ONLINE APPENDIX

This appendix continues the discussion of Z Energy's use of Swarm.AI® to converge on strategic priorities. Also presented are implications for future research.

Z Energy: Converging on Strategic Priorities

For interested readers, the following demonstrates the Swarm.AI® post-processing diagnostics available to Z Energy's leadership and employee teams.

The swarm decision space shows the swarm of 61 individual contributors deciding which initiative should receive the lowest priority. The post-processing diagnostics shown in the figures below are available immediately after each swarm concludes.

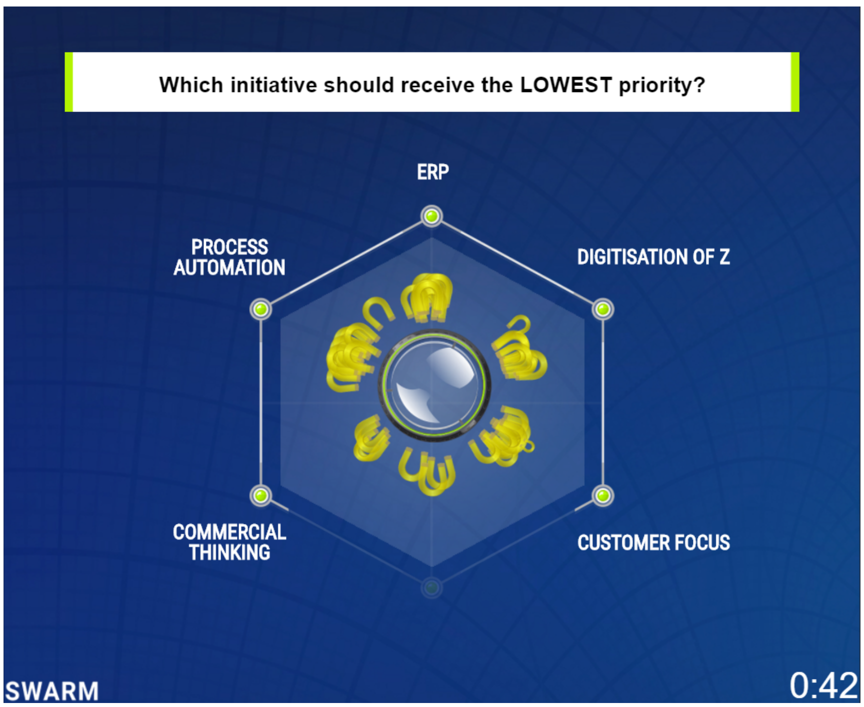


Figure A1: Snapshot of a networked "human swarm" answering a question as a unified system.

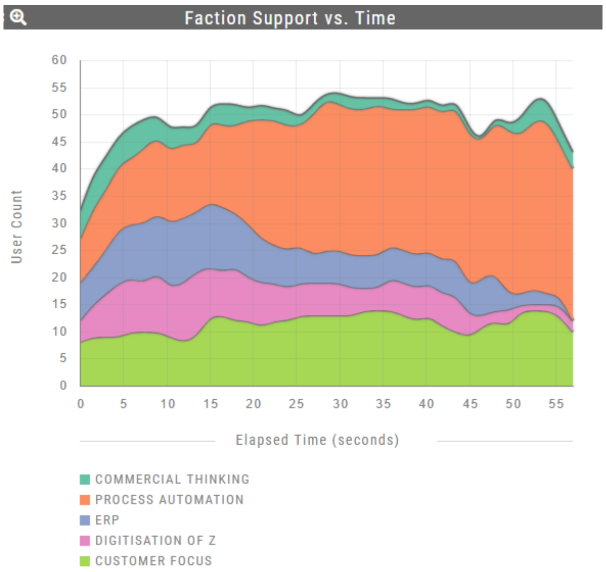
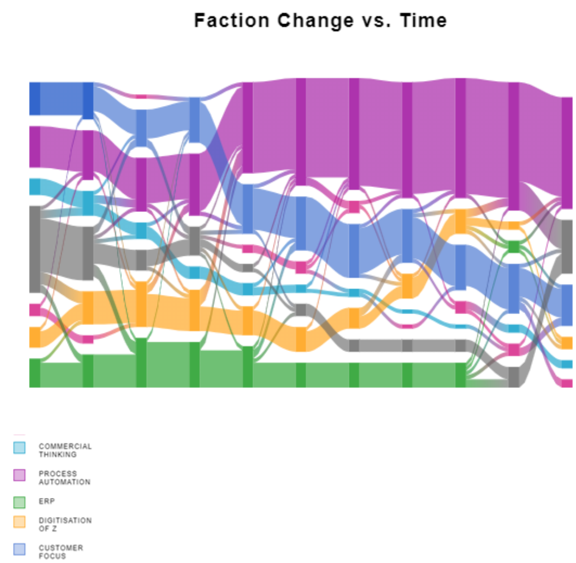
Participant Support vs Time	**Support Change vs Time**
	

Figure A2: Plots of the behavioral data collected during the real-time swarming process

The left side (by the y-axis) of the stacked area chart shows the initial starting positions of swarm participants. These starting positions (T=0 seconds) reflect top-of-mind opinions and, as the stacked area chart shows, there were sizable factions supporting each of the five strategic options as lowest priority. If a vote had been taken at T=0 seconds, Customer Focus would have received lowest priority. At T=10 seconds, as undecided participants joined the deliberation process, ERP and Process Automation emerged as close contenders for lowest priority. At T=15 seconds, there was a three-way tie for lowest priority between Process Automation, Customer Focus, and ERP. By T=20 seconds, the group began to converge on Process Automation and the end-state at T=59 seconds shows the swarm's decision on Process Automation as lowest priority. The stacked area chart shows that the group wrestled with the alternative initiatives for more than 15 seconds before finding a path to the priority they could converge on. The Sankey Chart provides another way of looking at the same information and shows how options gained or lost support over time, as well as the options from which the emergent loser (lowest priority) captured support.

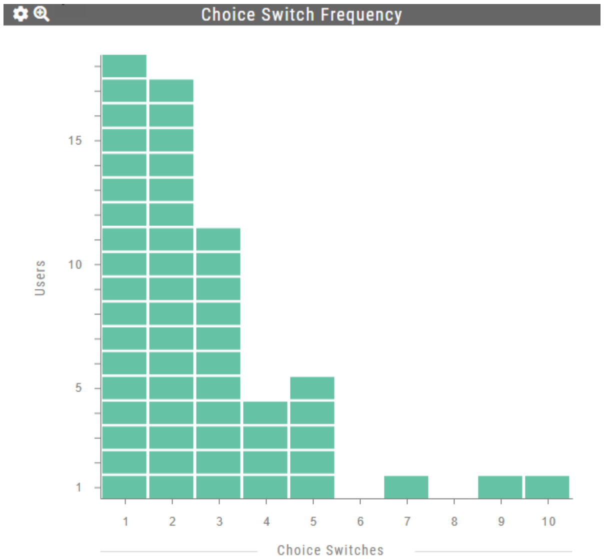
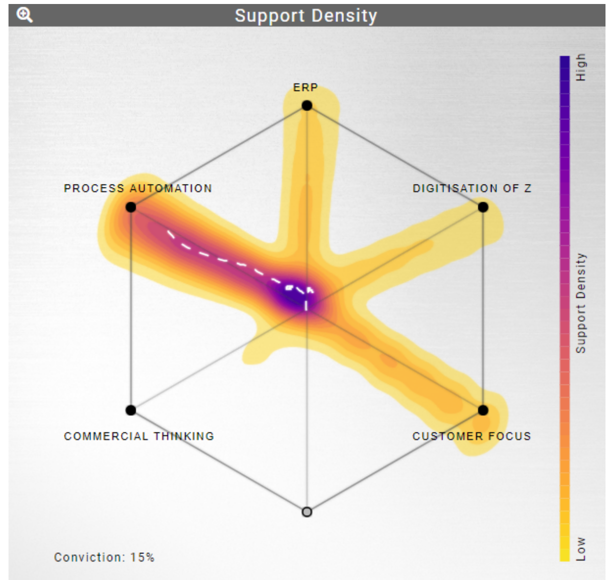
Histogram of Switch Frequency	Support Density Map
	

Figure A3: Plots of the behavioral data collected during the real-time swarming process

The histogram more clearly shows how many times participants switched their opinions during the 59 second deliberation. The data produced by the AI engine show that only 18 participants were "entrenched," meaning they never switched their priority ratings. On the other hand, the data show that 23 participants switched their opinion three or more times during the deliberation. The rapidly evolving factions shown in the Sankey chart illustrate how a human swarm can negotiate a consensus and reach a conclusion on which strategic option to eliminate.

At the same time, the rapidly evolving factions in the Sankey chart reflect uncertainty over which of the five initiatives should be identified as lowest priority. The Support Density map shows that Commercial Thinking was not a contender for lowest priority. While Process Automation was ultimately selected as lowest priority, it also shows that the swarm deliberated over Customer Focus, ERP, and Digitization, which produced a low confidence score of 15% for the swarm. The ability to visualize the uncertainty provides opportunity for discussion about the thinking behind the swarm’s deliberation.

Implications for Research

Our discussion of ASI points toward six broad areas of opportunity for future research.

Blending AI and Humans. There have been several calls for furthering an understanding of how to blend human and machine capabilities to improve decision making, to provide better predictions, and to enhance collective intelligence.[i] ASI provides an answer to these calls and also demonstrates the value of biomimicry. Future research into combining machine and human intelligence might likewise be inspired from other models of intelligent behavior found in nature, such as bacteria growth or root system development.[ii] Moreover, the present paper identifies two features of an AI decision-making platform—implicit confidence scores and parallel pooling of intelligence—that enable both large and small groups of participants to make more accurate predictions. Future research and development of decision support systems may incorporate and build upon these features. Finally, while current use of ASI has focused on pooling *human* intelligence, future research might examine ways in which machine agents informed by large datasets can become active participants in a swarm that also includes humans. The spatial-visual nature of ASI offers a compelling medium through which humans and machines might deliberate together to reach a joint decision.[iii]

Implicit Confidence Assessment. One contribution of this research is that it identifies a novel way of inferring participants' confidence in their swarm behavior (i.e. their pathway toward the decision option they ultimately support). Rather than soliciting a self-rating of their confidence in their judgments, ASI implicitly assigns a real-time confidence score to each participant in a swarm based on the movement of their magnet. This is significant because people are routinely overconfident when self-rating their confidence in their judgments.[iv] As a result, scholars advocate for identifying implicit or behavioral assessments of confidence, such as trade volume in prediction markets, that can be used to weight contributions from individuals when pooling intelligence.[v] Given the success of ASI in increasing the accuracy of judgments, more research is needed to understand the utility of behavioral measures of confidence. For example, there is potential in using implicit confidence scores to provide real-time feedback to individuals, which could aid them in calibrating their judgments.[vi] Another option is to examine the behavioral confidence cues that respondents exhibit when answering online polls or surveys. For example, the time spent on a question or the time spent oscillating between (i.e., selecting and deselecting) several different response options may be useful for generating implicit measures of respondent confidence.

Decision Satisfaction and Decision Acceptance. The Z Energy case reveals qualitative support for employee satisfaction with the use of ASI to support decision-making. Satisfaction with decision outcomes is critical to sustainable agreement and implementation.[vii] Whereas the use of the Delphi technique can generate higher quality decisions, group members are more likely to accept decisions reached by consensus.[viii] Similar to holding group discussions that lead to a consensus decision, ASI allows groups of people to converge upon a decision that the group can agree on. There is opportunity to explore whether the decisions produced by human swarms result in greater group satisfaction than decisions reached by more traditional means, such as surveys, votes, or group discussions. Additionally, research shows that decision acceptance is greater with direct participation.[ix] What is not known is whether or not the use of ASI results in greater decision acceptance than other forms of participatory or consensus-based decision-making, such as group discussions.

Questioning Communication. In the research on ASI referenced and the cases presented, swarm members did not explicitly communicate or share information with each other while they were swarming. Rather, they simply expressed their preferences—informed by their explicit and tacit knowledge—through their magnets. This indirect, technology-mediated form of communication was sufficient for human swarms to pool their intelligence and to provide more accurate predictions about known unknowns. This provokes an important question: What are conditions under which communication is helpful for or a hindrance to pooling knowledge? Future research is needed to identify the boundaries of when sharing information explicitly through spoken or written communication is more advantageous than expressing preferences behaviorally through a software interface. Similarly, there is an opportunity to explore further how participants make sense of their experience in a swarm. One approach might be to consider the ASI interface as a boundary object. Boundary objects are shared visual representations of knowledge that are shared among and editable by people. They provide a way to jointly represent and to transform their knowledge, which makes them particularly well-suited for making tacit knowledge more explicit.[x] Another approach might focus on investigating how seeing the swarm's collective preference through the puck fosters metacognition, which can make participants reevaluate their preference.[xi]

Integrating Intelligence Pooling Processes. Little work has considered how to integrate methods of pooling intelligence into an optimized process that leverages the strengths of different methods.[xii] For example, accurate predictions arise from cycles of divergent thinking, where ideas are generated, and convergent thinking, where ideas are evaluated and agreed upon.[xiii] Generating ideas is a particular strength of crowdsourcing platforms and ASI enables groups of all sizes to converge on options, suggesting the promise of a process that blends the generative capabilities of crowdsourcing with the convergent strengths of ASI. Alternatively, swarms of predictors, rather than individuals or groups, might be used to determine which contracts to buy in a prediction market. Finally, individual and group performance in a swarm might be used to identify talented participants to join a team of forecasters or a prediction market.

Swarm Dynamics. Additional research is needed to determine the enabling and limiting dynamics of human swarming. After all, honey bees fail to make optimal decisions 20% of the time.[xiv] Similarly, a greater understanding of the factors that suppress the effectiveness of swarms is needed. One area of exploration is the relationship between participant signals and herding behavior in the context of swarming. ASI is unique because it enables small and large numbers of participants to simultaneously and anonymously interact in parallel with each other. By contrast, herding behavior is commonly studied in contexts where participants interact sequentially. This reveals a need to understand how and if herding occurs in a swarming context. Indeed, prior research on ASI demonstrates that swarms have increased accuracy on decisions both when individual magnets are visible and when magnets are invisible to the swarm.[xv] While swarming is known to be resilient to noise in the system, there is a need to understand the extent of this resilience and how herding occurs in ASI. Future research is needed to explore whether or not making participants' magnets visible to all members of a swarm promotes herding behavior. Finally, other dynamics of swarming are ripe for further investigation, such as exploring the types of biases that might operate in swarms but that have not yet been identified.

---

1. Mohammad Hossein Jarrahi, "Artificial intelligence and the Future of Work: Human-AI Symbiosis in Organizational Decision Making," *Business Horizons,* 61/4 (July/August 2018): 577-586; Agrawal et al., op. cit.; Malone, op. cit.
2. For example, Ben Niu, Hong Wang, Qiqi Duan, and Li Li, "Biomimicry of Quorum Sensing using Bacterial Lifecycle Model," In *BMC Bioinformatics*, 14/8 (May 2013): S8; Lianbo Ma, Hanning Chen, Xu Li, Xiaoxian He, and Xiaodan Liang, "Root System Growth Biomimicry for Global Optimization Models and Emergent Behaviors," *Soft Computing,* 21/24 (December 2017): 7485-7502.
3. For example, the field of visual analytics keeps humans in the loop of big data analysis through the intuitive ability of humans to work with interactive visualizations. See Alex

Endert, M. Shahriar Hossain, Naren Ramakrishnan, Chris North, Patrick Fiaux, and Christopher Andrews, "The human is the loop: new directions for visual analytics," *Journal of Intelligent Information Systems*, 43/3 (December 2014): 411-435.

4. Philip E. Tetlock and Dan Gardner, *Superforecasting: The Art and Science of Prediction* (New York, NY: Random House, 2016).
5. Researchers suggest that prediction markets can improve their accuracy by automatically using trading volume from each participant to weight their judgment. See Graefe and Weinhardt, op. cit.
6. Overconfidence is considered the most significant cognitive biases in decision making. See Daniel Kahneman, *Thinking Fast and Slow* (New York, NY: Farrar, Straus and Giroux, 2011).
7. Kaner, op. cit.; Smith, op. cit.
8. Robert C. Erffmeyer and Irving M. Lane, "Quality and Acceptance of an Evaluative Task: The Effects of Four Group Decision-Making Formats," *Group & Organization Studies*, 9/4 (December 1984): 509-529.
9. De Dreu, et al., op cit.
10. Paul R. Carlile, "A Pragmatic View of Knowledge and Boundaries: Boundary Objects in New Product Development," *Organization Science*, 13/4 (August 2002): 442-455.
11. Nick Yeung and Christopher Summerfield, "Metacognition in Human Decision-Making: Confidence and Error Monitoring," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367/1594 (May 2012): 1310-1321; Stephen M. Fleming and Nathaniel D. Daw, "Self-Evaluation of Decision-Making: A General Bayesian Framework for Metacognitive Computation," *Psychological Review*, 124/1 (January 2017): 91-114.
12. For examples of integrative papers, see Tobias Prokesch, Heiko A. von der Gracht, and Holger Wohlenberg, "Integrating prediction market and Delphi methodology into a foresight support system—Insights from an online game," *Technological Forecasting and Social Change*, 97 (August 2015): 47-64; Simon Kloker, Tim Straub, and Christof Weinhardt, "Designing a Crowd Forecasting Tool to Combine Prediction Markets and Real-Time Delphi," in *Lecture Notes in Computer Science*, Vol. 10243 (Cham, Switzerland: Springer International, 2017): 468-473; Kloker, Simon, Frederik Klatt, Jan Höffer, and Christof Weinhardt, "Analyzing Prediction Market Trading Behaviour to Select Delphi-Experts," *Foresight*, 20/4 (August 2018): 364-374.
13. Schoemaker and Tetlock, op. cit.
14. Seeley and Buhrman, ibid.
15. For example, radiologists diagnosing pneumonia were able to see the magnets of all participants, yet the swarm outperformed both individuals and machine learning AI. In contrast, swarms of student business teams made more accurate decisions regarding the emotional state of a person when then magnets were invisible. See Rosenberg et al., "Artificial Swarm Intelligence Employed to Amplify Diagnostic Accuracy in Radiology."; David Askay, Lynn Metcalf, Louis Rosenberg, and Gregg Willcox, "Enhancing Group Social Perceptiveness through a Swarm-based Decision-Making Platform," In *Proceedings of the 52nd Hawaii International Conference on System Sciences*, (January 2019): 492-501.