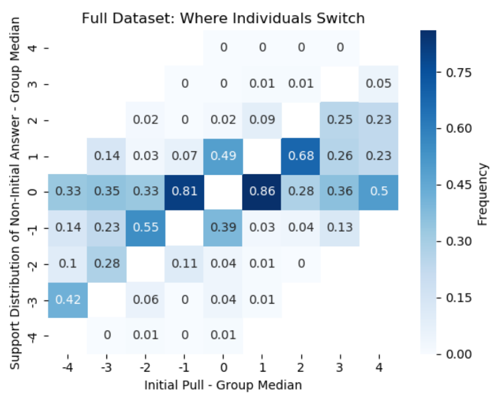

Fig. 10. Proportion of Individual Support for each answer, broken down by the distance from the initial pull to the initial group median (x axis).

Examining Figure 10 more closely, we see that when switching, the most common courses of action are twofold: (a) acceding one answer choice towards the group median, and (b) acceding all the way to the group median answer. Unless they were originally pulling for the group median answer, individuals very rarely switch to an answer that's not in the direction of the group median, which implies that the direction of the group's median answer at the start of the swarm is highly correlated with the switches that users decide to make later in the swarm.

This information collectively suggests that individual switching decisions may be influenced, but not explained completely, by the pulling behaviors of other individuals in the swarm.

VI. Datasets

To promote a better understanding of human behaviors in Artificial Swarm Intelligence systems, two datasets have been compiled and listed publicly. These datasets both contain anonymized individual behavioral data for authentic human swarms and contain over 350 individual examples of human swarms converging on solutions. The first dataset is named the Repeatability dataset, and contains the data used in this study. The second dataset is named the Fruit dataset, and contains the data used in another recent study [19]. Both datasets are visible on github at: https://github.com/unanimousai/rd-published-data. The Repeatability folder also contains pythonic code used to analyze the data as well as functions that were used in the analysis of the data and publication of this study.

The authors' aim in publishing this data is to enable other researchers to verify the existing findings about human decision-making in real-time swarm systems, and to encourage future research into real-time swarming systems.

VII. Conclusions

Over the course of the 25 general-opinion questions asked to each of 10 groups, the group's responses when swarming were significantly different than their responses when surveyed. The results of this study suggest that the time-varying, dynamic expression of individual answers, and the confidence in those answers, may be a key reason why groups that swarm produce different, and often more optimal, answers than groups that are surveyed. What determines whether this change occurs appears to be a complex, multivariable problem in which the individual must intuitively negotiate many factors, both internal and external, in a short period of time, including their own conviction in their answer and the real-time, changing distribution of answers in the group at large.

Future work may investigate other factors that influence individuals to switch their response, such as the changing location of the graphical puck, the amount of deliberation time left for the question, or the percentage of participant support for each of the answer options. Other future work may build predictive models to rigorously forecast switches based on this data, to test what degree of switching behavior can be understood through the behavioral data in human swarms alone. Other future work may investigate whether people switch differently in different questions: do some questions cause more entrenchment than others?

Acknowledgment

Thanks to Chris Hornbostel for his efforts in coordinating the swarms. Also, thanks to Unanimous AI for the use of the Swarm platform for this ongoing work. This work was partially funded by NSF Grant #1840937.

References

[1] Galton, F. (1907). Vox Populi. Nature, 75, 450-451.

[2] Steyvers, M., Lee, M.D., Miller, B., & Hemmer, P. (2009). The Wisdom of Crowds in the Recollection of Order Information. In Y. Bengio and D. Schuurmans and J. Lafferty and C. K. I. Williams

[3] Philip E. Tetlock and Dan Gardner. 2015. Superforecasting: The Art and Science of Prediction. Crown Publishing Group, New York, NY, USA.

[4] J Dana, P Atanasov, P Tetlock, B Mellers (2019), Are markets more accurate than polls. The surprising informational value of "just asking." Judgment and Decision Making 14 (2), 135-147

[5] Rosenberg, L.B., “Human Swarms, a real-time method for collective intelligence.” Proceedings of the European Conference on Artificial Life 2015, pp. 658-659

[6] Rosenberg, Louis. “Artificial Swarm Intelligence vs Human Experts,” Neural Networks (IJCNN), 2016 International Joint Conference on. IEEE.

[7] Rosenberg, Louis. Baltaxe, David and Pescetelli, Nicollo. "Crowds vs Swarms, a Comparison of Intelligence," IEEE 2016 Swarm/Human Blended Intelligence (SHBI), Cleveland, OH, 2016, pp. 1-4.

[8] Baltaxe, David, Rosenberg, Louis and N. Pescetelli, “Amplifying Prediction Accuracy using Human Swarms”, Collective Intelligence 2017. New York, NY ; 2017.

[9] Willcox G., Rosenberg L., Askay D., Metcalf L., Harris E., Domnauer C. (2020) Artificial Swarming Shown to Amplify Accuracy of Group Decisions in Subjective Judgment Tasks. In: Arai K., Bhatia R. (eds) Advances in Information and Communication. FICC 2019. Lecture Notes in Networks and Systems, vol 70. Springer, Cham

[10] L. Rosenberg, N. Pescetelli and G. Willcox, "Artificial Swarm Intelligence amplifies accuracy when predicting financial markets," 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), New York City, NY, 2017, pp. 58-62.

[11] L. Rosenberg and G. Willcox, "Artificial Swarm Intelligence vs Vegas Betting Markets," 2018 11th International Conference on Developments in eSystems Engineering (DeSE), Cambridge, United Kingdom, 2018, pp. 36-39

[12] L. Rosenberg, M. Lungren, S. Halabi, G. Willcox, D. Baltaxe and M. Lyons, "Artificial Swarm Intelligence employed to Amplify Diagnostic Accuracy in Radiology," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, 2018, pp. 1186-1191.

[13] Seeley T.D, Buhrman S.C 2001 “Nest-site selection in honey bees: how well do swarms implement the ‘best-of-N’ decision rule?” Behav. Ecol. Sociobiol. 49, 416–427

[14] Marshall, James. Bogacz, Rafal. Dornhaus, Anna. Planqué, Robert. Kovacs, Tim. Franks, Nigel. “On optimal decision-making in brains and social insect colonies.” Soc. Interface 2009.

[15] Seeley, Thomas D., et al. "Stop signals provide cross inhibition in collective decision-making by honeybee swarms." Science 335.6064 (2012): 108-111.

[16] Seeley, Thomas D. Honeybee Democracy. Princeton Univ. Press, 2010.

[17] Seeley, Thomas D., Visscher, P. Kirk. “Choosing a home: How the scouts in a honey bee swarm perceive the completion of their group decision making.” Behavioural Ecology and Sociobiology 54 (5) 511-520.

[18] Usher, M. McClelland J.L 2001 “The time course of perceptual choice: the leaky, competing accumulator model.” Psychol. Rev. 108, 550–592

[19] L. Rosenberg, G. Willcox, "Artificial Swarms find Social Optima: (Late Breaking Report)", 2018 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA), pp. 174-178, 2018.

[20] Willcox, Gregg and Rosenberg, Louis and Schumann, Hans, Sales Forecasting, Polls vs Swarms (March 14, 2019). http://dx.doi.org/10.2139/ssrn.3390043