Decoding the agility of artificial intelligence-assisted human design teams

Binyang Song, School of Engineering Design, Technology, and Professional Programs, The Pennsylvania State University, Pennsylvania, 16802, USA

Joshua T. Gyory and *Guanglu Zhang*, Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania, 15213, USA

Nicolas F. Soria Zurita, School of Engineering Design, Technology, and Professional Programs, The Pennsylvania State University, Pennsylvania, 16802, USA

Gary Stump, Jay Martin, Simon Miller, Corey Balon and *Michael Yukish*, Applied Research Laboratory, The Pennsylvania State University, Pennsylvania, 16802, USA

Christopher McComb and *Jonathan Cagan*, Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania, 15213, USA

Although necessary for complex problem solving, such as engineering design, team agility is often difficult to achieve in practice. The evolution of Artificial Intelligence (AI) affords unique opportunities for supporting team problem solving. While integrating assistive AI agents into human teams has at times improved team performance, it is still unclear if, how, and why AI affects team agility. A large-scale human experiment answers these questions, revealing that, with appropriately interfaced AIs, AI-assisted human teams enjoy improved coordination and communications, leading to better performance and adaptations to team disruptions, while devoting more effort to information handling and exploring the solution space more broadly. In sum, working with AI enables human team members to think more and act less.

Keywords: artificial intelligence, collaborative design, engineering design, human–computer interaction, complex problem solving

Corresponding authors:
Christopher McComb
ccm@cmu.edu
Jonathan Cagan
cagan@cmu.edu

With the accelerating development and confluence of technology, data, and human behavior, humankind is increasingly challenged by complex problems that involve multiple disciplines, unclear requirements, coupled parameters, and changing environments. Solving such

www.elsevier.com/locate/destud
0142-694X *Design Studies* **79** (2022) 101094
https://doi.org/10.1016/j.destud.2022.101094

problems efficiently and effectively requires efforts beyond human intelligence alone, such as creativity and intuition (Moravec, 1988; Lake et al., 2017). Although artificial intelligence (AI) has advanced far enough to be effective for task-specialized problem solving, it is still challenging for AI alone to solve such complex interdisciplinary problems purely through data analysis and powerful computing capacity (Jarrahi, 2018). Thus, a promising strategy is to fuse the complementary strengths of AI and humans by forming AI-assisted human teams that consist of humans and assistive AI agents (Kamar, 2016; Jarrahi, 2018; Dellermann et al., 2019; Moradi et al., 2019). This paper seeks to identify and understand the problem-solving process of such teams, and in particular, how AI changes team agility in problem solving.

In engineering design, AI agents have been trained to assist in multiple design phases, such as concept generation (Sarica et al., 2020), design space exploration (Law et al., 2019), and manufacturing (Williams et al., 2019). Well-trained AI agents have been shown competent for such specific design tasks (Lopez and Romary, 2019; Raina et al., 2019), but the effect of interfacing such assistive AI agents with human problem solving has not been broadly explored.

The position of this paper contrasts the recent paper by Zhang, et al. (2021), which highlights the challenges of human interaction with assistive AI agents. In that work, the AI, which itself is a high performer, is generally detrimental to the performance of the human team, especially the high-performing teams. Most importantly, that work highlights that the means for communicating the information from the AI has a critical effect on the performance of the humans, that the AI and its output needs to be cognitively easy to use and understand. Rather than providing implicit suggestions, assistive AI agents in the current paper are designed to return explicit solution options. It turns out that the assistive AI agents in this work are easy to use and interpret, which lessens the burden of the human designer and, as will be seen in the results, has a profound effect on problem solving.

In this paper, AI-assisted human teams include the presence of both humans and assistive AI agents, where the assistive AI agents assist the human in problem solving (Roll et al., 2014; Liew, 2018; Wilson & Daugherty, 2018; VanLehn et al., 2019; Zhang, et al., 2021). We refer to this as AI-assisted problem solving and adopt Pattie's (1995) definition for these AI-assistive agents as "computational systems that inhabit some complex dynamic environment, sense and act autonomously in this environment, and by doing so realize a set of goals or tasks for which they are designed." This is in contrast to assistive AI agents that collaborate as peers or guide human teammates (Demir et al., 2018; Kinne et al., 2021).

Teams as socio-technical systems consist of highly interdependent and often heterogeneous team members who share common goals and interact with each other to adapt to dynamic problem environments (Salas et al., 1995). This is true of engineering design teams. Studies have recognized teams' cognitive, motivational, and behavioral processes as influencing factors on team effectiveness (Kozlowski & Ilgen, 2006). Team structure determines the communication channels within a team and affects team performance (Wood et al., 2014; Stewart & Barrick, 2017). Previous work has used different communication patterns as abstractions of different team structures to simulate human teams using computational tools (Singh et al., 2013; Lapp et al., 2019). In the emerging area of human-AI teams, some early attempts have been made to understand the process differences between human-only teams and human-AI teams (Demir et al., 2018). These studies show that team members communicated with human teammates more often than with the AI assistant, but that the inclusion of assistive AI agents within a team stabilizes interaction speed in a changing environment more quickly (Demir et al., 2018, 2019). In addition, careful design of team structure and engaging in team building activities have been shown to have an impact on the processes of AI-assisted human teams (Walliser et al., 2019).

Despite being an important team property, the way in which agility is affected by incorporating assistive AI agents into human teams is still unclear. In particular, team agility is the ability of a team to keenly capture (if not obvious) and nimbly adapt to changing information, requirements, and strategically-relevant conditions from both inside and outside the team (Yusuf et al., 1999; Sherehiy et al., 2007; Appelbaum et al., 2017; Werder & Maedche, 2018). Agility is crucial for teams to address large, complex problems and sustain a competitive edge in dynamic environments, but also difficult to achieve in practice. This topic has attracted interdisciplinary research from varying perspectives, such as organization design (Yaghoubi et al., 2011; Appelbaum et al., 2017; Majid et al., 2018), operations management (Bernardes & Hanna, 2009; Qin & Nembhard, 2015), business (Mathiassen & Pries-Heje, 2006; Tallon, 2008; Vickery et al., 2010) and manufacturing (Gunasekaran, 1999; Sharifi & Zhang, 1999; Yusuf et al., 1999). Several team practices that boost agility have been identified, such as cooperation that makes team members stay alert to changes, constant learning that helps eliminate uncertainties, and supportive leadership that facilitates fast decision making and information exchange (Edmondson, 1999; Liu et al., 2015). Notably, compared to hierarchical team structures, flat team structures with open communication channels are considered to be beneficial for team agility (Sherehiy et al., 2007; Appelbaum et al., 2017). However, the understanding of how incorporating assistive AI agents in the team problem-solving process impacts team agility is quite limited.

This paper makes a step toward addressing that gap through a human subject experiment studying the agility of AI-assisted human teams challenged with a complex problem across several organizational structures. Complex problems often exist within dynamic environments, where teams may face two types of changes: evolving changes, which occur as interdisciplinary teams work on sub-goals and incrementally decode and communicate information and constraints; and abrupt changes, where changes in the fundamental problem or team structures can occur unexpectedly (Park et al., 2017; Rebentisch et al., 2018). To maintain or improve performance, teams must nimbly react and adapt to these changes (Sherehiy et al., 2007; Conforto et al., 2016). Accordingly, this paper models AI-assisted human teams through an experimental framework that models a complex design problem subject to evolving and abrupt changes, with the ultimate goal of answering the research question: *if*, *how*, and *why* does AI affect design team agility?

The experimental design study models a complex, open-ended drone design and operations problem. Teams are asked to maximize their profit by developing and operating a drone fleet within a given expense budget to deliver parcels routed to multiple locations of a target market. The content of the problem itself is not critical to the study; what is critical is that it is a complex problem requiring highly coupled decisions, across three very different disciplines. Among them, the design discipline is responsible for configuring, rescaling, and evaluating drones in terms of their range, velocity, payload, and cost. The operations discipline focuses on building drone fleets with currently-designed drones, generating operations plans for those fleets, and assessing them for profit. The business discipline works to determine the target market and select the final operations plan. In this way, the characteristics of the platform models evolving changes in interdisciplinary teams; team members in the three disciplines need to coordinate and work together in order to communicate the changes and evolution of their assigned sub-tasks and the overall problem. A novel open-source, online design research platform called HyForm[1] is introduced to implement the experiment and assist and record teams' design processes (Song et al., 2020) (see Figure 1 and Appendix 1 "Experimental platform"). In HyForm, text-based communication channels are available, enabling this information exchange amongst team members within and across disciplines.

Additionally, reconfiguration of communication channels, resulting in team structure changes, can be done in real-time. Team structural reconfigurations are done midway through the experiment to produce an abrupt change. Thus, the HyForm platform models both evolving and abrupt changes in complex problem solving, allowing for the study of team agility. Two assistive AI agents are also incorporated in HyForm: a design agent, which aids designers in creating drones, and an operations agent, which aids operations planners in building and operating drone fleets. Teams' access to the assistive AI agents is

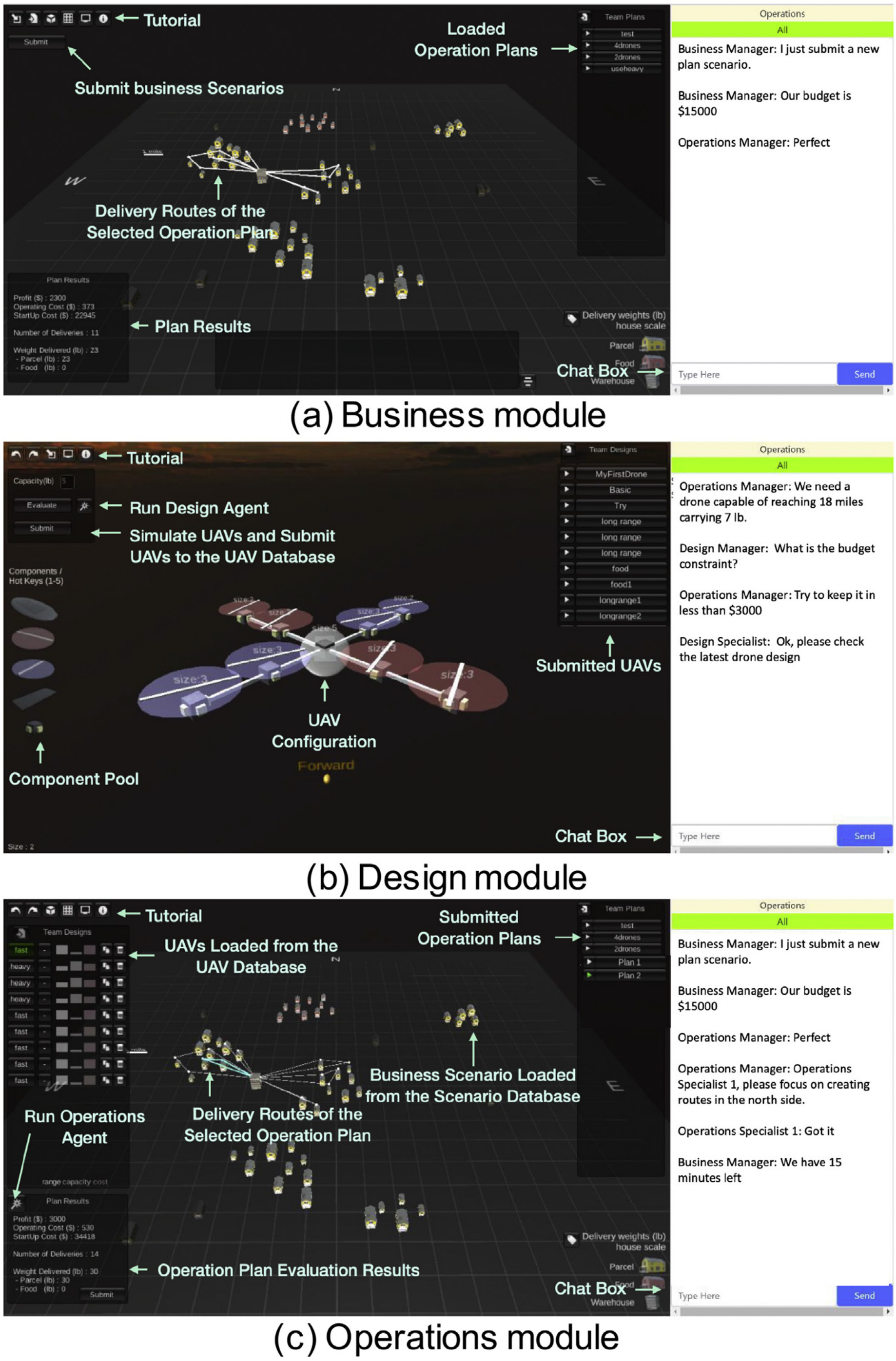


Figure 1 User interfaces of the three modules in HyForm. The drone designers and the operations planners can run the Design or Operations agents by clicking the corresponding buttons in the interfaces. The presence of these buttons is the only difference between the AI-assisted and non-AI versions of the platform

Decoding the agility of AI

controllable, enabling both human-only and AI-assisted human teams for different experimental conditions. During the experiment, every action performed through each module, every communication sent through each channel, and each generated solution is tracked by HyForm for later analysis.

This paper aims to study how evolving and abrupt changes affect team design process and performance. Evolving changes are characterized by the evolution of sub-problem requirements as teams learn more about the problem. Meanwhile, since team structure that determines communication channels within teams is important to team performance and agility according to previous literature (Appelbaum et al., 2017), we introduce immediate structural reconfigurations to induce abrupt changes to the design environment. On this basis, this paper further studies how the introduction of AI assistance into human teams impacts the influences of evolving and abrupt changes. By simulating the complex design process via HyForm, this paper shows that incorporating assistive AI agents into human teams improves team agility in evolving and abrupt changes, and explains the improvement in team agility through a comprehensive analysis on team design process.

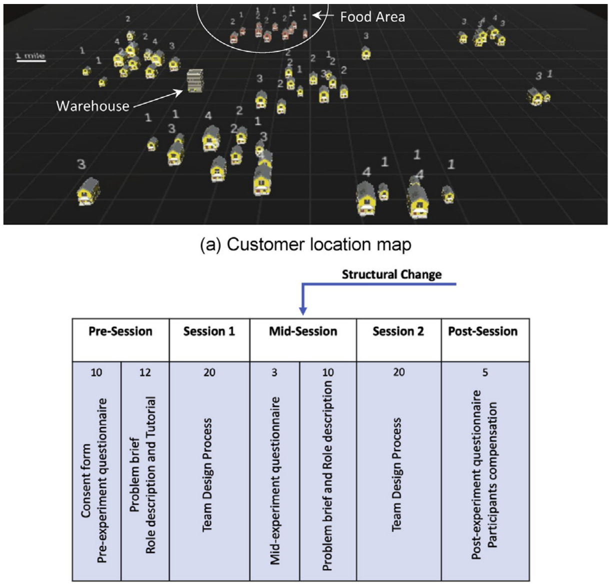
1 Human subject study overview

The main objective of this study is to reveal if, how, and why the integration of assistive AI agents into human teams affects team agility in the context of complex problem solving. The design process and outcomes of teams under distinct experimental conditions are compared.

1.1 Problem setting

A human subject experiment is run online with participants logging into the HyForm platform remotely. In the experiment, every team mimics a new company specializing in package and food delivery using drones. The company must maximize profit by designing and operating a drone fleet to succeed. The company can choose to provide a delivery service to any customers in a given area (as shown in Figure 2A). The houses represent customer locations, with a food area covering all food demands and other areas representing the package demands. The number on top of each house represents the food and package demands in the unit of pounds. As indicated in the figure, the rectangular building indicates the location of the company's warehouse, where drones pick up new packages from and must return to.

Each package order must be delivered within 24 h, and each food order must be delivered within 4−6 h from the start of the day. The company has an initial budget of \$15,000. According to the problem settings in HyForm, this budget allows the company to build and operate a fleet of a reasonable number of drones. The company receives \$100 in profit per pound of package orders delivered and \$200 in profit per pound of food orders delivered. Before the



(a) Customer location map

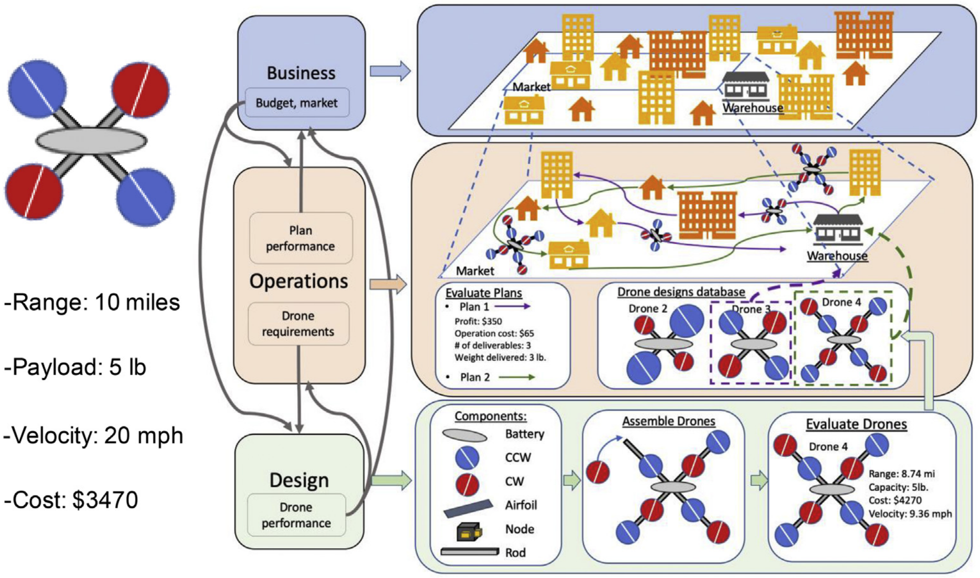
(b) Time allocation during experiment (numbers indicate duration in minutes)

Figure 2 Experimental design

experiment, each team member is assigned a specific role and discipline, accounting for drone design, operations planning, or business management. Throughout the experiment, each participant has access to a corresponding HyForm module (see more details of HyForm in Appendix 1) to complete the responsibility of their roles. Team members can communicate information and share their design outcomes through the deployed communication channels and databases.

The experiment takes approximately 1 h and 20 min, and a diagram of the time allocation is provided in Figure 2B. The experiment is broken into two equal periods of 20 min. Before each experiment session, participants have a training session to learn the problem information and the responsibility of their roles, and to go through the tutorial of the HyForm module they will use (the tutorial is accessible throughout the experiment). Team members can communicate through a text-only chat tool during the experiment, through which the communication channels are constrained based on the team structure being used. A description of the current team structure is provided to the participants before each session.

The design problem is inherently complex and achieving innovative solutions with high quality is difficult. As introduced above, each distributed team consists of the three disciplines: (1) the design discipline focuses on configuring drones, where a basic drone design is provided to users as a starting point (Figure 3A); (2) the operations discipline is responsible for operations



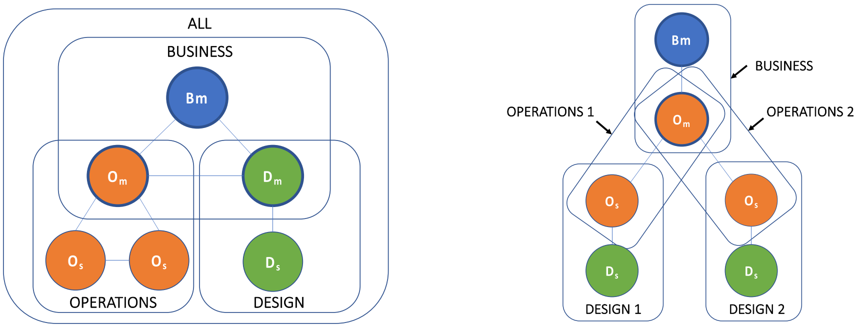
(a) Basic drone (b) Coupled relationships between the three different disciplines

Figure 3 Complex design problem

planning; (3) the business discipline works on market targeting and business management. These three disciplines are tightly coupled; the decisions made by one discipline directly influence the requirements of the others. Moreover, each discipline is privy to unique task information (e.g., the target market can be seen by the business and operations discipline, but not by the design discipline) and incrementally develop their own understanding of the coupled subproblems. As such, effective coordination is necessary to relay information, such as changing design constraints and goals, resulting in evolving changes for each discipline. Figure 3B illustrates this coupled relationship and information flow between the disciplines.

1.2 Design disciplines

Several pilot studies were conducted with different team sizes before the formal experiment. All co-authors either participated in or observed the pilot studies and most of them also interviewed the other participants, including students, professionals, and specialists from the military. This process informed the amount of effort required from each of the three disciplines (design, operations, and business) to achieve the given problem goals within two experiment sessions. Further, a team size is required that allows for meaningful reconfigurations during the study. On this basis, we determined that a team size of 6 (2 in design, 3 in operations, and 1 in business) is needed to accomplish the design task in the given time while also allowing for an exploration of team structural reconfigurations (Figure 4A,B).



(a) Open team structure including four communication channels: Business for the management team (Bm, Om, and Dm), Operations for the operations team (Om and Os), Design for the design team (Dm and Ds), and All for the whole team.

(b) Restrictive team structure including five communication channels: Business between (Bm and Om), Operations 1 and 2 (between Om and 2 Os), Design 1 and 2 (between the 2 pairs of Os and Ds).

Figure 4 Team structures. Bm-Business Manager, Om-Operations Manager, Os-Operations Specialist, Dm-Design Manager, Ds-Design Specialist. The managers are responsible for managing their disciplines and communicating their progress, while specialists focus on design-related tasks

Among all disciplines, the business discipline focuses on more cognitively-challenging tasks, such as strategic planning, which relies on extensive assimilation and synthesis of the output from the other disciplines. In contrast, the design and the operations disciplines work on more technically challenging subproblems. Accordingly, a design agent and an operations agent are incorporated and made accessible to each drone designer and operations planner, respectively, whereas no business agent is developed and instead that task remains a human meta-activity. The responsibilities of each role are described in Appendix 2 "Role responsibilities".

With regard to the two assistive AI agents, the design agent is based on a generative design algorithm using character-based long short-term memory recurrent neural networks (Char-LSTM) (Stump et al., 2019). Each drone design is represented using a string grammar (Stump et al., 2019). The string grammar defines all drone features, such as the configuration of the two-dimensional layout and the component type and size at each position. The Char-LSTM is initially trained on a set of random drone designs to generate new designs. The generated designs are evaluated through a multi-physics simulation engine, and only those with good performance enter the next iteration to re-train the Char-LSTM. In each iteration, the drone set is updated by including newly generated higher-performance designs to replace the lower-performance designs from previous iterations. After training, the algorithm can output a large set of drone designs with good performance as a pre-constructed solution space. Although the drone database is pre-constructed, it can be updated according to user preference collected from one experiment for the next experiment using generative design. When a designer clicks the AI button to invoke the design agent, the agent searches the solution space near the latest drone generated by the designer and returns a few drone options, each of which aims to optimize one of the four quality metrics (range, payload, velocity, and cost). The presentation of the design is explicit and complete, unlike in the Zhang, et al. (2021) work, where the AI output needed to be interpreted from a multi-hued color map. The designers can either select one drone from the options to work on or ignore them all to go back to their own designs.

The operations agent is based on a linear programming algorithm. The algorithm aims to find the optimal flying paths by maximizing the profit function within the solution space defined by the linear equations that indicate the various constraints, such as the available market, delivery time, and the carrying payload and flying range of the drone fleet. Explicit delivery paths are articulated. A participant can invoke their operations agent by clicking a button in their task interface. After the operations agent returns an operations plan, the planner can make further changes to it for budget control or other reasons. The participant decides when and how often to run the agents.

The AI agents in this study are considered assistants for human users for two reasons. Firstly, human users need to make crucial decisions before invoking the assistive agents so that the agents can provide idiosyncratic responses to each user. For example, a drone designer needs to specify a valid drone, which the design agent refers to as a starting point to search the surrounding solution space and returns optimized drone designs. An operations planner also needs to define a drone fleet (i.e., a group of promising drones selected from all available drones) as input to the operations agent. Secondly, according to the recorded data, many of the submitted solutions are not directly generated by the assistive AI agents. That means the team members are still actively engaged in generating their own solutions or modifying the AI returned solutions to fulfill their team-specific design requirements.

1.3 Participants

Participants complete the experiment online. The experiment is approved by the Carnegie Mellon University Institutional Review Board (IRB). The recruitment of participants for the experiment uses a registration link distributed to undergraduate and graduate students in the engineering schools at 18 universities in the United States. A total of 395 participants are recruited. The participants are randomly grouped into 54 teams of 6 members and 71 individual drone designers. Participants are assigned as individual drone designers in cases where whole teams could not be formed, generating data not relevant to this paper. We removed 10 teams from data analysis due to several technical issues: losing connection to the interface because of poor internet or other issues (6 teams), two team members logging in with the same username resulting in problematic log data (3 teams), or certain communication channels not working (1 team). Therefore, we obtain high-quality data for 44 teams (i.e., 264 participants) successfully. The data collected from the individual drone designers is explored in another study (Song et al., 2022) and not included in this paper. An analysis of participant demographics shows that among the 264 participants, 29.2% identify as female, 68.9% identify as male, and 1.9% prefer not to say. Participant age ranges from 18 to 48, with 81.2% being 24 years old or below. In terms of design experience, 85.2% and 79.6% of the participants report limited experience with drone design and drone operations, respectively; 66.7% of the participants report moderate or more experience with computer-aided design. Informed consent is obtained from all participants prior to participating in any aspect of the study. Each participant receives a payment of $20 after completing the experiment.

1.4 Experimental conditions

During the experiment, half of the teams are given access to the AI assistance, whereas the other half are not. They are referred to as AI-assisted human teams and human-only teams, respectively. Two distinct team structures are employed with different communication channel configurations: an open structure where all team members can communicate directly with each other, and a restrictive structure where only particular members can communicate with one another through explicit channels (Figure 4A,B, respectively). These two team structures resonate with the flat and hierarchical organization structures in practice, respectively. As described in Figure 2B, the experiment comprises two 20-min sessions to solve the problem. Each team undergoes a structural reconfiguration (i.e., a shifting of team structure from one to the other) during the transition between sessions to create an abrupt change. The transition from the open team structure to the restrictive team structure is referred to as a restricting reconfiguration (the degree of communication reduces), while the transition from the restrictive structure to the open structure is called an opening reconfiguration (the degree of communication increases). During these reconfigurations, team members maintain their assigned discipline. Figure 5 presents the four team conditions, following a 2 × 2 full-factorial experimental design. Data is obtained from 44 teams, resulting in 11 teams per condition.

This study focuses on the influence of AI assistance on team agility in evolving and abrupt team structure changes. The participants go through a training process prior to the start of the experiment, which equips them with a common skill basis to take on a specific role and contextualizes the team settings. The training process for AI-assisted team members only differs from that of the human-only team members by including an introduction to the corresponding assistive AI agents. The participants have sufficient time to review and assimilate the information, preventing performance bias in favor of either team. Moreover, an experimenter is available for assistance in case that the

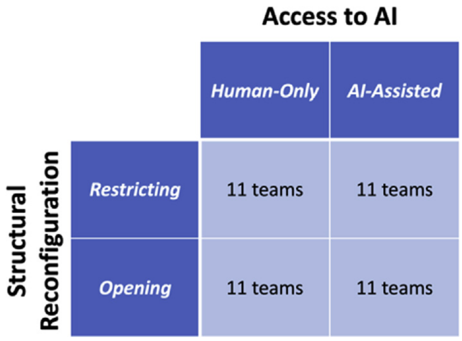


Figure 5 Scheme of experimental conditions. Two distinct team structures through restricting and opening reconfigurations, with and without the assistive AI agents

participants run into any questions about the training documents or technical issues with the interfaces during the entire experiment. Note that the experimenter only answers questions related to HyForm or problem settings and does not take the initiative to intervene with teams or provide any design or problem-solving insights. All members are randomly assigned to either AI-assisted human teams or human-only teams. These two groups of participants exhibit no difference in their demographics, experience, and educational background according to the assessment of data from a pre-study questionnaire. Given these measures, the comparison between conditions across a 2×2 experimental design allows for the comparison in team agility.

2 Evaluation of team performance and design process

2.1 Team performance

The performance of a team is measured at two levels: the team level and the discipline level. The team-level performance should reflect solution quality achieved by the team as a whole. According to the problem setting, the highest profit made by a team is a uniform measure that assesses the ability of the coupled disciplines, respectively focusing on drone design, operations planning, and market targeting, to solve the complex problem together. Therefore, the team-level performance is evaluated through each team's highest profit. At the discipline level, both the number and quality of solutions are considered to assess team performance. Since it is difficult to assess the quality of market targeting for the business discipline without considering how the target is fulfilled through design and operations, we only measure the discipline-level performance for the design and operations disciplines.

Specifically, the quality of a drone design is calculated according to its performance metrics. In HyForm, the performance of a valid drone is evaluated via four metrics: cost, range, payload, and velocity. Herein, a valid drone is one that can generate a balanced thrust enabling the drone to stay stable in the air and move forward. In order to make more profit (i.e., deliver more parcels with given time and budget constraints), drone designers need to decrease drone cost but increase drone flying range, velocity, and payload and balance these three metrics according to their target market properly. In general, drone designers aim to increase drone flying ranges, payloads, and velocities but decrease costs as much as possible. A utility function reflecting such design objectives is built to calculate the overall drone quality. We give the same weight to these four metrics. Equation (1) shows the function:

$$Quality_0 = \frac{Range \cdot Velocity \cdot Payload}{Cost}. \tag{1}$$

Then, it is normalized by Equation (2):

$$Quality = \frac{Quality_0 - \min(Quality_0)}{\max(Quality_0) - \min(Quality_0)}. \tag{2}$$

The average drone quality of each team measures team performance in the design discipline. This metric provides one way to measure the overall drone quality, and other metrics can be considered when users want to emphasize different aspects of drones.

We measure the quality of an operations plan as the profit made by the plan. In HyForm, the plan profit is calculated according to the weight of packages and food delivered by an operations plan, as shown by Equation (3):

$$Profit = \$100 \cdot Mass\ of\ package + \$200 \cdot Mass\ of\ food. \tag{3}$$

Herein, \$100 and \$200 are the profit made by a company for delivering 1 pound of package orders and food orders, respectively. The corresponding constraint is that the company has a budget of \$15,000 to design and operate its drone fleet.

2.2 Hidden Markov Model

To investigate how integrating the assistive AI agents reshapes teams' design process, the design activity data, including both action and communication records, is analyzed using a Hidden Markov Model (HMM) to simulate the aggregate design process. In this study, the design activities of each team is treated as a sequence of design actions and communications (i.e., communications are seen as a type of design activity for information exchange). The whole data set consists of 44 sequences of design activities (one for each team). A Hidden Markov Model (HMM) is a statistical model to capture hidden temporal patterns from sequential observations (e.g., in this work, team activities) (Baum et al., 1970; McComb et al., 2017). The HMM models design activities as a Markov process transitioning between a finite number of discrete states hidden from the observer with unknown parameters. The training process of the HMM determines the hidden parameters, the transition matrix, and the emission matrix, from the observable parameters. Specifically, the transition matrix has a size of $[m \times m]$, where m is the number of hidden states, containing the probability of transitioning from the current state to a future state; the emission matrix has a size of $[m \times n]$, where n is the number of unique actions, containing the probability of an action being emitted from a given state. Through the training process, the obtained hidden states represent the underlying cognitive or procedural states that the teams transition through during the experiment (McComb et al., 2017).

The Baum-Welch algorithm (Baum et al., 1970) is employed to train the HMM by maximizing the likelihood of the observations. Since the number of hidden states to use for modeling the aggregate design process is unknown, several models with the values of m varying from 1 to 15 (the number of unique actions) are trained and compared for selecting the best model. Higher values of m are not necessary for maintaining the independence of the emission probabilities of the states. For any m values within the given range, models are trained 44 times each with 43 sequences; leaving one sequence out as the testing sample to increase the generalizability of the results. The average testing log-likelihood (i.e., an indicator of the model's ability to describe the test sample) is calculated over the varying m values. Then, we identify the best model by selecting the lowest value of m for which the test log-loglikelihood is not significantly lower than the highest test log-likelihood at 5% level of significance. In such a way, the best model balances between model compactness and model accuracy to avoid overfitting. The captured hidden states are interpreted as design states in this study.

2.3 Latent semantic analysis

Latent Semantic Analysis (LSA) (Landauer et al., 1998) is utilized to compare the content of the team communications across the different communication channels. Specifically, we compare the channels between the AI-assisted human teams and the human-only teams to identify any differences in communication cohesion. LSA uses singular value decomposition (SVD) to reduce the dimensionality of the semantic space and compares the frequency of terms in the channels across the entire text corpus. With this approach, the channels can be represented as vectors of word frequencies, and the cosine similarity between document vectors computes how semantically close or semantically distant the communication channels are. The SVD of a matrix is defined in Equation (4) as:

$$X = USV^T, \tag{4}$$

where X is an $[p \times q]$ occurrence matrix with p number of words and q channel documents, U is an $[q \times r]$ concept vector matrix with rank r, S is an $[r \times r]$ singular values matrix, and V is an $[p \times r]$ channel matrix.

In the LSA model, each category of channels is treated as a distinct document. Accordingly, in the restrictive team structure, the 'Operations 1' and 'Operations 2' channels are combined into an *Operations* document while the 'Design 1' and 'Design 2' channels are combined into a *Design* document. The 'All' channel is unique to the open team structure. The figures (Figure 1B,C) of the team structures show which team members can communicate in the different channels. Thus, in the LSA model, there are four different document types: *All*, *Business*, *Designer*, and *Operations*. The chats sent by all the respective participants in a particular channel are included in a document; that is, the communication data in the documents is at the team level, not at the individual level.

In addition to the combination of channels for the different disciplines, the communication data goes through a few pre-processing steps. The documents are first tokenized, meaning that each document is represented as a vector of word frequencies. Stop words, specifically, the stop words identified in the Natural Language Toolkit (NLTK) (Bird et al., 2009), are removed, as well as punctuation. Infrequent words (those with frequencies less than 2) as well as short words (those less than two characters) are also removed to eliminate noise in the data. Words are then stemmed and lemmatized – the removal of prefixes and suffixes so that all words remain in the same tense and return to their dictionary forms.

3 Results and analysis[2]

3.1 Team agility

Agility represents the property of a system to implement necessary changes rapidly, measuring how efficiently and effectively the system can adapt to changes through a robust process (Dove, 1995; Schulz & Fricke, 1999). Team agility is challenging to measure directly and often defined and assessed case-by-case. In this paper, we use the rate of change of team performance over time as a proxy of team agility with respect to evolving changes. Herein, team performance is measured by the profit of the best operations plan found so far. We measure team agility with respect to abrupt team structure changes through action and communication change ratios. These change ratios are computed as the ratios of action or communication counts after and before the structural reconfigurations; lower ratios indicate less sensitivity to the structural change and thus higher agility, while higher ratios indicate lower agility. Team agility is compared between the AI-assisted human teams and the human-only teams. Overall, the results of this work show that team agility is significantly improved in AI-assisted human teams: AI-assisted human teams communicate more but take fewer actions while achieving better performance than human-only teams. In other words, AI-assisted human teams achieve better coordination, i.e., communicating and responding to teammates rapidly, while acting more effectively and efficiently, i.e., correctly fulfilling their individual tasks more rapidly.

3.1.1 Agility with evolving changes

Results show that the AI-assisted human teams exhibit greater agility than the human-only teams when facing evolving requirements, as measured by how rapidly the teams can improve their performance as they acquire more knowledge of the problem. In this case, evolving requirements result from team members' learning as they actively explore the solution space; team performance is assessed as the highest profit achieved by a team over time. As shown in Figure 6A, both teams start with similar team profits. However, the AI-assisted human teams improve their profits more rapidly over the experiment, as indicated by the steeper ($p < 0.001$) slope of the fitting line for team profit.[3] The higher rate of change of team performance indicates that the AI-assisted human teams are more responsive to the evolving requirements. As a result of their agility, the AI-assisted human teams perform better than the human-only teams. With the AI assistance, the AI-assisted human teams achieve significantly higher final performance than the human-only teams on average ($p < 0.001$, Figure 6B). Significant performance improvements are also observed at the discipline level in terms of drone quality ($p < 0.001$, Figure 6C) and plan profit ($p < 0.001$, Figure 6D), although the AI-assisted human teams and the human-only teams achieve a similar number of drone designs and operations plans ($p > 0.05$, Figure 6E,F).

3.1.2 Agility with abrupt changes

The AI assistance also improves team agility in response to abrupt structural changes. The open team structure and the restrictive team structure highlight the different communication channels, which directly affect team communication and may further affect team actions. We consider teams agile if the affected team members can adapt to changes at the discipline level, but the teams are less disrupted by the structural configurations at the team level. On the one hand, the AI assistance allows for necessary adaptions to the structural reconfigurations at the discipline level. When teams face the restricting reconfiguration, the AI assistance enables the operations planners to increase their communication count ($p = 0.005$, Figure 7A) with a similar action count (Figure 7B) in response to an improved need for passing information between the design and business disciplines. When teams face the opening reconfiguration, the AI assistance allows the drone designers to increase their communication count ($p = 0.029$, Figure 7C) in response to increased information sources while performing an unaffected number of actions (Figure 7D). On

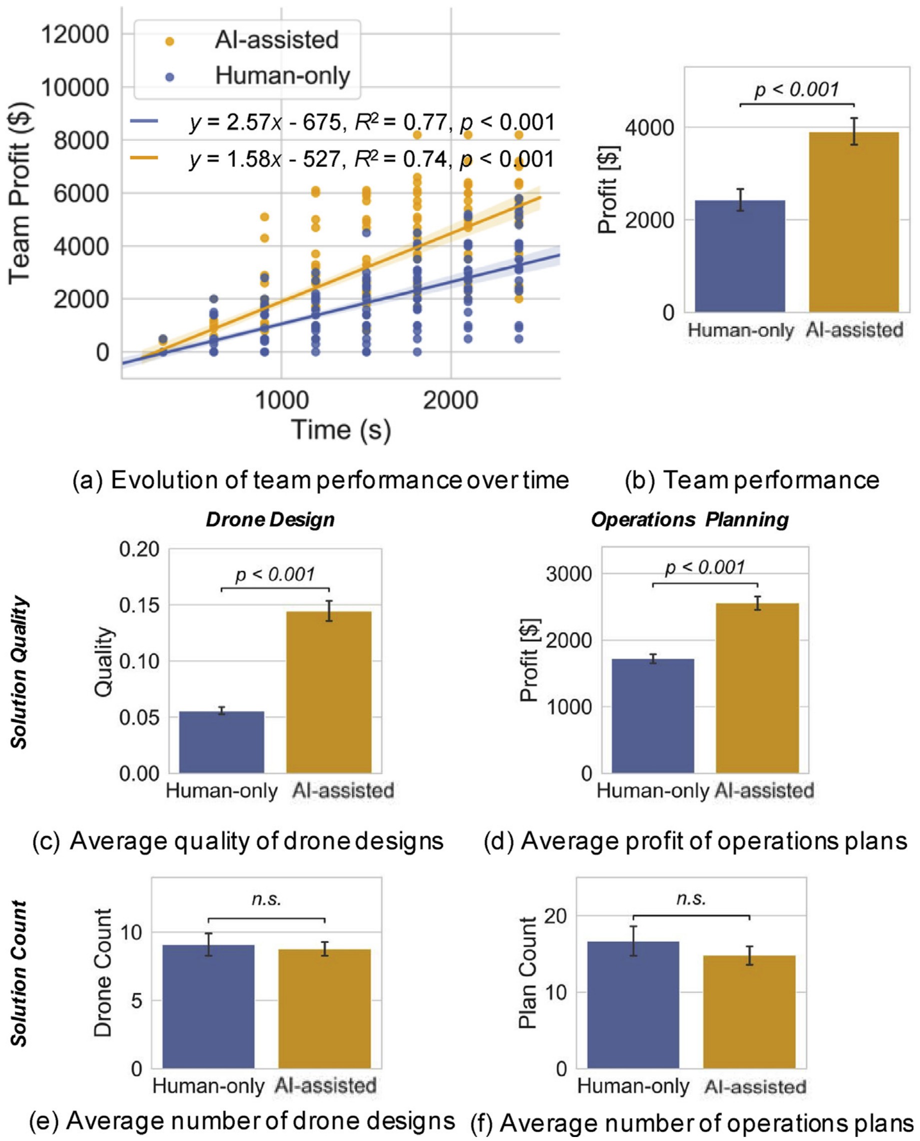
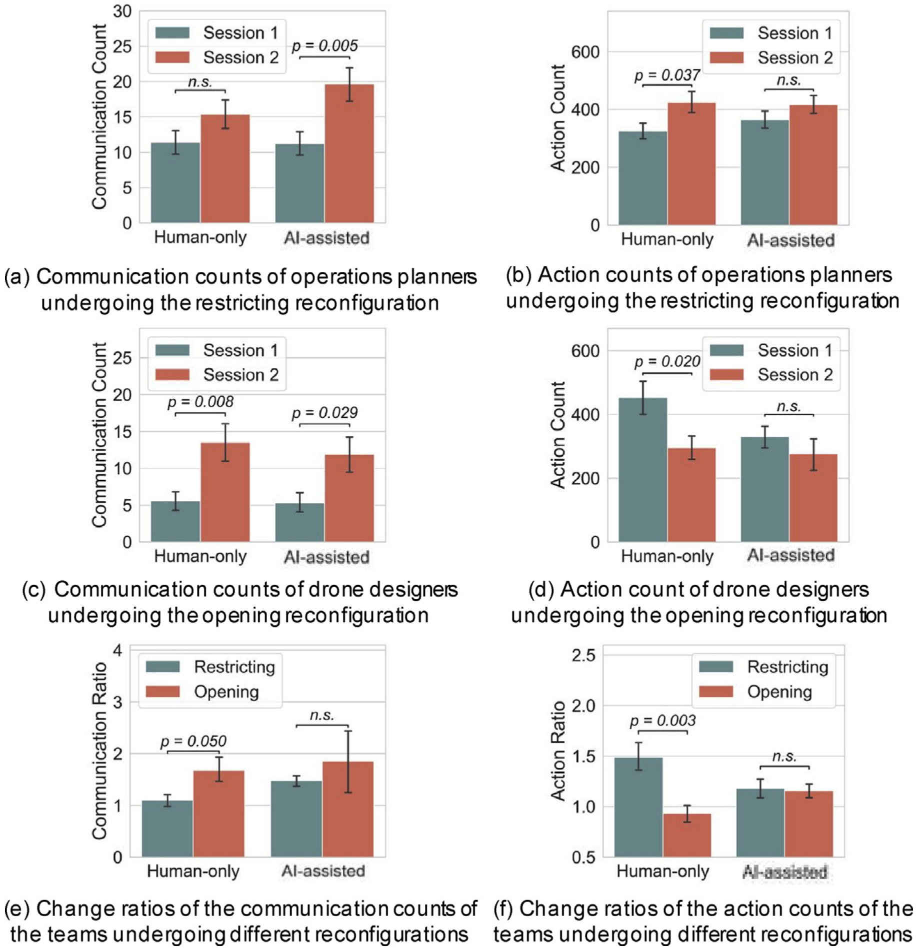


Figure 6 Comparison in performance between the human-only and AI-assisted human teams. Error bars show ±1 S.E. for the bar charts. P-values > 0.05 are annotated as "not significant" (n.s.)

the other hand, the AI-assisted human teams are less sensitive to distinct team structure reconfigurations at the team level, as depicted in Figure 7E,F by the action and communication change ratios. These ratios are derived by dividing the action/communication count after a reconfiguration by that before the reconfiguration. The AI-assisted human teams see no significant differences in change ratios, whereas the human-only teams exhibit significant differences in both action and communication change ratios between the opening and restricting reconfigurations.

Decoding the agility of AI



(a) Communication counts of operations planners undergoing the restricting reconfiguration

(b) Action counts of operations planners undergoing the restricting reconfiguration

(c) Communication counts of drone designers undergoing the opening reconfiguration

(d) Action count of drone designers undergoing the opening reconfiguration

(e) Change ratios of the communication counts of the teams undergoing different reconfigurations

(f) Change ratios of the action counts of the teams undergoing different reconfigurations

Figure 7 Comparison in action and communication between the human-only and AI-assisted human teams. Error bars show ±1 S.E. for the bar charts. P-values > 0.05 are annotated as "not significant" (n.s.)

Therefore, when facing structural changes, the affected team members in the AI-assisted teams can make more effective adaptions than those in the human-only teams at the discipline level, and the AI-assisted teams are less disrupted by different structural reconfigurations at the team level. In contrast, the less effective adaptions at the discipline level and the larger disruptions at the team level exhibited by the human-only teams may imply that a larger part of the behavioral changes in human-only teams is caused by disorders instead of necessary adaptions to the structural changes.

3.2 Design process

3.2.1 Design action and communication

We conduct several in-depth analyses to study how the AI assistance reshapes teams' design processes and understand why AI assistance improves team agility. First, the AI assistance improves the competence of individual team members, indicated by better performance and higher action effectiveness. As illustrated by the regressions shown in Figure 8A,B, the AI-assisted teams obtain more effective drones over the experiment, indicated by the higher intercept ($p < 0.001$) of the fitting line for drone quality; they also evolve their operations plans more efficiently, indicated through the steeper ($p < 0.001$) slope of the fitting line for plan profit. Meanwhile, the comparison of the action count and the communication count (Figure 8C,D) shows that the designers in the AI-assisted human teams perform fewer actions ($p = 0.031$, Figure 8D) compared to the human-only teams on average. Together, the results show that the AI-assisted human teams exhibit higher action effectiveness. Further, the improved individual competence mitigates the tradeoff between team actions and communications. The regression analysis in Figure 8E shows that integrating the assistive AI agents with humans also tempers the tradeoff between actions and communications. There is a negative correlation between the action count and communication count for the human-only teams ($p = 0.002$, Figure 8E), while such correlation is not significant for the AI-assisted human teams ($p > 0.05$). The insignificant correlation indicates that AI-assisted human teams do not need to sacrifice their communications for more actions or vice versa.

The AI-assisted human teams also allocate more design activity resources to exchange and manage information, achieving more cohesive and intensive coordination than the human-only teams. An analysis on team activity data using a Hidden Markov Model (Baum et al., 1970; McComb et al., 2017) illustrates the aggregate design process (Figure 9), resulting in five hidden states associated with path planning, drone configuration design, information handling, drone parameter design, and scenario generation, respectively. Figure 10 shows the percentage of team activities within each hidden state in 10-min periods. The AI-assisted human teams allocate more activities toward information handling and less toward drone parameter design. This change allows the AI-assisted human teams to sense and communicate the evolving situation in time.

3.2.2 Team coordination

In terms of team coordination, a latent semantic analysis measures the cohesion of teams' communications between each pair of the distinct communication channels. Herein, cohesion indicates the extent to which the two sets of communications from two channels are semantically related to each other and facilitates a unified knowledge basis for the team (McCarthy et al., 2007). The semantic distance matrices (Figure 11) illustrate that the distance value in each cell of the matrix for the AI-assisted human teams is smaller (i.e., higher cohesion) than its corresponding value for the human-only teams,

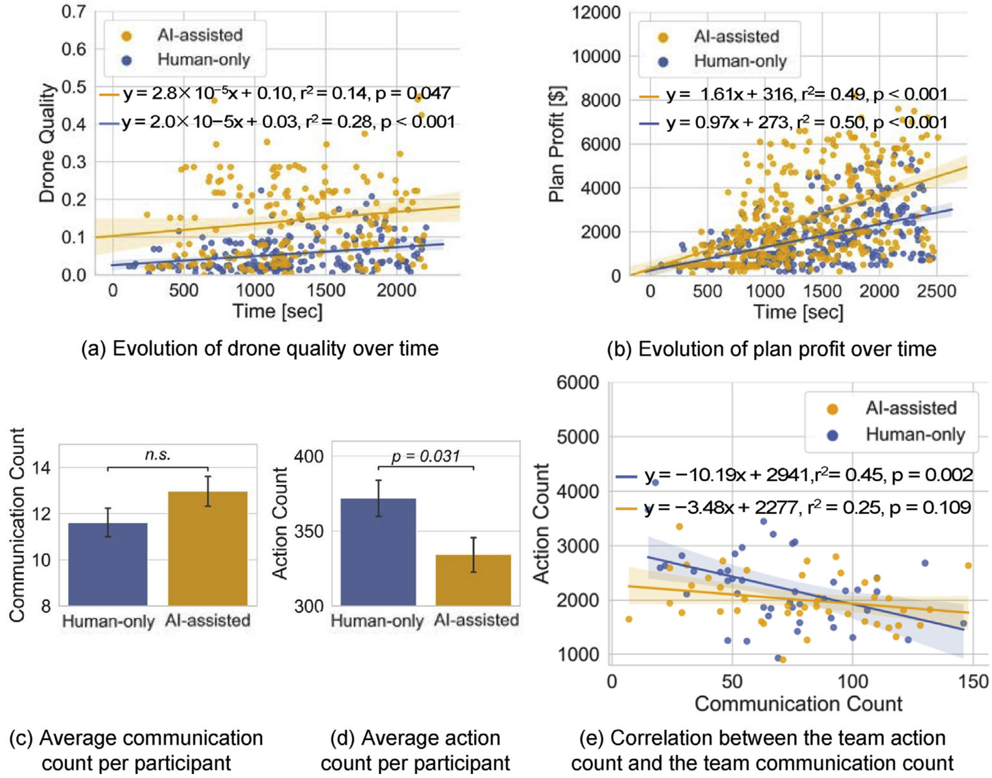


Figure 8 Comparison in design performance evolution and design action and communication between the human-only and AI-assisted human teams. Error bars show ±1 S.E. for the bar charts. P-values > 0.05 are annotated as "not significant" (n.s.)

indicating a more coherent understanding of the problem space within the AI-assisted human teams (Dong et al., 2004; Gorman et al., 2003).

Moreover, the content of team communications is analyzed. Based on insights obtained by reading the team communications, a set of keywords (see Table 1) are identified which capture information regarding the primary parameters of the problem-solving or the team strategies associated with the primary parameters. Team communications containing the corresponding keywords are identified as team coordination directly relevant to problem solving, separating them from communications regarding experiment logistics or technical issues related to the platform. Figure 12 suggests that the AI assistance enables more information sharing and assimilation directly related to problem solving for the AI-assisted teams, especially in the second experiment session.

3.2.3 Team exploration

Regarding team exploration, the AI-assisted human teams search the solution space more broadly in terms of both drone design and operations planning compared to the human-only teams. Figure 13A,B depict the aggregate spatial

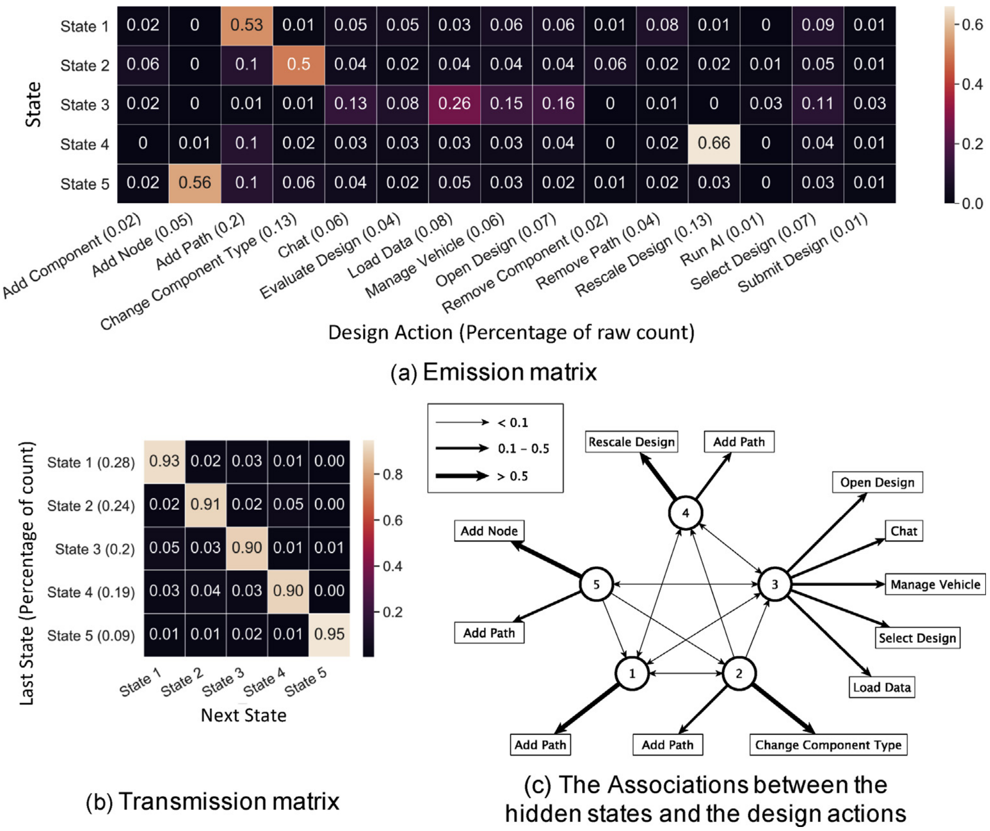


Figure 9 Hidden Markov Model describing the aggregate design process. The numbers in the parentheses in (a) and (b) indicate the percentages of the design actions or the percentages of design actions corresponding to specific states in the data set. Only associations with probabilities higher than 0.05 are shown in (c)

configurations of the drones designed by the human-only teams and the AI-assisted human teams. The size (i.e., diameter or width) of a shape indicates the total number of drone designs that deploy a component at that position. Figure 13A shows that the human-only teams concentrate on minor modifications from the initial drone design given to participants at the beginning of the experiment. In contrast, the AI-assisted human teams design more complex and often better-performing drones by adding components at distant locations. Incorporating airfoils adds another dimension for drone design space exploration since the basic drone design does not include airfoils. Figure 13B reveals that the AI-assisted human teams employ airfoils more frequently and add airfoils in more locations than the human-only teams.

For the operations planners, the aggregate network layout of the operations plans designed by the human-only teams and the AI-assisted human teams appear in Figure 14A,B. The location, size, and color intensity of each node represent a customer location, parcel weight at the location, and the frequency

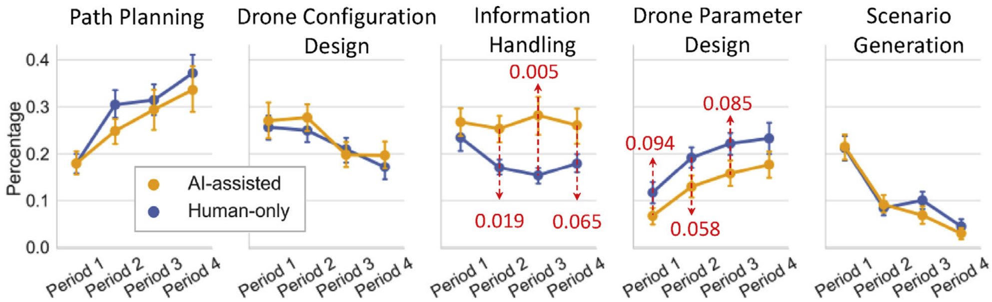


Figure 10 Comparison in the percentage of design activities within each hidden state between the human-only and AI-assisted human teams. The differences between the human-only and AI-assisted teams with p-values < 0.1 are shown in the plots

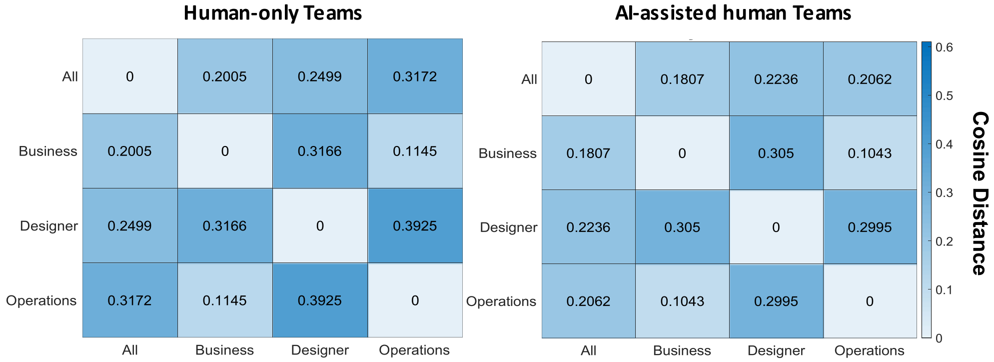


Figure 11 Semantic distances between conversations in the distinct communication channels. The value in each cell indicates the semantic distance between conversations in each pair of communication channels

Table 1 List of relevant keywords for identifying team coordination directly relevant to problem solving

Keyword Category	*List of keywords*	
Primary Parameter	range/distance	mile, mi
	payload/weight	pound, mass, lb, load, volume, capacity
	velocity	speed, mph
	cost/budget/profit	money, price, dollar, $, how much
	parcel type	food, package
	customer	consumer, people, client, house, area, deliver, warehouse
	time	hour, hr
Strategic Keyword	increase, reduce, sacrifice, shorten, cut, maximize, minimize, add, decrease, change, balance, focus, prioritize, optimize, combine, improve, cheap, expansive, heavy, high, large, long, low, quick, short, small, far, fast, big	

of the parcel being delivered by all AI-assisted or human-only teams, respectively. The width of each link between two nodes denotes the frequency of the two parcels being delivered in one delivery plan. Together, Figure 14A,B show that the AI-assisted human teams search more of the operations design

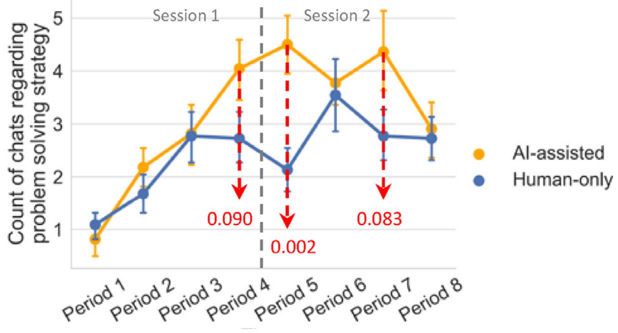


Figure 12 Evolution of team coordination directly relevant to problem solving over time. The differences between the human-only and AI-assisted teams with p-values < 0.1 are shown in the plots

space by serving more customers than the human-only teams. Notably, the AI-assisted human teams reach customers farther away from the warehouse with a higher frequency.

4 *Discussion*

This paper investigates if, how, and why integrating assistive AI agents into human design teams affects team agility. Through a human subject experiment, the results first answer *if* and *how*: the AI assistance affects team agility; in fact, the AI-assisted human teams exhibit significantly greater agility than human-only teams when dealing with both evolving and abrupt changes. Specifically, the AI-assisted human teams embody effective interdisciplinary collaboration, improving the quality of solutions in response to the evolving changes. In responding to abrupt changes, AI-assisted human teams are less sensitive to different types of structural changes at the team level. Furthermore, the AI-assisted human teams experience more beneficial adaptions at the discipline level, regardless of which type of reconfiguration they are faced with.

In-depth analyses into the design process of the teams help understand *why* the AI-assisted human teams exhibit this improved agility. The results reveal three factors:

1) The integration of the assistive AI agents enables human team members to think more but act less. The AI-assisted human teams perform fewer actions than the human-only teams in the experiment. The AI-assisted human teams submit similar numbers of drone designs and delivery plans as the human-only teams, but the average qualities of the drones and the operations plans are significantly higher. These results show that the AI assistance helps human designers reallocate their effort from time-intensive design activities to high-value cognitive activities, such as

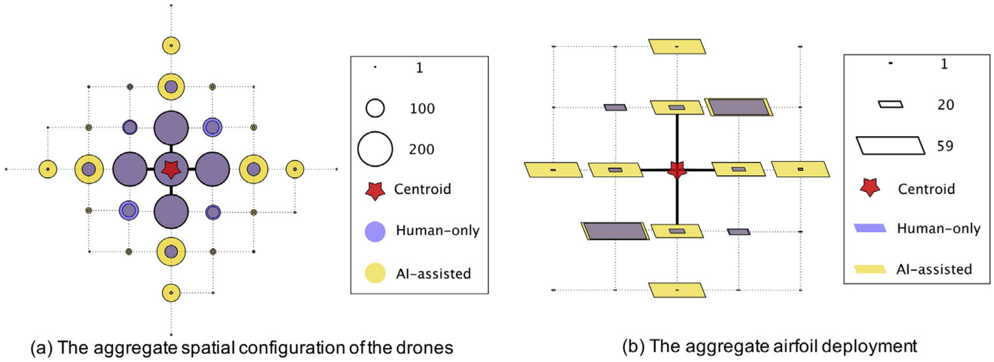


Figure 13 Drone design space exploration of the human-only and AI-assisted human teams. The solid black lines indicate the configuration of the basic drone. The size (i.e., diameter or width) of the shape indicates the total number of drone designs that deploy a component at that position. At a same position, the blue shape and the yellow shape respectively stand for the drones designed by the human-only teams and the AI-assisted teams

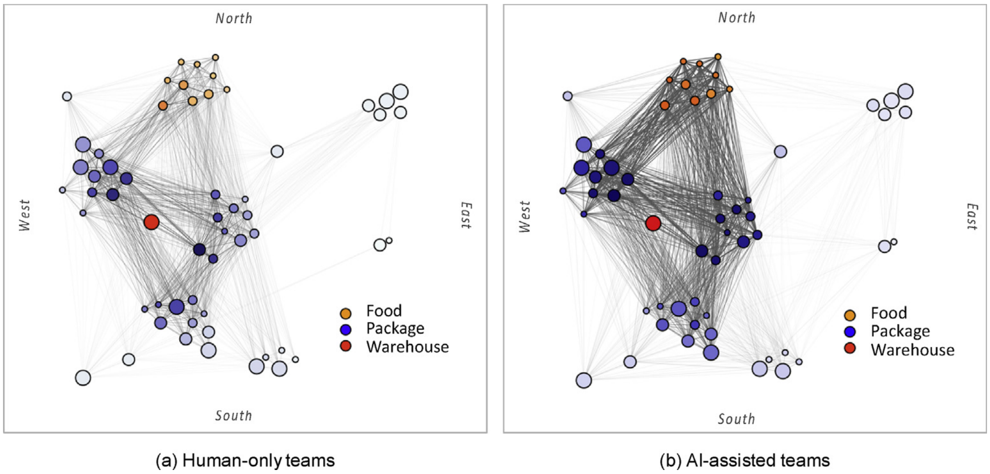


Figure 14 The aggregate network layouts of the operations plans of the human-only and AI-assisted human teams

decision making and coordinating with their teammates about how to solve the problem. These findings are in line with previous studies on agility that focused on human-only teams, showing that improved individual competence and flexibility in problem solving are beneficial for increased team agility (Fliedner & Vokurka, 1997; Yusuf et al., 1999; Vickery et al., 2010).

2) The AI-assisted human teams have higher quality communication. In the experiment, the AI-assisted human teams devote more effort to information handling. The communication between team members in the AI-assisted human teams is also more cohesive than the human-only teams. Meanwhile, the AI-assisted human teams conduct more intensive

problem-related coordination compared to the human-only teams. These results indicate that high-quality communication between disciplines may be a critical factor in determining the performance of AI-assisted human teams. Further, this aligns with literature showing that better information and knowledge management across disciplines within teams boosts team performance and agility in complex problem solving (den Otter & Emmitt, 2008; Senescu et al., 2013).

3) With the assistance of the AI agents, the AI-assisted human teams can explore the solution space more broadly. In general, an important factor that may hinder the ability of human-only teams to adapt to changing environments and lead to inefficiencies in problem solving is cognitive fixation (Jansson & Smith, 1991; Moss et al., 2007; Sio et al., 2015). The drone design space exploration analysis indicates that the human-only teams fixate on the basic drone design provided to them at the beginning of the experiment. In contrast, the design agent helps participants in the AI-assisted human teams overcome such design fixation and explore more broadly in the drone design space. Further, a lack of exploration is also found in operations planning in human-only teams. The operations agent assists the AI-assisted human teams in the time-consuming operations planning task, enabling participants in these teams to invest more effort towards delivering the food parcels with a higher profit rate and combining parcels from distinct areas into single operations plans. This broader exploration in both drone and operations design spaces equips the AI-assisted human teams with a more extensive set of alternative solutions when the teams face changes during the design process.

While this work is presented in the context of a drone design and operations problem, the resulting insights shed light on complex problem solving more generally. Complex problems occur frequently in domains that are subject to evolving and abrupt changes, such as planning, strategy, and design. For such problems, AI-assisted human teams may prove to be an effective and efficient means of achieving innovative outcomes. In such teams, the AI assistance leads to more intricate solution alternatives, while human team members can coordinate and steer team efforts. Ultimately this boosts team agility, resulting in more nimble and effective complex problem solving in dynamic environments.

The findings show that the design and operations agents studied in this paper are beneficial. However, although this may appear obvious or even unfair to the human-only teams, assistive AI agents are not always beneficial in design tasks which require creative search and insight, characteristic of many problems in the design domain. As reported in the Zhang, Raina, et al. (2021) study, an assistive AI agent, which on its own can generate high-performance solutions to a design problem, boosts the initial performance of low-performing teams when it is used as an assistant. However, those teams then fail to perform as well as the equivalent human-only teams after that initial burst; the assistive AI agent also always hurts the performance of high-performing teams. Therefore, assistive AI agents with distinct characteristics may result in varying effectiveness. Future research might provide an a priori assessment of if, how, and why a specific assistive AI agent is helpful in a given problem context. In the work of Zhang, Raina, et al. (2021), the interpretation of the AI output is cognitively and visually challenging, which likely led to less effective use of the technology by the design participants. It is not the case in the present work, where the AI agents provide clear suggestion articulation to the human participants. The findings from this paper help explain why the AI-assisted teams perform better and, more precisely, how adaptable the teams are to shocks or changes in the design process. The understanding can inform future assistive AI agent development.

While the findings reported in this paper are promising, this study is subject to a few limitations:

1) Control groups that do not experience structural changes are not included in the experimental design. Future studies will include such control groups to investigate how the structural changes affect the impact of AI.
2) To better understand the effectiveness of AI versus other types of assistance, we need to design and conduct experiments that explore how an assistive AI agent compares to a human expert as an aid to the team.
3) In this study, the complex design process is modeled through a drone design and operations experiment with a short time frame, so the findings presented in this paper still need to be verified in solving problems under broader and longer timeframe contexts. Further, team members in the current study are anonymous to each other, while future work might study how known team members impact the findings.
4) This study investigates two dimensions of team changes, while teams also face other changes in practice, such as external market changes, which can be investigated with respect to AI assistance in the future.
5) Feedback from participants on their experience in working with AI, which might impact their performance, is not collected during the experiment. Future studies will collect this information.
6) Since participants in this experiment are novice designers, the findings may not apply to experienced designers with rich domain-specific expertise.
7) Based on the current analyses of team performance, it is difficult to differentiate the direct contributions of the AI agents from the effect of symbiotic collaborations between the AI agents and the human designers. This is an area of future work.

Despite these limitations, the settings of this study retain the typical characteristics of complex design problems and AI-assisted human teams. More generally, this study provides insights into fundamental changes in the design process caused by integrating effective assistive AI agents into human teams as assistants and the corresponding influences on team performance.

5 *Conclusions*

The rise of AI affords unique opportunities to support team problem solving and design, yet it has been still unclear if, how, and why AI affects team agility. The current work studies the influence of assistive AI agents on the agility of design teams through a large-scale, 2×2, human subject experiment (opening versus restricting team structure reconfigurations and AI-assisted versus human-only team composition). During the experiment, teams experience both evolving requirement changes and abrupt structural changes, enabling the investigation of team agility in response to those changes. This work provides a counterpoint to prior research that highlights conditions where AI becomes detrimental to team problem solving. In this work, besides the teams performing better when partnered with assistive AI agents that communicate with humans in an effective way, the AI assistance boosts team agility with respect to both evolving and abrupt changes. The agile design teams become more thoughtful and perform more impactful actions, allowing for more cohesive communication between the team members. In turn, this leads to a broader exploration of the design space, ultimately resulting in higher performance at the team and individual levels.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We appreciate helpful discussions on this work with John Paschkewitz, Hannah Nolte, Adam Fouse, and Georgiy Levchuk. This work was supported by the Defense Advanced Research Projects Agency through cooperative agreement No. N66001-17-1-4064. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

Appendix 1. Experimental platform

HyForm is a collaborative design research platform providing a design environment that integrates assistive AI agents and humans (in combined, AI-assisted human teams) toward the design and operations of a drone delivery system. The platform allows experimenters to model and support different team structures and design strategies while merging AI capabilities into human design teams. Participants use three modules to solve the problem: the business module, the design module, and the operations module. The user interfaces of the three modules are presented in Figure 1A–C, respectively.

The business module permits business managers to define the target market for their teams by selecting a group of customers from a list of 56 unique customer locations to deliver packages and/or food. The business manager is also responsible for adapting the target market according to the development of team capabilities. Once updated, the selected target market can be sent to the operations planners. This module also allows business managers to compare all the operations plans submitted by the operations planners and select the best plan as the final solution for their teams.

The design module permits drone designers to assemble and evaluate drones. Designers can assemble drones using a group of pre-defined components, including battery, clockwise (CW) motor and rotor pair, counterclockwise (CCW) motor and rotor pair, airfoil, node, and rod connector. Multiple size options are available for the four main components: battery with 65 different size options, airfoil with 100 different size options, and CW and CCW motor and rotor pairs each with 50 different size options. In total, there are 265 unique components designers can use to build a drone. While building a drone, participants can explore a design space represented by a 169-node (13×13) two-dimensional grid by deploying different components at different positions. A drone example is illustrated in Figure 1B. At the beginning of the experiment, the design module provides participants with a basic drone design consisting of three components, one battery, two CW motor and rotor pairs, and two CCW motor and rotor pairs, as shown in Figure 3A. Participants can add components to the basic drone, remove components from a current design, or change component sizes to generate new drones. The generated drones can be evaluated in terms of four utility metrics, namely, range, capacity, velocity, and cost. An extra button in the AI-assisted version allows designers to invoke the design agent, as shown in Figure 1B. The created drones can be submitted to the operations planners in this module.

The operations module enables users to build drone fleets and create and evaluate drone delivery routes to maximize package and food delivery profit. Based on the target market defined by the business manager, the operations planners first need to select a group of promising drones from all available drones created by the drone designers to build a drone fleet. On this basis, the flying route of each drone in the fleet can be generated by linking a set of customer locations in a specific sequence. A warning is given when the route goes beyond the flying range or payload of the given drone. The flying routes of all drones in a fleet constitute an operations plan, which can be evaluated in terms of delivery profit, operating cost, and drone cost, etc. Likewise, an extra button in the AI-assisted version allows operations planners to invoke the operations agent, as shown in Figure 1C. The constructed operations plans can be submitted to the business manager in this module.

Appendix 2. Role responsibilities

In this experiment, each team is randomly assigned a starting structure, either the *open structure* or the *restrictive structure*. The two team structures are determined to allow for significant changes in team communication and problem solving within the 6-member team. In each structure, there are 6 team members consisting of five unique roles, including a Business Manager, an Operations Manager, two Operations Specialists, a Design Manager (who is also a Design Specialist when the teams are in the restrictive structure), and a Design Specialist. In this experiment, each role has access to a corresponding HyForm module according to their responsibility:

1) The Business Manager (Bm) is responsible for handling the company budget, maintaining communication channels with the team members, selecting the potential customers from the market, and approving or rejecting the operations plans provided by the Operations Team.
2) The Operations Manager (Om) is responsible for managing the operations team, maintaining communication channels with the Business Manager and the Design Manager, developing the drones' delivery routes, evaluating delivery routes and schedules, and submitting operations plans to the Business Manager.
3) The Operations Specialist (Os) is responsible for developing operations plans by generating delivery routes with designed drones to deliver parcels and providing operations plans to the Business Manager.
4) The Design Manager (Dm) is responsible for managing the Design Team, communicating the design requirements and design results to the Operations Manager and the Business Manager, informing the design requirements to the Design Specialist, designing drones, and submitting completed drone designs to the Operations Team.
5) The Design Specialist (Ds) is responsible for designing drones, informing the drones' status and capabilities to the Design Manager, and submitting completed drone designs to the Operations Team.

Both the managers and specialists in the design and operations disciplines have equal access to the shared databases for submitting and loading designs. However, the managers take on extra responsibility for managing their sub-teams and maintaining communication with the managers of the other disciplines. Accordingly, the managers would have a more comprehensive knowledge of team design progress; this enables them to engage in more meaningful cognitive and coordinating responsibilities, such as task allocation between sub-team members, sub-team goal setting, and strategic planning at a higher level, ensuring effective collaboration.

Notes

1. https://github.com/hyform.
2. Full data from this study has published in Zhang, Soria Zurita, et al. (2021).
3. The significance test is conducted using metrics available in (Paternoster et al., 1998).

References

Appelbaum, S. H., Calla, R., Desautels, D., & Hasan, L. (2017). The challenges of organizational agility (part 1). *Industrial & Commercial Training, 49*(1), 6–14.

Baum, L. E., Petrie, T., Soules, G., & Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics, 41*(1), 164–171.

Bernardes, E. S., & Hanna, M. D. (2009). A theoretical review of flexibility, agility and responsiveness in the operations management literature: Toward a conceptual definition of customer responsiveness. *International Journal of Operations & Production Management, 29*(1), 30–53.

Bird, S., Klein, E., & Loper, E. (2009). *Natural Language processing with Python: Analyzing text with the Natural Language toolkit*. O'Reilly Media Inc.

Conforto, E. C., Amaral, D. C., da Silva, S. L., Di Felippo, A., & Kamikawachi, D. S. L. (2016). The agility construct on project management theory. *International Journal of Project Management, 34*(4), 660–674.

Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. *Business and Information Systems Engineering, 61*(5), 637–643.

Demir, M., Cooke, N. J., & Amazeen, P. G. (2018). A conceptual model of team dynamical behaviors and performance in human-autonomy teaming. *Cognitive Systems Research, 52*, 497–507.

Demir, M., Likens, A. D., Cooke, N. J., Amazeen, P. G., & McNeese, N. J. (2019). Team coordination and effectiveness in human-autonomy teaming. *IEEE Transactions on Human-Machine Systems, 49*(2), 150–159.

den Otter, A., & Emmitt, S. (2008). Design team communication and design task complexity: The preference for dialogues. *Architectural Engineering and Design Management, 4*(2), 121–129.

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

Dove, R. (1995). Measuring agility: The toll of turmoil. *Production, 107*(1), 16–18.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly, 44*(2), 350.

Fliedner, G., & Vokurka, R. J. (1997). Agility: Competitive weapon of the 1990s and beyond? *Production and Inventory Management Journal, 38*(3), 19.

Gorman, J. C., Foltz, P. W., Kiekel, P. A., Martin, M. J., & 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* (pp. 424–428). Sage CA: Los Angeles, CA: SAGE Publications, No. 3.

Gunasekaran, A. (1999). Agile manufacturing: A framework for research and development. *International Journal of Production Economics, 62*(1–2), 87–105.

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies, 12*(1), 3–11.

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons, 61*(4), 577–586.

Kamar, E. (2016). Hybrid workplaces of the future. *XRDS: Crossroads, The ACM Magazine for Students, 23*(2), 22–25.

Kinne, S., Jost, J., Terharen, A., Feldmann, F., Fiolka, M., & Kirks, T. (2021). Process development for CPS design and integration in I4.0 systems with humans. In *Digital supply chains and the human factor* (pp. 17–32). Cham: Springer.

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest: A Journal of the American Psychological Society, 7*(3), 77–124.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences, 40*, E253.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*(2–3), 259–284.

Lapp, S., Jablokow, K., & McComb, C. (2019). KABOOM: An agent-based model for simulating cognitive style in team problem solving. *Design Science, 5*, 1–32.

Law, M. V., Dhawan, N., Bang, H., Yoon, S.-Y., Selva, D., & Hoffman, G. (2019). Side-by-Side human–computer design using a tangible user interface. *Design Computing and Cognition, 18*, 155–173.

Liew, C. (2018, May 1). The future of radiology augmented with artificial intelligence: A strategy for success. *European Journal of Radiology, 102*, 152–156.

Liu, M. L., Liu, N. T., Ding, C. G., & Lin, C. P. (2015). Exploring team performance in high-tech industries: Future trends of building up teamwork. *Technological Forecasting and Social Change, 91*, 295–310.

Lopez, C. E., Miller, S. R., & Tucker, C. S. (2019). Exploring biases between human and machine generated designs. *Journal of Mechanical Design, Transactions of the ASME, 141*(2), 021104.

Majid, N., Hossein, Z. M., Mehran, N., & Mahmood, Z. S. (2018). A hybrid approach to achieve organizational agility: An empirical study of a food company. *Benchmarking: An International Journal, 25*(1), 201–234.

Mathiassen, L., & Pries-Heje, J. (2006). Business agility and diffusion of information technology. *European Journal of Information Systems, 15*(2), 116–119.

McCarthy, P. M., Briner, S. W., Rus, V., & McNamara, D. S. (2007). Textual signatures: Identifying text-types using latent semantic analysis to measure the cohesion of text structures. In *Natural language processing and text mining* (pp. 107–122). London: Springer.

McComb, C., Cagan, J., & Kotovsky, K. (2017). Mining process heuristics from designer action data via hidden Markov models. *Journal of Mechanical Design, 139*(11), 111412.

Moradi, M., Moradi, M., Bayat, F., & Nadjaran Toosi, A. (2019). Collective hybrid intelligence: Towards a conceptual framework. *International Journal of Crowd Science, 3*(2), 198–220.

Moravec, H. (1988). *Mind children: The future of robot and human intelligence*. Harvard University Press.

Moss, J., Kotovsky, K., & Cagan, J. (2007). The influence of open goals on the acquisition of problem-relevant information. *Journal of Experimental Psychology: Learning Memory and Cognition, 33*(5), 876–891.

Park, Y., El Sawy, O., & Fiss, P. (2017). The role of business intelligence and communication technologies in organizational agility: A configurational approach. *Journal of the Association for Information Systems, 18*(9), 1.

Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology, 36*(4), 859–866, John Wiley & Sons, Ltd.

Pattie, M. (1995). Artificial life meets entertainment: Lifelike autonomous agents Pattie mat. *Communications of the ACM, 38*(11), 108–114.

Qin, R., & Nembhard, D. A. (2015, December 1). Workforce agility in operations management. *Surveys in Operations Research and Management Science, 20*(2), 55–69.

Raina, A., Cagan, J., & McComb, C. (2019). Transferring design strategies from human to computer and across design problems. *Journal of Mechanical Design, 141*(11), 114501.

Rebentisch, E., Conforto, E. C., Schuh, G., Riesener, M., Kantelberg, J., Amaral, D. C., & Januszek, S. (2018). Agility factors and their impact on product development performance. *Proceedings of International Design Conference, Design, 1*, 893–904.

Roll, I., Wiese, E. S., Long, Y., Aleven, V., & 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*, 169–182.

Salas, E., Bowers, C. A., & Cannon-Bowers, J. A. (1995). Military team research: 10 Years of progress. *Military Psychology, 7*(2), 55–75.

Sarica, S., Song, B., Luo, J., & Wood, K. L. (2020). Idea generation with technology semantic network. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 1–19.

Schulz, A. P., & Fricke, E. (1999). Incorporating flexibility, agility, robustness, and adaptability within the design of integrated systems - key to success? *AIAA/IEEE Digital Avionics Systems Conference - Proceedings, 1*. 1.A.2-1.A.2.

Senescu, R. R., Aranda-Mena, G., & Haymaker, J. R. (2013). Relationships between project complexity and communication. *Journal of Management in Engineering, 29*(2), 183–197.

Sharifi, H., & Zhang, Z. (1999). Methodology for achieving agility in manufacturing organisations: An introduction. *International Journal of Production Economics, 62*(1), 7–22.

Sherehiy, B., Karwowski, W., & Layer, J. K. (2007). A review of enterprise agility: Concepts, frameworks, and attributes. *International Journal of Industrial Ergonomics, 37*(5), 445–460.

Singh, V., Dong, A., & Gero, J. S. (2013). Social learning in design teams: The importance of direct and indirect communications. *AI EDAM, 27*(2), 167–182, Cambridge University Press.

Sio, U. N., Kotovsky, K., & Cagan, J. (2015). Fixation or inspiration? A meta-analytic review of the role of examples on design processes. *Design Studies, 39*, 70–99.

Song, B., Soria Zurita, N. F., Nolte, H., Singh, H., Cagan, J., & McComb, C. (2022). When faced with increasing complexity: The effectiveness of AI assistance for drone design. *Journal of Mechanical Design, 144*(2), 021701. https://doi.org/10.1115/1.4051871.

Song, B., Soria Zurita, N. F., Zhang, G., Stump, G., Balon, C., Miller, S. W., Yukish, M., Cagan, J., & McComb, C. (2020). Toward hybrid teams: A platform to understand human-computer collaboration during the design of complex engineered systems. *Proceedings of the Design Society: DESIGN Conference, 1*, 1551–1560.

Stewart, G. L., & Barrick, M. R. (2017). Team structure and performance: Assessing the mediating role of intrateam process and the moderating role of task type. *Academy of Management Briarcliff Manor, 43*(2), 135–148, NY 10510.
Stump, G. M., Miller, S. W., Yukish, M. A., Simpson, T. W., & Tucker, C. (2019). Spatial grammar-based recurrent neural network for design form and behavior optimization. *Journal of Mechanical Design, 141*(12), 124501.
Tallon, P. P. (2008). Inside the adaptive enterprise: An information technology capabilities perspective on business process agility. *Information Technology and Management, 9*(1), 21–36.
VanLehn, K., Burkhardt, H., Cheema, S., Kang, S., Pead, D., Schoenfeld, A., & Wetzel, J. (2019). Can an orchestration system increase collaborative, productive struggle in teaching-by-eliciting classrooms? *Interactive Learning Environments* 1–19.
Vickery, S. K., Droge, C., Setia, P., & Sambamurthy, V. (2010). Supply chain information technologies and organisational initiatives: Complementary versus independent effects on agility and firm performance. *International Journal of Production Research, 48*(23), 7025–7042.
Walliser, J. C., de Visser, E. J., Wiese, E., & Shaw, T. H. (2019). Team structure and team building improve human–machine teaming with autonomous agents. *Journal of Cognitive Engineering and Decision Making, 13*(4), 258–278.
Werder, K., & Maedche, A. (2018). Explaining the emergence of team agility: A complex adaptive systems perspective. *Information Technology and People, 31*(3), 819–844.
Williams, G., Meisel, N. A., Simpson, T. W., & McComb, C. (2019). Design repository effectiveness for 3D convolutional neural networks: Application to additive manufacturing. *Journal of Mechanical Design, 141*(11), 111701.
Wilson, H. J., & Daugherty, P. R. (2018). Collaborative intelligence: Humans and AI are joining forces. *Harvard Business Review, 96*(4), 114–123.
Wood, M., Chen, P., Fu, K., Cagan, J., & Kotovsky, K. (2014). The role of design team interaction structure on individual and shared mental models. In *Design Computing and cognition, Vol. 12* (pp. 209–226). Dordrecht: Springer.
Yaghoubi, N.-M., Kazemi, M., Dahmardeh, M. R., & Arhami, F. (2011). Organizational agility A way to import the advantageous aspects of competitive market (an analytical-comparative approach on agility models). *Interdisciplinary Journal of Contemporary Research in Business, 2*(12), 766–783.
Yusuf, Y. Y., Sarhadi, M., & Gunasekaran, A. (1999). Agile manufacturing: The drivers, concepts and attributes. *International Journal of Production Economics, 62*(1), 33–43.
Zhang, G., Raina, A., Cagan, J., & McComb, C. (2021). A cautionary tale about the impact of AI on human design teams. *Design Studies, 72*, 100990.
Zhang, G., Soria Zurita, N. F., Stump, G., Song, B., Cagan, J., & McComb, C. (2021). Data on the design and operation of drones by both individuals and teams. *Data in Brief, 36*, 107008.