Scalable Human Interaction with Robotic Swarms

Brian Pendleton[*] and Michael A. Goodrich[†]

Brigham Young University, Provo, UT 84604, USA

In this paper we evaluate the scalability of human-swarm interaction (HSI) in terms of operator workload and asses the impact of three control methods on swarm performance and operator workload. Specifically, we investigate the ability of HSI to (1) overcome fan-out limitations of traditional supervisory control and (2) manage a wide range of team sizes without altering the operator's control strategy. In order to evaluate our hypothesis, we conduct a user study which we back with validation on physical robots. We evaluate three high-level control methods – leader, predator, and stakeholders – on swarms of 20, 50, and 100 individuals. We find that larger swarm sizes allow increased swarm performance without increasing operator workload, overcoming the fan-out limitations of traditional supervisory control. We further find that control style substantially affects both swarm performance and operator workload, illustrating the impact design decisions can have on the human-swarm system.

I. Introduction

Human interaction with large-scale robot teams has received considerable attention in the literature. The number of robots in use is increasing and likely to continue increasing in the future as robots continue to assist humans in an increasing variety of roles. As the number of robots increases human operators will need to coordinate and control the actions of large teams of robots. Most approaches to multi-robot supervisory control require that an operator divide his or her attention between robots and is limited by the number of robots that a human or team of humans can reasonably control.[1] Olsen and Wood[1] define this number as the *fan-out* of the human-robot team and show that it can be calculated empirically from data.

Inspired by biological systems, swarm robotics and human-swarm interaction (HSI) offer an alternative approach to supervisory control.[2] In nature, large numbers of individuals frequently form large collectives or "swarms".[3] Natural swarms scale from dozens to tens of thousands of individuals (e.g. bird flocks) and allow many distinct individuals to function as a large cohesive unit even in the absence of leaders and centralized information (Figure 1). For thousands of years, humans have used a variety of methods to direct and influence these decentralized systems (e.g. cattle ranching and sheep herding). The field of swarm robotics seeks to understand the principles underlying these systems and encode them into robots to create large robust robot teams. The goal of human-swarm interaction (HSI) is to develop methods for directing and influencing these large decentralized systems.

Şahin[2] identifies three properties of swarms that are desirable for multi-robot systems. First, swarms are scalable. Swarm algorithms use interactions among individuals in a local neighborhood, allowing them to scale from ten or twenty individuals to tens of thousands.[2] Second, swarms are robust. Because swarms contain no explicit hierarchy or designated leader, individuals can join or drop out from the group without substantial impact to large scale group behavior.[4] Third, swarms are flexible. In swarm systems, group-level behaviors are built from simple individual-level building blocks.[5] Once understood, these building blocks can be adapted and combined to produce a wide variety of behaviors at the group level (see McLurkin[6]).

While swarms have been shown to be algorithmically scalable, the management of large robot teams also requires scalable *interaction* between humans and robots. The operator's mental workload – defined by Sanders[7] as "a measurable quantity of the information processing demands placed on an individual by

---

[*]MS Candidate, Computer Science.

[†]Professor, Computer Science.

[a]Photo by Kasi Metcalf and licensed under the Creative Commons Attribution Non-commercial No-derivatives 2.0 Generic License. Original photo: `http://www.flickr.com/photos/kasimetcalfe/118471837/`

[b]Photo courtesy of the SYMBRIAN project: `http://symbrion.eu`

American Institute of Aeronautics and Astronautics

(a) A swarm of ants form a bridge to cross a gap onto another leaf. Photo by Kasi Metcalf[a].

(b) A swarm of robots escapes its holding pen. The swarm can collectively accomplish goals that the individual robots could not[b].

Figure 1: Biological and robotic swarms

a task" – has been shown to be an important factor in determining the number of robots an operator can control.[1] In this paper, we evaluate the scalability of human-swarm interaction (HSI) in terms of operator workload and the control method used. Specifically, we investigate the ability of HSI to (1) overcome fan-out limitations of traditional supervisory control and (2) manage a wide range of team sizes without altering the operator's control strategy. In order to evaluate our hypothesis, we conduct a user study which we back with validation on physical robots.

II.   Related Literature

A wide variety of methods for interacting with large robot teams have been explored in the literature. Miller et al. advocate a playbook-style approach to human-robot teaming,[8] wherein a human calls plays that trigger predictable patterns of behavior. Simple plays, like grouping and searching, have been used to manage several large, simulated teams (50-200 robots).[9] More complicated plays, like coordinated rendezvous or formation-following, have been applied to smaller teams.[10] Bashyal and Venayagamoorthy provided users with a virtual avatar and simple behavior primitives for influencing swarm behavior.[11] Barnes et al.[12] used potential fields to maintain a convoy perimeter in response to operator input. Kira and Potter[13] discuss approaches based on controlling global parameters or a subset of agents. Our work builds on prior work by illustrating the importance of swarm design choices and the impact of three control methods on swarm performance and operator workload.

Several studies have evaluated the effectiveness of HSI in the context of real-world problems. The GUARDIANs project[14] demonstrated proximate swarming with robots specially designed to assist human firefighters. Bashyal and Venayagamoorthy conducted a user study to evaluate a swarming algorithm in an indoor search problem. Haas et al.[15,16] evaluated a swarm interface and control method using a simulated convoy protection task. Goodrich et al.[17,18] conducted several "Oz of Wizard"[19] studies to evaluate the influenced sustained by three swarm control methods. Our work demonstrates that HSI is not only feasible, but also scalable in terms of operator workload.

While this paper primarily focuses on interaction with robotic systems, the HSI literature includes research from a variety of communities including biology, physics, control theory, and psychology and human factors (e.g. crowd behaviors). Of special note is Couzin and associates' model of schooling fish,[20] which we use for UAV flocking, and the extensions provided by Conradt et al.,[21] which we use for control. Recent work by Marras and Porfiri[22] evaluates the use of robots to guide schools of fish. Work by Sumpter[23] provides principles for self-organization found in nature, which Goodrich et al.[17] use for modeling human-robot teams. Other work, such as that by Olfati-Saber,[24] has contributed algorithms and control theoretic guarantees. Our work demonstrates that existing abstract algorithms can be implemented on physical robots and used to effectively control a variety of swarm sizes.

American Institute of Aeronautics and Astronautics

III.   Modeling

In this section we present the models which we use in our user study and the parameters we selected. We use Couzin's model of schooling fish to represent a UAV flock under the control of a human operator. We then calibrate the model with parameters obtained from low-cost UAVs.

A.   Couzin's Model

The simulated agents use an existing swarm model published by Couzin et al[20] for dynamics and control. This model uses dynamics similar to Dubins airplane[25] where each agent i is defined by a position $[x_i, y_i]^T \in \mathcal{R}^2$ and a heading $\phi_i \in [-\pi, \pi]$ relative to the global x axis (east) and moves forward with a constant speed s and maximum turning rate ω. While Couzin's model allows for 3-dimensional movement, we consider only the 2-dimensional plane. This is done to simplify the display and control problem for study participants and provides a reasonable model for UAV area coverage tasks.

Flocking in Couzin's model is accomplished using a simple switching controller that maintains separation while staying close to the rest of the group. This is accomplished using three concentric zones of influence centered on the agent (Figure 2). The innermost zone (repulsion) represents a minimum separation or safety zone that the agent maintains between itself and other agents. The middle zone (orientation) represents alignment with the heading of nearby agents. The outermost zone (attraction) represents group cohesion and directs the agent back toward the group. Each agent also has a conical blind spot at angle α behind the agent. The controller sums the relative position vectors of the agents in each zone and prioritizes the zones as follows: If any agents are present in the repulsion zone, the agent turns to increase separation. If no agents are present in this zone, the agent averages the alignment and cohesion vectors and turns toward that heading. A full description is contained in Couzin et al.'s paper.[20]
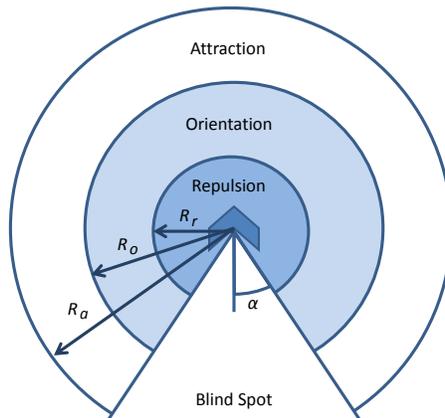

Figure 2: Couzin's model of schooling fish. Each simulated agent has three concentric zones of interaction and interacts with all other agents within these zones, provided they are not in a blind spot behind the agent. The agent is repelled by agents in the repulsion zone, attracted to those in the attraction zone, and tries to align itself with those in the orientation zone. Each agent moves forward with a constant velocity s and turns toward its desired heading at a turning rate ω.

B.   Adding Human Influence

Couzin's model represents a biological system and thus does not provide control inputs for the human operator. However, follow-on work by Conradt et al.[21] adds an additional term that directs the agents toward a goal location by weighting the existing model with a second vector that points toward the goal.

American Institute of Aeronautics and Astronautics

Parameter	Description	Units
s	Forward velocity	Units per second
ω	Turning rate	Degrees per second
α	Blind spot angle	Degrees
R_r	Zone of repulsion radius	Units
R_o	Zone of orientation radius	Units
R_a	Zone of attraction radius	Units
σ_g^2	Heading noise variance	None

Table 1: Summary of the parameters used by Couzin's model.

We use this method for adding in the three methods of human influence. Additionally, we add a bounding box term using the same method to keep the agents within the view of the operator.

Many existing approaches for swarm control are tied to specific algorithms or robot capabilities, while others, such as potential fields and machine learning, are general, but require that the desired behavior be known a priori. In this paper, we will instead focus on several simple control primitives that can be applied to a wide range of swarm algorithms. This approach provides foundational results that can be built on by future work and broadens the applicability of our results.

The three control primitives we investigate are: lead by attraction, lead by repulsion, and group influence. We refer to them as "Leader", "Predator", and "Stakeholders" respectively.

Leaders are physical or virtual agents that attract all other swarm members within a radius of influence towards themselves. Using leaders, human operators only need to manage one swarm member for any swarm size. However, this approach requires swarm members to recognize leaders as distinct from other swarm members and may introduce a single point of failure without careful implementation.

Predators are similar to leaders, but repel swarm members within their radius of influence rather than attracting them. Leaders and predators have similar advantages and disadvantages. Prior work by the authors[18] has shown that predators are able to split the swarm into subgroups more easily, but do not sustain influence as well as leaders.

Stakeholders are regular members of the swarm that are directly influenced by the human operator. Stakeholders are not recognized as different by other members of the group, meaning that stakeholders can be anonymous and need not be the same agents at each time step. Additionally, work by Kerman, Brown, and Goodrich[26] shows that stakeholders can be used to to change the collective behavior of the group. However, stakeholders must comprise a substantial fraction of the group (generally 20-40%).[26]

Control is added by weighting the each agent's desired heading vector d_i produced by Couzin's model with two additional vectors that represent (1) the goal location for stakeholders u_i^g, centered on the participant's cursor, and (2) the direction u_i^p toward or away from the leader or predator if it is in range. These additional

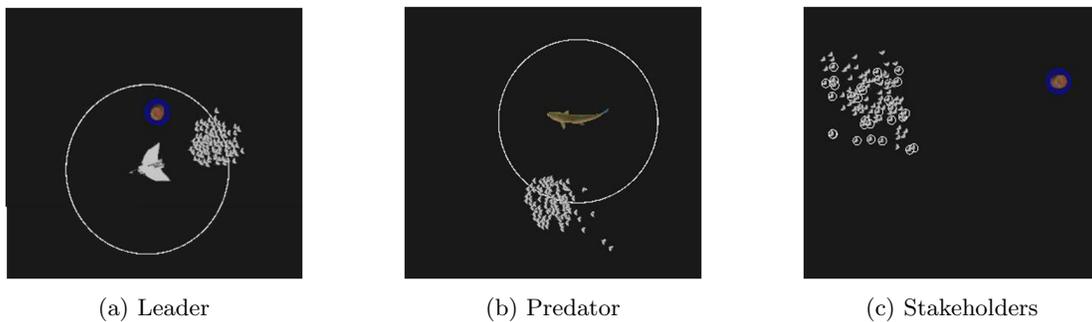(a) Leader   (b) Predator   (c) Stakeholders

Figure 3: shows the three control methods used in this thesis, as shown in the graphical user interface. Leaders (a) are operator-controlled agents that cause other swarm members to turn towards them. Predators (b) are similar to leaders, but cause other swarm members to turn away from them. Stakeholders (c) are members of the swarm that are influenced by the human operator. They interact normally with other swarm members.

influence vectors are given weights, w_g and w_p respectively, that balance flocking behaviors with human influence. Each agent then turns toward the weighted desired heading $\overline{d_i}$ with a fixed turning rate ω.

$$\overline{d_i} = d_i + w_g u_i^g + w_p u_i^p \tag{1}$$

C.   Model Parameters

The parameters for the model were chosen to match low-cost UAVs, including forward speed, turning rate, minimum recommended separation, object detection range using onboard cameras, and GPS localization error. UAVs match the dynamics of Couzin's model and provide additional experimental validity. Table 2 contains the parameter values for the augmented version of Couzin's model found in Section III. In the simulation, 1 unit is equal to 4 meters and the parameters from the UAV flight tests were scaled accordingly.

Parameter	Description	Value	SI
s	Forward velocity	4.0 units/s	16.0 m/s
θ	Turning rate	45.0 degrees/s	45.0 degrees/s
α	Blind spot angle	45.0 degrees	45.0 degrees
R_r	Zone of repulsion radius	1.5 units	6 meters
R_o	Zone of orientation radius	11.0 units	44 meters
R_a	Zone of attraction radius	14.0 units	56 meters
σ_{GPS}	GPS error stdev	0.25 units	1 meter
w_g	Stakeholder control gain	20.0	-
w_p	Leader/predator control gain	25.0	-
R_p	Leader/predator interaction radius	25.0 units	100 meters

Table 2: Simulation parameters for the model UAVs.

IV.   Validation

In this section, we discuss the experiments we performed to evaluate the feasibility of Couzin's swarm model and human-swarm interaction on two Willow Garage Turtlebots. Our goal is to provide evidence that the HSI evaluated in simulation later in this paper is applicable to real robots.

A.   Sensing and Human Interaction

We validated our model using two TurtleBots running Robot Operating System (ROS). We used the standard ROS navigation stack, which uses gmapping for SLAM, AMCL for localization against a known map, and an extended Kalman filter for sensor fusion. Communication was done using a wireless network.

The experiment consisted of a simulated force protection scenario in which the two TurtleBots escorted the operator as he moved down the hallway (see Figure 4). Turtlebot A was responsible for following the operator at a fixed distance and localizing the operator relative to itself using the Kinect sensor. Turtlebot B was responsible for patrolling in front of the operator as he moved down the hall. The human operator was equipped with a Wiimote to relay commands to the swarm. The operator's Wiimote could be used to start or stop the robots, or switch the behavior of the lead robot from patrolling to leading the operator by a fixed distance.

We successfully implemented the algorithm and demonstrated it on the second floor hallway in the Computer Science wing of the TMCB at BYU. This hallway was approximately 2.5m in width with small amount of clutter in the form of doorways and tables temporarily placed in the hallway. The human operator was able to advance down the hallway while the two robots maintained correct positions. Turtlebot A was able to keep itself and the operator localized relative to the map, and relay the information to Turtlebot B, which patrolled in front of the operator as he moved. The robots successfully responded to commands from the Wiimote during the experiment. We then repeated the experiment with a remote operator by relaying position commands to the robots rather than using Turtlebot A to provide operator localization.

Figure 4: Two TurtleBots escort a human operator in a simulated force protection scenario. Robot A is responsible for tracking and localizing the human operator, while robot B is responsible for patrolling ahead of the operator. The operator is equipped with a Wiimote used to issue commands to the robots.

B.   Couzin's Model

The first experiment showed that two TurtleBots using onboard sensing and communications can be managed by a human in both proximate and remote interactions. Building from this experiment, we demonstrated that it is possible to use Couzin's model on the TurtleBots. While the limited number of available robots (two TurtleBots) did not allow a physical demonstration of HSI, we were able to demonstrate that the controller used by agents in Couzin's model can be used on real robots with real sensors and communication. We were then able to use the data and parameters obtained from these experiments to verify group navigation and flocking behaviors using the Stage robot simulator.

To allow obstacle avoidance in a cluttered environment, we define an obstacle avoidance radius, R_{obs}, and compute the sum of the direction vectors for all objects in that range. This obstacle avoidance heading is then averaged with the desired heading from the model. Obstacles are defined as occupied cells in a dynamic local occupancy grid around the TurtleBot, which it creates using data from its RGB-D camera.

We conducted several runs in which we placed one of the TurtleBots in a hallway, initialized its approximate location, and allowed it to wander through the hallway. Note that in the absence of goals or additional robots, an agent being controlled by Couzin's model follows a random walk. We found that the model was sufficient to allow the TurtleBot to move through the hallway in a random walk and avoid obstacles.

V.   Methodology

This section describes the design and implementation of the user study presented in this work, along with the metrics we used to evaluate our hypothesis.

A.   Experiment Design

The objective of this user study was to measure the impact of (1) interaction style and (2) scale on the mental workload of the human operator while controlling a swarm. We use three simple interactions styles – leader, predator, and stakeholders – as outlined in the Section III. For scale, we use 20, 50, and 100 simulated robots, which represents the range of swarm sizes commonly found in the literature. The experiment format is a 3 x 3 within-subject experiment where each participant performs all nine experimental conditions. Performance is compared between conditions for each participant and averaged across participants.

B.   Study Scenario

The study scenario uses a canonical information foraging task (Figure 5) where participants guide simulated robots around the environment to gather information at specific points of interest (targets) graphically represented by barrels. The simulated robots do not seek targets on their own, but depend on the human operator to guide them to the targets. In each scenario, the study participant controls a leader, a predator,

American Institute of Aeronautics and Astronautics

or several stakeholders. The agents under the participant's control move toward the participant's cursor as it is moved around the screen with the mouse. The participant's task is to amass as much information as possible by guiding the swarm using agents under his or her control.

Initially, two points of interest are present, with new points of interest appearing randomly every 1 to 11 seconds during the 120 second scenario. Each point of interest contains an initial quantity of information drawn from a uniform distribution between $[10.0, 70.0]$. When a robot is within 6.0 units (24 meters) of a target it gathers information at a rate of one unit per second. Consequently, information gathering is proportional to the number of robots within range of a target. This represents a variety of tasks that benefit from multiple robots, such as surveillance or search and rescue. When all information has been gathered from the target, it disappears.
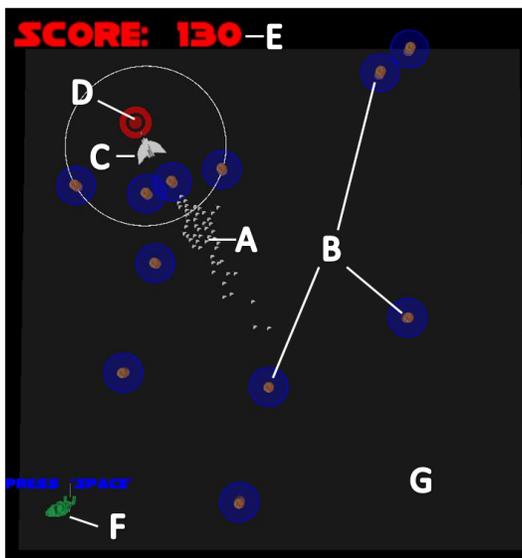

Figure 5: The information foraging task for the user study. The study participant is responsible for guiding a group of simulated robots (a) to areas of interest represented graphically by barrels (b), using leader, predator or stakeholder control (c). The leader, predator or stakeholders move towards the participant's cursor (d). The visual secondary task (f) is represented by a ninja that appears randomly on the playing field (g).

C.   Metrics

Multiple metrics are used for monitoring swarm performance and the participant's mental workload, as described in this section.

1.   Performance Metrics

The metrics in this category measure the performance of the swarm in each condition to measure the impact of scale and control style. The participant's score, total information gathered, and targets depleted are monitored. The participant's score for the scenario is shown in the upper left hand corner of the graphical user interface and primarily reflects the amount of information gathered, with small bonuses and penalties for secondary task performance. The score starts at zero when the scenario begins and points are added or subtracted as the scenario progresses according to the following equation:

$$score = v_{recognized} - 5 * v_{missed} + I_g + T_{depleted} \qquad (2)$$

where $v_{recognized}$ and v_{missed} are detections and missed detections on the visual secondary task, I_g is the total information gathered, and $T_{depleted}$ is the number of targets depleted.

Metric	Description	Units
Information gathered	Total information gathered during the scenario	agent-seconds
Targets depleted	Number of targets completely depleted	(Count)
Score	The participant's score for the scenario	(None)

Table 3: Performance metrics

D. Mental workload

Mental workload is monitored using secondary tasks, a post-scenario NASA-TLX questionnaire, and a post-scenario survey (Table 6). Two secondary tasks are used to monitor the participant's workload. The first is a visual response-time task that requires the participant to press a key whenever a small ninja is displayed on the screen. The second is an auditory counting and memory task that requires the participant to count the occurrence of one of two sounds. Participants report this number at the end of the scenario and rate their confidence that it is correct. Tables 4 and 5 list the metrics associated with these two tasks.

Metric	Description	Units
True Positives	The number of cow sounds played during the scenario	(count)
Guess	The participant's estimate of the number of cow sounds	(count)
Guess Ratio	Ratio of the participant's estimate to the true value	(none)
Percent Error	Percent error of the participant's estimate	percent
Confidence	The participant's confidence in his or her estimate	(survey)

Table 4: Auditory secondary task metrics

Metric	Description	Units
False Positives	The number of false positives on the visual task	(count)
True Positives	The number of correct responses on the visual task	(count)
Response Time (no MD)	Visual task response time (excluding missed detections)	seconds
Response Time	Visual task response time (including missed detections)	seconds

Table 5: Visual secondary task metrics

Question	Response Type
"I was able to control the agents easily"	"strongly agree" to "strongly disagree" (7 values)
"The agents went where I wanted them to"	"strongly agree" to "strongly disagree" (7 values)
"This scenario was frustrating"	"strongly agree" to "strongly disagree" (7 values)
"This scenario was difficult"	"strongly agree" to "strongly disagree" (7 values)

Table 6: Post-scenario survey

VI.   Results

This section presents the results of our user study, the data we gathered, and a discussion of our findings.

A.   Data Gathered

We recruited thirty-two participants from the campus population and conducted the user study in accordance with IRB protocol. The study participants consisted of 20 males and 12 females. The median age was 23 with an age range of 18 to 53. Of the thirty-two datasets we gathered, we excluded five due to study side

effects, computer issues, or deviation from the provided instructions. The remaining twenty-seven datasets are analyzed and presented in this section.

There were minimal side effects reported during the user study. One participant reported watery eyes from concentrating on the computer screen and discontinued the experiment in accordance with the IRB protocol. The symptoms were minor and abated once the experiment was discontinued. No other side effects were reported.

B. Statistical Analysis

The statistical analysis was conducted using mixed models analysis with Tukey-Cramer adjustment. We analyzed both fixed effects and pairwise within-subject differences across control style and scale. We also conducted an extensive analysis of confounding factors, demographics, and learning effects and failed to find any evidence that the study results were confounded by any of these factors.

All tables in this section show the Tukey-Kramer adjusted p-value of the fixed effect test. The first two columns of each table show the independent variables (scale and control style) the next four columns show demographic factors, and the last column shows learning effect. In this paper, we use a significance value of $p < .01$, while making note of effects where $p < .05$, which we refer to as weakly significant.

C. Performance

We found that both scale and control style significantly impacted swarm performance (Table 7). Larger groups sizes increased swarm performance, but provided diminishing returns as the size of the swarm increased (Figure 6). We observed a large statistically significant performance increase when increasing the swarm size from 20 to 50 agents. Further increasing the swarm size to 100 agents showed a much smaller performance increase that was not statistically significant. As with many real-world systems, the information foraging task showed performance gains as more agents were added, but diminishing returns as the number of agents approached a saturation point. These results matched our study hypothesis and show that properly design swarm systems can allow performance to scale without altering the control task of the human operator.

Swarm performance was also significantly influenced by control style. We found that leader-based control performed best on all performance measures, followed by stakeholder-based control, and then predator-based control (Figure 6). Leader-based and stakeholder-based control performed similarly and differences between them were not statistically significant. These results show that design choices, such as control method, can influence the performance of the swarm and should be taken into consideration in the design of swarm systems.

(a) Information gathered vs scale
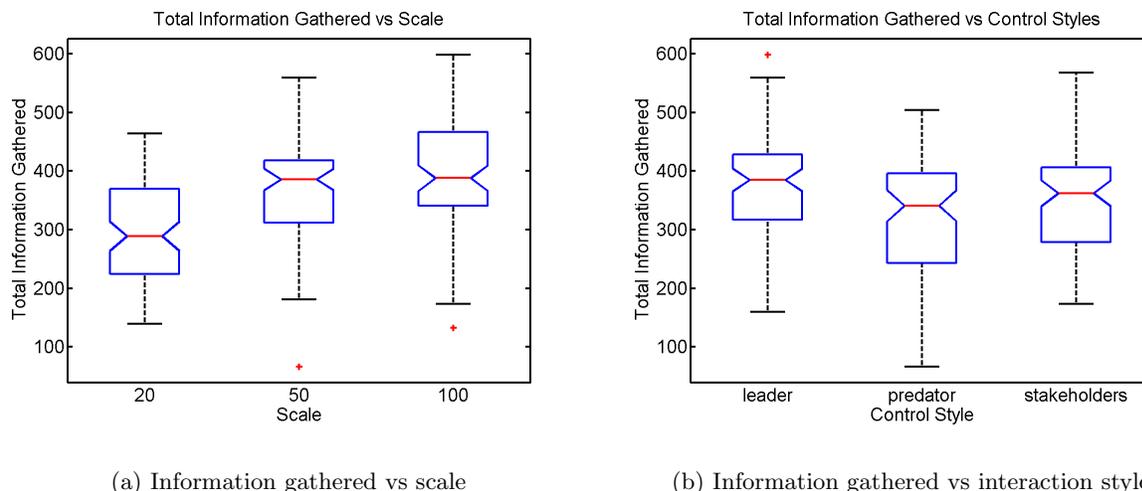
(b) Information gathered vs interaction style

Figure 6: The total information gathered as a function of (a) scale and (b) control style. The performance of the swarm was affected by both number of robots and control style ($p < .001$). Performance increased with swarm size up to a saturation point. Control style also significantly affected swarm performance.

American Institute of Aeronautics and Astronautics

Metric	Scale	Style	Sex	Vision	Robot Exp	Games Exp	Sim Order
Information Gathered	$< .0001$	**0.0002**	0.6078	**0.0085**	**0.0356**	**0.0212**	0.4298
Barrels Depleted	$< .0001$	$< .0001$	0.4840	0.0583	0.0937	0.0829	0.6665
Score	$< .0001$	$< .0001$	0.5140	0.1306	0.1923	0.0944	0.5211

Table 7: $Pr > F$ for fixed effect tests on the performance metrics used in the user study. We found that scale and control style both significantly affected performance ($p < .0001$) across all of our metrics. We also found that corrected-to-normal vision, robot experience, and video games experience were correlated with higher performance. See Table 3 for a description of the entries in this table. Bold text indicates statistical significance.

D.   Workload

The primary goal of the large-scale user study was to evaluate the effect of control style and scale on the workload of the human operator while controlling a swarm. We evaluated the participant's mental workload using a variety of objective and subjective methods (Figure 7), including multiple metrics on two secondary tasks (Table 10), NASA-TLX (Table 8), post-scenario surveys (Table 9), and post-experiment preference surveys. There was strong agreement between methods and the results showed that operator workload was affected by control style, but not scale. The control style results were strongly significant ($p < .001$) across all measures using both fixed effect and pairwise tests. Leader-based and stakeholder-based control performed similarly, and both outperformed predator-based control ($p < 0.005$) with a significant effect size (10%).

Scale was only significant on a single pairwise test[c], and most measures were far from significant ($p > .10$ or more). Furthermore, the effect size across scale on all measures was small and at most of limited practical significance had there been significance. The post-experiment survey results do not provide additional insight and are omitted.

Metric	Scale	Style	Sex	Vision	Robot Exp	Games Exp	Sim Order
Temporal Demand	0.4138	**0.0011**	0.4656	0.8135	0.5275	0.9372	0.5442
Effort	0.4964	**0.0164**	0.9819	0.5131	0.5682	0.4075	0.3938
Physical Demand	0.9443	**0.0092**	0.0719	0.1316	0.1775	0.8240	0.7123
Frustration	0.0937	$< .0001$	0.7191	0.7918	0.6520	0.3820	0.8882
Performance	0.1058	**0.0010**	0.2669	0.8341	**0.0288**	0.7014	0.1960
Mental Demand	0.7601	$< .0001$	0.8736	0.5972	0.4711	0.3382	0.3749
Overall	0.1133	$< .0001$	0.0714	0.9030	0.4216	0.0339	0.3105

Table 8: $Pr > F$ for fixed effects tests on the NASA-TLX survey. Significant differences were found across control style for the overall TLX rating and all subscales. No significant effects were observed across scale.

Metric	Scale	Style	Sex	Vision	Robot Exp	Games Exp	Sim Order
Controllability	0.1550	$< .0001$	0.7687	0.1592	0.1618	0.6439	0.0900
Frustration	**0.0075**	$< .0001$	0.7051	0.1726	0.8571	0.4196	0.2540
Difficulty	**0.0126**	$< .0001$	0.5892	0.2271	0.3215	0.5475	0.0766
Ease of control	0.2153	$< .0001$	0.4800	0.2216	0.4069	0.2523	0.7378

Table 9: $Pr > F$ for fixed effects tests on the post-scenario surveys. Bold text indicates statistical significance.

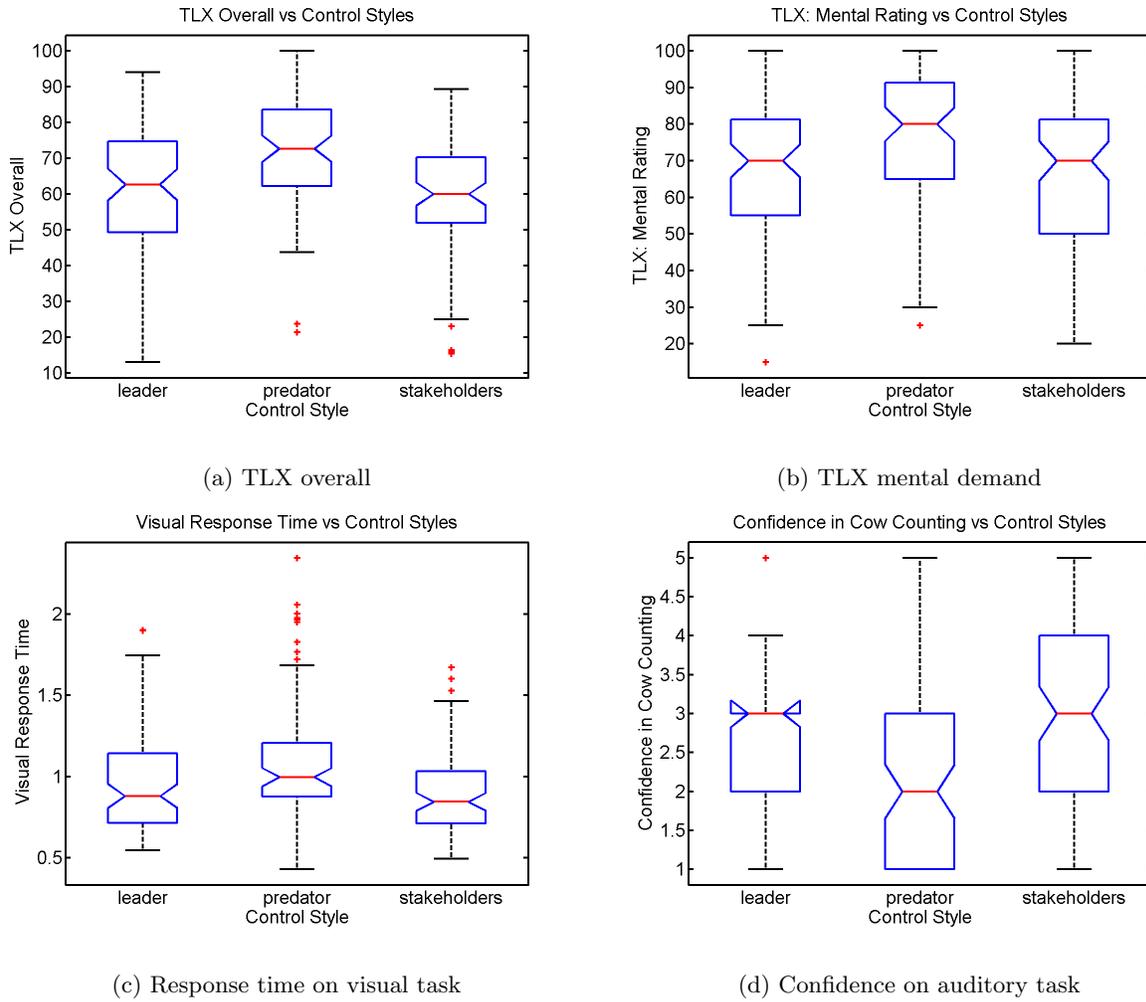[c]There was a difference of one half survey tick between 20 and 50 agents on the frustration component of NASA-TLX

(a) TLX overall

(b) TLX mental demand

(c) Response time on visual task

(d) Confidence on auditory task

Figure 7: Mental workload across control styles

Metric	Scale	Style	Sex	Vision	Robot	Games	Order
Audio Guess	0.2783	0.5195	0.4895	0.0924	0.3137	0.4057	0.4213
Audio Guess Error	0.9836	0.5891	**0.0118**	0.5916	**0.0279**	0.6197	0.0971
Audio Guess Error (absolute)	0.3008	**0.0336**	0.2572	0.8301	0.7109	0.7730	**0.0347**
Audio Guess Ratio	0.9464	0.5673	**0.0113**	0.2999	**0.0094**	0.6636	0.0547
Audio Percent Error	0.1784	**0.0331**	0.3685	0.6043	0.5806	0.6325	**0.0355**
Audio Confidence	0.4684	**0.0009**	0.6270	0.6170	0.3272	0.5046	0.0544
Visual False Positives	0.3192	0.5173	0.1158	0.5382	0.4833	0.0971	0.1538
Visual True Positives	0.4475	**< .0001**	0.6614	0.6197	0.7104	0.3518	**0.0029**
Visual Missed Detections	0.2338	**0.0001**	0.6664	0.6371	0.5867	0.3581	**0.0022**
Visual Response Time (no MD)	0.0821	**< .0001**	0.7379	0.6105	0.6596	0.5936	0.0739
Visual Response Time	0.1778	**< .0001**	0.9092	0.5976	0.5745	0.4525	0.0515

Table 10: $Pr > F$ for fixed effects tests on objective workload metrics. Control style significant affected control activity, performance on the visual secondary task, and confidence on the auditory secondary task. Scale was not statistically significant on any measure. See Tables 5 and 4 for a description of the entries in this table. Bold text indicates statistical significance.

VII. Discussion

Effective large-scale human-robot teams depend not only on scalable robotic algorithms, but also scalable *interaction* between humans and robots. Simply stated, the availability of large numbers of robots is only useful if the additional robots can be effectively leveraged to perform meaningful tasks. In our research, we found that swarms not only scaled algorithmically, but also in terms of performance and operator workload. The algorithms and control methods we evaluated in our study allowed study participants to use the same control method across a broad range of team sizes at relatively constant mental workload. The results also show that, even with the same control methods and constant workload, participants were able to use the additional robots in larger swarms to increase team performance. These results demonstrate that HSI can provide an effective and scalable approach for managing large human-robot teams.

In our study, we found that all three control methods – leader, predator, and stakeholders – were able to scale effectively and maintained relatively constant mental workload over a variety of swarm sizes. However, the amount of workload and the performance of the swarm showed large statistically significant differences across control styles. This means that, while HSI allows scalable interaction, the choices made in the design of the swarm significantly impact the performance of the human-swarm system. This illustrates the need to identify and quantify the basic building blocks of swarm algorithms in order to build reliable and scalable swarm systems.

VIII. Conclusion

In this paper we evaluated the scalability of swarm systems in terms of operator workload and the ability of HSI to overcome the fan-out limitations of traditional supervisory control. The results of our user study show that larger swarm sizes allow increased swarm performance without increasing operator workload, overcoming the fan-out limitations of traditional supervisory control. We found that all three control methods allowed scalable HSI, but that the control method selected had a substantial impact on both swarm performance and operator workload. Additionally, we successfully validated our algorithms on physical robots, providing ecological validity to our study. The results presented in this paper demonstrate the utility of swarm systems for controlling large numbers of robots. Our results also illustrate the impact design choices can have on the human-swarm system, illustrating the need for continued research.

Acknowledgments

References

[1] Olsen, D. and Wood, S., "Fan-out: Measuring human control of multiple robots," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2004, pp. 231–238.

[2] Şahin, E., "Swarm robotics: From sources of inspiration to domains of application," *Swarm Robotics*, edited by E. Şahin and W. Spears, Springer, New York, New York, 2005, pp. 10–20.

[3] Couzin, I. and Krause, J., "Self-organization and collective behavior in vertebrates," *Advances in the Study of Behavior*, Vol. 32, 2003, pp. 1–75.

[4] Brafman, O. and Beckstrom, R., *The starfish and the spider: The unstoppable power of leaderless organizations*, Penguin Group, New York, New York, 2006.

[5] Sumpter, D., *Collective animal behavior*, Princeton University Press, Princeton, New Jersey, 2010.

[6] McLurkin, J., *Stupid robot tricks: A behavior-based distributed algorithm library for programming swarms of robots*, Master's thesis, Massachusetts Institute of Technology, 2004, `http://people.csail.mit.edu/jamesm/publications/ McLurkin-SM-MIT-2004(72dpi).pdf`.

[7] Sanders, M. and McCormick, E., *Human factors in engineering and design*, Mcgraw-Hill Book Company, San Francisco, California, 7th ed., 1993.

[8] Miller, C., Funk, H., Dorneich, M., and Whitlow, S., "A playbook interface for mixed initiative control of multiple unmanned vehicle teams," *Proceedings of the 21st Digital Avionics Systems Conference*, Vol. 2, IEEE, 2002, pp. 7E4–1.

[9] Hardin, B. and Goodrich, M., "On using mixed-initiative control: A perspective for managing large-scale robotic teams," *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction*, ACM, 2009, pp. 165–172.

[10]McLain, T. and Beard, R., "Cooperative rendezvous of multiple unmanned air vehicles," *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, 2000.

[11]Bashyal, S. and Venayagamoorthy, G., "Human swarm interaction for radiation source search and localization," *IEEE Swarm Intelligence Symposium*, IEEE, 2008, pp. 1–8.

[12]Barnes, L., Fields, M., and Valavanis, K., "Swarm formation control utilizing elliptical surfaces and limiting functions," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, Vol. 39, No. 6, 2009, pp. 1434–1445.

[13]Kira, Z. and Potter, M., "Exerting human control over decentralized robot swarms," *4th International Conference on Autonomous Robots and Agents*, 2010, pp. 566–571.

[14]Penders, J., "GUARDIANS final report," Tech. rep., Sheffield Hallam University, 2010, `http://shura.shu.ac.uk/2340/1/FinalReportMarch.pdf`.

[15]Haas, E., "Extreme scalability: Designing interfaces and algorithms for soldier-robotic swarm interaction," Tech. rep., Army Research Lab Aberdeen Proving Ground MD, 2009, `http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA523930`.

[16]Haas, E., "Extreme scalability: Designing interfaces and algorithms for soldier-robotic swarm interaction, year 2," Tech. rep., Army Research Lab Aberdeen Proving Ground MD, 2010, `http://www.dtic.mil/dtic/tr/fulltext/u2/a498162.pdf`.

[17]Goodrich, M., Pendleton, B., Sujit, P., and Pinto, J., "Toward human interaction with bio-inspired robot teams," *IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, 2011, pp. 2859–2864.

[18]Goodrich, M., Pendleton, B., Kerman, S., and Sujit, P., "What types of interactions do bio-inspired robot swarms and flocks afford a human," *Proceedings of Robotics: Science and Systems*, 2012.

[19]Steinfeld, A., Jenkins, O., and Scassellati, B., "The Oz of Wizard: Simulating the human for interaction research," *4th ACM/IEEE International Conference on Human-Robot Interaction*, IEEE, 2009, pp. 101–107.

[20]Couzin, I., Krause, J., James, R., Ruxton, G., and Franks, N., "Collective memory and spatial sorting in animal groups," *Journal of Theoretical Biology*, Vol. 218, No. 1, 2002, pp. 1–11.

[21]Conradt, L., Krause, J., Couzin, I., and Roper, T., ""Leading according to need" in self-organizing groups," *The American Naturalist*, Vol. 173, No. 3, 2009, pp. 304–312.

[22]Marras, S. and Porfiri, M., "Fish and robots swimming together: Attraction towards the robot demands biomimetic locomotion," *Journal of The Royal Society Interface*, Vol. 9, No. 73, 2012, pp. 1856–1868.

[23]Sumpter, D., "The principles of collective animal behaviour," *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 361, No. 1465, 2006, pp. 5.

[24]Olfati-Saber, R., "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Transactions on Automatic Control*, Vol. 51, No. 3, 2006, pp. 401–420.

[25]Chitsaz, H. and LaValle, S., "Time-optimal paths for a Dubins airplane," *46th IEEE Conference on Decision and Control*, IEEE, 2007, pp. 2379–2384.

[26]Kerman, S., Brown, D., and Goodrich, M., "Supporting human interaction with robust robot swarms," *5th IEEE International Symposium on Resilient Control Systems*, IEEE, 2012, pp. 197–202.

[27]Pendleton, B., *Human-swarm interaction: Effects on operator workload, scale, and swarm topology*, Master's thesis, Brigham Young University, 2013, Under review for publication.