Swarm Intelligence Amplifies the IQ of Collaborating Teams

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

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

Abstract— In the natural world, Swarm Intelligence (SI) is a well-known phenomenon that enables groups of organisms to make collective decisions with significantly greater accuracy than the individuals could do on their own. In recent years, a new AI technology called Artificial Swarm Intelligence (ASI) has been developed that enables similar benefits for human teams. It works by connecting networked teams into real-time systems modeled on natural swarms. Referred to commonly as "human swarms" or "hive minds," these closed-loop systems have been shown to amplify group performance across a wide range of tasks, from financial forecasting to strategic decision-making. The current study explores the ability of ASI technology to amplify the IQ of small teams. Five small teams answered a series of questions from a commonly used intelligence test known as the Raven's Standard Progressive Matrices (RSPM) test. Participants took the test first as individuals, and then as groups moderated by swarming algorithms (i.e. "swarms"). The average individual achieved 53.7% correct, while the average swarm achieved 76.7% correct, corresponding to an estimated IQ increase of 14 points. When the individual responses were aggregated by majority vote, the groups scored 56.7% correct, still 12 IQ points less than the real-time swarming method.

Keywords— Swarm Intelligence, Artificial Swarm Intelligence, Collective Intelligence, Crowdsourcing, Wisdom of Crowds, IQ, AI.

I. INTRODUCTION

For over a century, biologists and ecologists have observed natural species that amplify their group intelligence by forming real-time systems among members. This process, commonly referred to as Swarm Intelligence (SI), enables a wide range of social organisms, from schools of fish and flocks of birds to swarms of honeybees, to solve problems in groups that are intractable to the individuals on their own. [1] In recent years, the technology of Artificial Swarm Intelligence (ASI) has enabled networked human teams to form similar systems and achieve similar benefits when making decisions. Referred to commonly as "human swarm" or "hive minds," these systems have been shown in many studies to significantly amplify the accuracy of human groups across a variety of decision-making tasks, from predicting financial markets and sporting events, to forecasting sales and marketing outcomes. [2-7,17, 18].

Research into ASI is often compared to traditional methods of harnessing the intelligence of human groups. Often referred to as crowdsourcing or tapping the "Wisdom of Crowds," these methods date back to the work of Galton (1907) and generally involve collecting survey responses from individuals which are aggregated statistically, often by plurality vote. [8-15].

While prior studies have shown that groups can increase their performance on standardized tests through statistical aggregation of answers, no prior study has compared statistical aggregation to real-time "human swarming" using a commonly administered IQ test. The objective is to explore if groups demonstrate better performance and thus higher IQ on the RSPM test when working as a real-time swarm, as compared to (i) taking the test as individuals, and (ii) reaching decisions as a group where RSPM answers are provided by majority vote.

II. AMPLIFYING INTELLIGENCE WITH ASI

The fundamental difference between crowd-based methods and swarm-based methods for harnessing team intelligence is that swarming creates a unified system in which the human participants work together in real-time, connected by feedback loops that allow them to converge on solutions together in synchrony. As shown in Figure 1 below, a typical ASI system includes a group of Human Users, each at their own remote location and each using their own computer. Each computer runs a software interface to continuously capture the user's real-time input, and continuously send it to a central processing engine that runs swarming algorithms in the cloud. This engine processes the real-time human input and sends back the evolving real-time collaborative output to each participant, thereby creating a closed-loop system among all users.
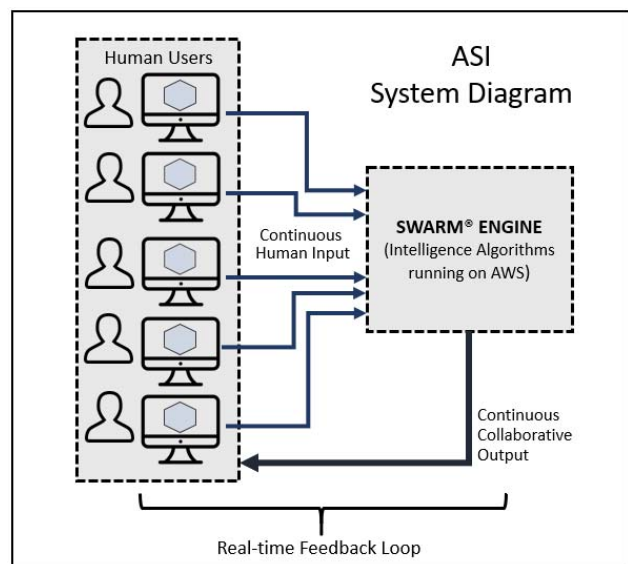

Fig.1. System Diagram for an ASI (Human Swarming) System

For this study, the ASI system used the Swarm® software platform from Unanimous AI. The Swarm platform enables users to connect from anywhere in the world using a standard web browser. Upon logging in, users access an animated client that captures real-time input from all participants simultaneously and feeds the data to the Swarm engine, which runs in the cloud on Amazon Web Services (AWS). The Swarm engine processes the data in real-time and streams the continuous output that represents the collective actions of the full group back to all participants. This creates a feedback loop between the users and the Swarm engine, enabling the group to quickly converge upon optimized solutions together in synchrony.

As shown below in Figure 2, the system used in this study enables networked teams to answer questions by collaboratively moving a graphical puck from a starting location to a target associated one of a set of available answer options. A question appears on the screens of all users at the same time, along with the answer options. Each user provides input by manipulating a graphical magnet with a mouse or touchscreen. By adjusting the position and orientation of their magnet with respect to the moving puck, participants express their input in real-time. Users find this highly intuitive, as they are "pulling" in the direction they want the puck to go, updating their input continuously as the puck moves across their screen. In this way, a team can deliberate, supporting or opposing the pull of others, until the group converges upon a direction and guides the puck to the one solution they can best agree upon. In the example below, the group evaluated the best invention of the 20th century by moving the puck from center of the screen to the answer Antibiotics. The process of deliberation and convergence took 22 seconds.
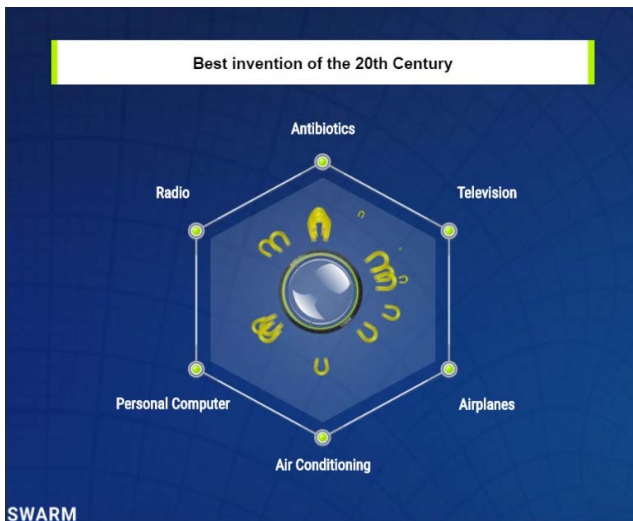

Fig.2. A human swarm choosing between options in real-time

It's important to stress that input from each user is not a discrete vote, but a stream of vectors that varies freely over time. Because all participants can adjust their intent continuously in real time, the group explores the decision-space, not based on the input of any individual member, but based on the emergent dynamics of the full system. This enables a complex negotiation, empowering the group to collectively wrestle with the issue and converge on the most agreeable solution in synchrony.

The complexity of the real-time deliberation can be shown visually using a technique called a Support Density Graph. It is a heat-map showing a time-integration of support (e.g. force) applied towards each of the six answer options over the deliberation period. Figure 3 below shows a Support Density graph for the question posed in Figure 2, the heat-map showing the aggregated force applied by the 35 participants over the 22 second deliberation period. While the group used in this example had thirty-five networked participants, the Swarm platform has been shown to successfully amplify the intelligence of groups as small as three persons and as large as hundreds.
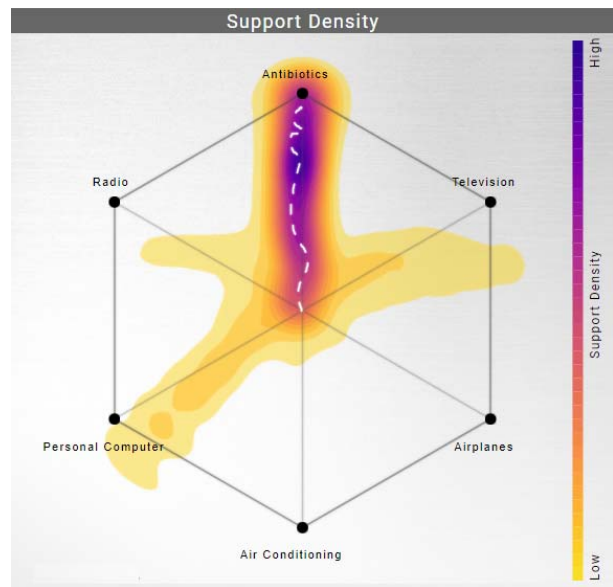

Fig.3. Support Density graph of swarm-based decision

It is important to note that participants do not only vary the direction of their pull using their magnet, but also modulate the magnitude of their pull by adjusting the distance between their magnet and the puck. Because the puck is in continuous motion across the decision-space, users need to continually adjust their magnet so that it stays close to the puck's rim. This is critical, as it requires that all participants to be continuously engaged during the deliberation process, evaluating and reevaluating 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 their applied sentiment 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 ASI system must continuously update their intent during the decision process, or lose their influence over the outcome.

III. EMPIRCAL STUDY OF TEAM IQ

To assess the ability of "human swarming" to amplify the IQ of networked human groups, a study was conducted across a set of five networked groups, each of 6 to 10 members. In total, 42 human subjects participated in this study. All were paid participants from Amazon Mechanical Turk. The study used a set of questions from a commonly used intelligence test known as the Raven's Standard Progressive Matrices (RSPM). This instrument measures the deductive reasoning ability in test-

takers. The RSPM test was chosen for this study because of its acceptance as a reputable measure of IQ as well as it's simple visual format – all questions are presented as a set of images with a missing image that completes a presented pattern. In addition, prior studies have shown the RSPM test gives consistent results when administered to paid participants from Amazon Mechanical Turk [15]. An example question from the RSPM test is shown below in Figure 4, modified to be presented as a five-option solution [16].
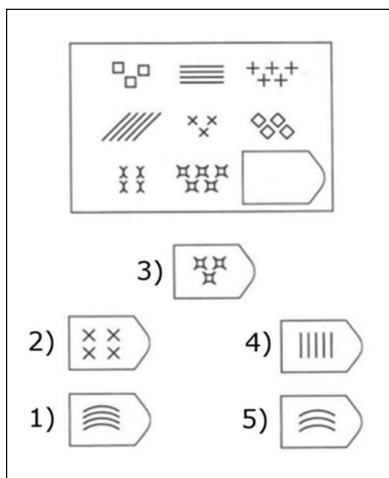

Fig. 4. Sample Question from RSPM test

Questions of the visual format shown above were used in both the individual and group assessments. For individuals, the questions were provided through a simple online survey. For teams using the Swarm platform, the graphical image was displayed to the group along with the swarming interface which allowed them to select among the five answer options. Figure 5 below shows a snapshot of a team in the process of pulling the graphical puck towards one of the answer options.
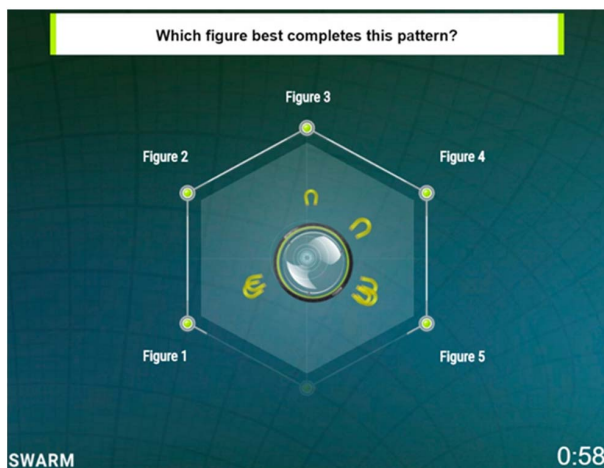

Fig. 5. Swarming Group responding to RSPM question

To prevent cheating on individual survey version of the test, all participants were allocated a maximum of 45 seconds to answer each question. This ensured that participants from Amazon Mechanical Turk would not have time to cheat by looking up answers. When using Swarm, the participants were also time limited to prevent cheating. All answers were recorded in 38 seconds or less.

IV. DATA AND ANALYSIS

Each of the 41 participants was asked to complete a six question RSPM test, providing 246 individual responses. Each of these participants was a member of one of the five real-time groups, consisting of between 6 to 10 members. These groups also completed the RSPM test using the Swarm platform. The RSPM tests were also aggregated by group such that the most popular answer a group was selected by plurality vote. For questions where the vote was split evenly across multiple answers, a "deadlock" was determined and classified as an incorrect response. 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.

The IQ of each response method was calculated as a function of the average and standard deviation of individual accuracies on the modified test, according to the standard IQ formula, where μ is the mean individual score on the test, σ is the standard deviation of individual scores on the test, and X is the score to convert to an IQ:

$$IQ(X) = 100 + 15 * \frac{(X - \mu)}{\sigma} \quad [Eq. 1]$$

V. RESULTS

The distribution of performance across all individuals who participated in this study was approximately normal, with a mean of 3.22 questions correct (53.7%), and a standard deviation of 1.51 questions correct (25.7%). The distribution of individual performances is shown in Figure 6 below.
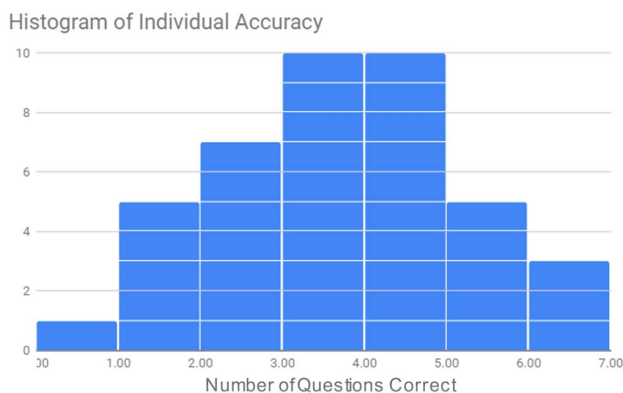

Fig. 6. Histogram of Individual Accuracies

A comparison was performed among each of the three response mechanisms:

1. Individuals taking an RSPM test alone
2. Groups taking an RSPM test by plurality vote
3. Groups taking an RSPM test as real-time swarms

As shown in the Table 1 below, the average individual achieved 53.7% correct on the RSPM test. When aggregating

responses for each working group by plurality vote, the average accuracy increased to 56.7% correct, which corresponds to a 2% increase in IQ score compared to the average individual. When enabling the teams to work together as real-time swarming systems, the performance increased to 76.7% correct, which corresponds to 14-point increase in IQ score compared to the average individual.

Response Method	Percent Correct	Measured IQ (points)	% IQ Increase over Average Individual
Individuals	53.7%	100	--
Group Vote	56.7%	102	2%
Swarm	76.7%	114	14%

Table 1: Test Performance by Response Method

To assess significance of this increase over individuals, a two-sample heteroscedastic t-test was performed that compares the Group Vote and Swarm percent correct to the 41 individuals. Using this method, we find that the swarm significantly outperforms the average individual in the study (p=0.025), but that the group vote does not (p=0.406), indicating that we can only be confident that swarms, and not votes, amplify the intelligence of teams, as measured by this modified RSPM test.

To assess whether the swarm outperformed the group by random chance, a paired t-test was conducted that compared each group's percent correct using each of the two response methods. Using this method, we find that this amplification of team intelligence is statistically significant (p=0.016), indicating that it's unlikely that the teams amplified their intelligence when swarming as compared to voting due to random chance alone.

VI. CONCLUSIONS

The results of this study are very promising for business and engineering teams that collaborate over computer networks. By using an ASI technology like the Swarm platform, networked groups were able to increase their effective IQ by 14 points as compared to the average individual and by 12 points as compared to teams that answered the questions by plurality vote. If teams can make themselves significantly smarter on an IQ test using an online technology such as the Swarm platform, then it's possible that teams can see similar benefits when making strategic decisions, numerical forecasts, and subjective judgements in real-world environments. Future studies should be performed across larger question sets and wider varieties of group sizes and population demographics.

VII. ACKNOWLEDGMENT

REFERENCES

[1] Krause, J., Ruxton, G.D., & Krause, S. (2010). Swarm intelligence in animals and humans. Trends in ecology & evolution, 25 1, 28-34 .

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

[3] Askay, D., Metcalf, L., Rosenberg, L., Willcox, G., "Enhancing Group Social Perceptiveness through a Swarm-based Decision-Making Platform." Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS-52), IEEE 2019

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

[5] 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.

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

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

[8] Galton, F. (1907). Vox populi. Nature, 75, 450.

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

[10] 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.

[11] Surowiecki, J. (2005). The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations. New York: Doubleday.

[12] Tetlock, P. E., Mellers, B. A., Rohrbaugh, N., & Chen, E. (2014). Forecasting tournaments: tools for increasing transparency and improving the quality of debate. Current Directions in Psychological Science, 23(4), 290–295. http://dx.doi.org/10.1177/0963721414534257.

[13] Wolf, M., Krause, J., Carney, P. A., Bogart, A., & Kurvers, R. H. J. M. (2015). Collective intelligence meets medical decision-making: The collective outperforms the best radiologist. PLOS ONE, 10(8), e0134269. http://dx.doi.org/10.1371/journal.pone.0134269.

[14] Malone, T. W., & Bernstein, M. S. (2015). Handbook of collective intelligence. Cambridge, MA: The MIT Press.

[15] Vercammen A, Ji Y, Burgman M, 2019, The collective intelligence of random small crowds: A partial replication of Kosinski et al. (2012), Judgement & Decision Making, Vol: 14, Pages: 91-+, ISSN: 1930-2975

[16] Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. Cognitive Psychology, 41(1), 1–48. http://dx.doi.org/10.1006/cogp.1999.0735.

[17] Rosenberg, L., Willcox, G., Askay, D., Metcalf, L., and Harris, E., "Amplifying the Social Intelligence of Teams Through Human Swarming," 2018 First International Conference on Artificial Intelligence for Industries (AI4I), Laguna Hills, CA, USA, 2018, pp. 23-26.

[18] Metcalf, L., Askay, D. A., & Rosenberg, L. B. (2019). Keeping Humans in the Loop: Pooling Knowledge through Artificial Swarm Intelligence to Improve Business Decision Making. California Management Review. https://doi.org/10.1177/0008125619862256