A review of collaborative guidance for collocated workers

Daniel A. Delgado
University of Florida
Gainesville, FL, United States
danieldel@ufl.edu

ABSTRACT
Collocated collaboration offers unique opportunities and disadvantages for teams working together. Groups communicate verbally, share physical objects, and work in parallel while aiming to complete a common goal. However, team members often struggle to coordinate with each other. Task guidance systems can enhance collaborative experiences by acting as a mediator between associates and altering group behavior. Even with this, deciding how to respond automatically to collaborative triggers is not trivial. The first step in providing guidance is knowing when to provide it. Fortunately, human-to-human interactions offer a rich source of observable data, such as facial expressions and perceived workload, that allows researchers to model collaborations. A perceived high workload in one collaborator enables a model to identify breakdowns in the collaboration and act accordingly. Previous literature reviews on task guidance systems have yet to present how collaborative modeling have supported adaptive task guidance for collaborators. We conducted a literature review on existing collaborative guidance systems to understand current methods for supporting of collaborators through diverse interfaces, intragroup collaborative modelling techniques, and system reactions. We then analyzed across contexts to identify the shared advantages and disadvantages of the approaches taken by systems. Finally, we offer opportunities for further exploration in the area of collaborative guidance.

Author Keywords
Structured Literature Review; Adaptive Task Guidance; Collocated Collaboration; Collaborative Modeling.

CCS Concepts
•**Human-centered computing** → **Human computer interaction (HCI);** *Haptic devices;* User studies; Please use the 2012 Classifiers and see this link to embed them in the text: https://dl.acm.org/ccs/ccs_flat.cfm

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI'20, April 25–30, 2020, Honolulu, HI, USA

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ISBN 978-1-4503-6708-0/20/04. . . $15.00

DOI: https://doi.org/10.1145/3313831.XXXXXXX

INTRODUCTION
Collocated collaboration (CC) refers to two or more people working together in a shared physical space to achieve a common goal or solution [42]. While other forms of collaboration exist (e.g., remote), we are focusing on only CC for this literature review. In group work, CC provides a unique set of advantages. Collaborators can seamlessly communicate, move between sub-activities, and remain aware of the task. Compared to remote collaboration, collocated workers are twice as productive [42]. However, collocation comes with its disadvantages which may lead to breakdowns. In this context, we define breakdowns as any event in the collaboration that can hinder it. For example, collaborators may become easily distracted by their working environment and physical limitations in their working space can limit productivity [42].

Task guidance systems can aid in collaboration by enhancing procedure following and overcoming some of the breakdowns in collocation. Existing systems can provide users with step-by-step instructions and improve coordination between collaborators [33, 41]. However, assisting without awareness of the current context can lead to an awkward and ineffective system [2].

Task guidance systems have taken an adaptive approach by employing user modeling techniques and contextual information to adapt the guidance to user needs. For example, Leelasawassuk et al. [32] used head motion to gauge user attention and trigger guides based on the objects detected. Lopresti et al. [34] allowed user input in the system to provide more instructions or time on tasks. When applied to CC, we refer to this work as collaborative modeling.

Technologies have become increasingly prevalent for mediating and navigating collocated group work. However, prior work has yet to present an analysis of existing methods for effectively understanding current group dynamics and providing appropriate interventions. An understanding of the relationships between modelling techniques and interventions is necessary for designing effective collaborative guidance systems. At the intersection of collaborative modeling and adaptive task systems, we investigate how these techniques have facilitated collocated collaboration. We then seek to identify how current systems identify breakdowns in CC and respond to them. Finally, we show how collaborative modeling can improve group efficiency. Our main contribution is an overview of the field in its current state and valuable insight into future research endeavors.

This literature review is organized as follows. First, we introduce previous literature reviews on related topics such as collaborative modeling and task guidance. The following section defines the methodology used to conduct the literature review. The results section presents our findings from the literature review. We conclude with a discussion section that analyzes the limitations of existing systems and open questions.

RELATED WORK: LITERATURE REVIEWS

Prior literature reviews on task guidance have looked into providing guidance for workers [41] and augmented reality (AR) assisted guidance [3, 51]. However, most of the existing work has only looked into individual guidance [3, 41, 51] or supporting remote collaboration [51]. Ockerman et al. [41] performed a literature review on task guidance for workers. From their review, they were able to define a set of system requirements. They included: task guidance systems should provide ease of use for mobile systems, leave the user with at least one hand free, be adaptable to various conditions, and be reasonably priced. Along with system requirements, the review found that current systems need to provide users with situational awareness. One limitation was the lack of technology to support the necessary sensors at the time (e.g., lightweight cameras). However, Ockerman et al. [41] conducted their review in 2000 and technologies have since improved in scalability and cost-effectiveness [46]. Additionally, Wang et al. [51] conducted a review on supporting assembly workers using AR devices. Their work presented challenges and opportunities in AR-aided assembly. However, their work mostly focuses on problems faced by the individual operator and does not take into group dynamics . Comparatively, Blattgerste et al. [3] investigated augmented reality devices for cognitively impaired users . Their work presented common trends in devices for supporting users. Augmented reality devices were shown to increase motivation and improve task completion times [3]. However, their work also mainly focused on individual guidance.

Previous literature reviews on collaborative modeling have focused on tailoring experiences to users in a collaborative interaction [40, 46]. Neumayr et al. [40] conducted a literature review on personalized and adaptive collaborative systems. They found a trend in the field towards user modeling and recommender systems. Recommender systems in social interactions can guide users through the experience, encouraging them to partake in conversations when they become isolated [40]. However, their work does not look into the relationships between collaborative modelling, interventions, and effective methods for aiding collaborators on a task. Furthermore, recommender systems were not included in this literature review since they mainly focus on web tools or remote collaboration [40]. Comparatively, persuasive computing systems are designed to alter user behavior and attitudes [19]. However, prior persuasive systems are mostly intended for changing individual user behavior and were not considered in a collaborative context [23]. Praharaj et al. [46] investigated multi-modal learning analytics for understanding the quality of a collaboration . Their work mostly focuses on the technology used to detect collaborative events and did not consider the number of collaborators involved [46].

Radu et al. [47] surveyed the needs of collocated collaborating users in an augmented reality setting [47]. One of the needs identified was the importance of guidance through step-by-step procedures. Their work found that task annotations in a mixed-reality environment reduced cognitive load and errors caused by users. Although, their work does not provide an analysis of existing methods for guiding collaborators, they present a need for developing a foundation to support collaborators.

Prior work has also looked at collaborative systems for encouraging social interactions [43, 49]. Silva-Calpa et al. [49] reviewed collaborative systems for aiding users on the autism spectrum. Their work presented strategies for stimulating and assisting collaboration and system design recommendations [49]. Olsson et al. [43] conducted a review of designs for enhancing collocated social interactions. They categorized existing literature into three abstract categories: facilitating, inviting, and encouraging. Their work presented limitations in existing methods for preventing social disruptions by technology [43]. Collocated collaboration is an inherently social phenomenon [42]. However, they may differ in their respective goals. Based on prior literature, prior work has mostly looked at enhancing social interactions [43] or encouraging participation [49] and not improving the overall performance of the task at hand.

Prior work has shown that groups can be monitored and tailored to using adaptive systems. However, prior work has not looked into understanding and actively aiding groups working on a joint task. We therefore seek to understand the current state of research, reveal limitations in current work, and provide future avenues of exploration.

METHODOLOGY

We conducted our literature review using the guidelines for planning, conducting, and presenting structured literature reviews in software engineering provided by Kitchenham and Charters [26]. Planning for a literature review involves specifying the research questions and developing a review protocol. Conducting the study calls for the selection of primary studies and data extraction. The final step entails formatting and presenting the main report [26].

Research Questions

- (R1) How have adaptive task guidance systems been used to facilitate co-located collaboration?
- (R2) How are breakdowns in collaboration identified in existing systems?
  - (R2.1) How do adaptive tasks systems currently react to breakdowns?
- (R3) How can collaborative modeling in adaptive task guidance systems be used to improve group efficiency?

Keywords

Keywords for the literature search were entered into Google Scholar, and the first ten pages of the results were used to find relevant papers. Google scholar was the only search engine used for this literature review. The search resulted in sixty-two papers/systems before filtering. Papers were filtered out based on an inclusion/exclusion criteria, which is discussed in the next section.

The keywords used for the literature search were: Adaptive task sharing, adaptive guidance system collocated collaboration, adaptive task guidance system collocated collaboration, co-located collaborative systems review, and adaptive guidance system collaboration review.

Inclusion/Exclusion

The abstracts of papers were initially read to assess whether they satisfied the inclusion criteria. We placed no constraints on the medium in which the task guidance is performed (e.g., mobile devices, head-mounted displays, paper prototypes).

Our inclusion criteria includes:

1. Papers focused on guidance system.
2. Systems designed for collocated collaborators.
3. Designs capable of providing task support for two or more users.

Papers were excluded if they were not relevant to the research questions.

Our exclusion criteria includes the following:

1. Papers irrelevant to the identified research questions.
2. Studies focused on remote collaboration.
3. Don't focus on human-to-human collaboration.

Several irrelevant papers came up in the literature search due to the keyword's semantics, but they were unrelated to the research questions. For example, the keyword "adaptive guidance system collocated collaboration" resulted in the publication by Lundgren et al. [35]. However, the paper did not involve a task guidance system, although it satisfied the second inclusion criteria.

Several papers on collaborative remote task guidance were found in the search. In one example, the keyword "adaptive task guidance system collocated collaborators" resulted in the publication by Le et al. [31]. The publication focused on remote expert guidance and therefore was excluded.

Snowballing

Often, a relevant paper was cited or referenced by another publication and satisfied the inclusion criteria, leading to inclusion in this literature review. We refer to this as snowballing [52]. Snowballing was used to identify more relevant papers that didn't initially appear with the original keywords. A total of ten papers were found using the snowballing technique.

Analysis

Firstly, papers were organized into conceptual designs or physical systems. We noted the interfaces used for implementation, collaborative modeling methods, and system responses or interventions for each system. We also took note of user evaluations, which usually included user reactions to interfaces, user studies, and their respective results. Systems were compared across similar attributes.

RESULTS

Out of all of the papers accessed in our literature search, fourteen were actual systems and four were conceptual designs. The difference between the number of actual systems and total papers accessed can be attributed to multiple papers presenting a system and usually a follow-up evaluation on the same system. For this review, we associate both papers with a single system for analysis.

Conceptual Designs

Industrial Guidance

Fang et al. [17] introduced a method for imposing virtual elements onto a physical space for augmented-reality based guidance in an industrial setting. Their work mostly explored localization techniques using images which captured the environment to train a scene cognition model and map the 3D environment. The model was then used to acquire the orientation of an operator based on a 2D image. Once the orientation is estimated, virtual instructions are overlaid on the physical space using an augmented reality headset via a simple text or 3D model.

The method was evaluated using the ChangeSim dataset [45], which includes images from an industrial environment. The model was trained using the dataset and then evaluated on its ability to recognize the scene, collaborative localization, and superimposing virtual objects onto the scene using the recognized positions. Additionally, the model was simulated on a scenario of five operators at a time. Results showed that the method was able to effectively recognize the scene and operator orientation with minimal error [17].

Dynamic Collaborative Emergency Response

Netten et al. [39] presented a method for dynamically delivering relevant task-information to collaborating emergency responders, described as an adaptive workflow simulation. For developing the adaptive workflow simulation, a Naive Bayes [1] text classification model was trained on the speech utterances of past emergency situations and was used to filter out the necessary information from communication channels. The information was then directed to the appropriate team members. By automatically filtering information, their method aimed to reduce information scarcity or overload of first responders. The system employed the adaptive workflow to estimate the likely next task, predicting future events and adjusting information flow to the current context and likely outcomes [39].

The system mainly utilized communication channels such as audio in the form of transcriptions. Audio features were extracted into word tokens which are used as input into the machine learning model. The machine learning model is used in conjunction with the adaptive workload model to guide the dynamic information distribution [39].

The system was evaluated using simulation data from the Koningskerk disaster scenario. The evaluation mostly looked at the performance of the machine learning model in identifying relevant task information. The model was evaluated in three different experiments: word tokens only; word tokens and first responder name; and word tokens, first responder name, and context. Out of all three experiments, the third one performed the best with the highest classification accuracies [39].

Human-Machine Function Allocation

Zhang et al. [54] used prior information on aircraft pilot crew skills and abilities to adjust levels of automation. A decision making approach for aggregating responsibilities and tasks to operators was developed using 2-tuple linguistic information. The 2-tuple linguistic model uses Weights which are assigned to the users' skills and abilities and are used to choose allocation schemes. Zhang et al. [54] showed the effectiveness of their allocation method using the ground proximity warning system in an aircraft cockpit example. In the example, five humans are simulated with unique abilities such as having good spatial reasoning and pattern recognition. The decision making system then uses the weight of their abilities to compare ability superiority Based on their analysis, their approach was a more reliable method for function allocation compared to alternatives [54].

Adaptive Instruction for Teams

Fletcher et al. [18] investigated how shared mental models could support adaptive team tutoring for collective tasks [18]. An intelligent tutor would allocate roles and responsibilities to team members and provide a context-specific evaluation. Behavioral and physiological sensors could detect individual user workloads and asses their engagement. Interdependencies between tasks would be taken into account when assessing team workload. A recommender engine would select instruction strategies based on detected user states and prior knowledge of the context [18]. They provided a framework for future work.

Systems

Systems were found to be designed for a variety of contexts, including assembly, emergency response, navigation, social engagement, medicine, and education. Based on the lack of prior work in understanding the relationships between number of collaborators, user interfaces, collaborative modelling, and interventions, we analyze across each dimension for systems where it is applicable. We provide an analysis of the relationships in the next section.

Hapticturk

Cheng et al. [7] presented *Hapticturk*, a people-based motion platform. *HapticTurk* employs workers to simulate a motion platform, aiming to create an immersive virtual reality simulation by coordinating worker movements and synchronization. The workers received a rhythm-based set of instructions to coordinate motions for the user in the virtual environment via mobile devices (iPhones, iPods, and Tablets). Worker's movements were not individually tracked, instead the performance of the group was evaluated as a whole through their perceptions of the task at the end of the user study [7].

Hapticturk was evaluated through an in-lab user study and an outdoor experimental deployment. For the in-lab study, a range of 2-4 participants were recruited per session and experimenters filled in the remaining positions to reach a total of 5 members. Participants would alternate through player/worker positions throughout the study. Participants rated being a physical worker as being less enjoyable than being a player in the game. However, participants appreciated seeing players joy and reactions. As they progressed through the study, participant fatigue was associated with less enjoyment with the task. Additionally, being grouped with strangers made the experience less favorable to workers [7]. For the experimental deployment, *Hapticturk* was deployed outside of the lab. Observations from the deployment showed the of ease of implementing *Hapticturk* outside of a controlled setting and in the wild [7].

Turkdeck

Based on the findings from *Hapticturk* [7], Cheng et al. [8] also presented *Turkdeck*, a prop-based virtual reality haptic-simulation environment. *Turkdeck* employs a group of physical workers to simulate an ever changing virtual environment for a user wearing a head mounted display, such as the Oculus Rift [6]. The *Turkdeck* system coordinates workers to move props around to match the player's virtual reality experience, creating an immersive environment the player can touch.

Turkdeck guides workers by displaying visual instructions on the environment via a lazer beam, auditory instructions, and sheets on the props. Tracking in *Turkdeck* is handled entirely by the workers, avoiding the need for any sensing technologies. When an action is completed, a worker presses a wireless button that triggers the system's predefined instruction and animation sequences [8].

The *Turkdeck* system was evaluated through a user study. The study had a baseline condition which used the *Turkdeck* system and one that did not use it. During a session, ten workers were employed for the movement of the props. Overall, participants rated the system as being useful in guiding them through the task. Comparable to the *Hapticturk* [9] system, participants reported fatigue during the task [8].

Crowdsourced Fabrication

Lafreniere et al. [28] explored *Crowdsourced Fabrication* through a three day exhibition in which a crowd of over one hundred volunteer workers were guided to build a bamboo structure. Aiming to expand on the concept of digital fabrication [22], their system employs an intelligent workspace that monitors and guides workers through physical tasks. The design goals of the system were to provide just-in-time learning, utilize unintrusive technology, promote worker safety, improve efficiency, and provide data analytics of the collaborative process [28].

As volunteers entered the workspace, they were given a smartwatch that would present them with unique instructions and provide the system with location tracking. Additionally, LED lights appearing ubiquitously would guide workers through step-by-step instructions. A main foreman engine would monitor the status of each worker, distribute tasks, and the build progress. Also, a staff member would monitor progress and provide help when needed [28].

The *Crowdsourced Fabrication* system was evaluated through a three-day long exhibit that had 108 participants. Their study did not compare the effects of the system to a control method. Their results were mostly observational and included participant feedback. The system received mostly positive reactions from participants rating it high in usefulness, overall building experience, self-rated confidence while using the system, and device experience [28].

WeBuild

Fraser et al. [20] introduced *WeBuild*, a multi-user adaptive task interface for assembly and construction. Their work builds upon Lafreniere et al.'s [28] *Crowdsourced Fabrication* system. Inspired by their approach, Fraser et al. [20] first studied collaborative assembly task interactions in an observational study which would later provide design guidelines for their adaptive task system. From their observations, they set design goals for providing users with continuous adaptions, task awareness, and step-by-step instructions [20].

The *WeBuild* interface consists of a dashboard display which provides an overview of the task and a mobile phone which displays graphical instructions. Collaborative modeling in the *WeBuild* system mostly focused on individual user characteristics and meta-data such as prior experience, tasks the user has already excelled at, and keeping the user in similar or a diverse set of tasks. An understanding of the group's cohesion is then built by combining all of these parameters together in a task tree which informs a task distribution algorithm. The task distribution algorithm decides which task a user is currently assigned to automatically. User's were also given the ability to ask for help when needed. The system was proposed to be scalable to any number of users but it was evaluated on groups of five [20].

A user evaluation of the *WeBuild* was performed on eight groups of five participants, totalling to forty participants in all. The study involved participants assembling a custom LEGO structure using tools such as a wrench and a screwdriver. In the experimental condition, participants performed the task with the *WeBuild* system and in the control condition, without the system. The user evaluation study looked at overall completion times and start-up times. Start-up times were described as the initial period before a task is begun where participants would go over the instructions and designate tasks between each other. The only significant result observed from the user study was the difference in start-up times between conditions. Participants were found to spend significantly less time in the start-up phase of the task while using the *WeBuild*. However, this came at a cost to overall task awareness. In the experimental condition, participants had significantly less perceived awareness of the overall task. Qualitative results from their user study showed that participants found the system to be helpful [20].

Automatics

Building upon *WeBuild* [20], Lakier et al. [29] present *Automatics*, a dynamic knowledge-resource tool capable of adapting to contexts. The main difference between *WeBuild* and *Automatics* is that the latter supports fabrication from raw materials and error recovery. Their work explores dynamic manuals for providing guidance and instructions to multiple workers [29].

A literature review conducted by Lakier et al. [29] showed deficiencies in existing instructional manuals concerning object handling, task sequence and representation. Their literature review was used to provide design guidelines for the *Automatics* system. For modeling individual performance, the *Automatics* system keeps track of how often a user abandons a task, the number of mistakes they make, user tool preferences, and similarities between the tasks the user performs. All of these characteristics combined are used to compute a possibility score that is used to assign tasks to each user by employing a greedy algorithm. The *Automatics* system is implemented on an 11" x 7" tablet. The interface provides a graphic visualization of the current build, a global overview of the task and allows users to modify their designs. Although the system automatically assigns tasks to the user, they can freely make changes to the process. The system is claimed to support any number of users popping in and out of the assembly process, but it was only evaluated on at most two people at a time [29].

The system was evaluated through an exploratory study that observed three groups of two people and six individuals, twelve participants in total, perform fabrication tasks. All participants experienced the paper instruction and *Automatics* condition. Their study was too small to evaluate statistically. However, participants completed more tasks and made fewer errors while using *Automatics*. Participants also commented that the *Automatics* system required less mental effort than paper instructions [29].

Overseer

Luqman et al. [36] presented a system for providing task instructions and coordinating ad-hoc disaster teams. Contextual information from mobile phones was gathered to decide which tasks are most appropriate for each person on a team.

In order to monitor contextual information, the *Overseer* [36] system utilizes mobile device sensors, which includes a global positioning system (GPS), microphones, and a camera. Using the sensor data, information about each persons attentiveness, current location and direction, and team density is calculated. The *Overseer* [36] system is implemented on a central command center that manages the information gathered from the mobile phones. The system is initialized through an application in which a user specifies their level of expertise in various disaster areas. For managing tasks, a dependency graph is created that uses the probability of task completion to determine the appropriate task allocations. The probability of task completion is evaluated based on the current mobile contexts of active users. The *Overseer* [36] system also utilizes experts-in-the-loop for task feedback and reallocation. Suppose the system detected a drop in performance via the dependency graph, then an expert would be alerted. The *Overseer* [36] algorithm would also consider alternative actions to improve the situation, such as assigning idle team members to the task. No user evaluation was provided for the system in the provided literature [36].

Crew Workload Manager for Aircraft Cockpits

Dorneich et al. [12, 13, 14] presented a crew workload manager for aircraft cockpits . The manager acts as a neutral external member to the pilot and co-pilot by monitoring their current workload states and providing suggestions for distributing task loads. The system aimed to overcome communication barriers in aircraft cockpits where crews may hesitate to share their current workload states [16]. Also, the system is considered open-loop since it allows the pilots to make the final decision in the adaptive process. A closed-loop system would make final decisions automatically without the user's input [13].

The pilots' workload was monitored using electrocardiogram (ECG) and electroencephalography (EEG) sensors, which measure heart and brain activity, respectively. A k-nearest neighbors algorithm was used to classify the workload states between low or high [15]. Each pilot's workload state was displayed on the main user interface on a continuous timeline that updated every 5 minutes. The goal was to allow both pilots to view each other's workload and their own. During workload imbalances, the manager suggests changes to task distributions through system alerts on the interface. From there, pilots can decide how to redistribute tasks [13].

The system was evaluated using the Shared Aviation Task battery (SAT-B), which was inspired by the Multiple Attribute Task Battery (MAT-B) [10]. The MAT-B was designed to evaluate a single pilot's performance through a set of aviation-related tasks [13]. The SAT-B allows for two pilots to work together on aviation-related tasks and share tasks with each other. The intensity of tasks is manipulated through the frequency of events which is used to further increase or decrease a pilot's workload. Participants took part in four 15-minute trials in which there were four opportunities for workload imbalance [13]. The evaluation had two conditions, one where the workload manager was activated and another where it was not. Results from the evaluation showed that there was a significant difference in the amount of time pilots spent in an unbalanced workload state between both conditions. The crew workload manager effectively informed users of workload imbalances and was able to help users overcome them. There was a significant different in the number of correct sharing requests from pilots. Also, the workload manager did not affect the task performance of the pilots. Overall, participants found the system to be easy to understand and helpful [13].

Workload-Adaptive Support for Helicopter Cockpit Crews

Brand et al. [5] introduced a workload-adaptive closed-loop system for helicopter cockpits where the crew controlled several other unmanned aircrafts from the cockpit. Compared to a traditional two-manned helicopter cockpit, the addition of controlling external unmanned aircrafts from the cockpit adds additional workload to an already strenuous environment. The system acts as an additional crew member monitoring and aiding in the crew's tasks while also maintaining its distance to prevent over-reliance on automation. The system only intervenes when the crew is not focused on the most important task, has a high workload, or is incapable of performing the task. System interventions aimed to prevent or correct human errors by guiding crews through the task and decreasing workload [5].

Periods of high workload from the crew were calculated using system predictions. Prior to beginning the mission, a task tree is established where the complexity and mental demand of each task are pre-determined. The sequence of tasks is used to predict potential high workload states. The adaptive system also takes into account whether the crew is neglecting or delayed on a critical task for determining potential workload peaks. Based on the predicted workload peaks, the system provides adaptive interventions. First, it determines whether the crew would be capable of handling the future task and adapts the interface to simplify the task if they aren't capable of doing so. A human-dialog interface would direct the attention of users. Additionally, if the system notices the crew poses the risk of complete failure, it automates the task and takes control [5].

The adaptive system was implemented on a simulated helicopter cockpit and was evaluated using mission scenarios. The subjects were trained military helicopter pilots. The number of subjects the system evaluated was not specified, and no statistical evaluation of their results was presented due to the small sample size. The number of participants was not specified in the literature, but they did mention testing their system through forty-five interventions. Participants felt supported by the system and felt the interventions were appropriate. However, potentially not all of the automatic interventions could be understood by the participants. Thus, showing the need for transparency in the system [5].

Healthcare Simulation Classroom

Martinez-Maldonado et al. [38] explored collaborative modeling for analyzing student interactions in healthcare simulation classrooms and providing appropriate feedback. Classrooms provide a rich environment of face-to-face interactions within groups and student tasks. However, a limited number of teachers are available to observe and provide valuable feedback to students. By automatically analyzing student behavioral interactions, feedback can be provided effectively. Martinez-Maldonado et al. [38] presented the challenges encountered when implementing classroom analytics around a mannequin simulation [38].

The system was implemented on a SimPad tablet where students could track their progress in a task. Additionally, a Kinect sensor was used to track student's positions around the mannequin. Finally, a microphone placed in the center of the mannequin was used to track students' conversation patterns and a video camera was used to capture the whole scenario. Their system did not provide interventions and focused only on collaborative modeling [38].

The simulation was evaluated through a study that analyzed five randomly chosen laboratory classes of nursing students. Each class consisted of five hospital beds with mannequins and the necessary sensors. The study had fifty-six student and four teacher participants. Each student was paired in groups of four. During the study, each student had a preset role in the task, such as the team leader or the nurse. The study mostly presented qualitative observations, which were analyzed using the *Activity-Centered Analysis & Design (ACAD)* framework [37]. Observations were analyzed across the social, setting, epistemic (the task), and runtime dimensions [38].

Results from the study present potential opportunities and challenges of the implemented system. There were opportunities for tracking their behavior and activity using their tracked location. However, there were challenges with the occlusion of participants when using sensors. Students would occlude each other by getting in front of one another, making it hard for the depth sensor to detect them. Audio tracking could be useful for tracking student activity. However, classroom noise can impede the quality of recordings. There were potential challenges in monitoring external devices that students use in the simulation, such as EMG or ECG. This constraint would make it difficult for a system to track task progress automatically. While monitoring the task, the system faces challenges in automatically tracking progress. For the study, students were asked to record their progress manually. However, not every student did so despite the teacher's request. During the study, it was observed that teachers implemented different styles of teaching, making it difficult to utilize analytics. Overall, students and teachers were comfortable with the system due to being monitored before. However, collaborative classroom analytics raises concerns about privacy due to the amount of personal data collected [38].

Meeting Mediator

Kim et al. [27] presented *Meeting Mediator*, a system that aims to facilitate group meetings by collecting data on real-time group dynamics and providing appropriate feedback. Their work considers the protagonistic understanding of dominant behavior in effecting group dynamics through the act of group mediation and encouraging participation. They hoped to gain an understanding on the relationship of the effects of dominance and group dynamics on task performance [27].

Meeting Mediator employs a sociometric badge for tracking individual behavior that captures speech features and 3-axis orientation using infrared (IR) sensors. Using the speech features, social cues were extracted, such as enthusiasm, interest, persuasion, and nervous energy. At the same time, the 3-axis orientation provided information on face-to-face interactions, body movements, and gestures. The user interface is implemented on a mobile phone application that displays current group activity. Group activity is represented through a circle and web visualization, where the circle is connected to nodes on the corner. The nodes represent a person in a group [27]. In a balanced social setting, the circle remains relatively in the center of the visualization, while in a setting where a group of people are dominating the conversation, the circle will lean closer to them [27].

Meeting Mediator was evaluated through a user study in which participants took part in a brainstorming and problem solving task. The study had two conditions, one with the *Meeting Mediator* and one without it. Their study also looked at the effects of *Meeting Mediator* on distributed collaboration. There were a total of thirty-six participants in groups of four. *Meeting Mediator* was effective in identifying dominant persons in a group. Results from their evaluation showed that *Meeting Mediator* had a significant effect in reducing the amount of overlapping speaking time and increased the frequency the main speaker changed compared to the other condition. *Meeting Mediator* was effective in balancing group dynamics. However, there was no significant difference between both conditions in task performance. Results from a post-study questionnaire showed that *Meeting Mediator* was not distracting to participants compared to the other condition [27].

Multimodal Support for Meetings

Terken et al. [50] introduced a multi-modal support system for collocated meetings aimed at balancing group dynamics by collecting participant speech and attention data to provide real-time feedback. The system acts as an external facilitator that provides each participant unique information about their social interactions [50].

For analyzing participants speaking activity, their audio was recorded using an eight channel audio controller and their eye-gaze was estimated using IR headbands that provided their head orientation. Participant speaking activity was sent to a server which informed the visualizations. The system was implemented on table-top visualization where three circles represent each participants' speaking activity, attention from the speaker, and attention from the listener. Each circle would expand depending on the level of activity from each participant [50].

Terken et al. [50] evaluated their multimodal system through user studies comprised of nineteen groups of four people and two groups of three for a total of eighty-two participants. Participants took part in a decision making task where each person was provided with a unique set of information, inspired by DiMicco et al.'s [11] hidden-profile task. The study had a feedback condition which involved using the system and a no feedback condition. Results from the study showed there was a significant difference in speaking times for under and over active participants between conditions. The system was capable of altering group dynamics. However, participants found the system to be distracting. Also, the eye-gaze estimations were found to be an inaccurate method of measuring attention [50].

Ambient Suite

Aiming to overcome the limitations encountered by *Meeting Mediator* [27], Fujita et al. [21] presented *Ambient Suite*, an immersive collaborative environment that enhances conversation through wall, floor, and personal displays. The system gathers the individual characteristics of each participant to provide feedback using the displays. The goal of the system was to reduce imbalanced conversation states and encourage participation. *Ambient Suite* is implemented in a standing-party group scenario. The standing-party group scenario has participants engage in a simulated social environment where they discuss predetermined topics [21].

For gaining participant information and providing feedback, each participant was provided a cup-shaped personal device with a microphone and an acceleration sensor. The device would monitor individual conversation activity and hand gestures. A marker was placed on participants to track 3D positioning in the room. Participant activity was calculated as a function of speaking time, hand acceleration, and head rotation. *Ambient Suite* was implemented using a wall display to provide viewable information to all participants. In the differing conditions, the wall display would show either a topic to spark conversation between participants or a graph of the overall conversation activity in the room. A floor projection under each participant would provide a heat map of each participant's individual conversation activity, further encouraging other participants to engage with them if they had a low activity state. The floor projection would update every 5 seconds [21].

Ambient Suite was evaluated through a standing-party study with information and no-information conditions. As mentioned, the information condition would show a topic to spark conversation between participants, and the no-information condition would show only the activity graph on the wall display. Over a hundred participants were recruited, creating seventeen groups of six people. Each condition lasted twelve minutes. The study aimed to determine whether the system would proficiently capture participant activity and their reactions to the environment. Results from the study showed that the system could find significant differences in head rotations between both conditions. Thus, showing the effectiveness of the sensors in detecting non-verbal behaviors. Participants rated the information presenting condition significantly higher in promoting shared interests and an active conversation space. Participants also felt an individual person was leading the conversation less in the information-presenting condition. Additionally, participants significantly favored the information-presenting condition over the no-information condition. Overall, the system promoted a rich social space while reducing the perception of conversation imbalances [21].

Multimodal Augmented Reality in medicine

Harders et al. [24] presented a framework for employing haptic feedback sensors in an augmented reality setting for medical training procedures. Haptic sensors would offer physical limitations on virtual artifacts to increase a sense of presence. Multiple user interactions are supported by utilizing location data to adjust visualizations [24].

The training system was implemented using an OPTOTRAK position-tracking device [48], a FireWire head-mounted camera, and a camera marker. The user's head position was estimated using the location of the camera marker detected by the position-tracking device. With the provided estimations, the system would adjust guided visualizations for collaborators [24]. No user evaluation was provided on the system.

Duplicated Reality

Yu et al. [53] introduced an augmented reality system for facilitating surgical mentoring. Their design was capable of supporting multi-disciplinary expertise collaboration. Similar to the approach of Harders et al. [24], visualizations on users' augmented reality headsets would change based on their position and angle [53].

The system was implemented using two Microsoft Hololens head-mounted displays for overlaying visualizations and tracking user position. Users could manipulate visualizations using gestures similar to those used on smartphones, such as pinching to grab and expanding or contracting fingers to scale. The position of visualizations was calibrated using QR markers placed in the working area. Two Kinect cameras were attached to the ceiling for viewing the working area. Changes made to visualizations would be represented on both users' Hololens [53].

The system was evaluated through a user study to understand its effects on user awareness, workload, and task performance. Twenty-six participants were recruited, creating thirteen pairs of two. The study task involved two surgeons, an editor, and an expert surgeon, in which they manipulating kinetic sand in which visualizations overlaid on the actual physical sand presented the desired outcome. An editor would perform the physical actions on the simulated patient while the expert stood beside as a consultant. The study employed two conditions, a baseline condition in which collaborators are forced to work in close proximity and a duplicated reality condition in which users could annotate objects from a distance. Results from the study showed there were no significant differences in the performance of collaboration between conditions. However, there was a significant loss of awareness from the expert collaborator in the duplicated reality condition [53].

ANALYSIS

Number of Collaborators

The number of collaborators supported by each system depended on the devices employed and the context. For example, systems designed for construction and fabrication contexts allowed multiple users to pop in and out of tasks [20, 28, 29]. The dynamic nature of task assignments enabled the efficient reorganization of roles and responsibilities as the current working members changed. Alternatively, guidance systems for aircraft cockpits supported two static users, a pilot and a co-pilot [5, 13]. Roles were pre-established and remained consistent throughout the task.

Additionally, a few systems claimed to be able to support any number of collaborators [20, 21, 28, 29, 53]. With the exception of *Crowdsourced Fabrication* [28], all of them were evaluated on a specific number of people. *Crowdsourced Fabrication* [28] was evaluated on a large group collaboration of over 100 people. However, the crowd was spread out over a period of three days. Also, no information was provided on how many could be supported at a single time.

Collocated collaboration (CC) is inherently a social phenomenon with people valuing being near each other [42]. Tasks are often seen as social activities instead of systematic processes [30]. Team members are able to spontaneously partake in relation work, strengthening bonds between members [9]. While some guidance systems have looked into supporting the social aspects of CC [21, 27, 50], social interactions have been largely ignored.

Social interactions were not controlled as a variable in evaluating collocated collaborative guidance systems. Fraser et al. [20] noted how study results could have been affected by social dynamics by participants that knew each other. Participants in *Automatics* were strangers and potentially may have influenced the social and collaborative behaviors observed [29]. Fujita et al. [21] noted how participants might cluster in groups larger than six people. Additionally, in the user evaluation of the *Hapticturk* [7] system, which requires participants to perform

Overview of Systems				
System	# of Collaborators Supported (Observed)	User Interfaces	Collaborative Modeling	Interventions
Hapticturk [7]	5	Mobile Devices	None	None
Turkdeck [8]	10	Ambient display	None	None
Crowdsourced Fabrication [28]	∞ (108)	Ambient display, smartwatches	Location	Task distributions
WeBuild [20]	∞ (5)	Ambient display, mobile phone	User skills	Task distributions
Automatics [29]	∞ (2)	Tablet	User skill	Task distributions
Overseer [36]	∞	Mobile device	Audio, location, user skill, video	Task distributions
Crew Workload Manager [13]	2	Cockpit	Bio-sensors	Adaptive interface
Workload-Adaptive Support for Cockpit Crews [5]	2	Cockpit	None	Task distributions, adaptive interface
Healthcare Simulation Classroom [38]	4	Tablet	Audio, location, video	None
Meeting Mediator [27]	4	Mobile device	Audio, gestures	Feedback
Multimodal Support for Meetings [50]	4	Ambient display	Audio, eye-gaze	Feedback
Ambient Suite [21]	∞ (6)	Ambient displays, mobile device	Audio, gestures, location	Feedback
Multimodal Augmented Reality in Medicine [24]	2	Augmented Reality Headset	Location	Feedback
Duplicated Reality [53]	∞ (2)	Augmented Reality Headset	Gestures, location	Feedback

Table 1. Overview of Current Systems.

physical tasks with others, participants reported unfavorably in using the system with strangers. Participants preferred to collaborate with friends or families [7]. Prior work has shown that collaborative tasks are social interactions, and participants must be socially aware of each other's presence [30, 47].

Furthermore, group dynamics shift with the introduction of more members, with clusters potentially forming as a result. Consequently, the strategies for providing collaborative support in these diverse formations may change. The effects of the number of collaborators on group dynamics when providing guidance remains largely unstudied.

Interfaces

Collaborative systems facilitated group interactions through a variety of means, including mobile devices [7, 20, 21, 27, 28, 36], tablets [7, 29, 38], augmented-reality headsets [17, 24, 53], aircraft cockpits [5, 13], and ambient displays [8, 20, 21, 28, 50]. From the identified literature, some include an evaluation of their systems, comparing it to a baseline condition. Table 2 presents an overview of interfaces used in existing collaborative systems. Positive effects in each category corresponds to a "+" symbol and negative effects to a "-". Neutral effects to each category correspond to a "=".

Mobile devices included phones [7, 20, 27, 36], smartwatches [28], and an iPod Touch [7, 21]. Smartphones were often chosen due to their ubiquitous presence [20, 36] and ability to present more information compared to smaller devices such as smartwatches [20, 28]. Phones can also be used in the user's peripheral view, reducing the attention demanded by the device while being capable of providing private information [27]. On the other hand, smartwatches provide a less obtrusive and hands-free interface [28]. Overall, users found mobile devices to be agreeable. A post-study questionnaire for the evaluation of the *WeBuild* [20] system showed that users appreciated the system, finding the step-by-step task interface useful. Also, an evaluation of the *Meeting Mediator* showed that there was no significant difference in the amount of distraction caused by the mobile phone in the control and experimental conditions [27]. A post-study questionnaire for the participants of the *Crowdsourced Fabrication* [28] evaluation showed that users found the size and unobtrusiveness convenient, but the small screen a disadvantage.

Compared to mobile phones, tablets offer similar benefits, being ubiquitous [29] and relatively portable. Additionally, tablets can present more information due to their larger screen area [29]. In Lakier et al's [29] evaluation of *Automatics*, results showed that users completed significantly more tasks, required less mental effort, and felt the system was easier to use compared to paper instructions. Alternatively, in Martinez-Maldonado et al.'s [38] study, participants overlooked reporting their task progress on tablets because they were used to reporting it on paper.

Augmented reality (AR) devices allow multiple users to view guiding visualizations overlaid on the physical world while simultaneously completing complex tasks [17, 24, 53]. More advanced headsets, such as the Microsoft HoloLens 2, feature hand-tracking and allow users to edit visual guides [53]. Out of the literature identified, only Yu et al. [53] provided a user evaluation of their AR guidance system, and both of the study conditions involved using an AR device.

Aircraft cockpits may involve diverse specific interfaces, and no two are the same [44]. From the two guidance systems identified in the literature for aircraft cockpits [5, 13], both used entirely different simulations for their evaluation. The interface presented by Dorneich et al. [13] presented a system capable of adapting to pilots' workload through a feedback loop. Results from their system evaluation showed that it significantly reduced unbalanced workload states in pilot crews between both conditions and an efficient method of overcoming communication barriers. The system was also an open-loop system and would suggest adaptations to pilots. Alternatively, Brand et al. [5] employed a closed-loop system that would automatically adapt to the interface based on predictions of future high workload events. Brand et al. [5] only provided an exploratory study of their closed-loop system. Still, they noted that the interface might require a notice of adaptation by the system to inform the user of an incoming change.

Ambient displays employed by collaborative guidance systems included a dashboard display [20], lights on the physical objects [28], wall and floor displays [8, 21], and a tabletop visualization [50]. Despite being different, they shared similar traits in being on the peripheral of users attention and being made public to everyone partaking on a task. The difference in user perception can be seen in how the ambient displays were used. For example, the *WeBuild* [20] system employed a dashboard to present overall task progress. However, a user study showed mixed-results on its usefulness. Similarly, a post-study survey of the *Crowdsourced Fabrication* [28] system showed that users felt the LED-lights were helpful in guiding them through an assembly task but not when they relied on them for navigating the space. Alternatively, through user evaluations, participants rated the auditory and visual instructions of the *Turkdeck* [8] system as being useful. Results were mixed on the use of ambient displays for construction tasks. In a social setting, Fujita et al. [21] showed that participants preferred to view common interests versus viewing their current activity on wall displays. The combined use of floor displays to express common interests and the wall visualizations was effective in encouraging conversation. Alternatively, Terken et al. [50] showed that participants felt the tabletop visualizations were distracting and removed their attention away from the task. Again, results were mixed on the use of ambient displays.

Collaborative Modeling

The most common methods collaborative systems modeled users' behavior were through user skill [20, 29, 36, 54], location [17, 24, 28, 36, 38, 53], audio [21, 27, 36, 38, 50], video [36, 38], body gestures [27, 53], eye-tracking [50], and biosensors [13]. From the identified literature, some present an evaluation of their collaborative modeling techniques, comparing the accuracy of their ability to recognize specific events to a human observer or manual account from participants. However, most do not. Table 3 presents an overview of methods used for collaborative modelling in existing systems.

Comparison of Interfaces					
Interface	Information Display	Distraction	Portability	Availability	User Interaction
Mobiles Devices:					
Phones [7, 20, 27, 36]	+	+	+	+	+
Smartwatches [28]	-	+	+	=	-
Tablets [7, 29, 38]	+	=	+	+	=
Other:					
Augmented Reality Headsets [17, 24, 53]	+	=	+	=	=
Aircraft Cockpit [5, 13]	+	=	=	=	+
Ambient Displays [8, 20, 21, 28, 50]	=	-	-	-	+

Table 2. Positive effects in each category corresponds to a "+" symbol and negative effects to a "-". Neutral effects to each category correspond to a "=".

Overview of Collaborative Modelling				
Collaborative Modelling	Invasiveness	Availability	Privacy	Rep. of Col.
User Skill [20, 29, 36, 54]	+	+	+	=
Location [17, 24, 28, 36, 38, 53]	+	+	-	+
Audio [21, 27, 36, 38, 50],	-	+	=	+
Video [36, 38]	=	+	-	=
Body Gestures [21, 27, 53]	-	=	=	=
Eye-Tracking [50]	-	=	=	=
Bio-Sensors [13]	-	-	=	+

Table 3. Positive effects in each category corresponds to a "+" symbol and negative effects to a "-". Neutral effects to each category correspond to a "=".

Most systems that employed user skill implemented a method where users would input their prior skill in a certain area prior to beginning a task [20, 29, 36]. In *WeBuild* [20], a task tree was created at the beginning of the task and whenever a user was available, the task assigner would give them a new task depending on their past performance and predefined skills. A similar approach was implemented in *Automatics* [29]. In both *Automatics* [29] and *WeBuild* [20], user skills were dynamically updated as the task progressed. Also, the *Overseer* [36] system required users to enter their expertise in a specific emergency profession (e.g., firefighter, medical).

User location provided a point of reference for triggering contextual actions [28], orienting visualizations in an augmented reality headset [24, 53], and monitoring activity patterns [36, 38]. In the *Crowdsourced Fabrication* [28] system, users' location was used to inform the main foreman engine of the ongoing process. Consequently, results from a post-study questionnaire of the *Crowdsourced Fabrication* [28] system showed that users found the location tracking useful. Using location to orient visualizations in an augmented reality setting, Yu et al. [53] could monitor collaborator activity around the duplicated reality. Additionally, the *Overseer* [36] system employed user location activity to create a dependency graph and calculate the probability of task completion . Similarly, Martinez-Maldonado et al. [38] presented how location can be used to observe student activity around a mannequin and observed different approaches to the task using a heat map.

Audio can be used to gauge the intensity of tasks [36], track activity patterns [38], and monitor conversation states [21, 27, 50]. Martinez-Maldonado et al. [38] integrated directional audio with location data to map user behavior patterns during a task. One challenge with implementing audio devices was isolating voices in a noisy environment [38]. For monitoring social activity, Fujita et al. [21] did not detect significant differences in the speaking times of participants between their study conditions. Alternatively, Kim et al. [27] reliably detected a significant reduction in overlapping speaking times between collaborators in their study conditions. Also, Terken et al. [50] showed consistent behavior in participant speaking times in differing study conditions. Audio effectively detects social behavior and collaborative cohesion based on the literature we identified.

Video was used to gauge the users' attention on a task [36] and post-task reflection [38]. Both systems presented by Luqman et al. [36] and Martinez-Maldonado et al. [38] utilized video differently. For example, Luqman et al. [36] measured attentiveness based on a user's current attention to their phone, which was measured using the front-facing camera. Alternatively, Martinez-Maldonado et al. [38] employed audio and video to capture the entire collaborating group and create a heat map of their activity. One issue raised in their study was the privacy risk of collecting audio and video data in which they cannot be anonymized [38].

Body gestures were used to measure nonverbal cues [21], movement patterns [27], and manipulating virtual objects in an AR setting [53]. Fujita et al. [21] implemented gesture recognition using a 3D marker and an acceleration sensor. Results from their study showed that it was able to detect significant differences in movement between the two study conditions. Prior work has shown that an increase in nonverbal cues correlates to an increase in a more active conversation [25]. Alternatively, Kim et al. [27] also recorded participant body movement in a social task. However, they did not find any differences in body movement between dominant and non-dominant speakers. Yet, there was a significant difference detected in speech. The lack of difference in detecting significant states of dominance could be attributed to how Kim et al. [27] implemented their gesture recognition compared to Fujita et al [21]. Kim et al. [27] utilized a sociometric device that was placed around the user's neck, while Fujita et al. [21] had acceleration sensors that tracked arm movements.

Eye-tracking was used to measure participants' direction of focus in a task. For example, Terken et al. [50] used eye-tracking to measure participants attention to each other during a conversation task. Results from their study showed that eye-gaze was not an accurate measure of attention due to the fact that someone could be listening to a conversation yet not looking at them at the moment [50]. However, this could be attributed to the method in which their approach was implemented. Terken et al. [50] estimated user eye-gaze based on their current head orientation.

Dorneich et al. [13] employed elelectrocardiogram (ECG) and electroencephalogram (EEG) sensors to measure pilot and co-pilot workload states. The effectiveness of their approach was presented in a prior evaluation which proved the effectiveness of the approach in significantly detecting low and high workload states [15]. When applied to a collaborative setting, ECG and EEG sensors were shown to be effective in detecting unbalanced workload states between collaborators [13].

Interventions

Collaborative system interventions to triggers included task distributions [5, 13, 20, 28, 29, 36, 39, 54], adapting the interface [5, 13, 17, 24, 53], and providing feedback on the task [27, 50]. Here we present the feedback initiated by a system to alter group behavior. Table 4 presents an overview of interventions in existing collaborative systems. Table labelling follows the same convention as used in previous tables.

Task distributions were a prominent method for balancing workload among collaborators [5, 13, 20, 29, 36, 39, 54]. The *WeBuild* [20] system allowed groups to begin tasks immediately and significantly reduce task start-up times between conditions in a user study. However, the automatic task distribution appeared to reduce users' awareness of the overall task because they relied on the system to make decisions for them [20]. Tasks were reassigned whenever a user quit or finished a task. Users also complained of the task assigner distributing tasks inappropriately at times, placing free users into already ongoing tasks [20]. Alternatively, *Automatics* [29] employed a similar task distribution approach as the *WeBuild* [20] system, and users felt they had a good understanding of the overall task due to the task overview presented to them [29]. The *Automatics* [29] system also showed to prevent conflict since users didn't have to decide amongst themselves how to manage tasks. Comparatively, Dorneich et al. [13] observed their adaptive task system reduced unbalanced workload states. Their system also increased the number of sharing requests at the cost of increasing the number of incorrect sharing requests [13]. Brand et al. [5] also presented similar results, with their adaptive task system providing appropriate interventions 93% of the time. However, not all of the interventions were completely understood by the participants [5]. There appears to be a missing link between automatic adaptations and user understanding of the interventions in closed-loop systems. Automatic task reassignments at inappropriate times can lead to user misunderstanding of the situation [5] and awkward placements where the reassignments were more inconvenient than helpful [20].

Adapting user interfaces to triggering events was used to adjust guided visualizations [17, 24, 53] and reduce workload amongst collaborators [4, 13]. Yu et al. [53] implemented a system where virtual guides and 3D artifacts in an AR headset would change based on the position of the collaborator. Users could also physically manipulate visualizations using hand gestures. One side effect of the duplicated reality was the loss of awareness between collaborators since visualizations would occlude their view [53]. Dorneich et al. [13] presented continuous collaborator workload states and task recommendations when necessary. A post-study user evaluation showed that participants found the system easy to understand and aided in overcoming communication barriers [13]. Additionally, Brand et al. [5] implemented a system where the user interface was simplified to reduce workload imbalances. The system would automate tasks that didn't require complete involvement [5].

Real-time task feedback to participants was given to adjust social behaviors [21, 27, 50]. Fujita et al. [21] showed that participants could understand the changes in the ambient displays and were a stimulus to group conversation participation. Visualizing gaze not as effective. Additionally, Kim et al. [27] found that visualizing group dominance as feedback changed group dynamics and created a friendlier environment. Based on the identified literature, presenting individual collaborators' feedback on their behavior can balance group dynamics.

DISCUSSION

In the last section, we presented an analysis of the literature in regards to guidance systems facilitating collaboration, analyzing and responding to collaborator groups, and improving group efficiency. Here we summarize our analysis and the implications of our findings. Collaborative guidance systems have guided user through assembly [8, 28, 29], medical [38, 53], navigational [5, 13], emergency [36], and social tasks [7, 27, 50].

Guidance systems are able to support a dynamic number of users, allowing for groups to interweave as tasks progress [20, 28, 29]. The dynamic movement of people allows for error recovery [29] and supporting large crowds [28]. Moreover, systems can adapt as the number of collaborators increases [28]. Consequently, group dynamics become more complex as the number of collaborators increases [21]. Most systems

Overview of Interventions				
Intervention	Awareness	Perception	Workload	Altering Behavior
Task Distributions [5, 13, 20, 28, 29, 36, 39, 54]	=	-	+	-
Adapting the Interface [5, 13, 17, 24, 53]	-	+	+	=
Feedback [21, 27, 50]	=	-	=	+

Table 4. Positive effects in each category corresponds to a "+" symbol and negative effects to a "-". Neutral effects to each category correspond to a "="‘.

identified in the literature were not controlled for social dynamics in groups when they were evaluated. Social relations between group members can affect their performance on tasks [29]. Additionally, users can even take on a competitive attitude to public displays of feedback [50]. The strategies for supporting collaboration may change as group size increases and dynamics become more complex. **Future work in collaborative guidance should investigate how to manage larger groups, considering their social dynamics remains largely unconsidered.**

Regardless of their context, systems shared similar user interfaces for facilitating collaboration. Mobile devices were favored for their widespread use and multimodal features [20, 21, 27, 28, 36]. Using a familiar device allows users to focus quickly on the task instead of learning to manage a new device. As shown by Martinez-Maldonado et al. [38], implementing a system with a device users are not familiar with can lead to them completely ignoring it. Moreover, results were mixed on the effectiveness of ambient displays. While ambient displays may effectively alter group behavior [21], they may become distracting [50] and disorienting [28]. Additionally, the devices employed in collaborative guidance are restricted to the context in which they are applied. Systems are limited by the number of devices that can be used in a single moment [53]. **Future work should consider what combination of interfaces best suite each context.** For example, in applications where a large number of collaborators are popping in and out, teams may benefit more from smaller devices that display quick information and less from interfaces that demand attention. However, a context with a static number of collaborators may benefit more from a system that presents more complex and situated information.

Systems monitored the ongoing collaboration by assessing user skills [20, 29, 36, 54], tracking their location [17, 24, 28, 36, 38, 53], listening to their conversations [21, 27, 36, 38, 50], capturing video recordings [36, 38], analyzing body gestures [21, 27, 53], tracking eye-gaze [50], and gauging workload through bio-sensors [13]. Determining which collaborative modelling methods are effective compared to others was difficult since most systems did not directly evaluate their techniques. Instead, they evaluated their systems as a whole. Additionally, which kinds of devices work better than others is completely contextual. Regardless, we have observed methods that have repeatedly shown success in tracking unique collaborative behavior. For example, audio has been shown to be a great indicator of social behavior and participation [27, 50]. Additionally, body gestures have been shown to be another indicator of social behavior [21], and prior work has correlated an increase in nonverbal gestures to an increase in more active conversations [25]. **Future work should investigate ways of representing collaborative modeling and determining whether it is an accurate representation of current group dynamics.**

After modeling collaborative modeling behavior, systems would implement some method of feedback or intervention to alter group behavior. Task distributions were a popular theme for organizing groups of collocated collaborators. Groups were assigned roles and responsibilities and could begin tasks as soon as they received a notice. Automatic distributions would allow for quick start-ups [20], reduce conflict among collaborators [29], and reduce the amount of time spent in unbalanced workload states [5, 13]. Furthermore, adaptive interfaces and real-time task feedback interfaces were shown to be beneficial in guidance. Adaptive interfaces simplified the presented content to guide users through high workload situations and allowed them to focus on the most physical tasks [5]. Additionally, real-time feedback can balance group dynamics in a social setting [21, 27, 50]. However, without having a good understanding of the current context, interventions cannot work effectively [5, 13, 20]. **Future work in collaborative guidance systems need an accurate assessment of current group dynamics in order to provide assistance at opportune times.**

In regards to improving group efficiency, guidance systems have shown promising results in aiding groups through various tasks, but improvements have come at a cost. For example, task distributions allow for parallel work and quicker start-up times but do not necessarily lead to more tasks being completed [20]. Furthermore, Dorneich et al. [13] showed that monitoring crew workload could be used to reduce unbalanced workload states but at the cost of an increase in inappropriate interventions. And presenting social feedback to individuals can balance group dynamics while also being distracting to the task itself [27]. Finally, the effectiveness of system interventions relies on appropriate timing and the collaborative modeling techniques utilized. **Future work in improving group efficiency using collaborative guidance systems need to investigate the trade-offs between methods in interfaces, collaborative modeling, and interventions.**

LIMITATIONS

The main limitation to our literature review would be the use of Google Scholar for accessing papers. Although sixty-two papers were found, our results were limited to the first ten pages accessed. Nevertheless, our method resulted in an extensive literature review across various contexts.

CONCLUSION

We investigated the intersections between collaborative modeling and adaptive task guidance to understand how it facilitates collocated collaboration, the techniques used to monitor and respond to context cues, and ways to improve efficiency. Our contribution is a cross-analysis of current work and a presentation of current limitations in systems and open questions. While existing systems and methods have shown promising results in modeling group dynamics and altering group behavior, improvements have been marginal in select contexts. More work is needed in understanding which collaborative modeling methods and interventions best suit each unique context.

ACKNOWLEDGMENTS

I would like to thank the members of the Ruiz HCI Lab for their unending support and guidance throughout the literature review. This work is partially supported by the National Science Foundation Grant Award #IIS-1750840. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect these agencies' views.

REFERENCES

[1] Thomas Bayes. 1968. Naive bayes classifier. *Article Sources and Contributors* (1968), 1–9.

[2] Nancy Baym, Limor Shifman, Christopher Persaud, and Kelly Wagman. 2019. Intelligent failures: Clippy memes and the limits of digital assistants. *AoIR Selected Papers of Internet Research* (2019).

[3] Jonas Blattgerste, Patrick Renner, and Thies Pfeiffer. 2019. Augmented reality action assistance and learning for cognitively impaired people: a systematic literature review. In *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. 270–279.

[4] Yannick Brand and Axel Schulte. 2017. Model-based prediction of workload for adaptive associate systems. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 1722–1727.

[5] Yannick Brand and Axel Schulte. 2021. Workload-adaptive and task-specific support for cockpit crews: design and evaluation of an adaptive associate system. *Human-Intelligent Systems Integration* 3, 2 (2021), 187–199.

[6] Ewan Branda. 2015. Oculus Rift. (2015).

[7] Lung-Pan Cheng, Patrick Lühne, Pedro Lopes, Christoph Sterz, and Patrick Baudisch. 2014. Haptic turk: a motion platform based on people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3463–3472.

[8] Lung-Pan Cheng, Thijs Roumen, Hannes Rantzsch, Sven Köhler, Patrick Schmidt, Robert Kovacs, Johannes Jasper, Jonas Kemper, and Patrick Baudisch. 2015. Turkdeck: Physical virtual reality based on people. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. 417–426.

[9] Lars Rune Christensen, Rasmus Eskild Jensen, and Pernille Bjørn. 2014. Relation work in collocated and distributed collaboration. In *COOP 2014-Proceedings of the 11th International Conference on the Design of Cooperative Systems, 27-30 May 2014, Nice (France)*. Springer, 87–101.

[10] J Raymond Comstock Jr and Ruth J Arnegard. 1992. *The multi-attribute task battery for human operator workload and strategic behavior research*. Technical Report.

[11] Joan Morris DiMicco, Anna Pandolfo, and Walter Bender. 2004. Influencing group participation with a shared display. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*. 614–623.

[12] Michael C Dorneich, Bretislav Passinger, Christopher Hamblin, Claudia Keinrath, Jiři Vašek, Stephen D Whitlow, and Martijn Beekhuyzen. 2011. The crew workload manager: an open-loop adaptive system design for next generation flight decks. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 55. SAGE Publications Sage CA: Los Angeles, CA, 16–20.

[13] Michael C Dorneich, Břetislav Passinger, Christopher Hamblin, Claudia Keinrath, Jiři Vašek, Stephen D Whitlow, and Martijn Beekhuyzen. 2017. Evaluation of the display of cognitive state feedback to drive adaptive task sharing. *Frontiers in neuroscience* 11 (2017), 144.

[14] Michael C Dorneich, William Rogers, Stephen D Whitlow, and Robert DeMers. 2016. Human performance risks and benefits of adaptive systems on the flight deck. *The International Journal of Aviation Psychology* 26, 1-2 (2016), 15–35.

[15] Michael C Dorneich, Stephen D Whitlow, Santosh Mathan, Patricia May Ververs, Deniz Erdogmus, Andre Adami, Misha Pavel, and Tian Lan. 2007. Supporting real-time cognitive state classification on a mobile individual. *Journal of Cognitive Engineering and Decision Making* 1, 3 (2007), 240–270.

[16] Michael Engle. 2000. Culture in the Cockpit-CRM in a Multicultural World. *Journal of Air Transportation World Wide* 5, 1 (2000).

[17] Wei Fang, Wei Fan, Wei Ji, Lei Han, Shuhong Xu, Lianyu Zheng, and Lihui Wang. 2022. Distributed cognition based localization for AR-aided collaborative assembly in industrial environments. *Robotics and Computer-Integrated Manufacturing* 75 (2022), 102292.

[18] JD Fletcher and Robert A Sottilare. 2018. Shared mental models in support of adaptive instruction for teams using the GIFT tutoring architecture. *International Journal of Artificial Intelligence in Education* 28, 2 (2018), 265–285.

[19] Brian J Fogg. 1998. Persuasive computers: perspectives and research directions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 225–232.

[20] C Ailie Fraser, Tovi Grossman, and George Fitzmaurice. 2017. Webuild: Automatically distributing assembly tasks among collocated workers to improve coordination. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 1817–1830.

[21] Kazuyuki Fujita, Yuichi Itoh, Hiroyuki Ohsaki, Naoaki Ono, Keiichiro Kagawa, Kazuki Takashima, Sho Tsugawa, Kosuke Nakajima, Yusuke Hayashi, and Fumio Kishino. 2011. Ambient suite: enhancing communication among multiple participants. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*. 1–8.

[22] Neil Gershenfeld. 2012. How to make almost anything: The digital fabrication revolution. *Foreign Aff.* 91 (2012), 43.

[23] Juho Hamari, Jonna Koivisto, and Tuomas Pakkanen. 2014. Do persuasive technologies persuade?-a review of empirical studies. In *Persuasive Technology: 9th International Conference, PERSUASIVE 2014, Padua, Italy, May 21-23, 2014. Proceedings 9*. Springer, 118–136.

[24] Matthias Harders, Gerald Bianchi, and Benjamin Knoerlein. 2007. Multimodal augmented reality in medicine. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, 652–658.

[25] Kosuke Kajimura, Kazuki Takashima, Tokuo Yamaguchi, Yoshifumi Kitamura, Fumio Kishino, Nao Maeda, Ken Fujiwara, Hitomi Yokoyama, and Ikuo Daibo. 2010. A Comparison of self and third party evaluation for atmosphere in triadic conversation. *IEICE Technical Report; IEICE Tech. Rep.* 110, 185 (2010), 43–48.

[26] Staffs Keele and others. 2007. *Guidelines for performing systematic literature reviews in software engineering*. Technical Report. Technical report, Ver. 2.3 EBSE Technical Report. EBSE.

[27] Taemie Kim, Agnes Chang, Lindsey Holland, and Alex Sandy Pentland. 2008. Meeting mediator: enhancing group collaborationusing sociometric feedback. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. 457–466.

[28] Benjamin Lafreniere, Tovi Grossman, Fraser Anderson, Justin Matejka, Heather Kerrick, Danil Nagy, Lauren Vasey, Evan Atherton, Nicholas Beirne, Marcelo H Coelho, and others. 2016. Crowdsourced fabrication. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. 15–28.

[29] Matthew Lakier, Michelle Annett, and Daniel Wigdor. 2018. Automatics: Dynamically generating fabrication tasks to adapt to varying contexts. *ACM Transactions on Computer-Human Interaction (TOCHI)* 25, 4 (2018), 1–44.

[30] Andreas Larsson. 2003. Making sense of collaboration: the challenge of thinking together in global design teams. In *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*. 153–160.

[31] Morgan Le Chénéchal, Thierry Duval, Valérie Gouranton, Jérôme Royan, and Bruno Arnaldi. 2016. Vishnu: virtual immersive support for helping users an interaction paradigm for collaborative remote guiding in mixed reality. In *2016 IEEE Third VR International Workshop on Collaborative Virtual Environments (3DCVE)*. IEEE, 9–12.

[32] Teesid Leelasawassuk, Dima Damen, and Walterio Mayol-Cuevas. 2017. Automated capture and delivery of assistive task guidance with an eyewear computer: the glaciar system. In *Proceedings of the 8th Augmented Human International Conference*. 1–9.

[33] Edmund LoPresti, Ned Kirsch, Richard Simpson, and Debra Schreckenghost. 2005. Solo: interactive task guidance. In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*. 190–191.

[34] Edmund F LoPresti, Richard C Simpson, Ned Kirsch, Debra Schreckenghost, and Stephen Hayashi. 2008. Distributed cognitive aid with scheduling and interactive task guidance. *Journal of Rehabilitation Research & Development* 45, 4 (2008).

[35] Sus Lundgren, Joel E Fischer, Stuart Reeves, and Olof Torgersson. 2015. Designing mobile experiences for collocated interaction. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 496–507.

[36] Faisal Luqman and Martin Griss. 2010. Overseer: a mobile context-aware collaboration and task management system for disaster response. In *2010 Eighth International Conference on Creating, Connecting and Collaborating through Computing*. IEEE, 76–82.

[37] Roberto Martinez-Maldonado, Peter Goodyear, Judy Kay, Kate Thompson, and Lucila Carvalho. 2016. An actionable approach to understand group experience in complex, multi-surface spaces. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 2062–2074.

[38] Roberto Martinez-Maldonado, Tamara Power, Carolyn Hayes, Adrian Abdiprano, Tony Vo, Carmen Axisa, and Simon Buckingham Shum. 2017. Analytics meet patient manikins: Challenges in an authentic small-group healthcare simulation classroom. In *Proceedings of the seventh international learning analytics & knowledge conference*. 90–94.

[39] Niels Netten, Guido Bruinsma, Maarten Van Someren, and Robert De Hoog. 2006. Task-adaptive information distribution for dynamic collaborative emergency response. *The International Journal of Intelligent Control and Systems* 11, 4 (2006), 238–247.

[40] Thomas Neumayr and Mirjam Augstein. 2020. A Systematic Review of Personalized Collaborative Systems. *Frontiers in Computer Science* (2020), 43.

[41] Jennifer Ockerman and Amy Pritchett. 2000. A review and reappraisal of task guidance: Aiding workers in procedure following. *International Journal of Cognitive Ergonomics* 4, 3 (2000), 191–212.

[42] Judith S Olson, Stephanie Teasley, Lisa Covi, and Gary Olson. 2002. The (Currently) Unique Advantages of Collocated. *Distributed work* (2002), 113.

[43] Thomas Olsson, Pradthana Jarusriboonchai, Paweł Woźniak, Susanna Paasovaara, Kaisa Väänänen, and Andrés Lucero. 2020. Technologies for enhancing collocated social interaction: review of design solutions and approaches. *Computer Supported Cooperative Work (CSCW)* 29, 1 (2020), 29–83.

[44] Robert J Ontiveros, Roman M Spangler, and Richard L Sulzer. 1978. *General aviation (FAR 23) cockpit standardization analysis*. Technical Report. US Department of Transportation, Federal Aviation Administration, Systems ....

[45] Jin-Man Park, Jae-Hyuk Jang, Sahng-Min Yoo, Sun-Kyung Lee, Ue-Hwan Kim, and Jong-Hwan Kim. 2021. ChangeSim: towards end-to-end online scene change detection in industrial indoor environments. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 8578–8585.

[46] Sambit Praharaj, Maren Scheffel, Hendrik Drachsler, and Marcus M Specht. 2021. Literature Review on Co-Located Collaboration Modeling Using Multimodal Learning AnalyticsCan We Go the Whole Nine Yards. *IEEE Transactions on Learning Technologies* (2021).

[47] Iulian Radu, Tugce Joy, Yiran Bowman, Ian Bott, and Bertrand Schneider. 2021. A Survey of Needs and Features for Augmented Reality Collaborations in Collocated Spaces. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–21.

[48] Jill Schmidt, Devin R Berg, Heidi-Lynn Ploeg, and Leone Ploeg. 2009. Precision, repeatability and accuracy of Optotrak® optical motion tracking systems. *International Journal of Experimental and Computational Biomechanics* 1, 1 (2009), 114–127.

[49] Greis Francy M Silva-Calpa, Alberto B Raposo, and Francisco R Ortega. 2021. Collaboration support in co-located collaborative systems for users with autism spectrum disorders: A systematic literature review. *International Journal of Human–Computer Interaction* 37, 1 (2021), 15–35.

[50] Jacques Terken and Janienke Sturm. 2010. Multimodal support for social dynamics in co-located meetings. *Personal and Ubiquitous Computing* 14, 8 (2010), 703–714.

[51] Zhuo Wang, Xiaoliang Bai, Shusheng Zhang, Mark Billinghurst, Weiping He, Peng Wang, Weiqi Lan, Haitao Min, and Yu Chen. 2022. A comprehensive review of augmented reality-based instruction in manual assembly, training and repair. *Robotics and Computer-Integrated Manufacturing* 78 (2022), 102407.

[52] Claes Wohlin. 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th international conference on evaluation and assessment in software engineering*. 1–10.

[53] Kevin Yu, Ulrich Eck, Frieder Pankratz, Marc Lazarovici, Dirk Wilhelm, and Nassir Navab. 2022. Duplicated Reality for Co-located Augmented Reality Collaboration. *IEEE Transactions on Visualization and Computer Graphics* 28, 5 (2022), 2190–2200.

[54] An Zhang, Wenhao Bi, Zhili Tang, and Chao Zhang. 2016. Human-machine function allocation method for aircraft cockpit based on interval 2-tuple linguistic information. *Journal of Systems Engineering and Electronics* 27, 6 (2016), 1291–1302.