COMMUNICATION IN AI-ASSISTED TEAMS DURING AN INTERDISCIPLINARY DRONE DESIGN PROBLEM

Gyory, Joshua T. (1);
Song, Binyang (2);
Cagan, Jonathan (1);
McComb, Christopher (2)

1: Carnegie Mellon University;
2: Pennsylvania State University

ABSTRACT

Human-artificial intelligent (AI) - assisted teaming is becoming a strategy for coalescing the complementary strengths of humans and computers to solve difficult tasks. Yet, there is still much to learn regarding how the integration of humans with AI agents into a team affects human behavior. Accordingly, this work begins to inform this research gap by focusing specifically on how the communication structure and interaction changes within AI-assisted human teams. The underlying discourse data for this work originates from a prior research study in which teams solve an interdisciplinary drone design and path-planning problem. Several metrics are employed in this work to study team discourse, including count, diversity, content richness, and semantic coherence. Results show significant differences in communication behavior in AI-assisted teams including more diversity and frequency in communication, more exchange of information regarding principal design parameters and problem-solving strategies, and more cohesion. Overall, this work takes meaningful steps towards understanding the effects of AI agents on human behavior in teams, critical for fully building effective human-AI hybrid teams in the future.

Keywords: Artificial intelligence, Communication, Human behaviour in design, Hybrid teaming

Contact:
Cagan, Jonathan
Carnegie Mellon University
Mechanical Engineering
United States of America
cagan@cmu.edu

1 INTRODUCTION

As technology and data fuse across boundaries of specialized domains, human society is increasingly challenged by multifaceted problems. It is difficult for humans to tackle these complexities alone. On one hand, advances in artificial intelligence (AI) have shown great potential in solving task-specialized, computation-intensive problems. On the other hand, AI is also not able to solve these interdisciplinary tasks alone as it still lacks the problem-solving behaviours to deal with uncertainty and creative thinking (Lake et al., 2017). To tackle this shortcoming, human-AI collaboration is recognized as a promising strategy to integrate the complementary capabilities of humans and AI to augment each other (Dellermann et al., 2019). Prior studies have shown human teams assisted by AI outperform human-only teams across a variety of domains (Roll et al., 2014; Liew, 2018). In fact, previous research by the authors corroborates these studies, having shown that AI-assisted human teams achieved considerably better performance in solving an interdisciplinary drone design and operations problem. To fully grasp these aforementioned results, there exists a need to understand how AI assistance changes team dynamics and behaviours. To date, little attention has been paid to the influence of AI on the problem-solving behaviours of the human team members assisted by AI.

Behavioural analyses in that previous work by the authors revealed that AI-assisted teams exhibited greater agility in adapting to structural shocks, performed fewer but more meaningful design actions, and explored more of the design space than human-only teams (Song *et al.*, 2020). While only a preliminary analysis on team communication was conducted, this current work aims to expand that analysis to deepen the understanding of this fundamental process of AI-assisted human teams, specifically focusing on the discourse among the human team members. Team communication is a critical component for efficient knowledge distribution, sharing, and integration, particularly when solving complex, multi-disciplinary tasks (Hendry, 2004). Analyses of team communication provide indicators to capture this knowledge exchange and common understanding between team members, which can significantly influence team performance (Zhang *et al.*, 2016). Thus, gaining a better understanding of how the process of communication is affected through AI boasts major insights to human-AI collaboration. This work includes a comprehensive set of analyses, including count, distribution, content richness, and semantic analyses, to dive deeper into this gap in the research.

2 BACKGROUND

2.1 Human-AI collaboration

Human-AI collaboration is expected to achieve superior outcomes by taking advantage of the complementary strengths of humans and AI. A strand of research in this area focuses on the effectiveness of human-AI collaboration in augmenting the ability of humans and/or AI in solving complex problems. Among them, Hu and Taylor (2016) show that a computer-aided design intelligent tutor helps students apply their learned knowledge to solve novel problems, providing suggestions and guiding their exploration. Instructional design agents support novice designers in exploring complex design spaces through a design study with a solar farm design problem (Schimpf *et al.*, 2019). Song *et al*., (2020) report that the assistive of AI agents can effectively improve the performance of AI-assisted human teams in the configurational and operational design of drones. Other contexts showing the effectiveness of AI assistance include tutoring and education (Roll et al., 2014), job shop scheduling (Higgins, 1999), and clinical imaging (Liew, 2018). However, it has also been reported that AI does not always boost the performance of human teams assisted by AI (Zhang, *et al.*, 2020)

Moreover, many efforts work towards the understanding, design, and improvement of human-AI interaction to facilitate collaboration (Dellermann et al., 2019; Gunning et al., 2019). For example, according to the information flow and role distribution between humans and AI, a group of researchers classify human-AI collaboration into two forms, human intelligence in the loop of artificial intelligence (human-in-the-loop) and AI in the loop of human intelligence (AI-in-the-loop) to clarify the required characters of interactive AI (Dellermann et al., 2019). In terms of AI design, the common language and the explainability and transparency of AI are considered essential to improve humans' understanding of and trust towards AI (Gunning et al., 2019; Ezer *et al.*, 2019). Van Den Bosch and Bronkhorst (2018) study human-AI collaboration in complex decision making, categorizing six levels according to the type of interaction and level of collaboration. Yang, Steinfeld, and Rosé (2020) identify design challenges of human-AI interaction and propose strategies for addressing them. Sill,

how AI specifically reshapes the interactions and behaviours between human team members remains an open question, and one this current work moves towards.

2.2 Team communication in design problem solving

Team communication is a fundamental component of complex engineering design processes. It enables team members to integrate specialized knowledge, bridge gaps, and negotiate differences across domains (Toft, Howard, and Jorgensen, 2003). Communication plays a boundary spanning role that supports design knowledge sharing, exploration, collaboration, and coordination (Sonnenwald, 1996; Hendry, 2004). Prior studies recognize communication as a prime success factor of shared leadership in teams (Kratzer, Leenders, and Van Engelen, 2010), impacting team creativity (Wu and Cormican, 2016). In design teams aiming at complex problem solving, the effectiveness of interpersonal communication influences team efficiency, performance, and progress toward their design (Zhang *et al.*, 2016). Inadequate or ineffective communication can hinder task and team achievements (Safoutin *et al.*, 1993). Moreover, studies on team communication facilitate our understanding of team cognitive processes. As such a critical component to team performance in complex engineering design, understanding how the integration of AI agents into human design teams affects team communication requires further attention.

A variety of techniques exist to analyse and model language-based communications, including content coding and sequential analysis (Bowers et al., 1998), hidden Markov models (Stolcke et al., 2000), classification (Sexton and Helmreich, 2000), Latent Semantic Analysis (LSA) (Dong, Hill, and Agogino, 2004; Gorman *et al.*, 2003), and topic modelling (Blei, Ng, and Jordan, 2003; Gyory, Kotovsky, and Cagan, 2020). As a fully automatic method of analysing, representing, and quantifying semantic information from text, LSA (Deerwester *et al.*, 1990) is a powerful tool to measure cognition and knowledge construction within teams through modelling the semantic meaning in sentences. For instance, Dong, Hill, and Agogino (2004) utilize LSA as a quantitative measure to model knowledge acquisition and representation through language-based communication within engineering design teams. Gorman *et al.* (2003) develop LSA-based methods of communication content modelling to categorize team communication, calculate its task relevance, and track task-relevant topic shifts. LSA has also been applied to modelling semantic meaning in communication in various contexts for assessing effectiveness of communication (Kane et al., 2020), trust between participants (Vrana et al., 2018), and emotions of participants (Inrak and Sinthupinyo, 2010). In this work, LSA compares the content of discourse across AI-assisted human teams and human-only teams and interactions among task-specialized disciplines.

3 METHODOLOGY

3.1 Drone Design Study

To study the overarching research goal of this work, a large-scale human design study was jointly conducted between researchers at Carnegie Mellon University and Pennsylvania State University. The study, approved by the Institutional Review Board at Carnegie Mellon University, recruited 395 participants representative across 18 different universities in the United States. The experiment, run via an online platform named HyForm™, simulates a real-world problem which merges business, drone design, and operations problems, in which teams aim to maximize their profit by operating designed drones to deliver parcels for customer markets. Teams consist of six individuals across three respective disciplines including business, operations, and design. The business discipline, consisting of a business manager, operations manager and design manager, focuses on market targeting; the operations discipline, consisting of two operations specialists and one operations manager, focuses on drone path planning; while the design discipline, consisting of one design specialist and one design manager focuses on drone design. More detailed descriptions regarding the experimental platform, problem statement, roles, and experimental methodology can be found in related work (Song *et al.*, 2020).

The team roles under each distinct discipline are provided with and have access to separate information regarding the overall problem (e.g., the operations discipline has access to the market information but no access to the drone configurations). Thus, to obtain maximum performance, team members must be able to effectively communicate with each other to share information and build common understanding of constraints and goals. In the experiment, text-only communication is enabled

through embedded communication channels. Figure 1 depicts two different team structures employed: the open team structure (Figure 1A) and the restrictive team structure (Figure 1B). The labelled communication channels in the figure designate who can directly interact with whom. For example, in the *Operations* channel in the open communication structure, the operations manager and operations specialists can all directly communicate with one another, but not directly with the design specialist.
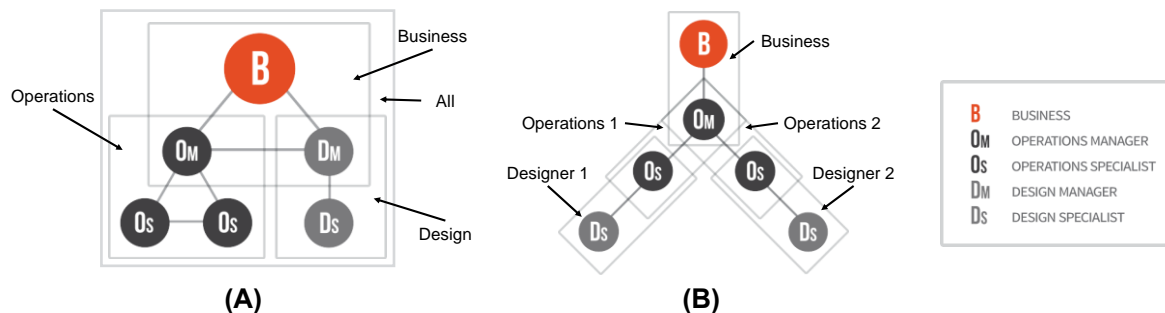

Figure 1. The two different team structures showing the channels of communication and the team roles: (A) Open team structure and (B) Restrictive team structure.

The experimental design split team problem-solving into two 20-minute sessions. Teams either started with the open team structure or the restrictive team structure and then switched between the two sessions, with changes only in the configuration of communication channels. Two AI agents, that can respectively design drones using character-recurrent neural networks (Stump *et al*., 2020) and generate path plans based on linear programming search, are also incorporated into the HyForm™ platform to collaborate with during problem solving. By controlling teams' access to the AI agents, the experiment consists of four experimental conditions: teams with no access to the AI agents (human-only teams) and those with access to the AI agents (AI-assisted human teams), and whether the teams start with the open or restrictive communication structures in the initial design session. In total, participants are randomly assigned to 44 teams with 11 teams in each of the experimental conditions (Zhang et al, 2021).

3.2 Extraction and pre-processing of raw discourse data

During the problem-solving sessions, team members communicate with each other through the channels dictated by the two team structures in Figure 1. The raw chat data is separated into documents according to these channels, each containing all communication within that channel. For example, the open team structure consists of an *All* channel in which all six team members can communicate directly with each other. Thus, the document for that channel contains all discourse between any of the members that utilize that channel to communicate. During the experiment, participants can communicate with the experimenter when needed through a *Help* channel. All discourse via that channel is removed from the text analyses.

The raw chat data goes through a few post-hoc pre-processing steps, common through natural language processing pipelines. The first step includes tokenization of the documents, in which each document is represented as a vector of words. The length of the vector is the number of words in the corpus and the value in the vector is the number of times a specific word appears in a document, i.e., word frequency. Words identified in the Natural Language Toolkit (NLTK) as stop words are removed from the corpus (Bird, Klein, and Loper, 2009). Punctuation are removed as well as short words (those words that have two or fewer characters) in order to eliminate noise in the data. Finally, words are also stemmed and lemmatized so that they return to the same tense and their dictionary forms.

3.3 Latent Semantic Analysis model

Following the pre-processing steps, a LSA model is trained on the discourse documents. Conceptually, LSA measures the similarity of text documents by comparing the frequency of terms across documents and across the entire text corpus (Landauer, Foltz, and Laham 1998). This natural language processing technique assumes that words semantically similar in meaning will appear in similar documents, thereby measuring the similarity of the documents themselves. Mathematically, LSA decomposes documents

into a lower-dimensional space via singular value decomposition. Once tokenized (represented as a vector of word frequencies), the singular value decomposition can be computed as defined in Equation 1:

$$X = U \cdot S \cdot V^T, \tag{1}$$

where X is an $[n \times m]$ occurrence matrix (describing the occurrence of words in the documents) with n number of words and m channel documents, U is an $[m \times r]$ concept vector matrix with rank r, S is an $[r \times r]$ singular values matrix, and V is an $[n \times r]$ channel matrix. The LSA model treats each channel as a distinct document, for example, the content from *Designer 1* would be a single document. This includes all discourse from the respective roles within the channel as dictated by the team structures in Figure 1. As such, the cohesion analysis is done at the discipline and team levels, not at the role level. Once reduced into this space, the cosine similarity between document vectors (d_i) computes the similarity between documents. Although in this work, the cosine distance, D, compares the cohesion (or lack thereof) across conditions, which is computed as the complement of the cosine similarity, as shown in Equation 2, for all i, k pairs of documents vectors:

$$D_{i,k} = 1 - \frac{d_i \cdot d_k}{\|d_i\| \, \|d_k\|} , \forall \; i, \; k \; \in [1, m]. \tag{2}$$

4   RESULTS & DISCUSSION

With the methods outlined in Section 3, this section investigates the influence of the incorporation of AI agents on team communication. Section 4.1 studies the communication count, distribution, and content richness of the team discourse as well as providing a temporal perspective. Then, in Section 4.2, an LSA model measures discourse coherence on the communication content. These analyses are done post-hoc to the collection of the data in real-time during the experiment.

4.1 Count, distribution and content richness of team communication

In this section, a set of statistical analyses are used to assess the count, distribution, and content richness of the team communication. First, the communication count is compared between the AI-assisted and the human-only teams. Here, any line of discourse, or turn, counts as a distinct interaction/count. Overall, the AI-assisted teams ($\bar{x} = 156 \pm 13$) do not communicate significantly more than the human-only teams ($\bar{x} = 137 \pm 12$) on average ($p = 0.16$). However, a temporal analysis shows that the counts do become significant over time. Figure 2 presents a one-minute sliding window average of team communication frequency. It shows that close to the entirety of the second experimental session, the AI-assisted teams have a significantly higher communication frequency than the human-only teams, occurring after the teams experience a team structure change.
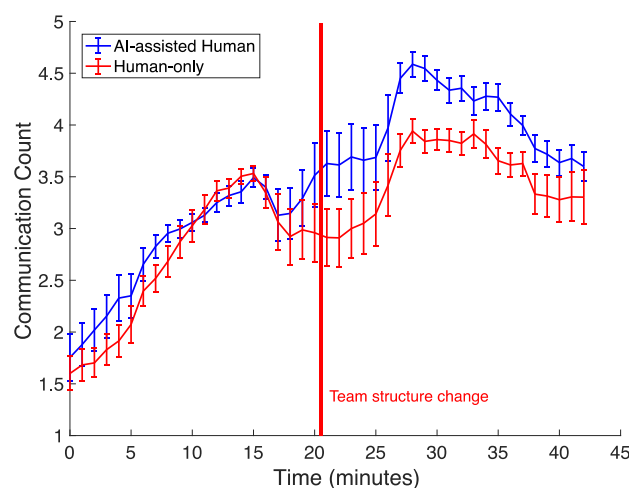
Figure 2. Temporal analysis of average communication frequency in one-minute intervals between the AI-assisted and the human-only teams. Error bars show ± 1 standard error.

The communication counts are further compared across the distinct channels. Figure 3 depicts the distribution of team communication across these channels for both team conditions. Corresponding to the open team structure (*All, Designer, Business,* and *Operations* channels), the AI-assisted teams

communicate in the *All* channel, relative to the discipline specific channels, more than the human-only teams do. That is, the participants working with AI communicate information with the whole team rather than only with the discipline specific members, indicating that the AIs enable the AI-assisted teams to process a larger amount of information. Corresponding to the restrictive team structure (*Business*, *Designer 1 & 2* and *Operations 1 & 2* channels), the AI-assisted teams present a different effect, with an even communication distribution between these channels compared with the human-only teams. Moreover, the *Operations 1 & 2* channels see significantly greater communication in the AI-assisted teams than in the human-only teams ($p < 0.01$). The experimental data shows that the operations specialists use the operations AI agent (42 times) significantly more than the design specialists use the design AI agent (14 times) ($p < 0.002$). This allows the roles in the *Operations 1 & 2* channels to devote more time to communication, which forms a critical pathway of communication between the management team and the design specialists in the restrictive team structure. In addition, Figure 3 shows a significantly larger proportion of the communication across both experimental conditions occurs in the *Business* channel ($p << 0.01$), which includes the managers across both team structures. While not different across team conditions, it does indicate that the inter-discipline discourse among the managers constitutes a significant proportion of the overall team discourse and thus flow of information throughout the problem-solving process.
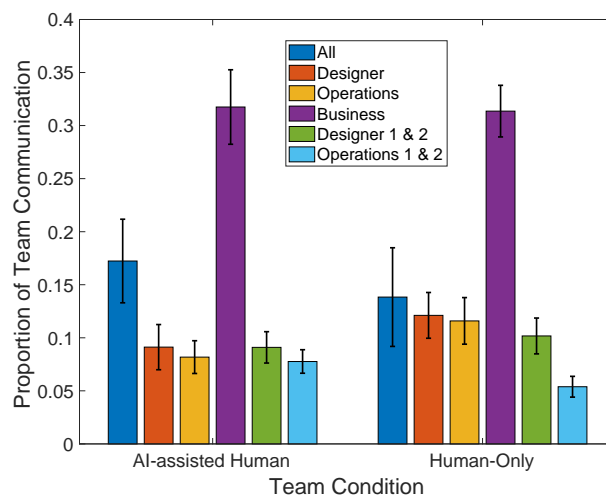

Figure 3. Average proportion of communication between the AI-assisted and the human-only teams across the communication channels. Error bars show ± 1 standard error.

Next, the content richness of team communication is measured in terms of both the total number of tokens and the total number of unique tokens. This richness provides an additional dimensionality for assessing communication than merely count, indicating the amount and diversity of the discourse content itself. Table 1 displays the average number of total tokens for both the AI-assisted and human-only teams as well as the average number of unique tokens. Both the total number of tokens and the number of unique tokens exhibit significant effects between the AI-assisted and human-only teams. Across all channels, the AI-assisted team discourse contains, on average, 513 total tokens while the human-only teams' discourse contains 409 total tokens ($p = 0.042$). For the number of unique tokens, the AI-assisted teams use, on average, 196 unique words while the human-only teams use 166 unique words ($p = 0.052$). That is, the AI-assisted teams present a higher communication richness than the human-only teams indicated by the total and unique number of tokens contained in their communications.

Table 1. The number of total and average tokens and total (left) and average of unique tokens (right) between the AI-assisted teams and human-only teams

	Total Tokens		Unique Tokens	
	AI-assisted Human	Human-Only	AI-assisted Human	Human-Only
Sum	11,268	9,002	4,310	3,658
Average	512.2	409.2	195.9	166.3
Standard Error	40.8	40.8	11.2	11.3

Motivated by this result, two categories of discourse information are further analysed to deepen the understanding of the difference in content between the AI-assisted and the human-only teams. The first category (Table 2) contains topic tokens and their counts representative of principal design parameters, such as drone metrics (e.g., flying range, velocity, cost), market and operations plan information (house, customer, food), and budget. The second category (Table 3) consists of tokens that indicate the descriptors of the strategies of teams' problem-solving process, including trend verbs and adjectives indicating qualitative requirements for the design parameters. Statistical tests show that the AI-assisted teams use these two categories of tokens significantly more frequently than the human-only teams ($p = 0.003$). This result illustrates that the AI-assisted teams exchange greater information regarding the principal topics of the complex problem and their problem-solving directions or strategies regarding the topics, and thereby providing evidence for more effective discourse.

Table 2. Frequency of occurrence of the principal design parameters topic tokens.

Tokens	AI-assisted Human	Human-only	Tokens	AI-assisted Human	Human-only	Tokens	AI-assisted Human	Human-only
budget	108	66	distance	49	38	range	304	258
capacity	134	113	food	147	102	speed	6	5
consumer	8	0	house	236	151	velocity	17	13
cost	218	164	package	32	22	weight	84	88
customer	50	17	parcel	58	34			
delivery	84	57	payload	12	8			

Table 3. Frequency of occurrence of the descriptive strategics topic tokens.

Tokens	AI-assisted Human	Human-only	Tokens	AI-assisted Human	Human-only	Tokens	AI-assisted Human	Human-only
increase	49	41	change	35	28	heavy	15	1
reduce	14	11	balance	7	1	high	85	52
sacrifice	4	0	focus	51	25	large	22	13
shorten	4	0	prioritize	14	7	long	50	43
cut	6	4	optimize	13	10	low	54	21
maximize	24	11	combine	7	3	quick	7	4
minimize	3	2	improve	7	17	short	30	26
add	25	32	cheap	17	20	small	14	15
decrease	6	2	expensive	9	11			

4.2 Discourse Coherence

An LSA model is then used to study the cohesion among the team discourse. The motivation behind this analysis is that the literature shows that more cohesive discourse indicates more shared knowledge representation, exchange, and can lead to better performance. To compare the impact of the AI agents across the disciplines themselves, the semantic distances between the AI-assisted and human-only teams' discipline documents are first computed. The documents for this analysis are combined in this way: "Design" combines all communication from the design-related channels, including *Designer*, *Designer 1*, and *Designer 2*; and "Operations" combines all communication from the corresponding operations-related channels. The computed cosine distances resulted in: *All*: $D = 0.22$, *Business*: $D = 0.098$, *Design*: $D = 0.078$, and *Operations*: $D = 0.11$ (where D is the cosine distance between the AI-assisted teams and human-only teams). This shows larger semantic disparities between the contents of the *All* channel between the AI-assisted and human-only teams rather than among the main disciplines themselves (i.e., business, design, and operations). Out of the main disciplines, the *Operations* channels experience the larger disparity between conditions, which as mentioned previously, relied more heavily on the AI agent.

Next, LSA is leveraged to also compare the semantic content between the channels. Figure 4 presents heatmap matrices of the semantic distances between the AI-assisted teams' channels (Figure 4A) and the human-only teams' channels (Figure 4B). These show all pairwise comparisons of the chat channels and therefore result in symmetric matrices. The main diagonals are zero as they define the semantic distance of a channel with itself. There exists high semantic similarity (small cosine distance) among the intra-discipline channels (i.e., *Designer*, *Designer 1*, and *Designer 2* in the design discipline and *Operations*, *Operations 1*, and *Operations 2* channels in the operations discipline). This

high similarity among the intra-discipline channels serves as a validation measure from the LSA model, as the discourse within the respective disciplines should be, and turns out to be, relatively more consonant with each other.
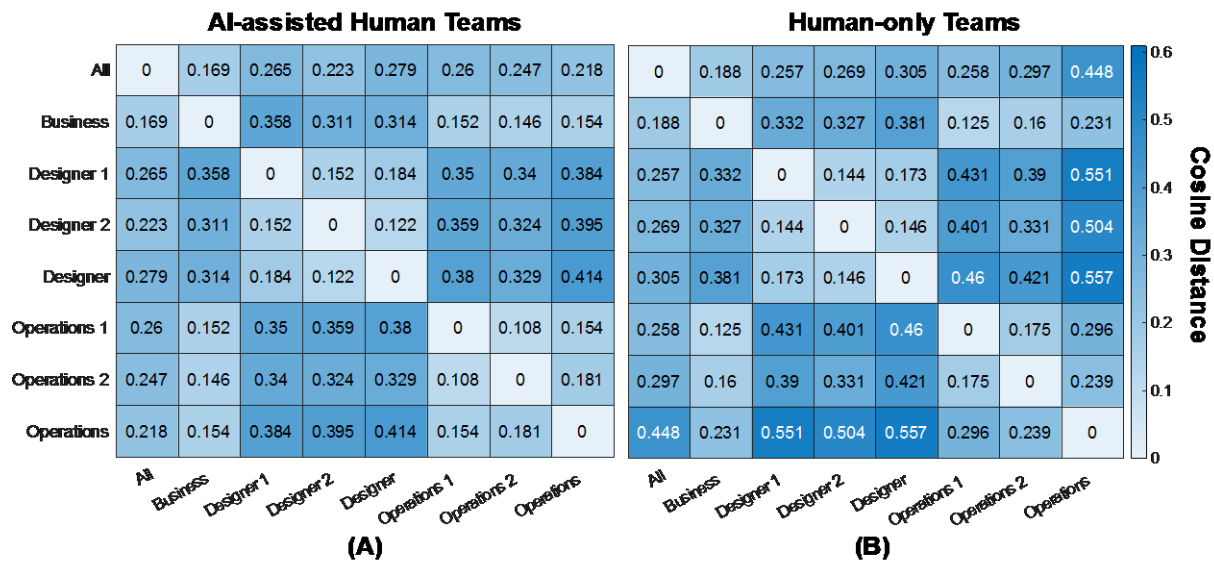

AI-assisted Human Teams

	AI	Business	Designer 1	Designer 2	Designer	Operations 1	Operations 2	Operations
All	0	0.169	0.265	0.223	0.279	0.26	0.247	0.218
Business	0.169	0	0.358	0.311	0.314	0.152	0.146	0.154
Designer 1	0.265	0.358	0	0.152	0.184	0.35	0.34	0.384
Designer 2	0.223	0.311	0.152	0	0.122	0.359	0.324	0.395
Designer	0.279	0.314	0.184	0.122	0	0.38	0.329	0.414
Operations 1	0.26	0.152	0.35	0.359	0.38	0	0.108	0.154
Operations 2	0.247	0.146	0.34	0.324	0.329	0.108	0	0.181
Operations	0.218	0.154	0.384	0.395	0.414	0.154	0.181	0

Human-only Teams

	AI	Business	Designer 1	Designer 2	Designer	Operations 1	Operations 2	Operations
All	0	0.188	0.257	0.269	0.305	0.258	0.297	0.448
Business	0.188	0	0.332	0.327	0.381	0.125	0.16	0.231
Designer 1	0.257	0.332	0	0.144	0.173	0.431	0.39	0.551
Designer 2	0.269	0.327	0.144	0	0.146	0.401	0.331	0.504
Designer	0.305	0.381	0.173	0.146	0	0.46	0.421	0.557
Operations 1	0.258	0.125	0.431	0.401	0.46	0	0.175	0.296
Operations 2	0.297	0.16	0.39	0.331	0.421	0.175	0	0.239
Operations	0.448	0.231	0.551	0.504	0.557	0.296	0.239	0

(A) (B)

Figure 4. The semantic distances between the communication channels of the AI-assisted teams (A) and the human-only teams (B).

The distance matrices reveal additional insights when analysing different pairs of channel interactions between the AI-assisted teams and the human-only teams. First, the global value between both sets of distance matrices is significantly different from each other ($p = 0.037$), with the AI-assisted having greater overall cohesion (smaller cosine distances). The upper right (and lower left), 3×3, design-operations blocks reveal significant differences between the two conditions ($p = 0.003$), with the AI-assisted teams exhibiting higher similarity between the operations and design channels. The 3×3 block of operations-operations channel similarities are also more highly similar in content ($p = 0.015$), but not so for the inter-similarity of the design discipline ($p = 0.68$). Again, this common theme of larger effects on the operations discipline (due to their higher reliance on the AI agents) even emerges in the semantic content of the discourse.

The team structures and experimental design of the study underscore the criticality of effective communication, particularly between the operations and design disciplines. Neither has access to the constraints and goals of the other, so communication is imperative. The higher cohesiveness found in the AI-assisted teams indicate this passing of information and knowledge transfer. Recall that the AI-assisted teams also present significantly higher content and diversity (more unique tokens) in communication (rather than just number of turns) as well as more discourse related to strategy and the design parameters. With the help of the AI agents, the human team members can expend more time to this information transfer and create more common knowledge and understanding of the design problem. Altogether, these results validate that incorporating the AI agents into human teams improves team communication effectiveness.

5 CONCLUSION

This work studies the effects of AI agents on the problem-solving behaviour of teams via team communication. The discourse data originates from a large-scale, human subjects study simulating a drone design and operations problem in which teams of six members need to coordinate across three task-specialized disciplines to produce as much profit as possible. Prior work by the authors revealed that the AI-assisted teams significantly outperformed the human-only teams in their overall team profit. As the sub-disciplines are highly coupled and team members need to communicate across these disciplines, this paper furthers previous work by taking a deeper dive into the team communication, gaining broader insight into how the AI agents reshape and improve team communication.

Results show that the integration of the AI agents into human teams enables increases in both the amount and richness of discourse content (not necessarily turns), with considerable effects exhibited

temporally as well. At the content level, AI-assisted teams also exchange more information regarding the principal design parameters and problem-solving strategies, providing evidence for increased communication effectiveness. Finally, LSA computationally models team discourse across team communication channels and shows increased semantic cohesiveness across discipline interactions. The operations discipline experiences the strongest effects on some of these measures, which affirms their significant reliance on the AI agents. The results of this work complement the prior work by the authors by uncovering detailed changes in team communication when incorporating AI agents. Overall, this work furthers the understanding of how the assistance of AI agents can reshape interactions between human team members in AI-assisted teams, which currently lacks in the literature on human-AI collaboration. Team communication, as a critical vehicle for effective information exchange in problem solving, needs to be better understood to fully harness the power of human-AI collaboration for the future. Future work by the authors addresses tracking some of these discourse measures continuously, to provide real-time updates into team problem-solving processes and recommendations for team efficiency.

ACKNOWLEDGMENTS

REFERENCES

Bird, S., Klein, E. and Loper, E., 2009. "Natural language processing with Python: analyzing text with the natural language toolkit." O'Reilly Media, Inc.

Blei, D.M., Ng, A.Y. and Jordan, M.I., 2003. "Latent dirichlet allocation." *Journal of Machine Learning Research*, 3(Jan), pp.993-1022.

Bowers, C.A., Jentsch, F., Salas, E. and Braun, C.C., 1998. "Analyzing communication sequences for team training needs assessment." *Human factors*, 40(4), pp.672-679.

Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K. and Harshman, R., 1990. "Indexing by latent semantic analysis." *Journal of the American society for information science*, 41(6), pp.391-407.

Dellermann, D., Ebel, P., Söllner, M. and Leimeister, J.M., 2019. "Hybrid intelligence." *Bus Inf Syst Eng*.

Dong, A., Hill, A.W. and Agogino, A.M., 2004. "A document analysis method for characterizing design team performance." *Journal of Mechanical Design*, 126(3), pp.378-385.

Ezer, N., Bruni, S., Cai, Y., Hepenstal, S.J., Miller, C.A. and Schmorrow, D.D., 2019, November. Trust Engineering for Human-AI Teams. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 63, No. 1, pp. 322-326). Sage CA: Los Angeles, CA: SAGE Publications.

Gorman, J.C., Foltz, P.W., Kiekel, P.A., Martin, M.J. and Cooke, N.J., 2003, October. "Evaluation of Latent Semantic Analysis-based measures of team communications content." In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 47, No. 3, pp. 424-428). Sage CA: Los Angeles, CA: SAGE Publications.

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S. and Yang, G.Z., 2019. XAI—Explainable artificial intelligence. Science Robotics, 4(37).

Gyory, J.T., Kotovsky, K. and Cagan, J., 2021. "The Influence of Process Management: Uncovering the Impact of Real-Time Managerial Interventions via a Topic Modelling Approach" *Journal of Mechanical Design*.

Hendry, D.G., 2004, August. "Communication functions and the adaptation of design representations in interdisciplinary teams." In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 123-132).

Higgins, P.G., 1999. "Job shop scheduling: Hybrid intelligent human-computer paradigm." University of Melbourne, Department of Mechanical and Manufacturing Engineering.

Hu, Y. and Taylor, M.E., A Computer-Aided Design Intelligent Tutoring System Teaching Strategic Flexibility. In: 2016 ASEE Annual Conference & Exposition Proceedings. ASEE Conferences. Epub ahead of print 2016. https://dx.doi.org/10.18260/p.27208.

Inrak, P. and Sinthupinyo, S., 2010, April. Applying latent semantic analysis to classify emotions in Thai text. In 2010 2nd International Conference on Computer Engineering and Technology (Vol. 6, pp. V6-450). IEEE.

Kane, L., Clayton, M.F., Baucom, B.R., Ellington, L. and Reblin, M., 2020. Measuring Communication Similarity Between Hospice Nurses and Cancer Caregivers Using Latent Semantic Analysis. Cancer nursing, 43(6), pp.506-513.

Kratzer, J., Leenders, R.T.A. and Van Engelen, J.M., 2010. "The social network among engineering design teams and their creativity: A case study among teams in two product development programs." *International Journal of Project Management*, 28(5), pp.428-436.

Lake, B.M., Ullman, T.D., Tenenbaum, J.B. and Gershman, S.J., 2017. "Building machines that learn and think like people." *Behavioral and brain sciences*, 40.

Landauer, T.K., Foltz, P.W. and Laham, D., 1998. "An introduction to latent semantic analysis." *Discourse processes*, 25(2-3), pp.259-284.

Liew, C., 2018. "The future of radiology augmented with artificial intelligence: a strategy for success." *European journal of radiology*, 102, pp.152-156.

Roll, I., Wiese, E.S., Long, Y., Aleven, V. and Koedinger, K.R., 2014. "Tutoring self-and co-regulation with intelligent tutoring systems to help students acquire better learning skills." *Design recommendations for intelligent tutoring systems*, 2, pp.169-182.

Sexton, J.B. and Helmreich, R.L., 2000. "Analyzing cockpit communications: the links between language, performance, error, and workload." *Human Performance in Extreme Environments*, 5(1), pp.63-68.

Song, B., Gyory, J.T., Zhang, G., Soria Zurita, N.F., Stump, G., Martin, J., Miller, S., Balon, B., Yukish, M., McComb, C., Cagan, J. "Decoding the agility of human-artificial intelligence hybrid teams in complex problem solving." *In submission*.

Song, B., Zurita, N.S., Zhang, G., Stump, G., Balon, C., Miller, S.W., Yukish, M., Cagan, J. and McComb, C., 2020, May. "Toward Hybrid Teams: A Platform to Understand Human-Computer Collaboration During the Design of Complex Engineered Systems." In *Proceedings of the Design Society: DESIGN Conference* (Vol. 1, pp. 1551-1560). Cambridge University Press.

Sonnenwald, D.H., 1996. "Communication roles that support collaboration during the design process." *Design studies*, 17(3), pp.277-301.

Stolcke, A., Ries, K., Coccaro, N., Shriberg, E., Bates, R., Jurafsky, D., Taylor, P., Martin, R., Ess-Dykema, C.V. and Meteer, M., 2000. "Dialogue act modeling for automatic tagging and recognition of conversational speech." *Computational linguistics*, 26(3), pp.339-373.

Stump, G.M., Miller, S.W., Yukish, M.A., Simpson, T.W. and Tucker, C., 2019. "Spatial Grammar-Based Recurrent Neural Network for Design Form and Behavior Optimization." *Journal of Mechanical Design*, 141(12).

Toft, Y., Howard, P. and Jorgensen, D., 2003. "Human-centered engineers—a model for holistic interdisciplinary communication and professional practice." *International journal of industrial ergonomics*, 31(3), pp.195-202.

Vrana, S.R., Vrana, D.T., Penner, L.A., Eggly, S., Slatcher, R.B. and Hagiwara, N., 2018. Latent Semantic Analysis: A new measure of patient-physician communication. Social Science & Medicine, 198, pp.22-26.

Wu, Q. and Cormican, K., 2016. "Shared leadership and team creativity: A social network analysis in engineering design teams." *Journal of technology management & innovation*, 11(2), pp.2-12.

Yang Q, Steinfeld A, Rosé C, et al., 2020. Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. https://dx.doi.org/10.1145/3313831.3376301.

Zhang, G., A. Raina, J. Cagan, C. McComb, 2021. " A Cautionary Tale about the Impact of AI on Human Design Teams." *Design Studies,* 72(2021), 100990.

Zhang, G., Soria Zurita, N.F., Song, B., McComb, C., Cagan, J., 2020. "Data on the design and operation of drones by both individuals and teams." *Data in Brief*. (Accepted)