



OPEN ACCESS

EDITED BY

Pradeep K. Murukannaiah,
Delft University of Technology, Netherlands

REVIEWED BY

Angelo Cangelosi,
The University of Manchester, United Kingdom
Francesco Semeraro,
The University of Manchester, United
Kingdom, in collaboration with reviewer AC
Myrthe Lotte Tielman,
Delft University of Technology, Netherlands
Carolina Centeio Jorge,
Delft University of Technology, Netherlands, in
collaboration with reviewer MT

*CORRESPONDENCE

Maitreyee Tewari
✉ maittewa@cs.umu.se

[†]These authors have contributed equally to
this work and share first authorship

RECEIVED 27 June 2024

ACCEPTED 28 October 2024

PUBLISHED 13 November 2024

CITATION

Kaelin VC, Tewari M, Benouar S and
Lindgren H (2024) Developing teamwork:
transitioning between stages in human-agent
collaboration. *Front. Comput. Sci.* 6:1455903.
doi: 10.3389/fcomp.2024.1455903

COPYRIGHT

© 2024 Kaelin, Tewari, Benouar and Lindgren.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Developing teamwork: transitioning between stages in human-agent collaboration

Vera C. Kaelin^{1†}, Maitreyee Tewari^{1*†}, Sara Benouar ¹ and
Helena Lindgren^{1,2}

¹Laboratory of Interactive and Intelligent Systems, Department of Computing Science, Umeå
University, Umeå, Sweden, ²Department of Community Medicine and Rehabilitation, Umeå University,
Umeå, Sweden

Introduction: Human-centric artificial intelligence (HCAI) focuses on systems that support and *collaborate* with humans to achieve their goals. To better understand how collaboration develops in human-AI teaming, further exploration grounded in a theoretical model is needed. Tuckman's model describes how team development among humans evolves by transitioning through the stages of forming, storming, norming, performing, and adjourning. The purpose of this pilot study was to explore transitions between the first three stages in a collaborative task involving a human and a human-centric agent.

Method: The collaborative task was selected based on commonly performed tasks in a therapeutic healthcare context. It involved planning activities for the upcoming week to achieve health-related goals. A calendar application served as a tool for this task. This application embedded a collaborative agent designed to interact with humans following Tuckman's stages of team development. Eight participants completed the collaborative calendar planning task, followed by a semi-structured interview. Interviews were transcribed and analyzed using inductive content analysis.

Results: The results revealed that the participants initiated the storming stage in most cases ($n = 7/8$) and that the agent initiated the norming stage in most cases ($n = 5/8$). Additionally, three main categories emerged from the content analyses of the interviews related to participants' transition through team development stages: (i) participants' experiences of Tuckman's first three stages of team development; (ii) their reactions to the agent's behavior in the three stages; and (iii) factors important to the participants to team up with a collaborative agent.

Conclusion: Results suggest ways to further personalize the agent to contribute to human-agent teamwork. In addition, this study revealed the need to further examine the integration of explicit conflict management into human-agent collaboration for human-agent teamwork.

KEYWORDS

human-agent teaming, human-AI collaboration, Tuckman's model, human-centered artificial intelligence, Activity Theory, health promotion, activities of daily living, occupational therapy

1 Introduction

Human-centric artificial intelligence (HCAI) focuses on developing systems that effectively *collaborate* with humans to achieve their goals (Nowak et al., 2018). Collaboration, in this context, can be defined as “a process of joint decision-making among interdependent parties, involving joint ownership of decisions and collective responsibility for outcomes” (Gray, 1989, as cited in) (Liedtka et al., 1998, p. 186). It is closely related to teamwork (Iftikhar et al., 2023), which refers to “the integration of individuals' efforts toward the accomplishment of a shared goal” (Mathieu et al., 2017, p. 458).

Human-AI teaming (HAT) has gained significant attention over the past years due to the rapid advancement of AI system's capabilities, which support their direct implementation in workplaces (Berretta et al., 2023; Prada et al., 2024). HAT is believed to facilitate the achievement of complex goals (Kluge et al., 2021) in effective, safe, and ethical ways (Berretta et al., 2023; Naser and Bhattacharya, 2023) by leveraging the unique and complementary strengths of AI and humans in jointly reasoning about and addressing task challenges.

HAT is particularly crucial in high-sensitivity, critical and complex domains such as healthcare where solutions are not straightforward (Berretta et al., 2023). The emphasis on AI systems has primarily been on task performances rather than team development process (Prada et al., 2024). However, research indicates that a team's collective intelligence is more influenced by the team dynamics than by the intelligence of its individuals (Charas, 2015; Woolley et al., 2010). There are mixed perspectives on whether HAT develops similarly to human-only teams (Prada et al., 2024; McNeese et al., 2021). Research has shown both similarities and differences between human-only and human-AI teams. For example, the iterative development of team cognition and the importance of communication appear to be similar in both types of team (Schelble et al., 2022), whereas sharing of mental models has been inconsistent, and trust in AI team members was generally lower, with communication quality often worse compared to human-only teams (Schmutz et al., 2024; Schelble et al., 2022). Nonetheless, a recent study found that humans tend to expect AI team members to exhibit human-like behaviors (Zhang et al., 2021), making traditional human-only teamwork models a promising starting point for investigating the different stages of team development in HAT.

Tuckman's model for team development is one such model (Tuckman, 1965), extensively studied and applied (Bonebright, 2010). According to this model, team development evolves by human team members transitioning through stages denoted as *forming*, *storming*, *norming*, *performing*, and *adjourning* (Tuckman, 1965). Forming, storming, and norming stages are arguably the most challenging and important stages for team development, as they set the foundation for the team's trajectory. The storming stage, characterized by tension and conflicts among human team members, is considered essential for teams to experience before they can engage in constructive teamwork (Tuckman, 1965). Thus, the transitions into and out of the storming stage are vital for effective team development.

Recent research in human-computer interaction has explored the introduction and resolution of tensions and conflicts in HAT, such as in the context of disobedient and rebellion AI (Mayol-Cuevas, 2022; Aha and Coman, 2017). This research examines AI systems that prevent and correct human errors by disagreeing with humans or, instead of complying with human intentions, fulfill overall goals by performing other sub-tasks (Mayol-Cuevas, 2022; Aha and Coman, 2017). Such disagreements can introduce tensions and conflicts in HAT, similar to challenges faced during the storming stage of Tuckman's model. However, the perception of these tensions and conflicts in the context of HAT development is not yet well understood. This is supported by a recent literature review by Iftikhar et al. (2023) which mapped 101 studies

about HAT to an all-human teamwork framework by Mathieu et al. (2019), highlighting a lack of research on human-AI team development in general and particularly on conflict resolution in HAT development processes. Understanding the stages of HAT development, including potential conflicts related to it, could provide insights into how AI systems might best navigate and support HAT to reach effective human-AI teamwork (Prada et al., 2024).

The purpose of the pilot study presented in this paper was to explore the transitions between the first three stages of forming, storming, and norming from Tuckman's model in a collaborative task involving a human and a human-centric agent which we refer to as *agent* throughout the rest of the article.

A scenario of HAT development was created in the context of healthcare, as one of the domains with high relevance for HAT (Berretta et al., 2023). We based our scenario for HAT on how a therapist (e.g., occupational therapist or psychologist) typically engages in a dialogue with a patient to plan activities for the upcoming week to achieve health-related goals. Following Tuckman's first three stages for team development, the patient and therapist start by getting to know each other (forming) and then move into a stage where conflicts and friction arise (storming) (Tuckman, 1965). After overcoming these conflicts, teams become more productive in performing tasks (norming) (Tuckman, 1965).

We explore the following research questions:

1. In what ways do participants and the agent transition through Tuckman's first three stages during the collaborative calendar planning activity?
2. How do participants experience the collaborative calendar planning activity with an agent designed to simulate transition through Tuckman's first three stages of team development?

To address our research questions we developed an agent who jointly plans activities with a patient aimed at improving their health and well-being. The agent may contribute with motives from a health perspective, mediating some health domain knowledge. A calendar application embedding an agent was developed, where the agent was designed to act as a team member transitioning through the forming, storming, and norming stages.

We make the following contributions:

1. A system embedding an agent that enacts strategies to support the development of HAT based on Tuckman's stages.
2. Insights on how participants transitioned between Tuckman's stages of team development and how they experienced the collaborative calendar planning activity.
3. Insights into how humans perceive conflicts within human-AI team development stages.

In the following, we describe the theoretical frameworks (Section 2) foundational to this work and outline the related work in Section 3. Section 4 introduces the methodology applied for developing the calendar application embedding the agent and conducting the empirical study. The results of the study are presented in Section 5 and discussed in Section 6, followed by conclusions in Section 7.

2 Background: theoretical frameworks

According to Tuckman's model, teamwork evolves as actors transition through various stages of team development (Tuckman, 1965). These stages are known as forming, storming, norming, performing, and adjourning. In the forming stage, the team familiarizes itself with the task and establishes initial relationships. After this forming stage, the team transitions into the storming stage, where conflicts may arise as team members resist the formation of a group structure. In this stage, emotional responses to task demands can become evident, particularly in tasks involving self-awareness and personal change, as opposed to impersonal and intellectual tasks. Moving into the norming stage, team members begin to accept each other's behavior, marking the emergence of a cohesive team. Finally, in the performing stage, the team optimizes its structure, resulting in enhanced task completion (Tuckman, 1965). In a revised version of the model, an adjourning stage was added, which reflects the separation of a team (Bonebright, 2010). Given the focus of this paper related to HAT transition into and out of the storming stage, this work specifically studies the three initial stages (i.e., forming, storming, and norming) of the team development process.

Activity Theory has evolved over the past 90 years as a systemic model of human activity that is inherently social in nature and continuously transforms as humans develop competence and skills within the social and cultural environment (Vygotsky, 1978; Kaptelinin and Nardi, 2006; Engeström, 1999). It describes human activity as purposeful, transformative, and evolving interaction between actors and the objective world through the use of mediating tools (Kaptelinin and Nardi, 2006). Cultural-Historical Activity Theory (CHAT) developed by Engeström (1999) expands upon this model to encompass social conditions, such as rules and norms and division of labor or work tasks, within a collective activity perspective. These components, including actors, community/team, tools, the object of activity, rules, and division of labor, form an activity system that generates *outcomes*. Interacting activity systems are typically interdependent, where the outcome of one activity system feeds into another, for example, the outcome of a design and development process such as an AI-based digital tool or a more knowledgeable professional returning from an educational activity.

According to Activity Theory, an activity is conducted through goal-oriented actions, often structured as a hierarchy of interdependent tasks (Kaptelinin and Nardi, 2006), where automated tasks such as communication or movement are performed without conscious thought. However, contradictions can arise in human activities, leading to conflicts within the actor(s), within an activity system, and between activity systems. For instance, a breakdown may occur when a concept is unfamiliar to an actor during a conversation. These contradictions can disrupt the activity and must be resolved before the activity can continue. In such cases, a transformation from one activity (e.g., task performance of a team) to another activity (e.g., team development) is necessary, often requiring automated operations to be brought to a conscious level to address the breakdown situation. Such transformations are considered the driving force

behind development of the activity, actors and other activity system components, including team development. Tuckman's (1965) model conceptualizes these contradictions that lead to team development, aligning with and complementing Engeström's CHAT. The combination of these two models forms the theoretical basis for the study presented in this article.

3 Related work

While the concept of HAT has been explored for several decades (McNeese, 1986), recent years have seen a significant increase in research on efficient and productive HAT (O'Neill et al., 2022; Iftikhar et al., 2023). Research closest related to investigating the stages of team development in HAT—the focus of this study - primarily explored relevant team characteristics (e.g., their shared mental models, communication, and trust) as well as human-AI conflicts within the context of HAT. In this section, we will first summarize studies related to these two aspects of HAT, outline our contributions compared to the found literature, followed by an overview of the application of Tuckman's model to study team development among humans as well as how our study adds to this body of literature.

3.1 Team characteristics in HAT

Research on team characteristics in HAT has examined shared mental models between humans and AI systems, trust in AI-teammate and/or human-AI teams, and perceptions of team communication in HAT. These characteristics were primarily investigated through experiments comparing different types of teams (e.g., human-only teams versus human-AI teams), with evaluations often based on survey data. For example, Schelble et al. (2022) compared teamwork across three scenarios: a triad of humans, two humans and an agent, and two agents and a human, all engaging in a team simulation activity called NeoCITIES. To evaluate teamwork, the authors compared the team's shared mental model using paired sentence comparison and the Pathfinder network scaling algorithm, as well as perceived team cognition, performance, and trust using surveys. Findings revealed similar iterative development of team cognition and importance of communication across humans-only and human-agent/s teams, however, human-agent teams explicitly communicated actions and shared goals. Trust in agent teammates was lower and shared mental models were less consistent in human-agent teams. Design implications for improving human-agent teamwork included: (1) the agent should act as an *exemplar* early in the teamwork, demonstrating how to achieve the goal effectively; (2) the agent teammate/s should provide concise updates on completed, ongoing, and upcoming actions; and (3) the agent teammates should clearly and concisely communicate its individual goals at the beginning of the team formation and emphasize how these align with the other teammates' goals (Schelble et al., 2022).

Walliser et al. (2019) conducted two experiments to explore whether perceiving an autonomous agent as a tool or a teammate and whether a team-building intervention could improve teamwork outcomes (i.e., team affects, team behavior,

team effectiveness). In the first experiment, participants were paired with either a human or an autonomous agent and performed a task under different impressions—either viewing the agent as a teammate or as a tool. Data were collected mainly through surveys capturing aspects of team affect (e.g., perceived interdependence), team behavior (e.g., team communication), and team effectiveness (e.g., satisfaction). The results showed that perceiving the autonomous agent as a teammate increased team affect and behavior but did not enhance team effectiveness. In the second experiment, the researchers introduced a team-building intervention, which included a formal role clarification and goal-setting exercise before task performance. This intervention significantly improved all teamwork outcomes, including team effectiveness for both human-only and human-autonomous agent teams.

Taken together, existing research on team characteristics in HAT primarily focuses on comparing human-only and human-AI teams in terms of relevant team characteristics through quantitative (quasi-)experimental study designs, mainly using survey data. While the results of these studies provide valuable new knowledge on differences between team types, they do not provide in-depth knowledge on how stages of team development in HAT may evolve and be perceived. Our study addresses this gap through qualitative interviews, which are often used to gain in-depth understanding and are particularly useful for generating knowledge in underexplored areas (Iftikhar et al., 2023; Prada et al., 2024), such as the one investigated in this paper.

3.2 Human-AI conflicts in HAT

Research on conflicts in HAT has mainly focused on differences in the number of conflicts and the effects of conflict resolution across different types of teams (i.e., human-only teams vs. human-AI teams). For example, McNeese et al. (2021) explored team situation awareness (TSA) and team conflict in human-machine teams operating remotely piloted aircraft systems (RPAS) where they needed to photograph critical way-points and overcome roadblocks. Teams composed of members taking over the role of a pilot, navigator, and photographer were tested in three team compositions: all-human teams, teams with a synthetic agent as the pilot, and teams with an experimenter simulating an effective synthetic pilot with all other team members being humans. The task involved RPAS navigation and communication through text. Data for TSA were captured via a task-specific measure (i.e., the proportion of roadblocks per-mission that needed to be overcome, team interaction across the team members), and teams' conflicts between photographers and navigators were measured via the numbers of conflicts obtained from text analyses. Results showed that TSA improved in synthetic and experimenter teams but remained unchanged in all-human teams, which also experienced the highest number of conflicts. The study concluded that effective synthetic agents could enhance TSA and reduce conflict, improving overall team performance.

Jung et al. (2015) aimed to explore whether robot agents could positively influence team dynamics by repairing team conflicts through emotion regulation strategies during a team-based problem-solving task. The teams, consisting of humans

and the robot agent, were subject to conflicts introduced by a confederate through task-related and personal attacks. The robot agent, using Wizard of Oz, suggested common emotional regulation strategies to help repair these conflicts. Four different scenarios were compared: task-directed attacks repaired by the robot agent, task-directed attacks not repaired, task-directed and personal attacks repaired by the robot agent, and task-directed and personal attacks not repaired. Data on team affect, perceived conflict, and the confederate's contribution were captured via surveys. The outcome on team performance was captured via a task-specific measure (i.e., number of moves the team made within 10 min). Results revealed that the robot agent's repair interventions heightened the group's awareness of conflict after personal attacks, counteracting humans' tendency to suppress conflict. This suggests that robot agents can aid in team conflict management, by making violations more salient and enhancing team functioning. The study indicates that agents can effectively contribute to teamwork by addressing conflicts, which is an important subject for further research and a focus of our study.

In summary, similar to current studies on team characteristics in HAT, research examining human-AI conflicts within HAT are mainly of quantitative nature, lacking a qualitative approach which can help examine human's subjective interpretations and experiences of HAT, which our research study presented in this paper aims to address. Additionally, the studies we found are not fully grounded in theories on team conflicts or team development and mainly view conflicts as negative. In our presented study, we are integrating well-established models and theories (i.e., Tuckman's model for team development and Activity Theory), to help understand and interpret findings. Compared to prior studies and aligned with the integrated model and theory, we approach conflicts in our research study as "opportunities" to learn and develop (Tewari and Lindgren, 2022) rather than the traditional view of errors and problems.

3.3 Application of Tuckman's model

Tuckman's theory of team development focuses on understanding the stages human-only teams go through in therapy groups, human relations training groups, and natural and laboratory task groups (Tuckman, 1965). Tuckman's theory has been widely applied in various settings, including organizations (Miller, 2003) and classroom (Runkel et al., 1971). For example, Sokman et al. (2023) examined students' perception of Tuckman's stages and the relationships between the storming stage and all other stages. They found that despite teams' uncertainties during the forming, storming, and norming stages, participants worked toward a common goal during the performing stage. Additionally, their results revealed that the storming stage positively impacted the forming and norming stages (Sokman et al., 2023).

Despite the common use of Tuckman's theory to understand and interpret the formation of human-only teams, to our knowledge, there is a lack of research applying Tuckman's (1965) model to HAT development, particularly in the context of supporting human health and wellbeing. Therefore, the present study provides new knowledge on HAT across the three stages of

Tuckman's model: *forming*, *storming*, and *norming*. Furthermore, this study provides findings on how human-agent collaboration evolves as part of a collaborative calendar planning activity aimed at improving a patient's health and wellbeing.

4 Methodology

The study was conducted within an interdisciplinary research context, focusing on health-promoting activities and health behavior change interventions. The chosen collaborative activity, where a purposefully designed agent collaborates with humans to plan activities for their upcoming week to achieve their health-related goals, holds relevance across several domains. For example, in occupational therapy and psychology, it is common practice to integrate meaningful activities to aid recovery from stress and promote a balanced, healthy lifestyle, such as for healthy aging and individuals with mental health challenges (Patt et al., 2023; Aronsson et al., 2024). Furthermore, addressing the general need for health behavior changes to prevent cardiovascular incidents is another relevant focus, directly tied to everyday choices of activities.

Based on the theoretical framework, specified in the Background Section 2, and occupational therapy concepts and theory, we developed a use case scenario (Section 4.1), for which the activity and its components were defined, as well as the conditions for the activity. We used a participatory design approach (Bødker et al., 2022) to develop the agent and the collaborative calendar planning activity (i.e., the calendar application), as detailed in Section 4.2. Section 4.3 presents the algorithm that guides the agent's behavior and decision-making, while Section 4.4 describes the resulting system.

This study employs a qualitative descriptive study design, which is commonly used for examining participants' perception of an application or intervention (Doyle et al., 2020). Section 4.5 describes participant recruitment and demographics, while Sections 4.6 and 4.7 detail data collection and analysis.

4.1 Design of use case scenario

We envisioned the activity in our use case scenario as a collaboration between a human and an agent that takes on roles such as assistant, therapist, or companion. The activity would incorporate a system with a graphical user interface (GUI), allowing the person to define and create activities they aim to conduct during the week, as well as log these activities for future reference. Additionally, the system would integrate the agent and methods for communication with it. Potential use contexts include monitoring stress through recovery activities and promoting health behavior change by incorporating meaningful health-related activities, such as social and physical activities, especially among older adults or people with mental health challenges, to prevent loneliness and improve their mental and physical health.

To conceptualize this activity, we defined the collaborative calendar planning activity according to Activity Theory (Engeström, 1999) as illustrated in Figure 1. The selected activity was intentionally kept generic to ensure broad relevance. The activity in focus was: "Plan the coming week using a set

of activities meaningful to the human that includes health-promoting activities." Both involved actors (i.e., the human and the collaborative agent) had an equal influence on the activity process and outcome, with the activities and schedule being personally relevant to the human.

The tools available to the actors for the task were (i) a repository of activities of different types, which have different purposes and levels of importance defined based on an ontology of activity (Lindgren and Weck, 2022) that was developed based on Activity theory (Engeström, 1999), and (ii) a text and/or voice-based communication channel. Further, their space to perform the activity consisted of a calendar view displaying days and time slots.

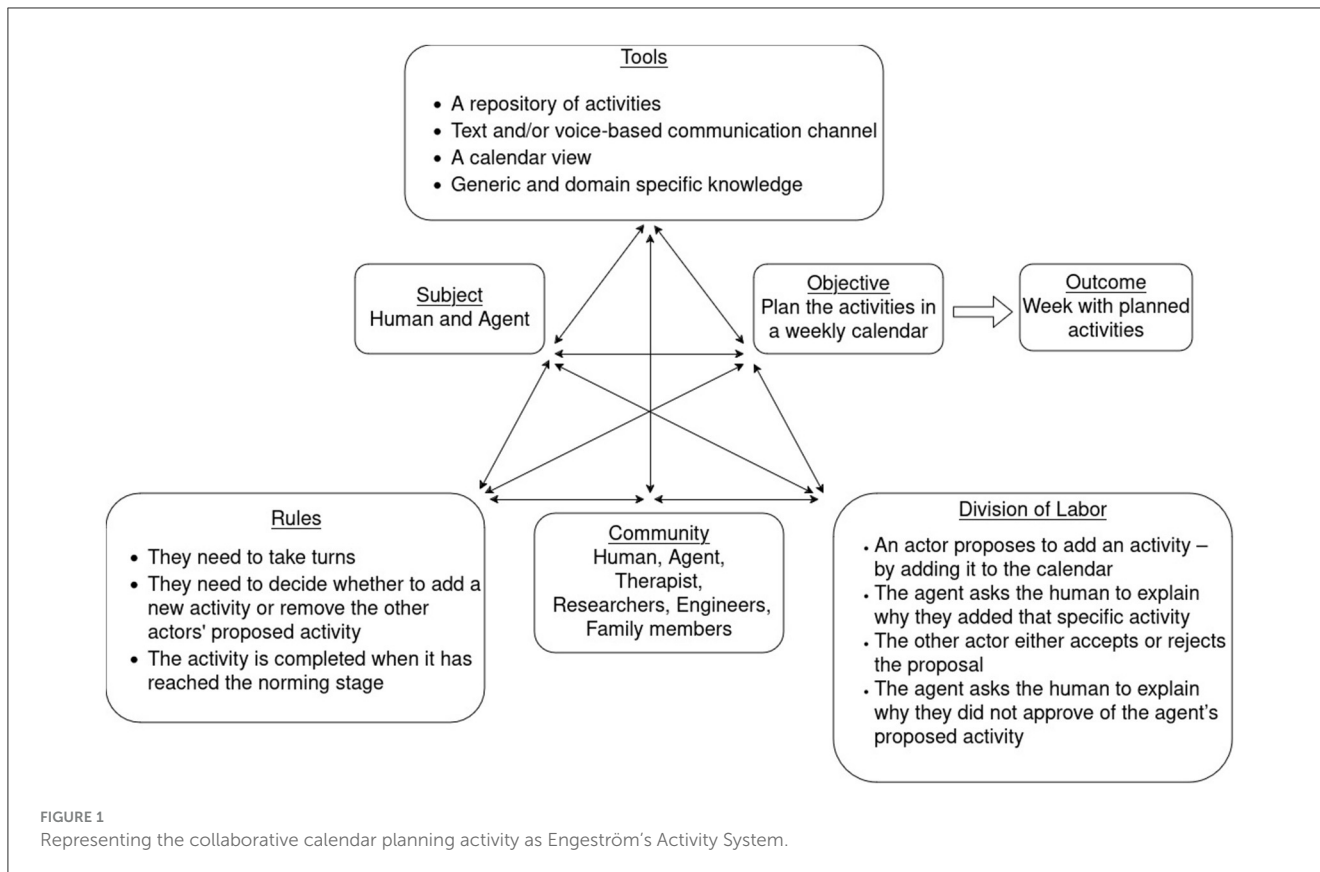
The actions that they could perform were the following: (i) propose to add an activity and provide a rationale; (ii) add the activity to a calendar; (iii) remove an activity from the calendar, explain the reason; and (iv) communicate to coordinate motivations.

The following division of tasks was applied: (i) an actor proposes to add an activity – by adding it to the calendar; (ii) the agent asks the human to explain why they added that specific activity; (iii) the other actor either accepts or rejects the proposal; and (iv) the agent asks the human to explain why they did not approve of the agent's proposed activity.

The conditions for acting were the following: (i) they need to take turns; (ii) they need to decide whether to add a new activity or remove the other actors' proposed activity; and (iii) the activity is completed when it has reached the norming stage as defined in the implementation.

We focused on the actors' knowledge as a crucial tool for the activity, identifying specific and generic knowledge domains as potential sources of contradictions that could lead to breakdown situations. The *specific knowledge domain* in our scenario is specific to the task and relates to health, preventing stress-related conditions, and recovery activities which informed the agent's motives for proposing certain activities. The *generic knowledge* relates to what is sometimes called common sense knowledge, or tacit knowledge that is typically not explicit in dialogues. In our scenario, the agent was expected to have a generic understanding of components in the "environment" and "activity and participation" domain of the International Classification of Functioning, Disability, and Health (ICF) (World Health Organization, 2001). More specifically, this included an understanding of time over day/night, days of a week (we denote this the *time space*); language-based interaction (we refer such knowledge to the *cognitive space*); social aspects when multiple agents interfering with/contributing to a task (falls into the *social space*). The social aspects that we pursued in this study are related to collaboration and the development of teamwork, specifically, the behavior that manifests in the forming, storming, and norming stages. To accomplish this, the components of the activity system, including the notion of contradictions, were applied.

The cognitive and social spaces were specified within dialogue scenarios. The cognitive space included the division of tasks and the conditions for acting. The social space embeds the cognitive space with the effects of tasks that the actors perform under specific conditions. Figure 2 illustrates one such dialogue scenario, where the social space is mapped out to show Tuckman's three stages: forming (highlighted in yellow), storming (in red), and norming



(in green), along with the transitions from forming to storming (in orange) and storming to norming (in blue). The actors perform their respective tasks and remain within a stage until a threshold is reached. For example, in the storming stage, the agent keeps rejecting the human's proposed activities and proposes previously declined activities until the human has either accepted or rejected the activities for a certain amount of time.

4.2 Designing the system for collaborative planning

We followed a participatory design approach to design and develop the system that embeds the agent for the collaborative calendar planning activity. In a participatory design process, stakeholders, including potential end users, designers, engineers, and researchers, work collaboratively together to explore, understand, develop, and reflect on prospective digital technologies (Bødker et al., 2022). In our case, this consisted of three consecutive phases: Phase 1 involved specifying the use case, phase 2 involved designing dialogue scenarios and developing the prototype, and phase 3 involved iterative refinement of the prototype and designing of the study.

Phase 1: We specified the use case during research meetings engaging several stakeholders: a junior AI researcher with lived experience as a long-term patient, an occupational therapist

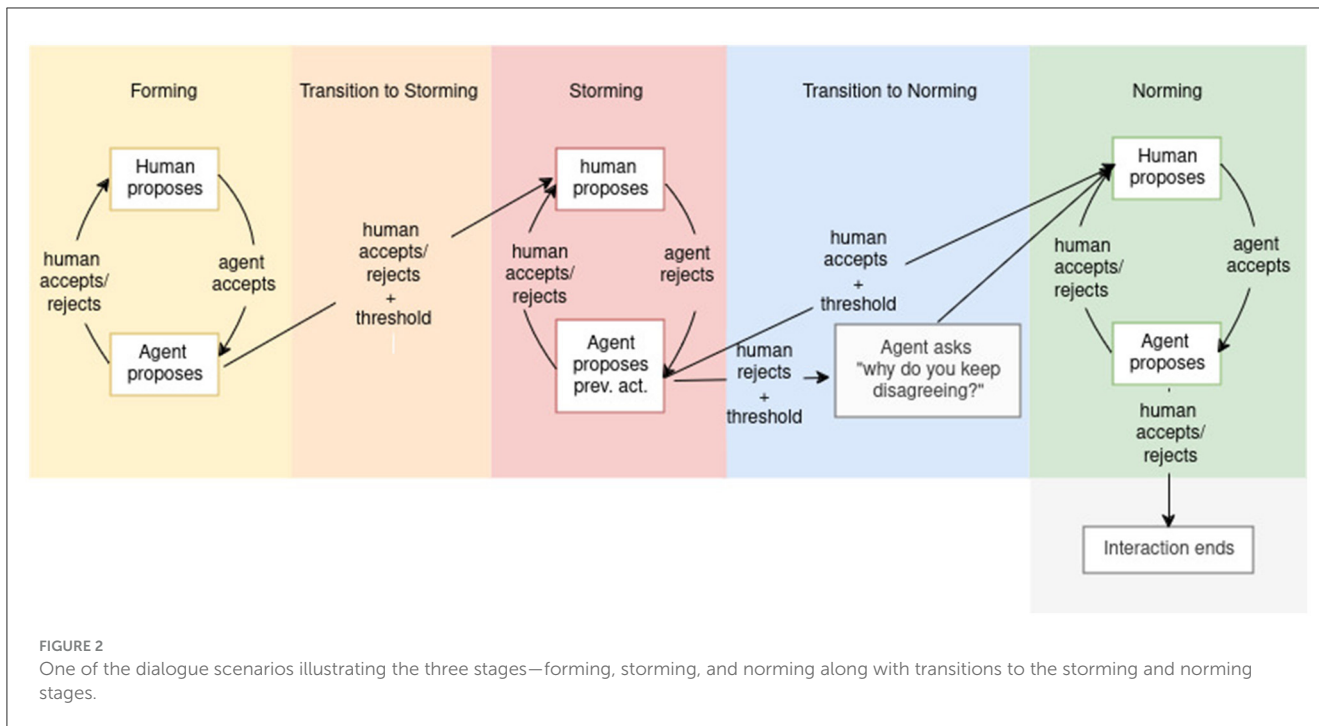
researcher with prior clinical experience, and a senior AI researcher with a background in clinical occupational therapy.

Phase 2: In separate research meetings, two stakeholders (i.e., the AI researcher with lived experience and the occupational therapist researcher) collaborated to draw the calendar planning activity scenario on a whiteboard, develop dialogue scenarios, and refine these scenarios by role-playing the calendar planning activity using paper and pen. The involved stakeholders discussed and finalized the calendar planning activity and the dialogue scenarios, which were then drawn on the whiteboard. Following this, the AI researcher with lived experience and the occupational therapist researcher held four additional meetings, during which they collaboratively drew, discussed, and finalized the dialogue scenarios using the draw.io¹ software (see Figure 2). Based on these finalized scenarios, the junior AI researcher developed a prototype application in Android Java.

Using the developed dialogue scenarios for the calendar planning activity as a foundation, the involved stakeholders discussed an algorithm, which was collaboratively finalized by the junior AI researcher and the occupation therapist researcher.

Phase 3: The AI researcher then refined the prototype application, integrating the developed algorithm and the finalized dialogue scenarios for the calendar planning activity. After five rounds of pilot testing the application with research team members and two additional potential end users, the AI researcher refined the prototype design and functionalities based on the received

1 <https://app.diagrams.net/>



feedback. The designed study where participants as potential end users engage with our developed prototype and provide reflections on their experiences and perceptions of the collaborative calendar planning task is detailed in Section 4.6.

4.3 Algorithm

We implemented the developed algorithm as a finite state machine, as illustrated in Figure 3. A finite state machine is a reactive system that operates through a series of states. It transitions from one state to another in response to an event, provided that the conditions controlling the transition are met. As illustrated in Figure 3, the group of states represents the stages (i.e., forming, storming, and norming) and the transitions between them. Actions and conditions trigger the transitions with and between the stages.

Our finite state machine consists of eight states ($S_0 - S_7$). Transitions between these states are triggered by specific conditions. These conditions consist of human input represented by H_n and a counter represented as C_m . There are four kinds of actions (i.e., representing events) a human can perform: (1) human proposes an activity $H_{proposes}$, (2) human likes the agent's proposed activity H_{likes} , (3) human dislikes the agent's proposed activity $H_{dislikes}$, and (4) human provides a reason for proposing a new activity or disliking an activity proposed by the agent $H_{providesReason}$.

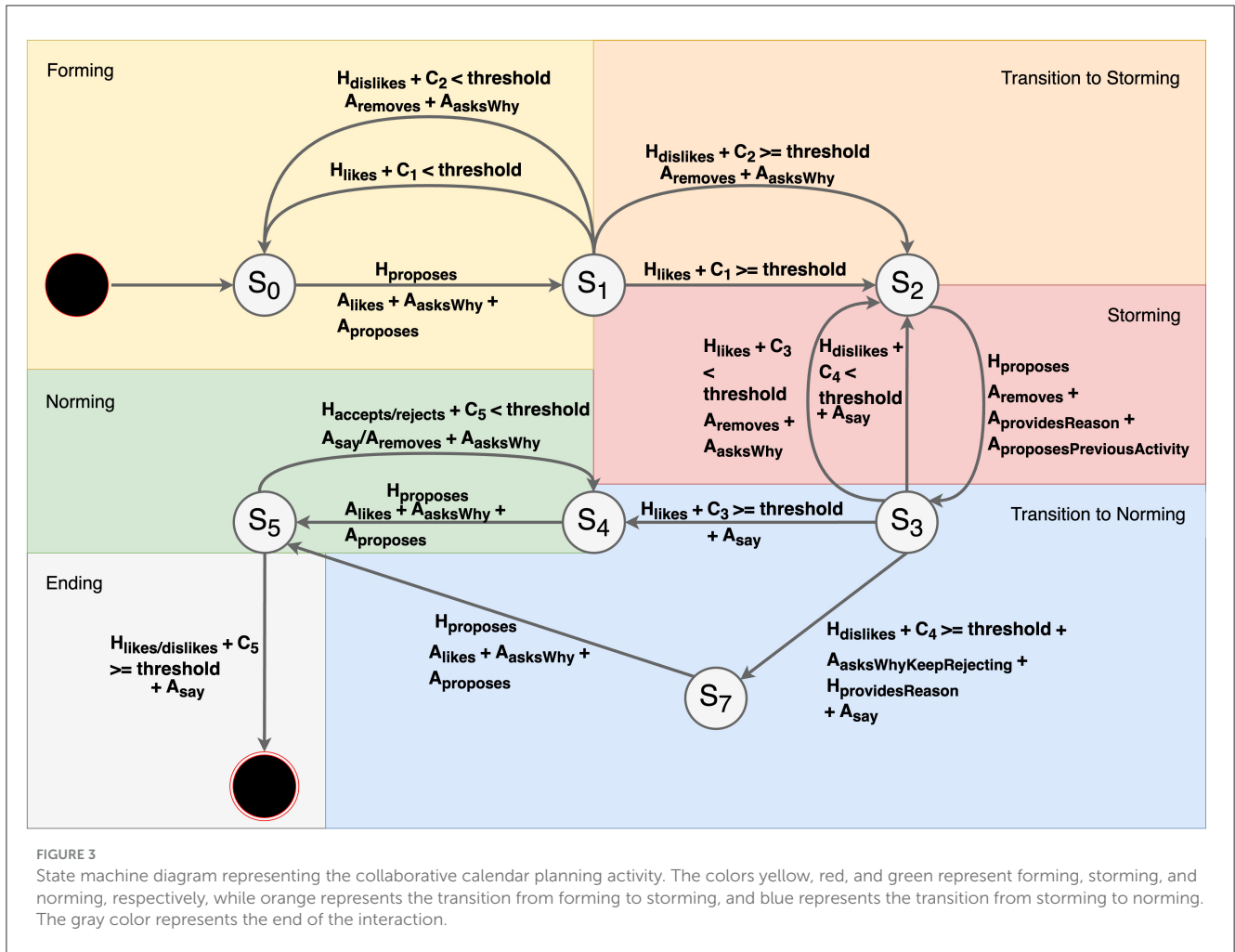
There are five counters ($C_1 - C_5$) that are used in the conditions with the two human inputs H_{likes} and $H_{dislikes}$ to guide the agent's behavior and determine when the human-agent team transitions through Tuckman's first three stages for team development, namely from the forming to the storming and from the storming to the norming stage. Each counter is further explained in the following sub-sections.

There are five kinds of actions the agent can perform: (i) the agent asks why the human chose a particular activity $A_{asksWhy}$, and proposes an activity $A_{proposes}$; (ii) when the human likes the activity proposed by the agent, the agent says, 'Great that you liked my proposed activity' $A_{saysGreat}$ and tells them it is their turn to add an activity A_{tells} ; (iii) when the human dislikes the activity proposed by the agent, the agent asks why they did not like the activity $A_{asksWhyDislike}$, removes the activity that the human disliked $A_{removes}$, and tells them it is their turn to add an activity A_{tells} ; (iv) when in storming stage, the agent removes the activity proposed by the human $A_{removes}$, provides a reason why it is removing the human's proposed activity $A_{providesReason}$ and proposes a previously removed activity $A_{proposesPreviousActivity}$; and (v) when the human keeps disliking the agent's proposed activity, the agent highlights that there is a persistent disagreement between them. The agent asks why the human keeps rejecting the activities it is proposing $A_{asksWhyKeepRejecting}$ and removes the activity the human disliked $A_{removes}$.

State transitions are triggered given one or more conditions. These transitions are designed to reflect the first three stages of team development by Tuckman (1965). Figure 3 shows transitions to storming and norming stages, both human and agent-initiated ones. These transitions are further described in the following sections.

4.3.1 Forming stage

The finite state machine represents the forming stage with transitions from $S_0 \rightarrow S_1$ and $S_1 \rightarrow S_0$ (highlighted in yellow in Figure 3). It is initiated by the human proposing an activity $H_{proposes}$ (action 1). The agent responds by asking why the human chooses that activity and then proposes a new one for the calendar $A_{asksWhy}$, $A_{proposes}$ (i.e., performs action i), leading



to the transition $S_0 \rightarrow S_1$. The transition $S_1 \rightarrow S_0$ happens regardless of whether the human likes or dislikes the agent's proposed activity as long as the thresholds C_1 or C_2 have not been reached. When the human performs action 2 (H_{likes}), the agent responds with action ii ($A_{saysGreat}, A_{tells}$). If the human performs action 3 ($H_{dislikes}$), the agent responds with action iii ($A_{asksWhyDislike}, A_{removes}, A_{tells}$). For both cases, the human performs action 4 ($H_{providesReason}$).

4.3.2 Transition to storming stage

The finite state machine represents the transition from the forming to the storming stage with the transition $S_1 \rightarrow S_2$ (highlighted in orange in Figure 3). A transition happens if the human likes the agent's proposed activity H_{likes} (i.e., performs action 2) and the counter C_1 value is greater than or equal to the threshold $C_1 \geq threshold$. Then the agent performs action ii ($A_{saysGreat}, A_{tells}$). A transition also happens if the human dislikes $H_{dislikes}$ the agent's proposed activity and the counter C_2 is greater than equal to the threshold $C_2 \geq threshold$. The agent performs action iii ($A_{asksWhyDislike}, A_{removes}, A_{tells}$).

4.3.3 Storming stage

The finite state machine represents the storming stage with the transitions $S_2 \rightarrow S_3$ and $S_3 \rightarrow S_2$ (highlighted in red in Figure 3). When in S_2 , the human proposes an activity $H_{proposes}$ (action 1), the agent removes the human's proposed activity, provides a reason why it did so, and proposes a previously removed activity (action iv) ($A_{dislikes}, A_{providesReason}, A_{proposesPreviousActivity}$). The state transitions to S_3 . In S_3 , the human may like (perform action 2, H_{likes}) or dislike (perform action 3, $H_{dislikes}$) the agent's proposed activity. The finite state machine loops until the thresholds for C_3 or C_4 have been reached. If the human performed action 2, the agent performs action ii ($A_{saysGreat}, A_{tells}$). If the human performed action 3 ($H_{dislikes}$), the agent performs action iii ($A_{asksWhyDislike}, A_{removes}, A_{tells}$).

4.3.4 Transition to norming stage

The transition from the storming to the norming stage is represented by the transition $S_3 \rightarrow S_7$. This transition happens when the human performs action 3, and the counter C_3 is greater than or equal to the threshold. The agent performs the action v ($A_{asksWhyKeepRejecting}, A_{removes}$). The transition $S_3 \rightarrow S_4$ happens when the human likes the agent's proposed activity (action 2), and

the counter C_4 is greater than or equal to the threshold. This stage is highlighted in blue in Figure 3.

4.3.5 Norming stage

The norming stage is represented by transitions $S_7 \rightarrow S_5$, $S_4 \rightarrow S_5$, $S_5 \rightarrow S_4$ (highlighted in green in Figure 3). When the human performs action 1 ($H_{proposes}$), the agent performs action i ($A_{asksWhy}$, $A_{proposes}$) and transitions from S_7 or S_4 to S_5 happens. The transition from S_5 to S_4 happens when the human performs action 2 or 3, and the counter C_5 has not reached the threshold. Corresponding to the human's action 2, the agent performs action ii ($A_{saysGreat}$, A_{tells}), and when the human performs action 3, the agent responds by action iii ($A_{asksWhyDislike}$, $A_{removes}$, A_{tells}).

4.3.6 Ending

The calendar planning activity ends, and the finite state machine makes the transition $S_5 \rightarrow S_6$ (highlighted in gray in Figure 3). This state is reached when the counter C_5 is greater than or equal to the threshold and regardless of the human's actions.

4.4 Collaborative planning system with an agent and a calendar

The finite state machine algorithm was developed for a tablet and implemented as an Android application in Java. We chose a tablet as the application interface for our research study because its large screen provides enough space to display the weekly calendar view, allowing participants to comfortably interact with the application. The application comprises four main modules, as illustrated in Figure 4: a Graphical User Interface (GUI), a controller, a finite state machine, and persistence modules.

The GUI includes the agent and two spaces for interaction: the communication space (the dialogue) and the activity space (the calendar and a repository of activities, Figure 5). The agent is represented as a robotic figure on the GUI, giving it some level of character and agency. The list of activities with the drag-and-drop functionality is displayed on the left side of the calendar. The calendar was developed using the Android library available at: <https://github.com/alamkanak/Android-Week-View>. The weekly view of the calendar can be scrolled both horizontally and vertically. At the top right of the interface, there is the option to save the activity information locally on the device (persistence).

The Controller is the main module of our prototype, responsible for managing the addition and removal of activities, the interaction between the human and the agent, and handling the added and the removed activities in the persistence. The controller allows the agent to propose an activity, remove an activity, and have dialogues with the human. The controller also allows the human to add an event either by dragging and dropping or by long-pressing an empty time slot on the calendar. The controller enables the persistence of information related to added and removed activities. The functions of the controller are used to define the finite state machine.

The State Machine Builder implements the algorithm that simulates Tuckman's first three stages of team development: forming, storming, and norming. The finite state machine governs the agent's behavior and the interaction between the agent and the human.

The Persistence Service manages the persistence of information related to added and removed activities. The information consists of the activity's name, added or removed (boolean values), timestamp, who added it, the start and end time of the activity, and the reason provided by the actors behind adding or removing the activity. The output is a comma-separated file stored locally on the tablet.

4.4.1 Human-agent interaction

The human and the agent take turns when conducting the planning task in the calendar. The human starts by reading the instructions provided by the agent, then selects and adds an activity to the calendar. This can be done by dragging and dropping an activity from a provided list in the application (Figure 5) or by long-pressing on an empty slot in the calendar.

When it is the agent's turn, the agent asks the human why they chose to add that particular activity (i.e., it asks the question *Why did you choose to add "x"?*). The human can write or use the built-in speech-to-text function on the tablet's keyboard to answer. Once the human provides their reason, the agent acknowledges it and then takes its turn to propose an activity by adding it to the calendar. The agent then asks if the human likes or dislikes the activity it proposed. The human responds whether they like or dislike the proposed activity by pressing either the happy face or sad face buttons provided with the question. If the human dislikes the agent's proposed activity, the agent asks the human to provide a reason for their dislike (Figure 5).

4.5 Participants

We recruited participants through convenience sampling with a focus on reaching diversity among participants. A total of eight (8) participants agreed to participate. Participants' age ranged from 20 to 71 years with a mean of 33 years 10 months (SD = 19 years 8 months). Three of the participants identified as male, and five as female. Four of the participants had a high school/secondary degree, three had a master's degree, and one had a doctorate degree at the time of the study (see Table 1).

4.6 Procedure and data collection

To explore how participants, representing potential end users, experience and react to the collaborative calendar planning task and the integrated stages of Tuckman's (1965) model, and to engage them in further design of the collaborative agent, we conducted a pilot study using a qualitative descriptive study design. Ethical approval was obtained through the Swedish Ethical Review Authority (Dnr 2019-02794).

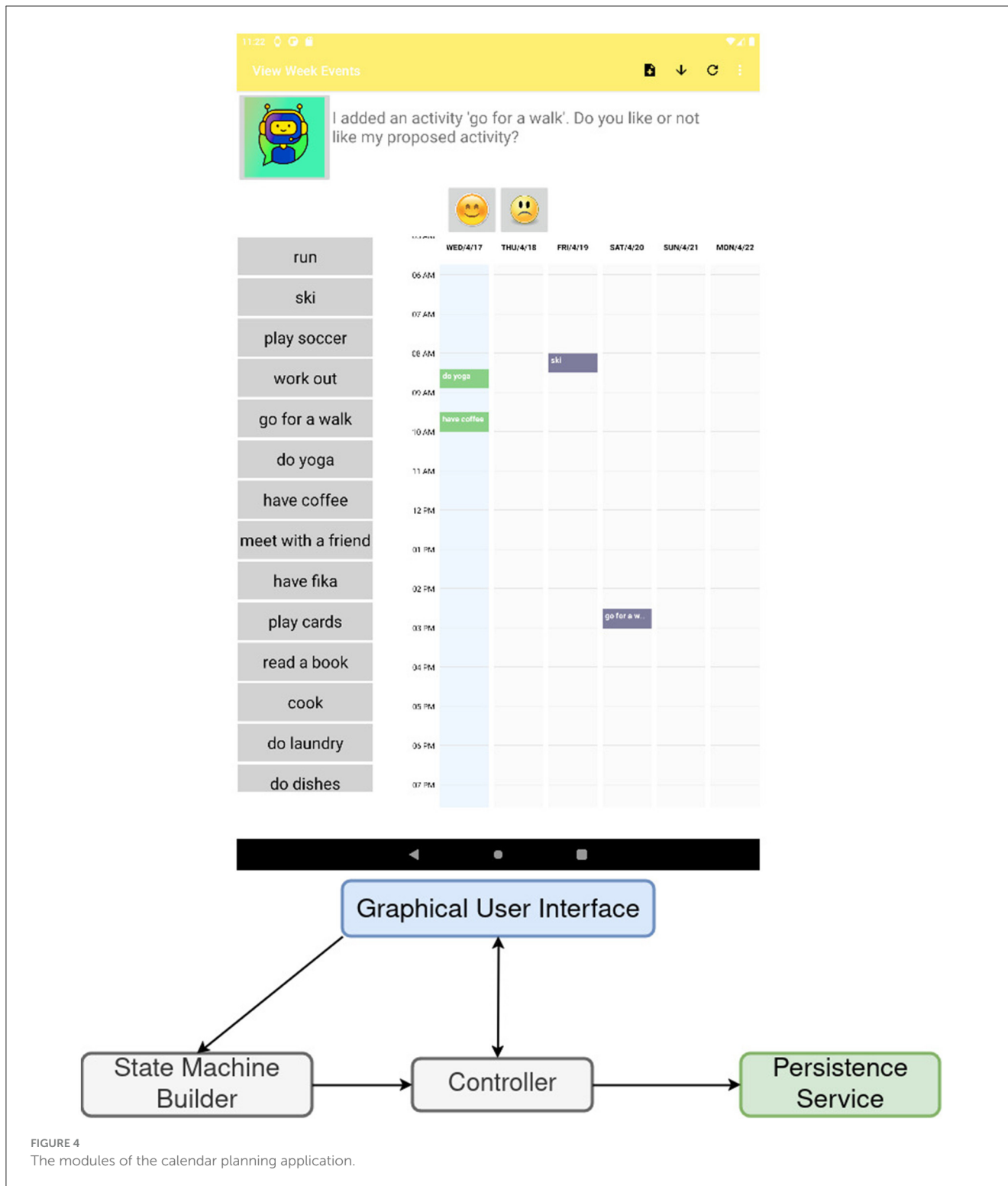
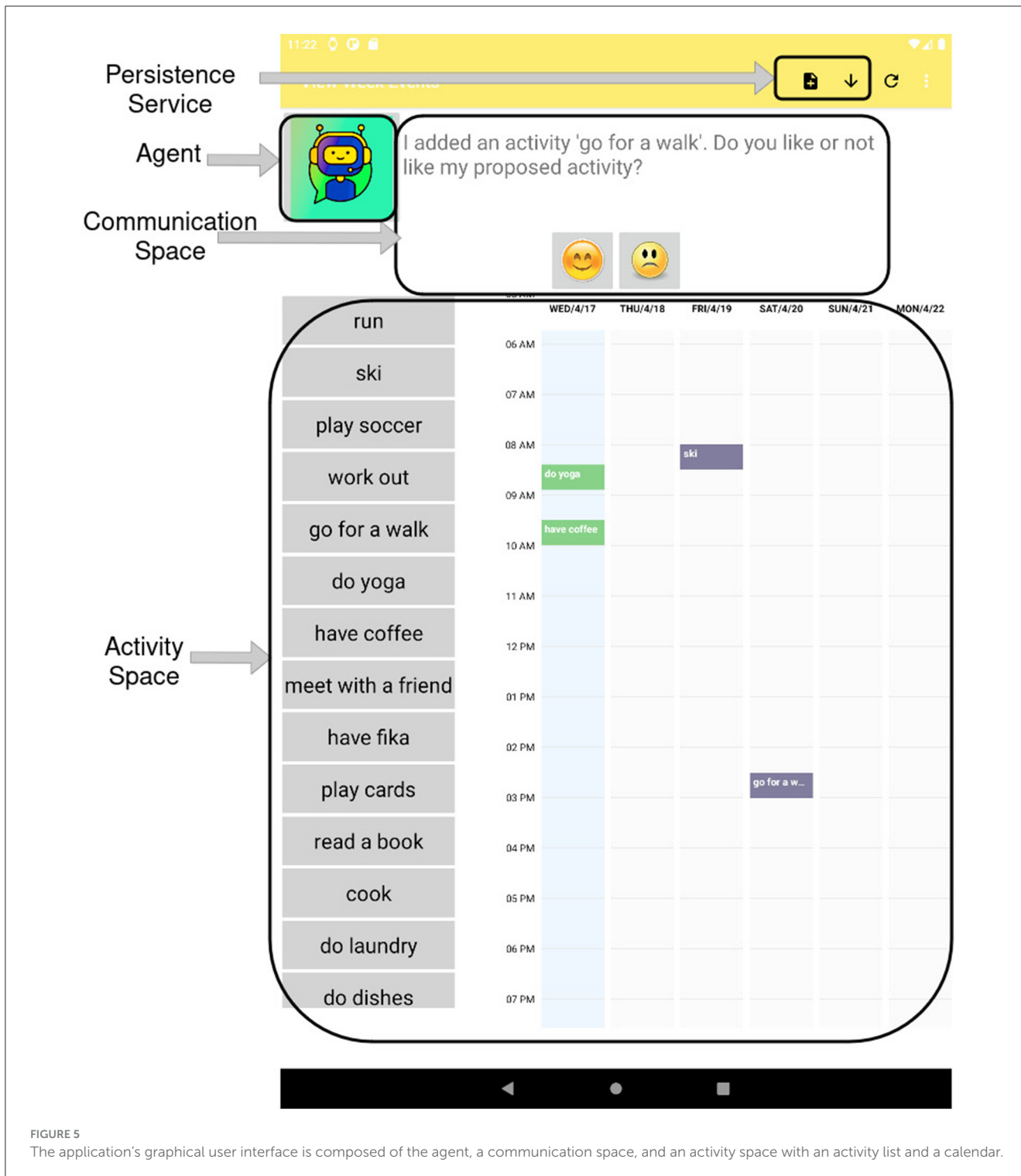


FIGURE 4 The modules of the calendar planning application.

Participants were informed about the research study and its purpose. After agreeing to participate, they were invited to a scheduled session, where they provided informed consent, completed the demographic questionnaire, and engaged in the collaborative calendar planning activity using an Android tablet. Data about participants' transitions through Tuckman's team development stages were collected by logging the interaction events

in a database and recording the tablet screen during the task (research question 1). The collaborative calendar planning task involving the human and the agent took between 8 and 23 min.

Captured data points included data on agents' and humans' moves, such as activity proposals by the human and agent, whether the human or agent agreed or disagreed with the proposal, and whether an activity was deleted. Data about



participants' experiences of the collaborative calendar planning task and the agent's behavior embedded in the different stages and their design proposals were collected through semi-structured interviews lasting between 14 and 25 minutes (research question 2). The interviews were conducted by two of the authors with prior experience in conducting semi-structured interviews. They were audio-recorded and transcribed verbatim.

4.7 Data analysis

Data analysis regarding participants' transitions through Tuckman's team development stages is described in Section 4.7.1 (research question 1). The analysis of the transcribed interview data on participants' experience of the designed collaborative calendar

TABLE 1 The table shows the participants' demographic information.

Participant	Age (years)	Gender	Educational background
P1	20	Male	High school/Gymnasium
P2	71	Female	Master (MSc, MA)
P3	22	Female	High school/Gymnasium
P4	28	Female	Master (MSc, MA)
P5	25	Male	Master (MSc, MA)
P6	24	Male	High school/Gymnasium
P7	21	Female	High school/Gymnasium
P8	59	Female	Doctorate (PhD, ScD)

planning activity and the embedded stages of Tuckman's model is detailed in Section 4.7.2 (research question 2).

4.7.1 Participants' transition through Tuckman's first three stages

To examine participants' transition through team development stages (research question 1), we fetched data on the agent's and participant's moves and summarized them for each stage (i.e., forming, storming, and norming stages). In addition, we determined whether the human or the agent initiated the transition into the storming and the norming stage. This was done by determining which threshold triggered the transition. Specifically, we considered the transition into the storming stage as human-initiated if it was triggered by reaching threshold C_2 , which includes the human disagreeing with the agent's proposals at least or equal to 2. If the human kept agreeing with the agents' proposal, the storming stage was considered as agent-initiated when the threshold $C_1 = 4$ was reached. Similarly, we determined the transition to the norming stage to be human-initiated if the human agreed with the agents' proposal at least or equal to 2 times (i.e., reaching the threshold for C_3), despite the agent removing all activities added by the human during this stage. An agent-initiated norming stage was determined as such when the human kept disagreeing with the agents' proposal during the storming stage (threshold $C_4 = 2$). Finally, we reported on the number of activities the human and the agent added in the final scheduled week.

4.7.2 Participants' experiences of the collaborative activity

The qualitative data about participants' experience of collaboration with the agent (research question 2) was analyzed using inductive content analysis (Elo and Kyngäs, 2008). First, one researcher with prior experience in inductive content analysis carefully read and re-read the transcript with participants' responses line-by-line. Then, notes denoting "condensed meaning units" were added alongside the corresponding narratives in the transcripts. Together with a second researcher serving as a key informant, we refined "condensed meaning units" and grouped similar "condensed meaning units" into sub-categories. Through

an iterative process, we added, adjusted, and compared sub-categories to ensure their distinctiveness. In the final abstraction phase, we consolidated sub-categories into categories and grouped these into main categories. We selected and extracted quotes to exemplify our categories and main categories and drafted a report of the findings on the participants' experiences of the collaborative activity.

5 Results

The results are organized as follows. To address the first research question, we report on participants' transition through the stages of team development. For the second research question, we present findings related to participants' experiences of the collaborative calendar planning activity, embedding an agent guiding participants through the three first stages of Tuckman's team-development process. Findings related to both research questions are further described in the following sections.

5.1 Participants' transition through Tuckman's first three stages

All participants transitioned through the three stages "forming," "storming," and "norming," based on how the staging was implemented.

Seven participants initiated the storming stage themselves by reaching the threshold C_2 , and only for one participant (P2) did the agent initiate the storming stage when the threshold C_1 was reached (see Figure 6). The transition to the norming stage was more often initiated by the agent (i.e., for five participants), and only three participants initiated the norming stage themselves.

The participants added between eight and 11 activities, with P4 adding the highest number of activities ($n = 11$), while P3, P7, and P8 each added eight activities to the calendar. At the end of the interaction, there were more activities in the calendar created by the participants compared to those added by the agent, with P5 showing the largest difference and P8 the smallest across participants (see Figure 7).

5.2 Participants' experiences of the collaborative activity

The content analyses revealed three main categories related to participants' experiences of the collaborative calendar planning activity: (i) participants' experiences of Tuckman's first three stages for team development; (ii) their reasoning and reactions to the agent's behavior in the three stages; and (iii) factors important to the participants to team up with a collaborative agent. Findings related to the three main categories are further described in the following sections.

















Participant	Forming		Storming initiator	Storming		Norming initiator	Norming		Norming recognition
	Liked	Disliked		Liked	Disliked		Liked	Disliked	
P1	2	2		1	2		2	4	Yes
P2	4	1		1	2		3	3	Yes
P3	1	2		2	0		1	4	No
P4	3	2		1	2		3	2	No
P5	3	2		2	1		0	5	No
P6	2	2		1	2		2	3	Yes
P7	1	2		1	2		3	2	No
P8	1	2		2	1		5	0	No

FIGURE 6 Illustrating who (agent or participant) initiated the storming and norming stages based on our designed algorithm guiding teamwork in the collaborative calendar planning activity. In addition, it indicates which participants recognized all three stages, including the transition into the norming stage.

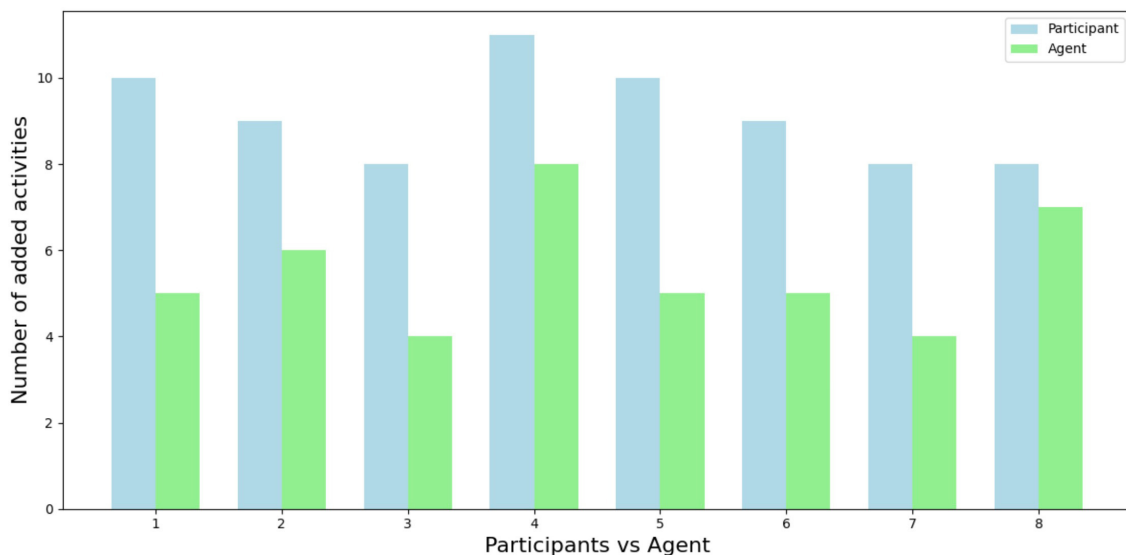


FIGURE 7 The number of activities remaining in the calendar at the end of the activity, created by the participants and the agent.

5.2.1 Participants' experiences of Tuckman's first three stages for team development

Three categories were identified related to participants' experiences of Tuckman's first three stages for team development, each of them representing one stage: (1) Forming stage: Familiarization with the agent, (2) Storming stage: Irritation and

frustration, and (3) Norming stage: Agent becoming more helpful. In the following sections, findings relating to these categories are presented.

Forming Stage - Familiarization with the Agent: The first stage (i.e., the forming stage) was characterized by participants' perception of a friendly atmosphere, such as feeling like the agent

was “nicer in the beginning” (P3). Participants felt valued by the agent and appreciated that it asked for their opinion, illustrated by P4: “in the beginning [the agent was taking] into consideration my input. So, he [the agent] was asking me, why did you not like, why I put this.”

Participants also experienced uncertainty during this initial stage, feeling that they did not yet know the agent and that the agent did not yet know them. More specifically, they felt like they lacked knowledge about each other’s goals and were unsure of where they were going as a team. P4, for example shared “where we’re going, I don’t know;” and P5 mentioned “initially I was a bit confused.” P2 described an initial irritation because she expected to complete the calendar planning activity by herself rather than in close collaboration with the agent, resulting in an initial hesitation with respect to collaborating with the agent: “I thought it was my activities, not the agent’s activities. [...] I wanted to make my own day. [...] That was, I didn’t realize that in the beginning, that he [the agent] would suggest [activities, too]” (P2).

Storming Stage - Irritation and Frustration: This second stage (i.e., the storming stage) was characterized by participants being surprised, irritated, and, after a while, frustrated by the agent’s behavior. Participants did not expect the agent to interfere with their planning as much as it did, including deleting activities they had added. P3, for example, expressed this surprise as “I’m adding, and then you’re adding, and then I noticed it [the agent] would say no as well.”

Over time, and particularly when the agent started to delete activities that were meaningful to the participant, they felt frustrated. P7, for example, mentioned: “I got a bit frustrated. I’m like, why do you do that? I want to do it” and P3 shared: “I was like really not okay it [the agent] did take away [that activity]”. Some participants realized that the storming stage was initiated by their disagreement with the agent. As P2 noted: “he [the agent] didn’t like that I disliked.” For some participants, this felt like a power shift, as described by P3: “It felt a bit like we were 50–50 in the beginning, and then we started to be a bit more like, [...] I say no, okay, then they [the agent] have to say no to me as well.”

Norming Stage - Agent Becoming More Helpful: Although the algorithm was designed to ensure that all participants transitioned through the first three stages of Tuckman’s model for team development, only three participants (P1, P2, and P6) described experiences related to all three stages when sharing their perceptions of the team development process with the agent. All other participants described only the first two stages (i.e., forming and storming). The three participants experiencing a norming stage described it as an improvement in the agent’s behavior, in the sense that the agent was better and more helpful toward the end. P1, for example, mentioned, “In the end, I find the agent more cooperative.”

5.2.2 Participants’ reasoning and reactions to the agent’s behavior in the three stages

Two categories emerged related to participants’ reasoning and reactions to the agent’s behavior in the three stages of Tuckman’s model for team development: (1) Reasoning about the stages, and (2) experiencing a transitioning or no transitioning from

the storming stage. Findings related to these two categories are described in the following sections.

Reasoning about the stages: Participants mainly reasoned and reflected about the storming stage and the disagreement they had with the agent. Even though “annoying,” as P2 described, most of the participants perceived the disagreement not only as negative. They mentioned how disagreements with an agent could be helpful and even needed for personal development. For example, P3 shared: “if [the agent] would just agree with me all the time, I don’t see the purpose of using it [the agent] maybe as much.” Some participants mentioned how disagreements between agents and humans might be needed for behavioral change among humans; as described by P3, “it challenges you to think in a different way and have a perspective, kind of.” Another participant compared the disagreement with the agent to disagreements she had encountered in the past with other humans and reflected how they might relate: “Sometimes you prefer a type of interaction, but you need another type of interaction” (P4).

Perceived transitioning or not transitioning from the storming stage: Although the algorithm that guided the agent’s behavior was built in order for all participant-agent teams to transition through the forming, storming, and norming stages, the interviews indicated that participants perceived the transition from the storming to the norming stage in two ways: participants either experienced a transition beyond the storming stage, or they perceived their participant-agent team to remain in the storming stage. Participants who experienced a transition beyond the storming stage described a clear norming stage, whereas participants not experiencing a transition from the storming stage stayed in the storming stage for the remainder of the collaborative calendar planning activity. This was independent of how they reasoned about the importance of disagreements. Transitioning from the storming stage was indicated by participants when they reasoned about why the agent may have had its reason for what it did and that the agent’s ideas could be helpful, too. For example, P6 shared: “I was fine with that, maybe because he [the agent] had other intentions, maybe looking out for not to get overtrained or something like that.” Other participants reflected on the transition, realizing the benefit of having an agent providing an additional perspective as P2 explained: “he [the agent] was kind of questioning my, my, then after a while when I realized that it’s, it could be helpful, then he [the agent] became an assistant instead, less annoying because I had made my plan and then, it changed there because I thought, but I haven’t thought about those things. It was helpful.” When participants stayed in the storming stage, they kept being frustrated with the agent, which resulted in a disruption of the collaboration or HAT, as exemplified by P5, who expressed that he gave up on the collaboration with the agent: “At some points I was like, okay, let me just do my things and let me just figure it out myself.”

5.2.3 Factors perceived as important for human-agent teamwork

We identified two categories related to factors participants perceived as important for human-agent teamwork: (1)

Preferring being in the lead, and (2) Factors for personalization. Findings related to these two categories are detailed in the next sections.

Preferring Being in the Lead: Most of the participants (5/8) felt they were in the lead during the collaborative calendar planning task and preferred it that way. One participant (P1) experienced a shift where the agent took the lead toward the end but still felt in charge of the overall task. Another participant (P5) mentioned that he would delegate the entire task to an agent if it was capable of understanding his needs and goals and said, “If I feel like the agent understands me, I am more likely and willing to give the planning to the agent.”

To accommodate for human spontaneity while planning, participants suggested that the agent should prepare a plan and then get it adjusted by the human. One participant reflected on the usefulness of the agent providing feedback on the proposed activities of the human. Another participant (P7) reflected on how the agent needed guidance but was not complying. The participant said “I had to lead him [the agent] a bit, but he [the agent] didn’t really listen.”

Factors for personalization: Participants specified person-specific and context- or task-specific factors important for an agent to capture in order to collaborate effectively. Some participants also reflected on mechanisms such as machine learning and recording sensor data that need to be embedded in an agent for a personalized experience.

The person-specific factors identified in the analysis were categorized into physical, and mental/ psychological factors. Physical factors reflected by the participants related to the amount of sleep and energy levels. One participant commented on how the agent could help the person organize their activities so that she could get enough sleep during the night. This could mean not scheduling many activities later during the day, as said by P7: “Then he [the agent] could help like, but if you plan this like this and tomorrow morning you want to get up at six, you won’t have your full amount of sleep.” Specific to mental/psychological factors were the knowledge of how the person’s mood was, how activities made them feel, and whether they were motivated to achieve their goals. One participant (P3) differentiated the approach that an agent could take depending on a person’s motivation and said, “if you’re not super motivated to change [...] then you would need a little bit [of a] nicer approach vs., I’m super determined to do this now, then you need a good push.”

Context- or task-specific factors included adapting activities based on a person’s family life, the flexibility of their work schedule and daily routines, and having a general sense of their plans. Additionally, it involved developing knowledge about the collaborative task. Several participants suggested that the agent could have a repertoire of activities that a person usually performs. Two participants emphasized that the agent could be programmed with knowledge of the person’s usual planning style (i.e., whether they plan on a daily or weekly basis). Information about routine activities and those that a person likes to do were considered important for an agent to know to collaborate in a calendar planning activity, as indicated by P4: “What I like what my daily routine is like when I wake up when I usually go to work when I usually do something like an activity after work or what I like maybe like my preferences.” Participants also suggested that the

agent could have activity categorizations such as physical, social, happy, relaxing, and self-care.

6 Discussion

The development of human-AI collaboration is a grand research challenge, as noted by various authors such as Crowley et al. (2022) and Nowak et al. (2018). In this study, we applied Tuckman’s theory for team development among humans (Tuckman, 1965) and Activity Theory (Vygotsky, 1978; Engeström, 1999; Kaptelinin and Nardi, 2006) to explore stages of human-agent team development during a collaborative calendar planning activity. This collaborative activity facilitated equal collaboration and represents an activity that is commonly performed in rehabilitation settings like occupational therapy or psychology. In this section, we discuss our results on participants’ transition through Tuckman’s stages of team development and participants’ experiences collaborating with the agent during the collaborative calendar planning activity. Then, we discuss some strengths and limitations of our study, and end with the conclusions of this study.

Our results revealed that seven out of eight participants went through a human-initiated storming stage, whereas the norming stage was agent-initiated in five cases and human-initiated in three cases. Regardless of whether stages were initiated by the human or agent, all participants reported similar experiences related to the forming and storming stage, which aligned with Tuckman’s team development theory (Tuckman, 1965; Bonebright, 2010). The forming stage was characterized by uncertainty, such as related to each other’s goals, while the storming stage was marked by feelings of surprise, irritation, and frustration, mainly due to the agent deleting activities that participants had added.

Interestingly, only three participants described and, thus, recognized a transition to the norming stage despite the agent being designed to support this transition for all participants. The remaining five participants appeared to stay in the storming stage until the end of the collaborative calendar planning activity, even though the agent had progressed to the norming stage. In other words, these five participants may not have been able to transition out of the storming stage as intended for the human-agent team by design. This result could be explained by the short time allocated for each stage, which may not have provided sufficient time for the human and the agent to resolve conflicts and transition to the norming stage. Alternatively, this result may indicate a need for more explicit conflict management between the human and the agent. This is supported by the fact that all three participants who described a transition to the norming stage, and thus described all three stages, went through an agent-initiated norming stage. The agent-initiated norming stage included an additional step in the collaborative process, where the agent acknowledged the participant’s significant disagreement with the agent and asked the participant to explain why they thought this might be. This extra step functioned as a means to elicit and address disagreement, shifting the focus to resolving it. However, this extra step was not included when the transition to the norming stage was human-initiated (i.e., if the participant liked at least two activity proposals by the agent despite the agent deleting all activities suggested by the participant).

From an Activity Theory perspective, this extra step of acknowledging human-agent disagreement represented a temporarily shift in activities and their objectives, namely from planning activities to resolving human-agent disagreements to facilitate team development. This highlights the rapid shifts that can occur between activities and the associated changes in objectives, tools, division of labor, rules, community, and subjects. The importance of addressing conflicts in team development in HAT has also been emphasized in studies by Jung et al. (2015) and Walliser et al. (2019), which showed that making human-AI conflicts explicit can enhance team performance. Jung et al. (2015) further revealed that when a robot agent helped resolve conflicts, it improved human participants' perception of the actor that introduced the conflict. In our study, the conflict management step in the collaborative planning activity may have similarly contributed to a positive perception of the agent, despite the agent's prior task-directed conflict behavior during the storming phase. However, the conflict management strategy in our collaborative planning activity was relatively simple, and our results based on qualitative data cannot be generalized to broader populations. Thus, further research is needed to explore conflict management for team development in HAT, including when and how conflicts should be addressed for team development in HAT. Team preferences may vary, with some teams preferring to minimize task disruptions and avoid discussing all tensions, whereas others might favor addressing even minor disagreements. For some participants in our study, disagreements with an agent were also described as beneficial and even necessary for personal development, such as behavior change. This aligns with Activity Theory, which views conflicts (i.e., breakdown situations) as a driving force for development and, therefore, as a positive element (Kaptelinin and Nardi, 2006).

The human-AI conflicts encountered by participants in our study could also reflect an internalization process related to using a digital tool, which has been the focus in previous research on human-computer interaction within the framework of Activity Theory (Clemmensen et al., 2016; Lindgren et al., 2018; Guerrero et al., 2019). This process is characterized by users' transformation through stages of conscious use of tools. Initially, the users must consciously engage with unfamiliar software (i.e., tool) as they learn how it functions. This learning process of mastering new skills naturally leads to conflicts (i.e., contradictions within the activity system) (Engeström, 2001). Over time, as users become more accustomed to the tool, a transition occurs: the software (i.e., tool) shifts from being the focused object of the activity to a mediator within the activity that no longer requires conscious thought (Kaptelinin and Nardi, 2006). In our study, participants who reached the norming stage may have experienced a similar transformation. However, for those who remaining in the storming stage, a different shift might have occurred. These participants may have begun to perceive the agent not as a collaborator but as a tool designed to execute tasks according to their instructions, rather than engaging in a collaborative process. In other words, for those participants, the activity system may have transformed into a more developed version, where the agent is given a different role in the activity (i.e., a tool) (Engeström, 1999). Exploring how participants perceive the agent's role within the activity and how

this perception might influence whether and how users transition out of the storming stage is an interesting focus for future research.

In some instances, conflicts may also be necessary to prevent negative outcomes, as also evident in research related to obedience and rebellion AI (Mayol-Cuevas, 2022; Aha and Coman, 2017). For example, prior research looked at design mechanisms that trigger conflicts to ensure that users are aware of critical information such as life-critical aspects in a medical health record system or a cockpit. From the activity-theoretical perspective, this focuses on bringing the tool into the focus of the activity. In such situations and depending on the activity, the role of the agent, and the criticality of the information the agent provides, it may not always be desirable to advance into the norming stage. For example, if the agent is designed to enforce regulations (represented by the rules-node of the activity system model in Figure 1), it should not conform to a human who attempts to violate these rules. In our use case scenario, if the human would disagree with the agent's proposal of a physical activity for improving health, with the claim that physical activity in general is unhealthy, the agent should try to understand the basis of this perception. However, it should remain steadfast in its domain knowledge, which includes knowledge about the positive effects of physical activities on health, even if this may risk not reaching the norming stage as a human-AI team.

In our study, participants who did not transition to the norming stage remained frustrated with the agent, leading to disruptions in collaboration, such as giving up on the agent and the human-agent teamwork. Similarly, research by Lindgren et al. (2024) exploring collaboration between older adults and an unknowledgeable agent during a common daily life activity found that older adults "put the agent to sleep" when it interrupted social situations. Exiting a collaboration with an AI system might be easier than with a human, introducing new dynamics to how stages of team development may unfold in HAT. The tendency to exiting collaboration and thereby "resolving" conflicts, may also explain findings from McNeese et al. (2021), where AI-inclusive teams displayed fewer conflicts than human-only teams. This could indicate that humans engage in fewer or different types of conflicts with AI systems compared to humans, potentially altering the nature of the storming stage in human-AI teams. The fact that some participants in our study experienced the norming stage indicates that conflicts can be resolved. However, whether this conflict resolution helps the team development in the same way it does in human-only teams, and whether the stages of team development in HAT mirror those in human-only team, are subject to future research.

For human-only teams, the success of team development depends on the involved actors (Morgeson et al., 2005). In other words, team development can be more challenging with an actor that lacks social competence and the ability to adapt to other team members' preferences and needs (Morgeson et al., 2005; Fisher and Marterella, 2019). In the context of HAT, this has also been noted by Walliser et al. (2019), who pointed out that critiques of HAT often arise in the context of less developed agents than those available of today. The need to adapt to team members' needs is closely related to the necessity of further personalizing an agent, as also suggested by participants. They proposed several person- and context- or task-specific factors, such as knowledge of activities that individuals' usually perform, how those activities

make them feel, and information about their family and work life (e.g., flexibility of work schedule and daily routines), which could support the planning process. Given the importance of addressing individuals' current states and needs for effective teamwork and conflict management (McKibben, 2017), personalization is crucial for supporting the transition to norming and should be a focus for future research. To incorporate the identified personalization factors, a hybrid approach is needed, where both the agent and human provide complementary perspectives on the task. In future versions of the implemented system, the agent could offer more person-tailored and context-tailored domain knowledge and knowledge about the human and the task, including an extended knowledge graph to capture person- and task-specific factors for personalization. Machine learning could be employed to predict and suggest preferred activities during preferred timeslots, and/or a large language model could be integrated for a more personalized experience when humans interact and collaborate with the agent.

6.1 Strengths and limitations

This study presents several strengths and limitations that should be considered when interpreting the results. Strengths include the interdisciplinary approach, which integrated expertise from long-term patient experiences throughout the whole research process. Additionally, the study is grounded in well-established theories, namely, the Tuckman's model of team development (Bonebright, 2010) and Activity Theory (Kaptelinin and Nardi, 2006; Vygotsky, 1978; Engeström, 1999), which provide a solid foundation for the research. Limitations include the focus on only one scenario (i.e., collaborative calendar planning task), the limited number of participants for evaluating the agent embedded in the collaborative task, the inclusion of only two team members, and the short task for transitioning through Tuckman's stages of team development. Tuckman's model includes five stages that typically occur over longer periods of time. For feasibility reasons, we selected three stages instead of five and a short interaction time (within 30 min). Although we did not study all the stages of team development, our aim was fulfilled as the focus of this study was on the initial stages, particularly the storming stage.

The mechanisms embedded in the agent application that directed its behavior included thresholds for transitioning between Tuckman's stages for team development. Depending on how these thresholds were set, the timing of the agent's behavior may have been more or less well-tuned to different individuals. It can be assumed that future personalization would also include tailoring these thresholds to accommodate different collaboration styles. Future studies will explore whether such a tailored mechanism could enable more participants to notice a clear transition of the human-agent teams from the storming to the norming stage.

Another limitation of the current application version is the extent of the agent's knowledge regarding the health aspects of activities. While the embedded knowledge was considered sufficient to study the principles of team development, some objections to the agent's proposals may have stemmed from participants perceiving the agent's motivation as questionable. Future studies will aim to develop a more dynamic dialogue system that leverages large

language models and knowledge graphs, embedding knowledge specific to the domain of health behavior change.

7 Conclusions

This study explored human-agent team development through the first three stages of Tuckman's model. We designed a collaborative activity and developed an Android calendar application with an embedded agent whose design is guided by Tuckman's team-development model and Activity Theory. Our findings indicate first attempts for team development among participants and the agent. They suggest extending this agent to include more comprehensive knowledge of activities, more person-tailored mechanisms for eliciting disagreements, and managing conflicts to facilitate a transition of the human-agent team from the storming to the norming stage. Future studies should focus on these enhancements, as well as on further developing transitions to additional stages in Tuckman's model for team development.

Data availability statement

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author/s.

Ethics statement

The studies involving humans were approved by Swedish Ethical Review Authority (Dnr 2019-02794). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

VK: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing. MT: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing. SB: Formal analysis, Visualization, Writing – review & editing, Data curation, Writing – original draft. HL: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by the WASP-HS project Digital Companions as Social Actors: Employing Socially Intelligent Systems for Managing Stress and Improving Emotional Wellbeing, funded by Marianne and Marcus Wallenberg Foundation (grant

number MMW 2019.0220); the project Collaborative Storytelling in Argument-based Micro-dialogues to Improve Health funded by the KEMPE Foundation (grant number JCSMK22-0158); and the HumanE-AI-Net project, which received funds from the European Union's Horizon 2020 research and innovation programme (grant agreement 952026).

Acknowledgments

We sincerely thank all participants for their participation in this research study. We thank Michele Persiani for his support with programming and Kaan Kilic, a member of the Interactive and Intelligent Systems Lab, for his input and critical feedback on aspects related to this study. We used ChatGPT, version 4.0 for grammar checks.

References

- Aha, D., and Coman, A. (2017). "The AI rebellion: changing the narrative," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 31. doi: 10.1609/aaai.v31i1.11141
- Aronsson, I., Neely, A. S., Boraxbekk, C.-J., Eskilsson, T., and Gavelin, H. M. (2024). "Recovery activities are needed every step of the way" - exploring the process of long-term recovery in people previously diagnosed with exhaustion disorder. *BMC Psychol.* 12, 248. doi: 10.1186/s40359-024-01756-z
- Berretta, S., Tausch, A., Ontrup, G., Gilles, B., Peifer, C., and Kluge, A. (2023). Defining human-AI teaming the human-centered way: a scoping review and network analysis. *Front. Artif. Intell.* 6:1250725. doi: 10.3389/frai.2023.1250725
- Bødker, S., Dindler, C., Iversen, O. S., and Smith, R. C. (2022). *Can You Define Participatory Design For Me?* Cham: Springer International Publishing, 15–29. doi: 10.1007/978-3-031-02235-7_1
- Bonebright, D. A. (2010). 40 years of storming: a historical review of Tuckman's model of small group development. *Hum. Resour. Dev. Int.* 13, 111–120. doi: 10.1080/13678861003589099
- Charas, S. (2015). Improving corporate performance by enhancing team dynamics at the board level. *Int. J. Disclosure Govern.* 12, 107–131. doi: 10.1057/jdg.2013.35
- Clemmensen, T., Kaptelinin, V., and Nardi, B. (2016). Making HCI theory work: an analysis of the use of activity theory in HCI research. *Behav. Inf. Technol.* 35, 608–627. doi: 10.1080/0144929X.2016.1175507
- Crowley, J. L., Coutaz, J., Grosinger, J., Vazquez-Salceda, J., Angulo, C., Sanfeliu, A., et al. (2022). A hierarchical framework for collaborative artificial intelligence. *IEEE Perv. Comput.* 22, 9–18. doi: 10.1109/MPRV.2022.3208321
- Doyle, L., McCabe, C., Keogh, B., Brady, A., and McCann, M. (2020). An overview of the qualitative descriptive design within nursing research. *J. Res. Nurs.* 25, 443–455. doi: 10.1177/1744987119880234
- Elo, S., and Kyngäs, H. (2008). The qualitative content analysis process. *J. Adv. Nurs.* 62, 107–115. doi: 10.1111/j.1365-2648.2007.04569.x
- Engeström, Y. (1999). *Activity Theory and Individual and Social Transformation*. Cambridge: Cambridge University Press, 19–38. doi: 10.1017/CBO9780511812774.003
- Engeström, Y. (2001). Expansive learning at work: toward an activity theoretical reconceptualization. *J. Educ. Work* 14, 133–156. doi: 10.1080/13639080020028747
- Fisher, A., and Marterella, A. (2019). *Powerful Practice: A Model for Authentic Occupational Therapy*. Ft. Collins, USA: Center for Innovative OT Solutions.
- Gray, B. (1989). *Collaborating: Finding Common Ground for Multiparty Problems*. CA, USA: Jossey-Bass Publishers.
- Guerrero, E., Lu, M.-H., Yueh, H.-P., and Lindgren, H. (2019). Designing and evaluating an intelligent augmented reality system for assisting older adults' medication management. *Cogn. Syst. Res.* 58, 278–291. doi: 10.1016/j.cogsys.2019.07.001
- Iftikhar, R., Chiu, Y.-T., Khan, M. S., and Caudwell, C. (2023). Human-agent team dynamics: a review and future research opportunities. *IEEE Trans. Eng. Manag.* 71, 10139–10154. doi: 10.1109/TEM.2023.3331369
- Jung, M. F., Martelaro, N., and Hinds, P. J. (2015). "Using robots to moderate team conflict: the case of repairing violations," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 229–236. doi: 10.1145/2696454.2696460
- Kaptelinin, V., and Nardi, B. A. (2006). *Acting With Technology: Activity Theory and Interaction Design*. Cambridge, USA: MIT Press. doi: 10.5210/fm.v12i4.1772
- Kluge, A., Ontrup, G., Langhoff, V., and Wilkens, U. (2021). Mensch-KI-teaming: Mensch und künstliche Intelligenz in der Arbeitswelt von morgen. *Zeitschrift Wirtschaftlichen Fabrikbetrieb* 116, 728–734. doi: 10.1515/zwf-2021-0112
- Liedtka, J. M., Whitten, E., and Sorrells-Jones, J. (1998). Enhancing care delivery through cross-disciplinary collaboration: a case study/practitioner response. *J. Healthcare Manag.* 43, 185–203. doi: 10.1097/00115514-199803000-00013
- Lindgren, H., Kaelin, V. C., Ljusback, A.-M., Tewari, M., Persiani, M., and Nilsson, I. (2024). "To adapt or not to adapt? older adults enacting agency in dialogues with an unknowledgeable agent," in *Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization* (New York, USA: Association for Computing Machinery), 307–316. doi: 10.1145/3627043.3659562
- Lindgren, H., Lu, M.-H., Hong, Y., and Yan, C. (2018). Applying the zone of proximal development when evaluating clinical decision support systems: a case study. *Stud. Health Technol. Inform.* 247, 131–135.
- Lindgren, H., and Weck, S. (2022). "Contextualising goal setting for behaviour change - from baby steps to value directions," in *Proceedings of the 33rd European Conference on Cognitive Ergonomics, ECCE '22* (New York, USA: Association for Computing Machinery), 1–7. doi: 10.1145/3552327.3552342
- Mathieu, J. E., Gallagher, P. T., Domingo, M. A., and Klock, E. A. (2019). Embracing complexity: reviewing the past decade of team effectiveness research. *Annu. Rev. Organ. Psychol. Organ. Behav.* 6:17–46. doi: 10.1146/annurev-orgpsych-012218-015106
- Mathieu, J. E., Hollenbeck, J. R., van Knippenberg, D., and Ilgen, D. R. (2017). A century of work teams in the journal of applied psychology. *J. Appl. Psychol.* 102, 452–467. doi: 10.1037/apl0000128
- Mayol-Cuevas, W. W. (2022). Rebellion and disobedience as useful tools in human-robot interaction research-the handheld robotics case. *arxiv.org/abs/2205.03968*.
- McKibben, L. (2017). Conflict management: importance and implications. *Br. J. Nurs.* 26, 100–103. doi: 10.12968/bjon.2017.26.2.100
- McNeese, M. D. (1986). Humane intelligence: a human factors perspective for developing intelligent cockpits. *IEEE Aerospace Electr. Syst. Magaz.* 1, 6–12. doi: 10.1109/MAES.1986.5005199
- McNeese, N. J., Demir, M., Cooke, N. J., and She, M. (2021). Team situation awareness and conflict: a study of human-machine teaming. *J. Cogn. Eng. Decis. Mak.* 15, 83–96. doi: 10.1177/15553434211017354
- Miller, D. L. (2003). The stages of group development: a retrospective study of dynamic team processes. *Canadian J. Admin. Sci.* 20, 121–134. doi: 10.1111/j.1936-4490.2003.tb00698.x
- Morgeson, F. P., Reider, M. H., and Campion, M. A. (2005). Selecting individuals in team settings: the importance of social skills, personality characteristics, and teamwork knowledge. *Pers. Psychol.* 58, 583–611. doi: 10.1111/j.1744-6570.2005.655.x
- Naser, M. Y., and Bhattacharya, S. (2023). Empowering human-AI teams via intentional behavioral synchrony. *Front. Neuroergon.* 4, 1181827. doi: 10.3389/fnrgo.2023.1181827

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Nowak, A., Lukowicz, P., and Horodecki, P. (2018). Assessing artificial intelligence for humanity: will AI be the our biggest ever advance? Or the biggest threat [opinion]. *IEEE Technol. Soc. Magaz.* 37, 26–34. doi: 10.1109/MTS.2018.2876105
- O'Neill, T. A., Mcneese, N. J., Barron, A., and Schelble, B. G. (2022). Human-autonomy teaming: a review and analysis of the empirical literature. *Hum. Factors* 64, 904–938. doi: 10.1177/0018720820960865
- Patt, N., Kupjetz, M., Kool, J., Hersche, R., Oberste, M., Joisten, N., et al. (2023). Effects of inpatient energy management education and high-intensity interval training on health-related quality of life in persons with multiple sclerosis: a randomized controlled superiority trial with six-month follow-up. *Mult. Scler. Relat. Disord.* 78, 104929. doi: 10.1016/j.msard.2023.104929
- Prada, R., Homan, A. C., and van Kleef, G. A. (2024). “Towards sustainable human-agent teams: A framework for understanding human-agent team dynamics,” in *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems, AAMAS '24* (Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems), 2696–2700.
- Runkel, P. J., Lawrence, M., Oldfield, S., Rider, M., and Clark, C. (1971). Stages of group development: an empirical test of Tuckman's hypothesis. *J. Appl. Behav. Sci.* 7, 180–193. doi: 10.1177/002188637100700204
- Schelble, B. G., Flathmann, C., McNeese, N. J., Freeman, G., and Mallick, R. (2022). “Let's think together! Assessing shared mental models, performance, and trust in human-agent teams,” in *Proceedings of the ACM on Human-Computer Interaction* (Association for Computing Machinery), 1–29. doi: 10.1145/3492832
- Schmutz, J. B., Outland, N., Kerstan, S., Georganta, E., and Ulfert, A.-S. (2024). AI-teaming: Redefining collaboration in the digital era. *Curr. Opin. Psychol.* 58, 101837. doi: 10.1016/j.copsyc.2024.101837
- Sokman, Y., Othman, A., Abdul Aziz, A., Musa, M., Azizan, N., and Rahmat, N. (2023). Stages in group work: is there a relationship among them? *Int. J. Acad. Res. Bus. Soc. Sci.* 13, 19674. doi: 10.6007/IJARBSS/v13-i11/19764
- Tewari, M., and Lindgren, H. (2022). Expecting, understanding, relating, and interacting-older, middle-aged and younger adults' perspectives on breakdown situations in human-robot dialogues. *Front. Robot. AI* 9, 956709. doi: 10.3389/frobt.2022.956709
- Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychol. Bull.* 63, 384. doi: 10.1037/h0022100
- Vygotsky, L. S. (1978). *Mind in Society: Development of Higher Psychological Processes*. Harvard, USA: Harvard University Press.
- Walliser, J. C., de Visser, E. J., Wiese, E., and Shaw, T. H. (2019). Team structure and team building improve human-machine teaming with autonomous agents. *J. Cogn. Eng. Dec. Making* 13, 258–278. doi: 10.1177/1555343419867563
- Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., and Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science* 330, 686–688. doi: 10.1126/science.1193147
- World Health Organization (2001). *International Classification of Functioning, Disability and Health*. Geneva, Switzerland: World Health Organization.
- Zhang, R., McNeese, N. J., Freeman, G., and Musick, G. (2021). “An ideal human” expectations of AI teammates in human-AI teaming. *Proc. ACM Hum. Comput. Inter.* 4, 1–25. doi: 10.1145/3432945