Amplifying the Social Intelligence of Teams Through Human Swarming

Louis Rosenberg and Gregg Willcox
Unanimous AI
San Francisco, CA USA

David Askay, Lynn Metcalf and Erick Harris
California Polytechnic State University
San Luis Obispo, CA USA

Abstract— **Artificial Swarm Intelligence (ASI) is a method for amplifying the collective intelligence of human groups by connecting networked participants into real-time systems modeled after natural swarms and moderated by AI algorithms. ASI has been shown to amplify performance in a wide range of tasks, from forecasting financial markets to prioritizing conflicting objectives. This study explores the ability of ASI systems to amplify the social intelligence of small teams. A set of 61 teams, each of 3 to 6 members, was administered a standard social sensitivity test —"Reading the Mind in the Eyes" or RME. Subjects took the test both as individuals and as ASI systems (i.e. "swarms"). The average individual scored 24 of 35 correct (32% error) on the RME test, while the average ASI swarm scored 30 of 35 correct (15% error). Statistical analysis found that the groups working as ASI swarms had significantly higher social sensitivity than individuals working alone or groups working together by plurality vote (p<0.001). This suggests that when groups reach decisions as real-time ASI swarms, they make better use of their social intelligence than when working alone or by traditional group vote.**

Keywords— *Swarm Intelligence, Collective Intelligence, Artificial Swarm Intelligence, Human Swarming, Artificial Intelligence, Social Sensitivity, Emotional Intelligence.*

I. INTRODUCTION

In the natural world, many species amplify their collective intelligence by forming real-time closed-loop systems. Referred to as Swarm Intelligence (SI), this process enables schools of fish, flocks of birds and swarms of bees to solve problems with amplified accuracy. In human groups, the technology of Artificial Swarm Intelligence (ASI) enables similar benefits by connecting networked groups as real-time closed-loop systems. Often referred to as "human swarms" or "hive minds", these systems have been shown to significantly increase accuracy in a variety of tasks, from predicting sports and equity markets to dispute resolution and medical diagnosis [1-9].

While ASI has been shown to amplify the accuracy of human groups in analytical tasks like forecasting, prioritizing, estimating, and diagnosing [1-8], formal studies investigating the potential of ASI to amplify the social intelligence of teams have not been conducted. This is important to scholarship as a group's mean social intelligence has been found to be a strong indicator of a team's overall performance [10, 11].

Social intelligence, also referred to as social sensitivity, is often measured in teams by averaging each member's individual performance on the "Reading the Mind in the Eyes" (RME) test – an instrument designed to quantify how well individuals "can put themselves into the mind" of another person and assess their mental state." [12]. Because prior research has shown that the effectiveness of teams is significantly correlated with the mean social intelligence of group members, it stands to reason that if "human swarming" can amplify the effective social intelligence of small teams on a standard RME test, it may indicate that swarming can also increase group effectiveness across a wide range of collaborative tasks. For example, if a business team was tasked with making critical hiring decisions, amplification of the team's social intelligence through swarming could enable the group to converge upon more effective and insightful decisions. Similarly, if business teams are tasked with predicting how consumers will react to marketing messages, product features, or sales tactics, an amplification of the team's social intelligence could enable more accurate and insightful forecasts.

To explore whether the real-time swarming process can amplify the social intelligence of small working groups, the present study explored if teams perform with higher social intelligence on a standard RME test when working as a real-time swarm, as compared to (i) taking the RME test as individuals and (ii) reaching decisions by plurality vote.

II. BUILDING "HUMAN SWARMS"

Artificial Swarm Intelligence (ASI) is modeled after natural systems such as schools of fish, flocks of birds, and swarms of bees. But unlike birds, bees and fish, humans have not evolved the natural ability to form real-time closed-loop swarms, as they lack the subtle connections that other organisms use to establish feedback-loops among members. Schooling fish detect subtle vibrations in the water around them. Flocking birds detect high-speed motions propagating through the formation. Swarming bees generate complex body vibrations called a "waggle dance" that encodes information. To enable networked human groups to form similar real-time systems, a software platform called swarm.ai was developed by Unanimous AI, Inc. It enables distributed groups, connected from remote locations around the world, to answer questions, make predictions, and reach decisions by working together as closed-loop swarms.

As shown in Figure 1 below, the swarm.ai platform enables groups of networked participants to answer questions by collaboratively moving a graphical puck to select from among a set of alternatives. Each participant provides individual input by

manipulating a graphical magnet with a mouse, touchpad, or touchscreen. By adjusting the position and orientation of their magnet with respect to the moving puck, participants express their personal intent on the system. The input from each user is not a discrete vote, but a stream of vectors that varies freely over time. Because all members of the group can adjust their intent continuously in real-time, the swarm explores the decision-space, not based on the input of any individual, but based on the emergent dynamics of the full system. This enables synchronous deliberations among all members, empowering the group to consider the options and converge on the optimal solution.
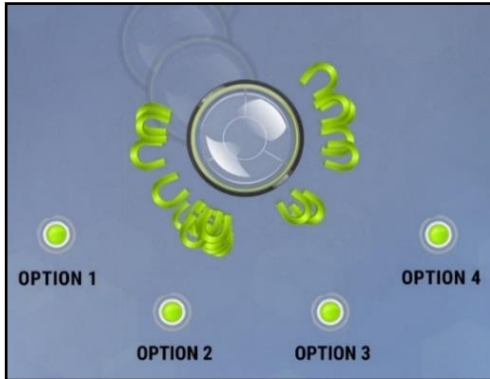

Fig.1. A human swarm choosing between options in real-time

While the swarm shown above is composed of twenty networked participants, each of whom are connected from a remote location, the swarm.ai platform has been used successfully with groups with as few as three members and as many as 150 participants. It is important to note that participants not only vary the direction of their intent but also modulate the magnitude of their intent by adjusting the distance between their magnets and the puck. Because the graphical puck is in continuous motion across the decision-space, users need to continually move their magnets so that they stay close to the puck's rim. This is significant, for it requires that all participants, regardless of group size or composition, to be engaged continuously throughout the decision process, evaluating and re-evaluating their intent in real-time. If a participant stops adjusting their magnet with respect to the changing position of the puck, the distance grows and the participant's influence on the group's decision wanes.

Thus, like bees vibrating their bodies to express sentiment in a biological swarm, or neurons firing to express conviction levels within a biological neural-network, the participants in an artificial swarm must continuously update and express their changing preferences during the decision process, or lose their influence over the collective outcome. This is generally referred to as a "leaky integrator" structure and common to both swarm-based and neuron-based systems. In addition, intelligence algorithms monitor the behaviors of swarm members in real-time, inferring their relative conviction based upon their actions and interactions over time. This reveals a range of behavioral characteristics within the swarm population and weights their contributions accordingly, from entrenched participants to flexible participants to fickle participants.

III. SOCIAL INTELLIGENCE STUDY

To assess the ability of human swarms to amplify the social intelligence of working groups, a study was conducted across a set of 61 teams, each of 3 to 6 members, totaling 302 subjects. All were college students in communications, engineering and business courses, for which a team project was required. To measure social intelligence, a widely used instrument, "Reading the Mind in the Eyes" (RME) test, was employed [8]. The test includes 35 questions, each showing a narrow facial image restricted to a region around the eyes and a set of four options that describe the emotion expressed. Participants were tasked with reading the emotional state of facial image based only on the eyes. An example question from a standard RME test is shown below in Figure 2, with the four options provided.
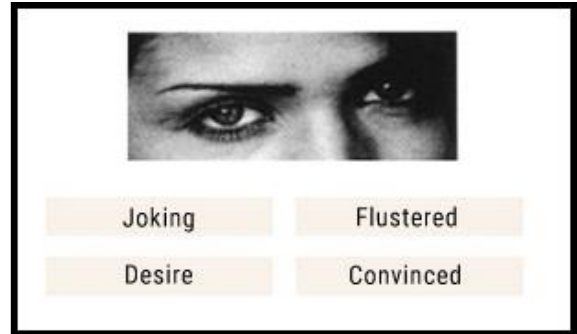

Fig.2. Sample Question from Standard RME Test.

Prior studies have shown that the RME test is a reliable measure of social intelligence, with strong internal consistency and test-retest stability [13]. Social intelligence is often described as a person's ability to perceive, interpret, and respond to the intentions, dispositions, and behaviors of others [14, 15]. These skills are extremely important for effective decision making, especially by problem-solving teams, as understanding and/or empathizing with the needs, goals, intentions, and beliefs of others is a fundamental skill required of many critical decisions made by organizations of all sizes [16].

To test whether real-time swarming enabled working groups to amplify their effective social intelligence when making group decisions, a two-stage process was employed. First, each of the 302 study participants were administered a 35-question RME assessment individually through an online survey. To limit bias and knowledge of correct answers, individual scores were not shared, and discussion of the assessment was discouraged.

In the second stage, each of the 61 teams were administered the RME test through the swarm.ai platform such that the group was tasked with answering each question as a real-time swarm. Team members were discouraged from communicating with each other during the assessment, instead relying only on the closed-loop interaction afforded by the platform (i.e., via pulling the puck). The platform presented the image of the face to everyone along with the four potential responses. Each team had 60-seconds to collaboratively coverage upon an answer. Figure 3 below is a snapshot of a participant's screen during a response, which represents the pull of each teammate through a magnet. It should be noted that to discourage conformity, participants did not see the magnets during the actual swarming session.
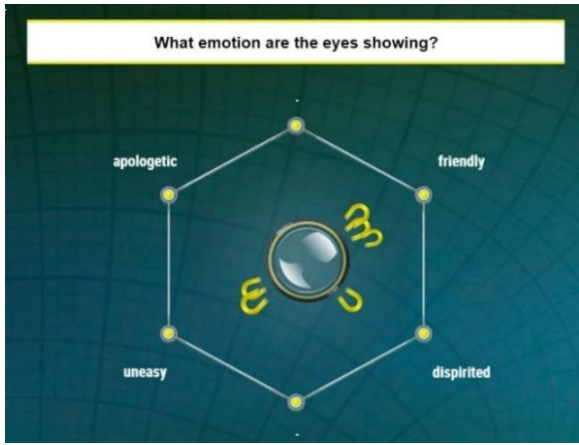
Fig. 3. Swarming Group responding to RME question

IV. DATA AND ANALYSIS

The RME was administered to 302 individuals across 61 teams and produced three unique datasets. First, we received fully completed individual assessments from 266 participants (88% response rate), totaling over 9,000 item responses. These responses were used to calculate *individual RME scores* for each participant. Second, these same responses were aggregated by team to generate a *plurality RME score*, which was calculated by plurality vote (the most popular answer within a group) for each of the 61 teams. For questions where the vote was split evenly across multiple answers, a "deadlock" was determined and classified as an incorrect response. This provided a dataset of over 2,500 plurality vote responses to RME assessment questions. Finally, a *swarm RME score* for each team was calculated from the responses collected through the swarm.ai platform. For questions where the swarm could not converge upon an answer within the 60 second time limit, a "deadlock" was determined and classified as an incorrect response.

V. RESULTS

Mean scores and error rates for RME tests were calculated for the individual, plurality vote, and swarm generated scores. As shown in Table 1 below, the average *individual RME* score was 23.96, which corresponds to an error rate of 31.5%. The average of each team's *plurality RME* score was 25.92, which corresponds to an average error rate of 25.9%. When teams worked together as a real-time closed-loop *swarm*, the average RME score increased to 29.65, which corresponds to an average error rate of 15.3%. In other words, by working together as an ASI system, the 61 groups, on average, reduced their error rates by more than half. This supports the notion that working as a swarming system can increase the social intelligence of teams.

Testing Method (Deadlocks as Errors)	Mean # Correct	Error Rate	95% Error Rate Confidence Interval	95% Error Rate Difference to Swarm CI
Individual Average	23.96	31.54%	[29.9%, 33.2%]	[14.0%, 18.6%]
Plurality Voting	25.92	25.94%	[22.7%, 29.2%]	[7.11%, 14.4%]
Swarm AI	29.65	15.29%	[13.1%, 17.6%]	N/A

Table 1: Decision Method Error Rate and Confidence Interval

Next, the statistical significance of three RME assessment methods were calculated using a 10,000-trial bootstrap analysis of the error rate for each method. The 95% confidence intervals

and p-values were then calculated for the difference between individual REM scores, plurality RME scores, and swarm RME scores. The results show that the swarm significantly outperforms both individual ($\mu_{difference} = 16.3\%$ error, $p < 0.001$) and plurality scores ($\mu_{difference} = 10.7\%$ error, $p < 0.001$). The bootstrapped error comparison is shown below in Figure 4.
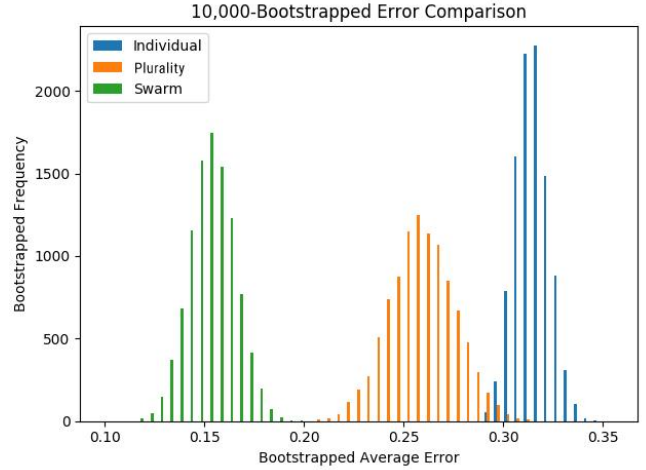

Figure 4: Bootstrapped Average Error Rate

With respect to deadlocks, a comparison was made between the rate of deadlocks determined by plurality vote as compared to the rate of deadlocks reached by swarms. Across the 61 working groups, plurality voting resulted in deadlocks in 12% of questions. Across those same groups, when working together as swarms, the rate of deadlocks dropped substantially to 0.6% of questions. This is a significant improvement, reducing the need for further steps to resolve undecided groups.

In addition, an analysis was performed that assumed that deadlocked votes were resolved by giving partial credit for tied answers that included a correct response: one-half credit for a two-way tie, one-third credit for a three-way tie, etc. To balance this, deadlocked swarms were given the chance to resolve immediately following a deadlock in another 60-second swarm, with the answer chosen in this second round selected as the final answer. There were no swarms that deadlocked twice in a row.

As shown in the Table 2 below, when deadlocks were resolved using partial credit, plurality vote had an average RME score of 28.23, or an error rate of 19.3%. When enabling the swarms to work together as real-time systems and resolve their deadlocks in a follow-up swarm, the swarm RME score increased to 29.64, or an error rate of 15.3%. In other words, even when giving partial credit for deadlocks in group responses determined by plurality vote, the swarm outperformed.

Testing Method (Deadlocks Resolved)	Mean # Correct	Error Rate	95% Confidence Interval	95% Difference to Swarm CI
Individual Average	23.96	31.54%	[29.9%, 33.2%]	[14.0%, 18.6%]
Plurality Voting	28.23	19.33%	[17.0%, 21.6%]	[1.41%, 7.12%]
Swarm AI	29.64	15.29%	[12.9%, 17.5%]	N/A

Table 2: Decision Method Error Rates with Deadlocks Resolved.

To assess statistical significance, a bootstrap analysis of the error rate for each method was again performed across 10,000 trials. We find that the swarm outperforms both the plurality

vote ($\mu_{difference} = 4.0\%$ error, p < .002) and individuals ($\mu_{difference} = 16.3\%$ error, p < .001). The bootstrapping of the error rate confidence intervals is shown below in Figure 5.
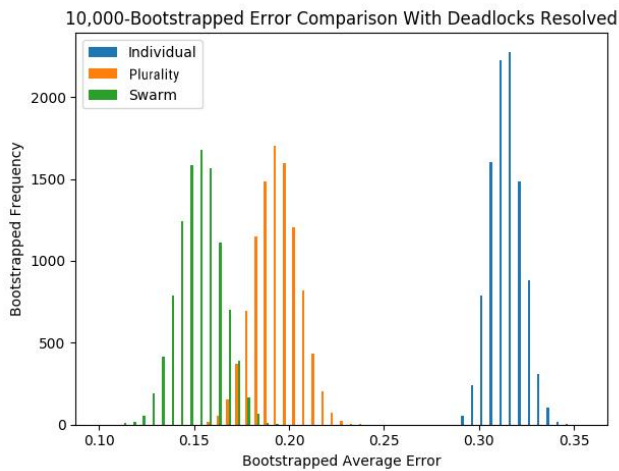

Figure 5: Bootstrapped Average Error Rate

In addition to comparing against the average individual, the swarm can be compared against all individuals. On average, swarms are in the 93rd percentile of individuals, indicating that an average swarm scores better than 93% of individuals taking the test alone. The histogram of user performance and average swarm performance is shown below in Figure 6.
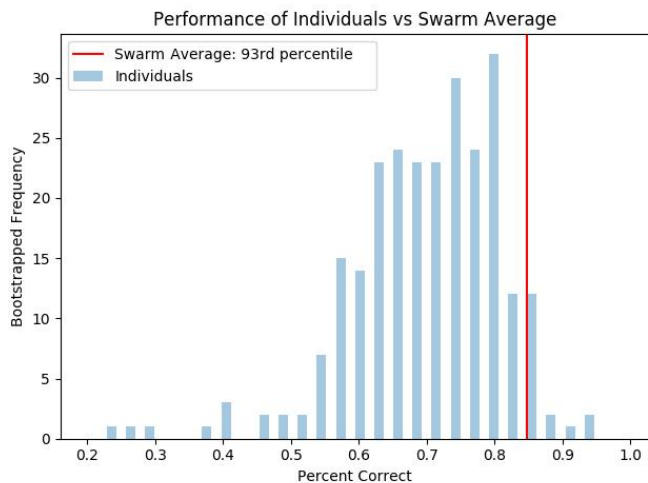

Figure 6: Bootstrapped Average Error Rate

VI. CONCLUSIONS

Can small teams, working together as real-time ASI swarms, amplify their effective Social Intelligence? The results of this study suggest this is the case. As shown across 61 working groups, each with 3 to 6 members, the average social intelligence increased significantly as compared to working (i) individually or (ii) by plurality vote. In fact, teams collaborating on an ASI platform reduced the error rate of the RME by half compared to individuals. The probability that the swarm outperformed both the individuals and the group vote by chance was low (p < 0.001 and p < 0.002 respectively). The swarms performed on average in the 93rd percentile of users taking the RME test, indicating a significant amplification of social intelligence. In addition,

swarms deadlocked substantially less frequently than when voting, which may lead to improved decision times and greater buy-in among members. Together, this indicates that teams functioning as swarms through an ASI platform amplify their performance on social perception and emotional reasoning tasks. Finally, because prior research shows that social intelligence is significantly correlated with overall team performance, it stands to reason that enabling business teams and other working groups to make critical decisions as real-time swarms could significantly improve their overall team effectiveness. Further research is recommended to explore this.

REFERENCES

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

[2] Rosenberg, L. "Artificial Swarm Intelligence vs Human Experts," Neural Networks (IJCNN), 2016 International Joint Conference on. IEEE.

[3] Befort, K., Proffitt, C., Baltaxe, D., and Durbin, D., "Artificial Swarm Intelligence Technology Enables Better Subjective Rating Judgment in Pilots Compared to Traditional Data Collection Methods." In Proceedings of the Human Factors and Ergonomics Society 2017 International Annual Meeting, HFES 2018

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

[5] Halabi,S, et al. "Radiology SWARM: Novel Crowdsourcing Tool for CheXNet Algorithm Validation" SIIM Conference on Machine Intelligence in Medical Imaging, 2018 (San Francisco, CA).

[6] Baltaxe, D, Rosenberg, L., and N. Pescetelli, "Amplifying Prediction Accuracy using Human Swarms", Collective Intelligence 2017. New York, NY ; 2017.

[7] Rosenberg, L, Pescetelli, N, and Willcox, G. "Human Swarms Amplify Accuracy in Financial Predictions," Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), IEEE Annual, New York, NY, 2017.

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

[9] Rosenberg, L. and Willcox, G. "Artificial Swarm Intelligence Vs Vegas Betting Markets". 2018 11th International Conference on Developments in eSystems Engineering (DeSE), Cambridge, UK 2018

[10] Woolley AW, Aggarwal I, Malone TW (2015) Collective intelligence and group performance. Curr Dir Psychol Sci 24(6):420–424.

[11] Woolley AW , Chabris CF , Pentland A , et al. 2010. Evidence for a collective intelligence factor in the performance of human groups. Science 330 : 686–88.

[12] S. Baron-Cohen, S. Wheelwright, J. Hill, Y. Raste, I. Plumb, The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. J. Child Psychol. Psychiatry 42, 241 (2001).

[13] Vellante M, Baron-Cohen S, Melis M, Marrone M, Petretto DR, Masala C, Preti A: The "reading the mind in the eyes" test: systematic review of psychometric properties and a validation study in italy. Cogn Neuropsychiatry. 2012, 18: 326-354.

[14] Fiske ST, Taylor SE: Social cognition: From brains to culture. 2013, London, UK: SAGE Publications Limited

[15] Kunda Z: Social cognition: Making sense of people. 1999, Cambridge, MA: The MIT Press

[16] Frith CD, Singer T (2008) The role of social cognition in decision making. Philos Trans R Soc Lond B Biol Sci 363:3875–3886