Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 46 (2017) 3–12

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Team situation awareness within the context of human-autonomy teaming

Mustafa Demir *, Nathan J. McNeese, Nancy J. Cooke

Human Systems Engineering, Arizona State University, Mesa, AZ, USA

Received 12 September 2016; accepted 18 November 2016
Available online 23 November 2016

Abstract

Effective team communication, a fundamental part of team coordination, is crucial for both effective Team Situation Awareness (TSA) and team performance. In this study, we looked at the role that team interaction (i.e., more specifically team verbal behaviors) played in TSA and team performance in order to better understand Human-Autonomy Teaming (HAT). We first analyzed team verbal behaviors (i.e., pushing and pulling information) across conditions of human-autonomy teams and human-human teams, and then analyzed their relationship with TSA and team performance via Growth Curve Modelling (GCM). Good teamwork involves anticipating the needs of teammates and that means pushing information before it is requested. Therefore, if things are going well, there should be little need for pulling information. In this study's task, participants were instructed to push information to others, and over time master the specific timing of information sharing to the intended recipient. Findings indicate that pushing information was positively associated with TSA and team performance, and human-autonomy teams had lower levels of both pushing and pulling information than all-human teams. Through this study, we have learned that anticipation of other team member behaviors and information requirements in human-autonomy teams are important for effective TSA and team performance. In order to make HAT more effective in terms of teamwork, we need to develop mechanisms to enhance pushing information within HAT.

Keywords: Human-autonomy teaming; Team situation awareness; Teamwork; Team cognition; Synthetic agent; Growth curve modelling

1. Introduction

Human team members increasingly work in highly dynamic environments involving interaction with highly automated systems, such as synthetic agents and robots. Although, these highly automated systems provide many benefits to Human-Autonomy Teaming (HAT), they can also place high cognitive demands on human teammates (because of the overall systems complexity and the agents' lack of human-like behavior). Therefore, there are costs and benefits when human team members interact with highly automated systems as a team member or as supervisory control.

Autonomy is a system independent of human control (i.e., it makes its own decisions about its actions during the task), whereas automation is a system that will do exactly what it is programmed without independent action (Vagia, Transeth, & Fjerdingen, 2016). Autonomy can be defined as any system which can, through some intelligence-based capabilities, respond to situations that are outside the range of those considered when the system was designed; to some degree, such systems can direct their own behavior with the human's proxy for decisions (Endsley, 2015). Therefore, autonomy can be considered

* Corresponding author at: Human Systems Engineering, Arizona State University, 150A Santa Catalina, Mesa, AZ 85212, USA.
E-mail address: mdemir@asu.edu (M. Demir).

http://dx.doi.org/10.1016/j.cogsys.2016.11.003
1389-0417/

an extension of automation; the former being a well-designed, highly capable version of the latter (Endsley, 2015). There are several characteristics that make an agent more autonomous, including: independence of goal achievement and better performance under dynamical environments via communication or non-communication (even being able to adjust for system failures); all of this done without outside intervention (Endsley, 2015; Krogmann, 1999; Schooley, Zeigler, Cellier, & Wang, 1993).

From the perspective of autonomy (and not automation), HAT is a team in which human team members work with autonomy with varying levels of intelligence (Schulte, Donath, & Lange, 2016). In HAT, developing a highly autonomous agent as a team member is quite challenging, especially in terms of the agent's interaction with human team members and the current dynamic environment (Klein, Woods, Bradshaw, Hoffman, & Feltovich, 2004).

1.1. Team situation awareness and team interaction for human autonomy teams

In general, situation awareness (SA) is "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" (Endsley, 1988, p. 97). Within the team perspective, SA can be defined as "the degree to which every team member possesses the situation awareness required for his or her responsibilities" (Endsley & Robertson, 2000, p. 303; see also Endsley, 1989).

At any point in time, individuals on the team need to understand the current situation and adapt their understanding to the dynamic situation. As such, in a dynamic situation, teams need to understand the characteristics of a specific situation, such as cues and patterns, and situation-dependent characteristics (Cooke, Salas, Cannon-Bowers, & Stout, 2000). In addition, Gorman, Cooke, and Winner (2006) highlight that, especially in highly dynamic environments (e.g., Decentralized Command and Control) where the dynamics themselves can also undergo change, adaptive team cognition requires continuous coordination via effective communication among team members in order to achieve a common and valued goal. In this case, anticipation of other team member behaviors and information requirements in the teams are important for effective Team Situation Awareness (TSA) (Endsley & Robertson, 2000).

Human-autonomy teams may be able to successfully achieve their goals in a dynamic environment if they have high levels of TSA among the team members, an optimal level of trust, manageable workload levels during their mission (Endsley, 2015), and effective team interaction (i.e., team communication and coordination) (Gorman, Cooke, Winner et al., 2006). In HAT, the autonomy itself often lacks certain perceptional capabilities, limiting its ability to ascertain situation awareness. For this reason, humans may need to compensate for the autonomy's shortcomings by adapting to novel situations based on pattern recognition, mental models, analogical reasoning, and unpredictable action (especially in command-and-control environment) (Endsley, 2015). Although, in an attempt to overcome the autonomy's deficiencies, the human's own situational awareness may be compromised, leading to a lack of overall TSA.

Much like TSA, the concept of team cognition is important to the effectiveness of HAT. Team cognition is cognitive activity at the team level, often leading to a shared understanding (Cooke, Gorman, Myers, & Duran, 2013). There are several studies (Cooke & Gorman, 2009; Cooke, Gorman, & Kiekel, 2008; Cooke, Gorman, Pedersen et al., 2007; Cooke et al., 2013; Gorman, Amazeen, & Cooke, 2010; Gorman, Cooke, Pedersen et al., 2006) which consider communication and coordination being critical to team cognition. Cooke et al. (2013)'s theory of Interactive Team Cognition highlights through these previous studies, effective communication and coordination is positively related with team performance (p. 262). Therefore, carrying this theory into the context of HAT, one may hypothesize that in order for human-autonomy teams to be effective they must communicate and coordinate at a high level. Through the literature, we know that both TSA and interactive team cognition are important to traditional human-human teams. Yet, little is empirically known regarding TSA and interactive team cognition in human-machine teams, or in this case, human-autonomy teams. In this article, we seek to empirically explore: (1) how human-autonomy teams utilize their interactions to develop TSA, and (2) how these interactions are associated with team performance.

1.2. Current study and hypotheses

In this study, three heterogeneous teammates comprised an unmanned aerial vehicle (UAV) team: (1) Air Vehicle Operator (pilot) – controls the UAV's heading, altitude, and airspeed; (2) Data Exploitation, Mission Planning and Communications (navigator) – provides a dynamic flight plan with speed and altitude restrictions of the critical waypoints; and (3) Payload Operator (photographer) – obtains correct altitude and airspeed restrictions of the current target waypoint from the pilot, and takes photographs of the target by adjusting camera settings. The main goal for this team was to take good photos of target waypoints.

There were three conditions which were created by manipulating the pilot position: (1) the *synthetic condition*, the pilot was a synthetic teammate, (2) the *control condition*, the pilot was a randomly assigned human participant, and (3) the *experimenter condition*, the pilot was one of the experimenters who was an expert at the pilot's task. In the experimenter condition, the expert pilot focused on structured coordination in which the pilot asked questions to other team members to promote adaptive passing of information at target waypoints in a timely manner. The other two roles (i.e., the navigator and the pilot) were randomly assigned participants for all three conditions. Communication between these three team members occurred over a text-based communications system. In the synthetic condition, it was critical that communication between the synthetic teammate and human team members occurred in the correct way in order to achieve the team's main goal (taking a good photo of the target), because the synthetic teammate's limited language capability may cause it to take longer (or even fail) to understand the current situation. Said differently, the messages sent from human teammates to the synthetic teammate must not be ambiguous or cryptic (Demir et al., 2015).

In this research, team members coordinate with each other based on three special kinds of communication–coor dination events during the task which is depicted in Fig. 1 (Cooke, Gorman, Duran, & Taylor, 2007): first, information about the upcoming target waypoint is sent by the navigator to the pilot; then, negotiation, occurs between the photographer and the pilot regarding an appropriate altitude and airspeed for the target's required camera settings in order to take a good photograph; and lastly, the feedback about the status of the target photo is sent by the photographer to other team members. Successful communication is required for successful coordination; if communication fails, then team performance will suffer.

In this study, each team member is in charge of adequately pulling and pushing information among the team members in a timely manner. To get an idea of the mechanics of this process, a set of verbal behaviors associated with team process were measured, including: repeated requests, general status update, inquiry about status of others, planning, suggestions to others, unclear communications, negative and positive communication. Because the confederate pilot, in the experimenter condition, takes on more responsibilities regarding teamwork and team interaction, team members will have more opportunities to practice and develop team verbal behaviors (pushing and pulling information). Thus, in this study, we seek to answer the following research questions via testing the following hypotheses:

Research Question 1. Do the team verbal behaviors (i.e., pushing and pulling) significantly differ between the three conditions (the synthetic versus the control versus the experimenter)?

Hypothesis 1. Teams in experimenter condition will have more team verbal behaviors (both pushing and pulling) than the teams in the control condition, and the control teams will have more team verbal behaviors (both pushing and pulling) than the synthetic teams.

Research Question 2. How are team verbal behaviors (i.e., pushing and pulling information) related to team situation awareness (i.e., completing roadblocks) and team performance over time?

Hypothesis 2. The effect of both *pushing* and *pulling* information will be positively related to team situation awareness (i.e., completing roadblocks). That is, more pushing and pulling information is associated with better team situation awareness.

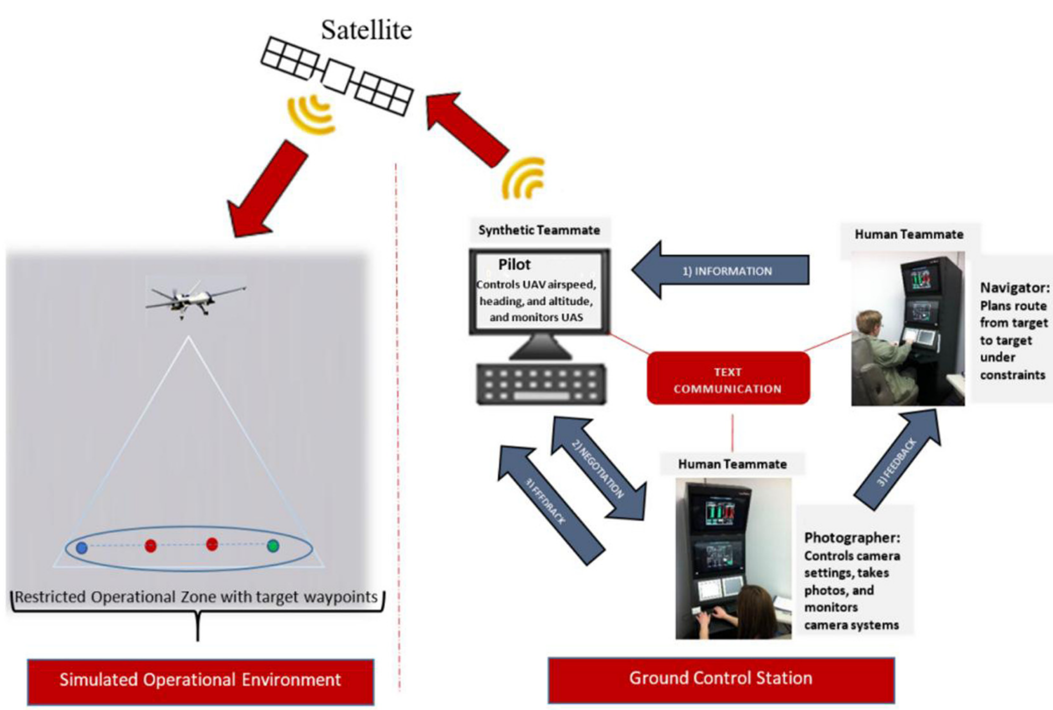

Fig. 1. Human-autonomy team (the red dashed line indicates two different environments: the simulated operational environment and the ground control station) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

Hypothesis 3. The effect of both *pushing* and *pulling* information will be positively related to team performance. That is, more pushing and pulling information is associated with better team performance.

With that in mind, in this study, we looked at the role that team interaction (i.e., more specifically team verbal behaviors) played in TSA and team performance. In the analytic procedure, we first analyzed team verbal behaviors (i.e., pushing and pulling information) across conditions, and then analyzed their relationship with team situation awareness and team performance via growth curve modelling.

2. Method

2.1. Participants

Seventy participants (ages ranged from 18 to 38 years, $M_{\mathrm{age}} = 23.7$, $SD_{\mathrm{age}} = 3.3$, across 60 males and 10 females) were recruited from Arizona State University and surrounding areas in Mesa, Arizona, each completing an eight-hour experiment as a team. Each team had a different number of participants based on the conditions: in the control condition, each team had three randomly assigned participants; in the synthetic and experimenter conditions, each team had two randomly assigned participants (one each for the photographer and navigator roles). In the synthetic condition, the pilot was always filled by a synthetic teammate, but in experimenter condition, the pilot role was filled by a trained confederate. Overall, 30 teams completed the experiment. Participants were required to have normal or corrected-to-normal vision and be fluent in English. Each participant was compensated $10 per hour for participation.

2.2. The synthetic teammate

In the current study, a synthetic teammate was used as a team member and communicated with other human teammates by using text chat during the task. The synthetic teammate, which was developed using the ACT-R cognitive modelling architecture (Anderson, 2007), consists of five components (Ball et al., 2010): (1) *language analys*is for handling a broad range of English constructions, (2) *language generation* for selecting possible utterances, (3) *dialog modelling* for management of the communication obligations, (4) *situation modelling* for task and situation awareness, and (5) *agent-environment interaction* for flying the Unmanned Air Vehicle (UAV) between the locations in a cognitively plausible manner.

2.3. Materials

The Cognitive Engineering Research on Team Tasks Laboratory, UAV - Synthetic Task Environment (CERTT UAV-STE), is used to simulate many teamwork related aspects of UAV operations (Cooke & Shope, 2004, 2005). CERTT-II, which is the updated version of the CERTT-UAV-STE software, was used for this experiment and it has two important features: (1) a text-based chat program for team members to communicate with; and (2) eight new hardware consoles: the first four of them for participants (see Fig. 1 for participants' consoles) and the second four for two experimenters to observe and manage the simulation (and insertion of roadblocks) (Cooke & Shope, 2004).

Specifically, the experimenters used two of the experimenter consoles: (1) "texting experimenter" console, which is used to give ratings for taskwork relatedness, behavior, and situational awareness. This console also communicates with the other three roles (pilot, navigator, and photographer) consoles by sending text messages (in a chat window); and (2) the "non-texting experimenter" console, which has a camera to track the participants and also allows the experimenters to rate coordination, taskwork relatedness, and behavior.

In addition to the consoles, we used PowerPoint to design a sequence of training tutorials for the participants, utilized custom software package for collecting individual data, and provided role-specific "cheat sheets" and supplementary materials to each participant about the rules of the task.

2.4. Task and roles

Three heterogeneous and interdependent team members were required to photograph critical target waypoints: (1) the navigator produces a dynamic flight plan by providing information about target location, speed and altitude restrictions to other team members; (2) the pilot controls the UAV's heading, altitude, and airspeed; and (3) the photographer takes photos of ground targets by adjusting sensor equipment.

Each participant was randomly assigned to a role before arriving at the session. The task itself consisted of five missions during which teams had to photograph 11–20 targets in 40 min or less; each mission ended after 40 min or whenever the team felt that they had achieved their mission goals, whichever came first. Instructions were to capture the greatest number of "good" photos without violating rules or setting off alarms.

2.5. Procedure

During the experiment, the navigator and the photographer were seated in one room, but separated by partitions to prevent face-face contact, and the pilot was isolated in another room. Each team completed five missions during the eight-hour experimental session. The complete experimental session and its order is show in Table 1.

2.6. Measures

In this experiment, the following measures were collected: *team performance measures* - mission and target

Table 1
Experimental sessions.

	Sessions
(I) Before Training	Consent forms
(II) Training	PowerPoint training, and hands on training
(III) Task	Mission 1, NASA TLX I/Team Knowledge I, Mission 2, Mission 3, Mission 4, Mission 5, and NASA TLX II/Team Knowledge II
(IV) After Task	Demographic questions/debriefing, and Post-checklist by the experimenters

level performance scores; *team process measures* - team verbal behaviors, team coordination, team situation awareness, and process rating; and *other measures* - taskwork and teamwork knowledge, NASA-TLX workload measure and post-experiment questionnaire. Due to the focus of this paper, we only focus on team situation awareness, mission level team performance, and team verbal behaviors.

2.6.1. Team performance score

The team performance score is a composite score which is calculated using the following mission variables: the rate of successfully photographed targets, the rate of acquiring critical waypoints, and how long each individual spends in alarm and warning states. In the beginning of each mission, teams began with a score of 1000, which is then deducted based on the final values of the mission sub-scores (Cooke, Gorman, Pedersen et al., 2007).

2.6.2. Team situation awareness

Team Situation Awareness (TSA) roadblocks were driven by *ad hoc* target waypoints that take place within the scenario at set times within each mission. During each mission, teams were sometimes presented with "roadblocks", such as the introduction of a new target waypoint. Completion of the roadblocks was used as a measure of team situation awareness. The triggering mechanism for these roadblocks is based on each team's position relative to the waypoints in the mission; as such, the number of waypoints triggered in each mission may vary from team to team.

2.6.3. Team verbal behaviors

Eight different team verbal behaviors are associated with team coordination (Demir, McNeese, & Cooke, 2016). Of these, five were considered for this study and classified into two groups: *pushing* or *pulling* of information among the team members (see Table 2) (Cooke, Gorman, Duran et al., 2007; Demir & Cooke, 2014). The remaining three behaviors (positive and negative behaviors, and unclear information) could not be categorized as pulling or pushing information, and they were not considered for this study. In this task, it is important to note that more pushing than pulling information is typically required, and thus, pulling information is not as frequent an occurrence.

3. Analytic procedure

Two analysis techniques were applied for the research questions and their hypotheses. The first technique was Multivariate Analysis of Variance (MANOVA) to answer the first research question of whether team verbal behaviors significantly differ between the three conditions (synthetic versus the control versus the experimenter).

For the second research question and its hypotheses (Hypotheses 2 and 3), Growth Curve Modelling (GCM) (multilevel regression) was selected for analysis due to the repeated measures nested structure of the data set. GCM attempts to estimate between team differences in within-team change, and these within-team patterns of change are referred to as time trends or growth curves (Curran, Obeidat, & Losardo, 2010). These trajectories may change team by team such as they can be flat (i.e., no change over time) or linear (i.e., increasing or decreasing over time. Capturing effects of time trends (e.g., team's learning effect over time) makes GCM more advantageous than traditional techniques such as correlation analysis or linear regression. For instance, correlation analysis and linear regression do not capture the effects of time trends. Also, while GCM meaningfully quantifies individual differences, traditional ANOVAs do not. In GCM, we have built a statistical model step by step and added an additional variable by considering time, and we have only reported the likelihood ratio test statistics for comparisons of our nested models (Cillessen & Borch, 2006; Mirman, 2014).

In the current study, two GCMs were specified to test the relationships of pushing and pulling information on the outcome measures of number of roadblocks completed (TSA metric) and team performance over time. We have followed all modelling steps and reported the test statistics. All analyses were carried out in R version 3.2.3 ("R: The R Project for Statistical Computing 2015) using the "lmerTest" package version 2.0-32 for GCM (Kuznetsova, Brockhoff, & Christensen, 2016) and "ggplot2" package for creating the figures (Wickham, 2010).

4. Results

In a separate article, we have reported significant condition effects for team performance and team situation awareness (Demir, McNeese, Cooke, & Myers, 2016; McNeese, Demir, Cooke, & Myers, submitted for publication). For both team situation awareness and team performance, the synthetic teams performed significantly worse than the experimenter teams, but equal to the control teams (i.e. Synthetic = Control < Experimenter). In this article's current analysis, we build on our findings in our other studies (Demir, McNeese, Cooke et al., 2016;

Table 2
Classification of the team verbal behaviors.

Behaviors	Classification (Push vs Pull)	Description of the behaviors
General Status Updates	*Push*	Informing other team members about current status (*e.g., from pilot to photographer: The airspeed for the current target is 200*)
Suggestions	*Push*	Making suggestions to the other team members (*e.g., from photographer to pilot: Decrease the altitude above 2000 for the current target*)
Planning Ahead	*Push*	Anticipating next steps and creating rules for future encounters (*e.g., From navigator to pilot: The next waypoint is K-AREA. It is a target. The altitude is 1900 and the airspeed 200. The radius is 5*)
Repeated Requests	*Pull*	Requesting the same information or action from other team member(s) (*e.g., From pilot to navigator: What is the next waypoint?*)
Inquiry about status of others	*Pull*	Inquiring about current status of others, and expressing concerns (*e.g., From pilot to photographer: Do we have a good photo for the current target?)*

McNeese et al., submitted for publication) to better understand how teams pushed and pulled information across conditions. In addition, we seek to better under the relationships between both pushing and pulling on TSA and team performance.

4.1. Inter-rater reliability test

There was a substantial agreement between the two experimenters' observations on recording the team verbal behaviors based on the Cohen's Kappa (κ) result, $\kappa = 0.774$ (95% CI, 0.754–0.794), $p < 0.001$. Thus, the mean of both experimenters' observations of the verbal behavior was used, and were then analyzed as follows.

4.2. Pushing and pulling information across the conditions

In order to analyze the frequency of both pushing and pulling information, we performed a 3 (condition) × 5 (mission) × 2 (pushing and pulling information) repeated measures Multivariate ANOVA on the counts of pushing and pulling information. Mauchly's test indicates that the assumption of sphericity for the repeated measure (mission) and mission by behavior interaction were satisfied (χ^2 (9) = 10.9, $p = 0.28$ and χ^2 (9) = 12.7, $p = 0.18$, respectively). Table 3 presents the results of the MANOVA. There were significant two-way interaction effects for behavior by condition and mission by behavior. However, there was no interaction effect of mission by condition.

The significant interaction effect of behavior by condition (i.e., the difference between the conditions for the pushing and pulling information; see Fig. 2) indicates that the synthetic teams pushed significantly less information than the teams in the control condition ($p < 0.05$), and also pulled less (but not significant) information than the control teams ($p = 0.11$; Table 4). In addition, the experimenter teams pushed and pulled significantly more information than the synthetic teams ($p < 0.001$; $p < 0.05$, respectively). On the other hand, teams in the experimenter condition pushed more information than the teams in control condition ($p < 0.05$), and pulled more information (but not significant) than control teams ($p = 0.12$). These effects show that the team members in synthetic condition did less information sharing (pushing information) and also less information requesting (pulling information) than the human teams.

Table 3
MANOVA results for pushing and pulling information.

Source	*df*	*F*	*p*	η^2
Condition	2	16.8	0.001	0.55
Mission	4	3.81	0.006	0.12
Mission by Condition	8	1.15	0.337	0.08
Behavior	1	208	0.001	0.89
Behavior by Condition	2	6.23	0.006	0.32
Mission by Behavior	4	6.79	0.001	0.20
Mission by Behavior by Condition	8	2.01	0.052	0.13

Additionally, the significant interaction effect of mission by behavior shows that the trends for pushing and pulling information differ across the missions. Pushing information increased significantly from Mission 1 to Mission 4 ($M_{M1} = 35.2$, $SD_{M1} = 12.1$, $M_{M4} = 43.1$, $SD_{M4} = 15.9$, $p < 0.001$), but decreased significantly from Mission 4 to Mission 5 ($M_{M5} = 37.1$, $SD_{M5} = 14$, $p < 0.05$). However, pulling information continuously decreased across the missions ($M_{M1} = 14.3$, $SD_{M1} = 8.25$, $M_{M5} = 9.8$, $SD_{M5} = 6.55$, $p < 0.001$: see Fig. 3). To summarize, while team members tended to push more information across missions (except for the last one), they tended to pull less information across missions.

4.3. Prediction of team situation awareness via team verbal behaviors

Growth curve modelling was used to analyze completion of the roadblocks (i.e., team situation awareness) over the five-missions. The overall completion of the roadblocks was modeled by a linear growth model with fixed effects of pulling and pushing information variables in a linear time term. Statistical significance (p-values) for individual parameter estimates was assessed using the normal approximation (i.e., treating the t-value as a z-value).

The model test results (fit statistics) are shown in Table 5. Model 0 represents the null model which only includes the intercept. In Model 1, linear time was added as a predictor of roadblock completion, and it was tested with linear time.

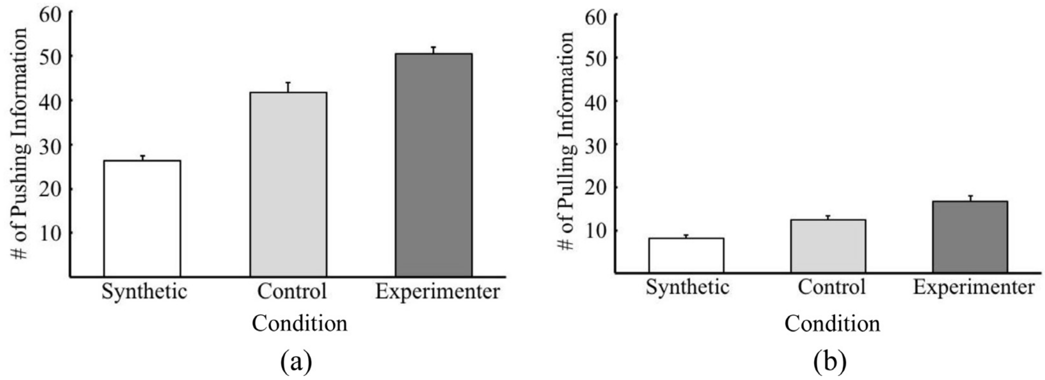

Fig. 2. Team verbal behaviors: (a) pushing and (b) pulling information across the conditions (vertical lines indicate +SE).

Table 4
Means and Standard Deviations of the team verbal behaviors.

Conditions	Pushing information		Pulling information	
	Mean	SD	Mean	SD
Synthetic	26.23	8.97	8.16	5.59
Control	41.73	16.3	12.46	6.77
Experimenter	50.51	9.76	16.69	1.25

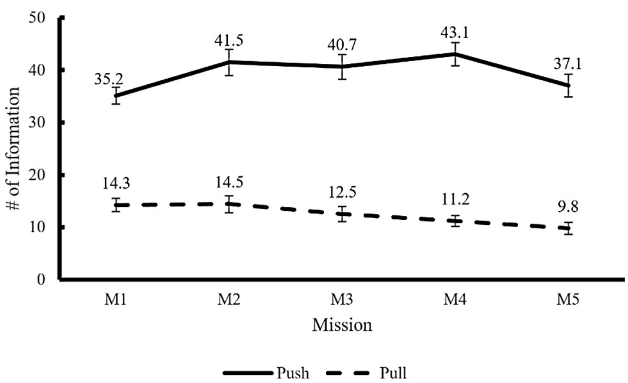

Fig. 3. Team verbal behaviors: (a) pushing and (b) pulling information across the missions (vertical lines indicate ± SE).

After testing linear time, we also added quadratic time in the model. The findings indicate that the linear time improved model fit (χ^2 (3) = 26.19, $p < 0.001$), but not adding quadratic time (χ^2 (1) = 2.75, $p = 0.09$). Therefore, the model with linear time was retained, and linear change of roadblock completion was modeled in all subsequent analysis. In Model 2, the effects of *Pushing* and *Pulling* information were included to the best fitting model, and adding these variables improved the model fit (Model 1 – Model 2 comparisons: χ^2 (2) = 12.95, $p < 0.05$). This result indicates that Model 2 fits better than Model 1. In Model 3, two more effects were added: how both *Pushing* and *Pulling* information influenced the rate of change of roadblock completion over time (interaction effect of pushing and pulling information with linear term). Adding these effects improved the model fit (based on Model 2 – Model 3 comparisons: χ^2 (2) = 7.37, $p < 0.05$). The test statistics of Model 3 are summarized in this study, however, we have still provided fixed effect parameter estimates and their standard errors with *p*-values estimated using normal approximation for the *t*-values for each model during the step by step model development process in Table 5.

In Model 3, initial status ($\gamma_{00} = 1.07$, $SE = 0.12$, $p < 0.001$) and the rate of change, i.e., slope for time ($\gamma_{10} = 0.65$, $SE = 0.14$, $p < 0.001$) were reliably different from zero, which indicates that the teams had training before starting the task. The positive and significant slope indicates that the roadblock completion linearly increased over the course for the teams, meaning that the teams learned how to overcome roadblocks over time.

The main effect of pushing information had a statistically significant positive association with the initial status of roadblock completion ($\gamma_{20} = 0.02$, $SE = 0.01$, $p < 0.05$) in Model 3, meaning that higher levels of pushing information during the task were associated with higher roadblock completion. Also, this positive relationship between pushing information and roadblock completion was constant as the missions progressed ($\gamma_{11} = -0.01$, $SE = 0.01$, $p = 0.85$). On the other hand, the main effect of pulling information on the initial status of roadblock completion was not significant, however, the slope for time was significantly increased depending on pulling information ($\gamma_{12} = 0.34$, $SE = 0.14$, $p < 0.05$). That is, as the missions progressed, the relationship between pulling information and roadblock completion became significantly stronger.

4.4. Prediction of the team performance via team verbal behaviors

We have also applied similar model building steps to predict team performance via pushing and pulling information. The model test results (fit statistics) are shown in Table 6. Model 0 represents a null model. In Model 1, first, linear time was added as a predictor of team performance, and it was tested with linear time. After that quadratic time was added in the model and tested. The findings indicate that linear time and also quadratic time did not improve model fit (χ^2 (3) = 3.64, $p = 0.30$ and χ^2 (1) = 1.59, $p = 0.21$, respectively). However, we still retained the linear

Table 5
Results of model tests for roadblock completion.

	Model 0	Model 1	Model 2	Model 3
Fixed effects				
Initial status (β_{0i})				
Intercept (γ_{00})	1.03*** (0.15)	1.03*** (0.15)	1.03*** (0.12)	1.07*** (0.12)
Push (γ_{20})			0.02*** (0.01)	0.02**(0.01)
Pull (γ_{30})			0.05 (0.08)	0.08 (0.08)
Rate of Change (β_{1i})				
Time (linear) (γ_{10})		0.63*** (0.15)	0.62*** (0.15)	0.65*** (0.14)
Linear by Push (γ_{11})				−0.01 (0.01)
Linear by Pull (γ_{12})				0.34** (0.14)
Fit statistics				
LL	−196.28	−183.19	−176.71	−173.03
$\chi^2 df$		3	2	2
χ^2		26.19***	12.95**	7.36**

Note. Standard errors are between parentheses.
*** $p < 0.001$.
** $p < 0.05$.

Table 6
Results of model tests for team performance.

	Model 0	Model 1	Model 2	Model 3
Fixed effects				
Initial status β_{0i}				
Intercept (γ_{00})	310.1*** (17.40)	310.09*** (17.40)	310.10*** (17.66)	312.93*** (17.58)
Push (γ_{20})			2.03** (0.69)	1.99** (0.69)
Pull (γ_{30})			−16.48 (9.06)	−13.65 (9.23)
Rate of Change β_{1i}				
Time (linear) (γ_{10})		−20.52 (14.33)	−33.15** (15.01)	−30.45** (15.13)
Linear by Push (γ_{11})				0.24 (1.19)
Linear by Pull (γ_{12})				17.28 (14.85)
Fit statistics				
LL	−894.76	−892.94	−888.28	−887.22
$\chi^2 df$		3	2	2
χ^2		3.64	9.32**	2.13

Note. Standard errors are between parentheses.
*** $p < 0.001$.
** $p < 0.05$.

time term during the rest of the model building steps. This was done to test if when the verbal behaviors are added to the above model that linear time explains change of team performance (Snijders & Bosker, 2011). In Model 2, the effects of pushing and pulling information were included to the best fitting model, and adding these variables improved model fit (Model 1 – Model 2 comparisons: χ^2 (2) = 9.32, $p < 0.05$). In Model 3, two more effects were added: how both *Pushing* and *Pulling* information influenced the rate of change of team performance over time (interaction effect of pushing and pulling information with linear term). Adding these effects did not improve the model fit (based on Model 2 – Model 3 comparisons: χ^2 (2) = 2.13, $p = 0.35$). Because there was no model improvement in Model 3 and no significant interaction effects of time by push and pull (see Model 3 in Table 6), we explain the results from Model 2.

In Model 2, initial status ($\gamma_{00} = 310$, $SE = 17.7$, $p < 0.001$) and the rate of change, i.e., slope for time ($\gamma_{10} = -33.2$, $SE = 0.15$, $p < 0.05$) were reliably different from zero. Initial status indicates that the teams had training before starting the task. The negative and significant slope indicates that the team performance linearly decreased over the course of time (likely due to the complexity of the last mission). In Model 2, the pairwise comparison indicates that pushing information was positively and significantly associated with team performance ($\gamma_{20} = 2.03$, $SE = 0.69$, $p < 0.05$). In addition to that, there was no main effect of pulling information on team performance ($\gamma_{30} = -16.5$, $SE = 9.06$, $p = 0.07$). Therefore, when the teams provide information (pushing) during the task, the teams tend to perform better. However, requesting information (pulling) among the team members was unrelated to overall team performance.

Table 7
The results of the hypotheses.

Hypotheses	Test results
H1: Teams in experimenter condition will have more team verbal behaviors (both pushing and pulling) than the teams in the control condition, and the control teams will have more team verbal behaviors (both pushing and pulling) than the synthetic teams	• *Pushing Information:* Synthetic < Control < Experimenter • *Pulling Information:*Synthetic < Control < Experimenter
H2: The effect of both *pushing* and *pulling* information will be positively related with TSA	• *Pushing* and TSA = positive & significant • *Pulling by Time* and TSA = positive & significant
H3: The effect of both *pushing* and *pulling* information will be positively related with Team Performance (TP)	• *Pushing* and TP = positive & significant • *Pulling* and TP = negative & not significant

5. Discussion and conclusion

Effective team communication, a fundamental part of team coordination, is crucial for both effective Team Situation Awareness (TSA) and team performance. In this study's task, participants were instructed to push information to others, and over time master the specific timing of information sharing to the intended recipient. Therefore, if things are going well, there should be little need for pulling information. In addition, good teamwork involves anticipating the needs of teammates and that means pushing information before it is requested. In this study, we examined the relation between team verbal behaviors (i.e., pushing and pulling information) on TSA and team performance in the context of human-autonomy and human-human teaming. Therefore, we tested the following hypotheses, which are depicted in Table 7. Pushing information was positively associated with TSA and team performance, and the synthetic teams had lower levels of both pushing and pulling information than the all-human teams.

The synthetic teams' lower levels of pushing information indicate an inadequate ability to anticipate teammate needs, as opposed to members of the all-human teams—and the experimenter condition in particular. Interestingly, the lower levels of pushing exhibited by the synthetic teams were not compensated by increased pulling or requesting of information. Therefore, the synthetic teams were sharing less information overall, which may explain a lack of TSA and lower performance when compared to the all-human teams. In the experimenter condition, in which the confederate pilot could take on more responsibilities regarding teamwork and team interaction, team members may have had more opportunities to practice and develop team verbal behaviors (i.e., pulling and pushing information). This, in turn, might have led to more effective anticipation of their teammates' needs and, ultimately, higher TSA and team performance.

Within the HAT concept, having effective team verbal behaviors among the team members is related to high levels of TSA and team performance. In this study, we have outlined the relationship between the team-level verbal behaviors (as opposed to individual-level behaviors), TSA, and team performance within the context of human-autonomy and human-human teaming. Through our research with the synthetic teammate, we have learned that anticipation of other team member behaviors and information requirements in the synthetic teams are important for effective TSA and team performance. In order to make HAT more effective in terms of teamwork, we need to develop mechanisms to enhance pushing information within HAT. Futures studies might examine this same relationship, but using individual-level pushing and pulling behaviors in order to better understand the source of behavioral limitations in HATs.

Acknowledgements

This work was partially supported by the Office of Naval Research Award – 'United States' – N000141110844446 (Program Managers: Marc Steinberg, Paul Bello). We also acknowledge the assistance of Christopher Myers, Jerry Ball, Mary Friedman, Steven M. Shope, Cade Bartlett, Ashley Knobloch, and Lourdes Reyes.

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?*. Oxford New York: Oxford University Press.

Ball, J., Myers, C., Heiberg, A., Cooke, N. J., Matessa, M., Freiman, M., & Rodgers, S. (2010). The synthetic teammate project. *Computational and Mathematical Organization Theory, 16*(3), 271–299. http://dx.doi.org/10.1007/s10588-010-9065-3.

Cillessen, A. H. N., & Borch, C. (2006). Developmental trajectories of adolescent popularity: A growth curve modelling analysis. *Journal of Adolescence, 29*(6), 935–959. http://dx.doi.org/10.1016/j.adolescence.2006.05.005.

Cooke, N. J., & Gorman, J. C. (2009). Interaction-based measures of cognitive systems. *Journal of Cognitive Engineering and Decision Making, 3*(1), 27–46. http://dx.doi.org/10.1518/155534309X433302.

Cooke, N. J., Gorman, J. C., Duran, J. L., & Taylor, A. R. (2007). Team cognition in experienced command-and-control teams. *Journal of Experimental Psychology: Applied, 13*(3), 146–157<https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1037/1076-898X.13.3.146>.

Cooke, N. J., Gorman, J. C., & Kiekel, P. A. (2008). Communication as team-level cognitive processing. In M. P. Letsky, N. W. Warner, S. M. Fiore, & C. A. P. Smith (Eds.), *Macrocognition in teams: Theories and methodologies* (pp. 51–64). Abingdon, Oxon, GBR: Ashgate Publishing Group, Retrieved from<http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10250476>.

Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013). Interactive team cognition. *Cognitive Science, 37*(2), 255–285. http://dx.doi.org/10.1111/cogs.12009.

Cooke, N. J., Gorman, J., Pedersen, H., Winner, J., Duran, J., Taylor, A., … Rowe, L. (2007). *Acquisition and retention of team coordination in command-and-control*. Retrieved from http://www.dtic.mil/docs/citations/ADA475567.

Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. J. (2000). Measuring team knowledge. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 42*(1), 151–173. http://dx.doi.org/10.1518/001872000779656561.

Cooke, N. J., & Shope, S. M. (2004). Designing a synthetic task environment. In L. R. E. Schiflett, E. Salas, & M. D. Coovert (Eds.), *Scaled worlds: Development, validation, and application* (pp. 263–278). Surrey, England: Ashgate Publishing, Retrieved from<http://www.cerici.org/documents/Publications/scaled%20worlds%20paper3.pdf>.

Cooke, N. J., & Shope, S. M. (2005). Synthetic task environments for teams: CERTT's UAV-STE. In N. Stanton, A. Hedge, K. Brookhuis, E. Salas, & & H. Hendrick (Eds.), *Handbook of human factors and ergonomics methods* (pp. 41–46). Boca Raton: FL: CRC Press.

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society, 11*(2), 121–136. http://dx.doi.org/10.1080/15248371003699969.

Demir, M., & Cooke, N. J. (2014). Human teaming changes driven by expectations of a synthetic teammate. *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 58(1), pp. 16–20). . http://dx.doi.org/10.1177/1541931214581004.

Demir, M., McNeese, N. J., & Cooke, N. J. (2016). Team communication behaviors of the human-automation teaming. In *2016 IEEE international multi-disciplinary conference on cognitive methods in situation awareness and decision support (CogSIMA)* (pp. 28–34). http://dx.doi.org/10.1109/COGSIMA.2016.7497782.

Demir, M., McNeese, N. J., Cooke, N. J., Ball, J. T., Myers, C., & Freiman, M. (2015). Synthetic teammate communication and coordination with humans. *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 59(1), pp. 951–955). . http://dx.doi.org/10.1177/1541931215591275.

Demir, M., McNeese, N. J., Cooke, N. J., & Myers, C. (2016). The synthetic teammate as a team player in command-and-control teams. *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 60(1), pp. 116). . http://dx.doi.org/10.1177/1541931213601026.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 32(2), pp. 97–101). . http://dx.doi.org/10.1177/154193128803200221.

Endsley, M. R. (1989). *Situation Awareness in an Advanced Strategic Mission (No. NOR DOC 89–32)*. Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (2015). *Autonomous Horizons: System Autonomy in the Air Force - A Path to the Future* (Autonomous Horizons No. AF/ST TR 15-01). Washington DC: Department of the Air Force Headquarters of the Air Force. Retrieved from <http://www.af.mil/Portals/1/documents/SECAF/AutonomousHorizons.pdf?timestamp=1435068339702>.

Endsley, M. R., & Robertson, M. M. (2000). Situation awareness in aircraft maintenance teams. *International Journal of Industrial Ergonomics, 26*(2), 301–325. http://dx.doi.org/10.1016/S0169-8141(99)00073-6.

Gorman, J. C., Amazeen, P. G., & Cooke, N. J. (2010). Team coordination dynamics. *Nonlinear Dynamics, Psychology, and Life Sciences, 14*(3), 265–289.

Gorman, J. C., Cooke, N. J., Pedersen, H. K., Winner, J., Andrews, D., & Amazeen, P. G. (2006). Changes in team composition after a break: building adapative command-and-control teams. *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 50(3), pp. 487–491). . http://dx.doi.org/10.1177/154193120605000358.

Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics, 49*(12–13), 1312–1325. http://dx.doi.org/10.1080/00140130600612788.

Klein, G., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., & Feltovich, P. J. (2004). Ten challenges for making automation a "team player" in joint human-agent activity. *IEEE Intelligent Systems, 19*(6), 91–95.

Krogmann, U. (1999). *From Automation to Autonomy-Trends Towards Autonomous Combat Systems* (Unclassified No. RTO MP-44). France: NATO Science and Technology Organization. Retrieved from <http://www.dtic.mil/dtic/tr/fulltext/u2/p010300.pdf>.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *Package "lmerTest"* [R Software Package]. Retrieved from <https://cran.r-project.org/web/packages/lmerTest/lmerTest.pdf>.

McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. W. (submitted for publication). Teaming with a synthetic teammate: insights into human-autonomy teaming. *The Journal of the Human Factors and Ergonomics Society*.

Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R* (1st ed.). Boca Raton: Chapman and Hall/CRC.

R: The R Project for Statistical Computing (2015). Retrieved August 24, 2016, from <https://www.r-project.org/>.

Schooley, L. C., Zeigler, B. P., Cellier, F. E., & Wang, F. Y. (1993). High-autonomy control of space resource processing plants. *IEEE Control Systems, 13*(3), 29–39. http://dx.doi.org/10.1109/37.214942.

Schulte, A., Donath, D., & Lange, D. S. (2016). Design patterns for human-cognitive agent teaming. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics* (pp. 231–243). Springer International Publishing, Retrieved from<http://link.springer.com/chapter/10.1007/978-3-319-40030-3_24>.

Snijders, T., & Bosker, R. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Los Angeles: SAGE Publications Ltd..

Vagia, M., Transeth, A. A., & Fjerdingen, S. A. (2016). A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed? *Applied Ergonomics, 53*(Part A), 190–202. http://dx.doi.org/10.1016/j.apergo.2015.09.013.

Wickham, H. (2010). *ggplot2: Elegant graphics for data analysis* (1st ed. 2009. Corr. 3rd printing 2010 edition). New York: Springer.