Crowds vs Swarms, a Comparison of Intelligence

Louis Rosenberg and David Baltaxe
Unanimous A.I.
2443 Fillmore Street, #116
San Francisco, CA. USA
david@unanimousai.com

Niccolo Pescetelli
University of Oxford
Christ Church College
Clarendon, UK
niccolo.pescetelli@chch.ox.ac.uk

***Abstract*— for well over a century, researchers in the field of Collective Intelligence have shown that groups can outperform individuals when making decisions, predictions, and forecasts. The most common methods for harnessing the intelligence of groups treats the population as a "crowd" of independent agents that provide input in isolation in the form of polls, surveys, and market transactions. While such crowd-based methods can be effective, they are markedly different from how natural systems harness group intelligence. In the natural world, groups commonly form real-time closed-loop systems (i.e. "swarms") that converge on solutions in synchrony. The present study compares the predictive ability of crowds and swarms when tapping the intelligence of human groups. More specifically, the present study tasked a crowd of 469 football fans and a swarm of 29 football fans in a challenge to predict 20 Prop Bets during the 2016 Super Bowl. Results revealed that the crowd, although 16 times larger in size, was significantly less accurate (at 47% correct) than the swarm (at 68% correct). Further, the swarm outperformed 98% of the individuals in the full study. These results suggest that swarming, with closed-loop feedback, is potentially a more effective method for tapping the insights of groups than traditional polling.**

Keywords— Swarm Intelligence, Artificial Intelligence, Human Swarming, Wisdom of Crowds, Collective Intelligence

I. Introduction

In the field of Collective Intelligence, one story is told more often than any other – the tale of Sir Francis Galton and his pioneering use of a crowd to estimate the weight of an ox at a county fair in England in 1906. 800 people tried their luck at guessing the ox's weight – not experts, but a mix of farmers and butchers and regular fairgoers. Galton's goal was to show that average voters were not very smart, but he was shocked to find that the average of the 787 estimates (800 minus 13 that were illegible) was nearly perfect. This was true despite the fact that the vast majority of the individual estimates were way off [1]. And thus was born the field of Collective Intelligence.

In the century since, the most common methods used for tapping the wisdom of crowds have not changed much, still focusing on the averages of independently contributed votes and estimates. In some cases, the methods for tallying votes are dependent upon the prior accuracy of individual participants. This has been used to improve crowd averages [2,3,4] but such methods require task-relevant history and then use that history to alter the make-up of the crowd. This is not the same as improving the methods by which the crowd's intelligence is tapped. This begs the question – *is there a fundamentally better way to harness the intelligence of human groups?*

To answer this question, researchers have looked to Mother Nature for guidance, finding that many species have evolved methods for tapping the intelligence of groups [5,6,7,8,9,13]. What nature does not do is collect independent samples and then aggregate the data after the fact, the way Galton did in his famous experiment. Instead, nature forms real-time closed-loop systems with continuous feedback, enabling large groups to work in synchrony and converge on solutions together. Known as Swarm Intelligence (SI), this is the reason why birds flock, fish school, and bees swarm – they make better decisions together than the independent organisms could make on their own [10, 11, 12]. Of course we humans didn't evolve the ability to swarm, for we lack the innate connections that other species use to establish feedback-loops among members. Ants use chemical traces. Fish detect ripples in the water around them. Bees use vibrational gestures. Birds detect motions propagating through the flock. This said, new technologies have enabled groups of online human users to form real-time closed-loop systems modeled after natural swarms.

Known as Artificial Swarm Intelligence (ASI), or more simply "Human Swarming", these computational methods enable online human groups to work together in real-time by forming a unified system that can answer questions, make predictions, reach decisions, or take actions. As a system, human swarms can collectively explore a decision-space and converge upon preferred solutions. Prior studies have shown that by working in swarms, human groups can outperform their individual members as well as outperform groups taking traditional votes or polls [5,6]. For example, in a prior study, a randomly selected human group was tasked with predicting the 2015 Oscars, both by taking a poll and by forming a swarm. Across 48 participants, the average poll result achieved 6 of 15 correct predictions (40% success). When taking the most popular prediction in the poll as a crowd aggregate, the group improved, achieving 7 of 15 correct predictions (47% success). But when working together as a swarm, the group improved far more, achieving 11 of 15 correct predictions (73% success). This suggests that human swarming may be a superior method for tapping the wisdom of crowds as compared to traditional votes, polls, and surveys. The present study aims to explore the power of crowds vs. swarms further, by comparing a poll of nearly 500 people with a swarm of only 29 participants in a more challenging prediction task.

II. Enabling Human Swarms

To evoke a real-time Artificial Swarm Intelligence (ASI) among groups of networked humans, technology is required to

978-1-5090-3502-1/16/$31.00 ©2016 IEEE

close the loop among members. To address this need, an online platform called **UNU** was developed to allow distributed groups of users to login from anywhere around the world and participate in a closed loop swarming process. As shown in Figure 1, users answer questions by collectively moving a graphical puck to select among a set of alternatives. The puck is modeled as a physical system with a defined mass, damping and friction. Users provide input by manipulating a graphical magnet with a mouse or touchscreen. By positioning their magnet, users impart their personal intent as a force vector on the puck. The input from each user is not a discrete vote, but a stream of vectors that varies freely over time. Because the full set of users can adjust their intent at every time-step, the puck moves, not based on the input of any individual, but based on the dynamics of the full system. This results is a real-time physical negotiation among the members of the swarm, the group collectively exploring the decision-space and converging on the most agreeable answer.

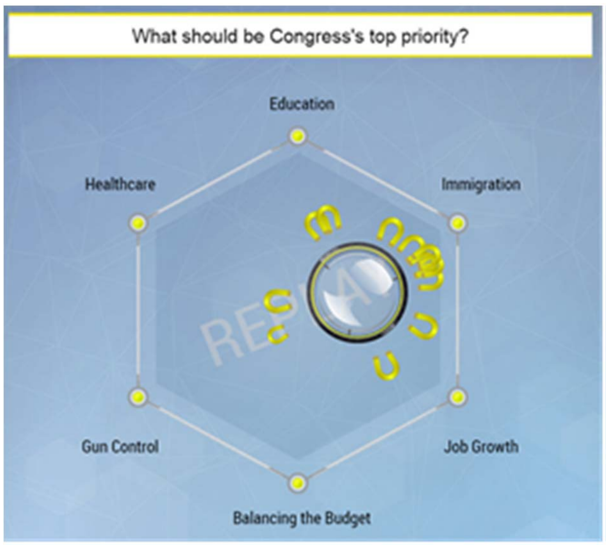

Fig 1. A human swarm comprised of user-controlled magnets.

It's important to note that users can only see their own magnet during the decision, not the magnets of others users. Thus, although they can view the puck's motion in real time, which represents the emerging will of the swarm, they are not influenced by the specific breakdown of support across the available options. This limits social biasing. It's also important to note that users don't just vary the direction of their input, but also the magnitude by adjusting the distance between the magnet and the puck. Because the puck is in motion, to apply full force users need to continually move their magnet so that it stays close to the puck's rim. This is significant, for it requires all users to be engaged during the decision process. If they stop adjusting their magnet to the changing position of puck, the distance grows and their applied force wanes.

III. Testing Crowds vs Swarms

To compare the predictive ability of crowds and swarms, a formal prediction experiment was conducted using an easily verifiable set of prediction events – a set of twenty independent "Proposition Wagers" on Super Bowl 50, which took place on Sunday, February 7, 2016. Known generally as "Prop Bets" these are simple binary wagers aimed at the general public rather than sophisticated sports fans, and are defined by Vegas odds-makers to have equal probabilities. In other words, if Vegas sets the bets correctly, the average person placing the 20 wagers would get 10 correct and 10 incorrect. Of course, the attraction for betting on Prop Bets is that most people believe they can outsmart the odds makers and choose the more likely alternative. In practice, this not the case – most people perform at the expected odds which is why Vegas makes a healthy profit.

To compare the predictive ability of Crowds and Swarms, two experimental groups were fielded – Group A and Group B. Group A was comprised of 469 self-identified football fans who were randomly assembled and asked to make predictions for each of the 20 prop bets by working together as a crowd. This entailed each of the 469 participants filling out an online survey to indicate their individual picks for each of the 20 bets. A set of "crowd-based wagers" were then generated by taking the most popular answers across the full set of participants. Group B was comprised of 29 self-identified football fans who were randomly assembled and asked to make predictions for each of the 20 prop bets by working together in real time as a swarm using the UNU software platform. A set of "swarm-based wagers" were then generated by the group converging in synchrony as a real-time dynamic system. Prior to participating as a swarm, the 29 members of Group B also recorded their individual predictions by completing surveys.

It should be noted that Prop Bets for the Super Bowl are a mix of sports predictions and pop-culture predictions, making them a good target for a random sampling of casual sports fans. For example, one of the target questions was firmly sports related – *"Which team will score first, the Broncos or the Patriots?"* while another of the target questions was more about predicting the culture of football – "*Who will the MVP thank first when interviewed after the game?*" Figure 2 below shows the swarm in the process of answering that question.

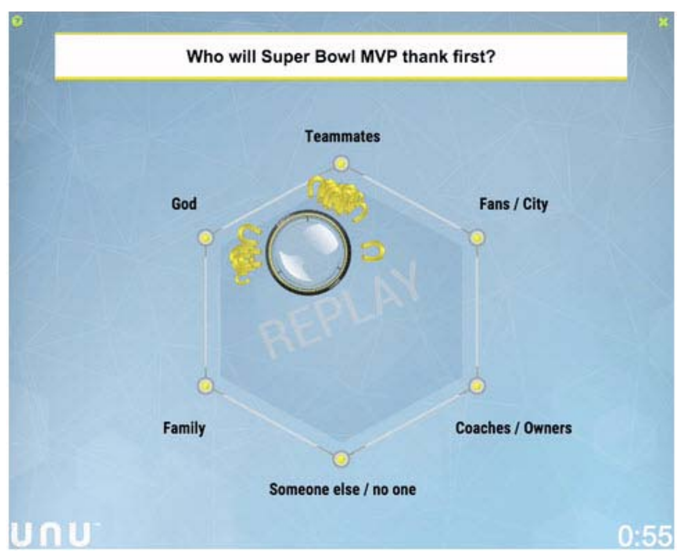

Fig 2. A human swarm answers a Super Bowl prop bet.

IV. RESULTS

Looking first at crowd-based wagers, the group of 469 football fans collectively achieved 9 correct picks out of 19 wagers placed. It should be noted that the 20th wager was canceled by Vegas at the end of the game because it involved an event that did not transpire. Thus, the 9 out of 19 success rate for the crowd-based wagers translated into a **47%** accuracy and a small gambling loss. In this case, tapping the wisdom of the crowd by taking an average of poll results did not allow the participants to beat the Vegas odds makers and did not achieve a profit on their wagers. In fact, the crowd-based bets performed at the expected odds distribution for individual bets.

Looking next at the swarm-based wagers, the group of 29 randomly selected football fans who worked together as a real-time closed-loop system, collectively achieved 13 correct picks of the 19 wagers placed. Thus by working together as a swarm, this group of 29 individuals produced a **68%** accuracy rate. This translated into a 36% gambling gain on the placed wagers. In other words, the swarm of randomly selected football fans were able to defy Vegas odds by pooling their knowledge and intuition in real-time as a Swarm Intelligence. This conforms to prior studies that show similar results. [6, 7]

It appears that by forming a Swarm Intelligence using the UNU platform, the group of 29 randomly selected football fans were able to significantly amplify their group intelligence with respect to forecasting Super Bowl prop bets and thereby provide deeper insights. The question remains, was the amplification of gambling insight shown by the Swarm statistically significant as compared to the 29 individuals, if they had each worked alone and made their own wagers. Similarly, the question remains, was the amplification of gambling insights shown by the Swarm statistically significant as compared to the much larger pool of 469 individuals who participated in the crowd.

To answer these questions, the swarm's performance was compared to the statistical distribution of performance of individual members of Group A (the crowd) and Group B (the swarm). In Figure 3 below, the swarm's performance is represented by the red line on each graph. The x-axis represent how many questions the individuals answered correctly. The y-axis represent how many people answered correctly to that particular number of questions. For both groups, most people answered correctly 10 questions out of 19.

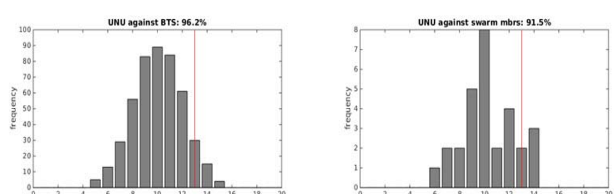

Fig 3. Swarm Performance vs Group Members

On the left side of Fig 3, the performance distribution of the 469 member crowd is shown. As indicated by the red line, the swarm outperformed the individuals in the crowd by 2 standard deviations (Z=1.99). In fact, out of the 469 randomly selected people, only 4 individuals did better than the swarm. In other words, the swarm outperformed 99% of the individual members of the crowd. This suggests a significant amplification of intelligence resulting from the swarming process.

On the right side of Fig 3, the performance distribution of the 29 members of the swarm is shown. In other words, this compares the swarm as a unified system with the individual participants in the swarm itself. As shown by the red line, the swarm outperformed the individuals in the group by 1.7 standard deviations (Z=1.72). In fact, out of the 29 randomly selected people, only 3 individuals did better than the swarm. This corresponds with the swarm outperforming 90% of the individual swarm members. This suggests a clear amplification of intelligence resulting from the swarming process.

As a further statistical test, the swarm's performance was compared to the distribution of correct answers expected by chance. Using a random sampler, 19 individuals were selected at random (with replacement) and the n-th individual answer taken as the answer of the sample for the n-th question; that is, the first randomly picked individual provides the answer for the 1st question, the second individual provides the answer for the 2nd question and so on. This process was repeated 10,000 times until 10,000 nominal groups were created. Figure 4, below, represents how many times these groups answered correctly exactly x questions.

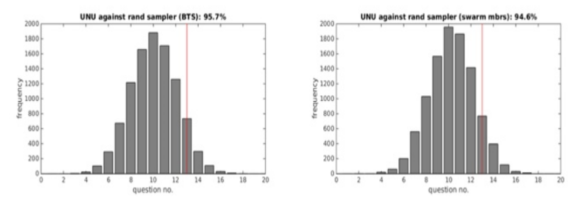

Fig 4. Swarm Performance vs Groups of Randomly Selected Members

In Figure 4, the chart on the left depicts the distribution of correct answers when individuals were randomly selected from among the 469 member crowd; the chart on the right shows the distribution when selections came from the members of the swarm itself. Again, the swarm's performance is represented by the red line. This analysis shows that the swarm outperformed chance significantly, falling almost 2 standard deviations from the mean in both comparison groups (Z=1.90 and 1.81, respectively).

V. DISCUSSION AND CONCLUSIONS

Can human swarming amplify the intelligence of groups, enabling a population of individual forecasters to perform better together than the vast majority could perform alone? The results of this study, along with prior studies, suggest this is the case. This bolsters the premise that by working together in closed-loop systems, with real-time feedback control, groups can more effectively explore a decision-space and converge on optimal solutions. Although football wagers were used as the testing framework for this study, we believe the result are applicable to a wide range of applications where groups of individuals can contribute a diverse set of opinions, insights, and intuitions.

Is human swarming a more effective method for harnessing the intelligence of groups than traditional "Wisdom of Crowd" methods for aggregating forecast data? The results of this study, along with prior studies, suggest this is the case. The 29 members of the swarm performed significantly better as a synchronous system than as a collection of independently surveyed participants. In addition, the swarm outperformed the crowd-based poll despite the fact that the poll had a sample size that was 16-times larger. This suggests that swarming not only provides more accurate insights, it enables insights to be attained from human groups with a much smaller sample sizes than polls or surveys. Future work is needed to quantify the effective size differences between crowds and swarms.

Should we be surprised by the effectiveness of Artificial Swarm Intelligence for harnessing group insights? If we look to Mother Nature as our guide – probably not. After all, the results of this study parallel the benefits of swarming among honeybees and other social organisms, where the decisions are reached in real-time synchrony as closed-loop dynamic systems [5, 12]. In fact, the swarming algorithms used by the UNU platform on which this study was run, were modeled specifically after the decision making processes of honeybees [7,13], so it's reasonable to expect a similar amplifications of intelligence.

ACKNOWLEDGMENT

This work was directly supported by Unanimous A.I., the maker of the UNU platform for real-time human swarming. For more information about UNU, visit http://UNU.ai.

REFERENCES

[1] Galton F (1907) Vox populi. Nature 75, 450–451

[2] Whitehill J, Wu TF, Bergsma J, Movellan JR, Ruvolo PL (2009) Whose vote should count more: Optimal integration of labels from labelers of unknown expertise. Adv. Neur. In. 2035–2043

[3] Lee MD, Steyvers M, De Young M, Miller B (2012) Inferring expertise in knowledge and prediction ranking tasks. Top. Cogn. Sci. 4(1), 151–163. doi: 10.1111/j.1756-8765.2011.01175.x PMID: 22253187

[4] Madirolas G, de Polavieja GG (2015) Improving Collective Estimations Using Resistance to Social Influence. PLoS Comput Biol 11(11): e1004594. doi:10.1371/journal.pcbi.1004594

[5] Seeley, Thomas D. Honeybee Democracy. Princeton University Press, 2010.

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

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

[8] Palmer, Daniel W., et al. "Emergent Diagnoses from a Collective of Radiologists: Algorithmic versus Social Consensus Strategies." Swarm Intelligence. Springer International Publishing, 2014. 222-229.

[9] Eberhart, Russell, Daniel Palmer, and Marc Kirschenbaum. "Beyond computational intelligence: blended intelligence." Swarm/Human Blended Intelligence Workshop (SHBI), 2015. IEEE, 2015.

[10] J.A.R. Marchall, R. Bogacz, A. Dornhaus, R. Planque, T.Kovacs, N.R. Franks, On optimal decision making in brains and social insect colonies, J.R. Soc Interface 6,1065 (2009).

[11] I.D. Couzin, Collective Cognition in Animal Groups, Trends Cogn. Sci. 13,36 (2008).

[12] 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. Behavioral Ecology and Sociobiology 54 (5) 511-520.

[13] L. B. Rosenberg, "Human swarming, a real-time method for parallel distributed intelligence," Swarm/Human Blended Intelligence Workshop (SHBI), 2015, Cleveland, OH, 2015, pp. 1-7.