

***Empathosphere*: Promoting Constructive Communication in Ad-hoc Virtual Teams through Perspective-taking Spaces**

PRANAV KHADPE, Carnegie Mellon University, USA

CHINMAY KULKARNI, Carnegie Mellon University, USA

GEOFF KAUFMAN, Carnegie Mellon University, USA

When members of ad-hoc virtual teams need to collectively ideate or deliberate, they often fail to engage with each others' perspectives in a constructive manner. At best, this leads to sub-optimal outcomes and, at worst, it can cause conflicts that lead to teams not wanting to continue working together. Prior work has attempted to facilitate constructive communication by highlighting problematic communication patterns and nudging teams to alter their interaction norms. However, these approaches achieve limited success because they fail to acknowledge two social barriers: (1) it is hard to reset team norms mid-interaction, and (2) corrective nudges have limited utility unless team members believe it is safe to voice their opinion and that their opinion will be heard. This paper introduces *Empathosphere*, a chat-embedded intervention to mitigate these barriers and foster constructive communication in teams. To mitigate the first barrier, *Empathosphere* leverages the known benefits of "experimental spaces" in dampening existing norms and creating a climate conducive to change. *Empathosphere* instantiates this "space" as a separate communication channel in a team's workspace. To mitigate the second barrier, *Empathosphere* harnesses the benefits of perspective-taking to cultivate a group climate that promotes a norm of members speaking up and engaging with each other. *Empathosphere* achieves this by orchestrating authentic socio-emotional exchanges designed to induce perspective-taking. A controlled study ($N = 110$) compared *Empathosphere* to an alternate intervention strategy of prompting teams to reflect on their team experience. We found that *Empathosphere* led to higher work satisfaction, encouraged more open communication and feedback within teams, and boosted teams' desire to continue working together. This work demonstrates that "experimental spaces," particularly those that integrate methods of encouraging perspective-taking, can be a powerful means of improving communication in virtual teams.

CCS Concepts: • **Human-centered computing** → **Empirical studies in collaborative and social computing**; **Empirical studies in HCI**; **Collaborative and social computing systems and tools**.

Additional Key Words and Phrases: virtual teams; team communication; social norms; perspective-taking; team dynamics

ACM Reference Format:

Pranav Khadpe, Chinmay Kulkarni, and Geoff Kaufman. 2022. *Empathosphere*: Promoting Constructive Communication in Ad-hoc Virtual Teams through Perspective-taking Spaces. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 55 (April 2022), 27 pages. <https://doi.org/10.1145/3512902>

1 INTRODUCTION

Online platforms empower people to come together, collaborate, and contribute towards common goals, often in an ad-hoc fashion - in effect forming "virtual teams." Such ad-hoc virtual teams- often

Authors' addresses: Pranav Khadpe, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA, pkhadpe@cs.cmu.edu; Chinmay Kulkarni, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA, chinmayk@cs.cmu.edu; Geoff Kaufman, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA, gfk@cs.cmu.edu.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2022 Copyright held by the owner/author(s).

2573-0142/2022/4-ART55

<https://doi.org/10.1145/3512902>

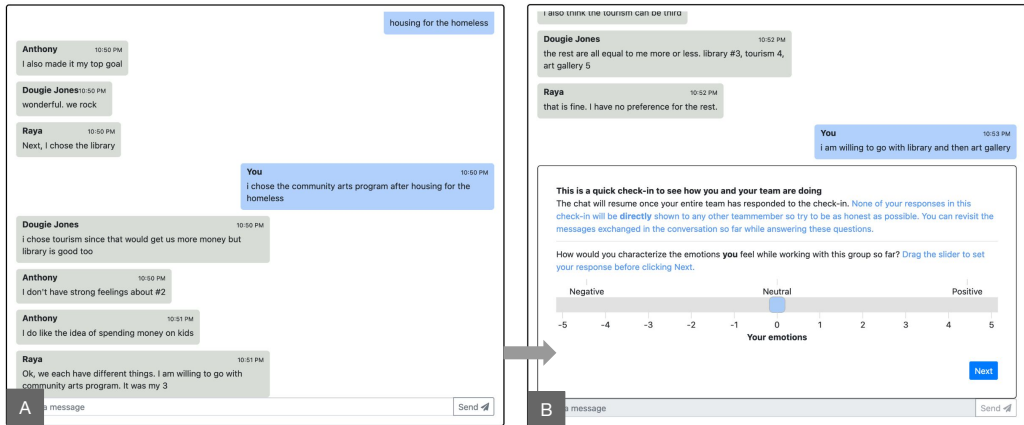


Fig. 1. (A) Virtual teams often co-ordinate and communicate over chat. *Empathosphere* helps teams establish positive interaction norms that promote sharing and engaging with each other's perspectives. (B) When triggered, *Empathosphere* appears as a widget in the chat interface and disables the chat while teams carry out a perspective-taking exercise. The perspective-taking exercise in *Empathosphere* supports making team members more attuned to each others' emotions. It does this by prompting members to reflect on how others might be feeling and by providing them with socio-metric feedback on group climate and feedback on their perceptiveness towards their teammates emotions. We find that *Empathosphere* improves work satisfaction and encourages more open communication and feedback within the team.

composed of strangers brought together by a shared purpose- have demonstrated the potential to accomplish complex tasks including authoring open-source software [76], authoring open-access encyclopedia articles [11], interface prototyping [85], collective design [72], and even research [84]. Virtual teams have indeed achieved complex goals, but only when supported by tools and processes that decompose these goals into tasks, workflows, or well defined roles, ultimately dividing labor into tasks that *individual* workers can accomplish. When virtual teams engage in tasks that cannot be decomposed, where team members need to *collectively* ideate, engage with each others' perspectives, and make decisions, they often fail [64]: at best, their outcomes are suboptimal [9]; at worst they struggle to manage inter-team conflict [46], which can end in team members refusing to continue collaborating [91]. Therefore, despite their promise, virtual teams are currently limited in their ability to collaborate effectively.

A possible response to these challenges of ad-hoc teams is to design systems that make disagreements and conflicts between team members less likely, so teams are more apt to succeed. Unfortunately, disagreements are often the result of diverse perspectives and ideas of individual members [75], and so eliminating disagreements may well eliminate these diverse contributions as well. Prior work shows that disagreement is in fact a double-edged sword: when managed well, open communication of disagreement and alternate ideas can improve collaboration [24, 60, 61]. When managed poorly, disagreement can spiral out of control, triggering interpersonal conflict [46, 91].

Thus, when teams need to accomplish tasks that benefit from the diverse perspectives of members and require collective decision-making, surfacing and managing team friction is crucial. Specifically, members need to be able to communicate opinions openly but without conflicts of opinion impairing teams. While all teams struggle to strike this balance, the lack of richness of computer-mediated communication compared to face-to-face communication exacerbates this problem for virtual teams. Without such balance, virtual teams are prone either to social loafing where team members

avoid speaking up and voicing their opinions [44], or situations where team members ignore other members' perspectives, triggering conflict and threatening the team's performance and sustainability [32].

To help members of virtual teams engage constructively with each others' perspectives, this paper introduces *Empathosphere*, a chat-embedded intervention that attempts to mitigate some of the social and psychological barriers to open team communication. While the artifacts that ad-hoc virtual teams work on, appear on one set of platforms (e.g. Github in the case of open source teams), the work of discussion and deliberation often occurs through real-time discussion channels like chat [71, 96]: open source teams have Discord servers [1], Wikipedians use IRC [3], and some teams of crowdworkers use Slack [71, 91]. The primary role of chatrooms in facilitating team deliberation make them an important setting for intervention.

Empathosphere identifies and addresses two potential barriers to healthy communication in virtual teams based on literature in organizational psychology. The first barrier is that groups require a climate of *safety* and *efficacy* for members to speak up- team members speak up only when they feel that it is safe to voice their opinion (*safety*) and that their opinion will be taken into account (*efficacy*) [62]. The second barrier is that that team norms can be hard to reset mid-interaction [58, 92]. In response, *Empathosphere* induces perspective-taking as a useful tool to maintain the *safety* and *efficacy* essential for team members to engage with each other. To transform established norms and cultivate *safety* and *efficacy* by inducing perspective-taking, *Empathosphere* leverages the power of "experimental spaces" [25, 97] which are social settings separated by physical, temporal or symbolic boundaries from everyday work that diminish the salience of existing interaction norms and catalyze the emergence of new dynamics [12, 25, 50, 97].

Empathosphere instantiates a norm-diminished "experimental space" through a temporally and spatially separated channel within a chat-based collaboration environment (Fig. 1). Within *Empathosphere*, we orchestrate an exercise aimed at inducing perspective-taking. By uniting the benefits of experimental spaces and perspective-taking, we find that *Empathosphere* improves work satisfaction, encourages more open communication and feedback within teams, and boosts team members' desire to continue working together, i.e. team viability.

In doing so, *Empathosphere* shifts the focus from easily observable team communication patterns to the subtle (but vitally important) social and psychological barriers preventing team members from changing these patterns. For instance, prior work has highlighted problematic communication patterns by providing sociometric feedback [45, 52, 53, 82] and by making explicit recommendations for behavior change [44, 82]. These systems assume that highlighting problematic patterns and suggesting changes is sufficient to transform behaviors. However, because they ignore the social and psychological barriers preventing team members from speaking up, they often have adverse outcomes: a system that displayed group-level agreement led to team members expressing agreement with a majority opinion even if they did not agree with it, in attempts to improve displayed agreement [53].

To test if *Empathosphere* is effective, we conducted a controlled-study ($N = 110$ in 24 teams) where we compared the benefits of *Empathosphere* against an alternative intervention in which participants simply reflected on their team experience. In both conditions, participants recruited from Amazon Mechanical Turk worked synchronously on a decision-making task, in teams of 4-6 via a chat-based collaboration system.

We found that participants in the *Empathosphere* condition were significantly more satisfied with their team's solution and they judged their teams to be significantly more viable. Furthermore, we find that participants in the *Empathosphere* condition were approximately 25.4% more likely to be willing to provide feedback to teammates than in the control condition, and a larger fraction of participants were willing to receive feedback from teammates. Analyzing linguistic patterns

exhibited by teams during deliberation in the two conditions, we find that *Empathosphere* increased teams' usage of second-person pronouns, suggesting that the conversation was more other-focused: members engaged with what other team members were saying and drew them into the conversation. *Empathosphere* also led to an increase in informality and netspeak in the teams' chatlogs suggesting that team members formed stronger social bonds.

This work contributes embedded interventions that use “experimental spaces” coupled with protocols to enhance perspective-taking as a powerful tool in establishing positive interaction norms in virtual teams. This opens up possibilities to use such spaces to inspire counternormative behaviors that can improve group climate and even, amplify minority voices in teams. With *Empathosphere*, we suggest that mitigating the social and psychological barriers to changing interaction norms is a crucial step in actually transforming them.

2 RELATED WORK

This research contributes to a growing body of work that aims to improve collaboration experiences of computer-mediated teams, tying together research in organizational behavior with research in HCI. We start by reviewing existing HCI research on software support for teams and highlight the need for systems that foster positive interaction norms. We then characterize interaction patterns of performant and sustainable teams by drawing on research in organizational behavior. Finally, we outline potential approaches for designing effective interventions to cultivate these constructive interaction patterns in teams.

2.1 Beyond team assembly and task support: designing systems to nurture teams

To improve outcomes of computer-mediated teams, computer-supported collaborative work (CSCW) researchers have developed systems that aim to assemble the “right team” for the task. Prior work has shown that teams composed of members with higher skill diversity [36], demographic diversity [6], and differing personality types [55] perform better and are more satisfied. Consequently, researchers have developed systems that optimize for these criteria and assemble teams with the most promise for success. Researchers have also built “team-dating” systems to improve collaboration. These systems allow people to experience work with different potential collaborators before team formation [56]. Other systems rotate members strategically to promote mixing of ideas [71]. To allow teams to cope with evolving work, researchers have also designed systems to on-board team members with the right expertise in near real-time [69, 85]. However, while these systems have demonstrated success in identifying good collaborators, team composition is naturally constrained by availability of talent. How can we help *any* team achieve their full potential and work effectively?

Once teams are formed, one way to help them work effectively is to provide support for task completion and coordination of effort. CSCW researchers have designed several systems to support team tasks. This includes systems to create and manage workflows [68, 85], systems to structure interaction between collaborators [65, 68]. These systems decompose work into tasks that can be handled individually by members. As such, they reduce the amount of task planning that teams need to do and lower the need for members to communicate with each other. Consequently, when tasks can be decomposed into independent fragments and members' roles can be well-defined, these systems allow virtual teams to thrive. However, when tasks cannot be completely compartmentalized, and members need to come co-ideate, deliberate, and negotiate, communication becomes crucial again, and computer-mediated teams are once again vulnerable to conflict and misunderstanding [91, 92]. For instance, one prior study showed that even when task-support interventions improved the performance of participant teams on average, interpersonal conflict in some teams negated the positive effects of the intervention [10].

Healthy team communication is foundational: teams with healthy interpersonal processes provide a fertile ground for the positive impact of task-support interventions to improve performance [10]. Without effective communication, team performance can suffer and can even lead to teams not wanting to continue working together. Furthermore, team performance alone doesn't paint a complete picture of how a collaboration is going. Even high performing teams might have toxic culture, unresolved conflict, divisive interactions, and member burnout [19, 20]. For a more holistic evaluation of a team, it is important to look beyond just *team performance* and evaluate *team viability*: the team's sustainability and capacity for continued success [8, 16, 91]. Team viability is an emergent state that captures the potential of the team to continue performing well, the desire of the team to continue working together, and the extent to which the team can adapt to changing circumstances [8]. Antecedents of team viability include perceived performance and interpersonal friction [16]. Consequently, viability requires that in addition to being performant, teams also have healthy communication and interpersonal processes in place [16].

2.2 Managing team-friction: identifying interaction patterns of viable teams

Research on what constitutes constructive team-communication patterns suggests that teams need to have the right amount of friction, which in turn must be managed well. Too much disagreement, unresolved conflict, and disregard for others' opinions and criticism can trigger negative emotions, causing members to get angry and distressed [41]. This distress can spiral out of control and lead to a loss of viability [41]. Too little disagreement and critical communication is also detrimental and can impair team performance [4, 67], team satisfaction, and consequently, viability. However, for teams to cultivate and benefit from a "just right" amount of friction - where members communicate opinions openly but opinion conflicts don't hinder teams - teams must manage conflict so that members feel it is both *safe* and *effective* to voice their opinions [62].

Team members feel *safe* expressing themselves when they do not anticipate potential negative consequences of speaking up [62]: they can voice their opinion without fear of being attacked or reprimanded for speaking up. However, safety is not enough; to improve outcomes, team members must also engage with others' perspectives, foster inclusivity, and affirm each other [15, 34, 71]. This not only allows teams to incorporate diverse opinions of members but also signals to members that their opinions will be taken into account and acted upon. This can persuade members that voicing their opinions is an *effective* way to enact change and, in turn, motivate them continue to voice their opinions.

The diverse perspectives and unique expertise of individual members make teams potent and so, for teams to be effective, it is crucial that team-members both speak up expressing their perspectives and that they engage with others' perspectives, in a respectful manner [62]. Eliciting diverse perspectives by fostering open communication of ideas, opinions, and issues can help uncover new solutions [60], improve decision making, facilitate team learning [61, 62], and help identify problems.

CSCW researchers have developed methods to detect whether the interaction patterns in a team are healthy [78, 88]. They have developed models that can leverage information about facial expression, eye gaze, head movements, speech, text, physiology, and interaction to detect the presence of positive traits [14, 78, 88] such as turn-taking [40], interactivity [5], high-levels of empathy [38], and team cohesion [63]. These models have also enabled predicting team performance and team viability from a team's interactions [14]. While this line of work has furthered our understanding of which interpersonal processes and behaviors promote team viability, cultivating these prosocial behaviors in teams remains challenging [50]. Our work attempts to move from these descriptive insights towards prescriptive interventions to help teams develop positive communication patterns.

2.3 Fostering constructive interaction patterns in teams

Prior work has attributed the difficulty of establishing healthy communication to the low levels of information richness, lack of social cues, context, and shared emotion in computer-mediated teams [33]. In response, researchers have developed more information-rich communication channels to improve social cues and emotional expression in these messaging systems. This includes systems to communicate affect by altering the font and animation of text [13, 49, 89], and via abstract visuals [79, 80, 87]. Researchers have also explored the design of computer-mediated interventions to nudge people towards more positive interaction patterns by providing them socio-metric feedback on group communication patterns [45, 52, 53, 82]. Some tools go beyond making extant behaviors salient and also make specific recommendations on how teams should alter their practices to move them closer to desired behaviors [44, 82]. However, these systems have only had limited success, because they assume that highlighting problematic patterns and providing information-rich communication channels is sufficient to nudge team members to express their authentic opinions and change their interaction norms. In doing so, they ignore the social and psychological barriers preventing team members from speaking up: specifically, they do nothing to alter team members' perceived *safety* and *efficacy* of speaking up, necessary for open and healthy team communication. This can lead to undesirable outcomes: prior work found that visualizing group-level agreement led to a form of social loafing, where team members expressed agreement with the majority opinion even if they did not agree with it with it, ultimately resulting in lower quality work [53].

Transforming and shaping interaction norms to promote open and respectful communication is challenging. It requires team members to take initiative and proactively engage in open and frank communication but team members often do not speak up. They may view expressing disagreement, voicing concerns, or asking questions as risky (low levels of perceived *safety*), or they may believe that their inputs will not be taken seriously [62] (low levels of perceived *efficacy*), pushing them to avoid disagreeing entirely. These beliefs about speaking up are shaped by the interaction history of the team [62]. For newly convened teams, an uncertainty about group norms and potential consequences of speaking up could lead to low perceived *safety* and discourage individuals from speaking up. Team members who do not strongly identify with their team, or who are unsatisfied with their team, might perceive low levels of *efficacy* and think speaking up is futile [62]. In such cases, a "climate of silence" [61] becomes the norm, which can undermine team performance and viability [4, 67].

2.4 Psychological theories informing the design of our intervention

In designing our intervention, we leverage perspective-taking processes as a useful tool to maintain the *safety* and *efficacy* necessary for healthy team communication. To support the development of perspective-taking effectively, we design an environment that is conducive to the emergence of new practices. This environment is designed based on the concept of "spaces" (sometimes called "counter-normative", "experimental", or "transformative" spaces) that provide a zone of *safety* for experimental behaviors, which can then develop into a perspective-taking exercise.

2.4.1 Creating an experimental space to induce perspective-taking. To suspend pre-existing interaction norms and induce perspective-taking, we draw on work that has explored the benefits of creating "spaces" for transformative dialogue [50, 97]. Spaces are social settings separated by physical, temporal or symbolic boundaries from everyday work within which the salience of existing norms and interaction patterns diminishes. This creates an environment conducive to risky and counternormative behaviors that can catalyze the emergence of new dynamics [12, 25, 50, 97]. Examples of such spaces include creating meetings to specifically engage in discussion about work challenges or organizing a retreat to help team members connect outside of work environments.

Such “experimental spaces” have been effective at reducing barriers to open communication in teams and consequently, fostering positive relational dynamics [50]. In this paper, we explore what such a space might look like in the context of online team communication environments and how we might use such a space to induce perspective-taking towards the goal of improving team interactions.

2.4.2 Inducing perspective-taking to maintain safety and efficacy. Perspective-taking can be a useful tool to maintain perceived *safety* of speaking up. By taking another’s perspective, team members are more likely to anticipate disagreement, recognizing that other people will have different views [75]. This can both reduce initial opposition to others’ ideas, as well as mentally prepare individuals to handle opposition to their ideas. Both of these together contribute to *safety*.

Perspective-taking encompasses considering others’ evaluative standards which can lead to better evaluation of their ideas and more constructive feedback [34]. This can help team members feel heard and contribute to improved perceptions of *efficacy*. Trying to understand someone else’s perspective can also trigger clarifying communication that can ultimately facilitate better understanding of everyone’s ideas [34]. Perspective-taking can lead to a cognitive reframing that leads to better integration of others’ ideas [27]. Taken together, this suggests that perspective-taking can help teams take everyone’s opinions into account and make everyone feel heard, both of which improve perceived *efficacy* of speaking up.

3 EMPATHOSPHERE

To prompt team members to engage in perspective-taking, we designed and implemented *Empathosphere*, a “space” embedded within a chat-based collaboration interface within which participants engage in a perspective-taking exercise. To explore different possible designs within a completely configurable chat environment, we built a custom chat-based collaboration environment (Figure 1). This was a client-server web application built using Meteor.js. *Empathosphere* attempts to improve team communication and perspective taking by providing an experimental space that encourages members to express their authentic emotions and prompts reflection on others’ emotions. In doing so, *Empathosphere* uses the power of “spaces” to facilitate counternormative behaviors. When triggered, *Empathosphere* appears as a widget in the chat interface and disables the chat while teams carry out the perspective-taking exercise. The perspective-taking exercise (Figure 2) in *Empathosphere* was designed by drawing on known strategies to regulate perspective-taking.

3.1 Design of the perspective-taking exercise

A person’s ability to empathize or take-perspectives is not an immutable trait; instead it can be regulated by *attention modulation*: drawing attention to others’ emotions and affective states can increase perspective-taking [18, 94]. *Empathosphere* prompts this process by asking each member to reflect on how other members might be feeling and by highlighting discrepancies between their evaluation of the group and the group’s actual emotions.

To aid reflection and discrepancy evaluation, while still maintaining the ambiguity and anonymity necessary to prompt authentic expression, *Empathosphere* executes this process in three stages. First, *Empathosphere* privately elicits how team members actually feel about working with their group (Figure 2A) on a scale ranging from -5 to 5; -5 being the most negative and 5 the most positive. Next, it asks each member, in private, to guess how each of the other members in the team might be feeling on the same scale, to nudge them to direct their attention towards others in the team (Figure 2B). Finally, *Empathosphere* calculates the mean of responses from the first stage to present each participant with feedback about the aggregate group climate, i.e. how positive/negative the group as a whole is feeling, without revealing individual responses (Figure 2C). (While emotions i.e., how

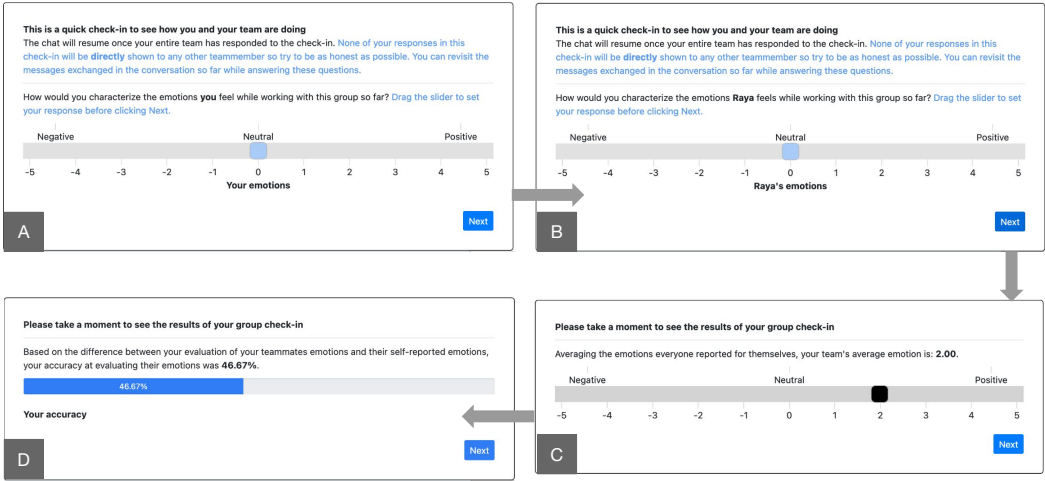


Fig. 2. *Empathosphere* creates a separated space within a team’s chat within which members engage in a perspective-taking exercise. They A) share their own socio-emotional states, B) think about the socio-emotional states of their teammates and finally, *Empathosphere* presents them with C) a measure of the group’s cumulative socio-emotional state and D) a measure of how accurate they were at evaluating their teammates’ true socio-emotional states.

people feel, can be multi-faceted, we focus on a simple positive/negative model in our intervention to ease reflection and evaluation for participants.) Forcing reflection on others’ emotions and presenting the average group climate are the first way *Empathosphere* draws members’ attention to the possibility that some team members might not be having a positive team experience.

Beyond attention modulation, perspective-taking is also regulated by *appraisal* or an individual’s evaluation of the authenticity and intensity of others’ emotions [94]. Emotions expressed privately in the “space” are likely to be more authentic and uninhibited than those expressed in the chat. So, team members are more likely to recognize potential grievances and act on them when they are confronted with them within the space. To further aid in perspective-taking, *Empathosphere* also presents every member with feedback on how accurate they were at guessing others’ emotions (Figure 2D) in their responses in the second stage above. Specifically, we use a representation of mean absolute error that facilitates easier interpretation. For a team member i , the accuracy \mathcal{A}_i is calculated as below, where G_{ji} is i ’s guess about j ’s emotional state, and S_j is j ’s self-reported emotional state:

$$\mathcal{A}_i = \max\left(0, 1 - \frac{1}{5} \sum_{i, i \neq j}^n \frac{|G_{ji} - S_j|}{n-1}\right)$$

In an earlier version of *Empathosphere*, we presented the raw mean absolute error, however, participants in our pilot studies found it hard to interpret this number. In response, we chose to transform the mean absolute error to an accuracy metric such that the accuracy is at 100% if a member guessed all other members’ emotions perfectly, and at 0% (in expectation) for random guesses. This measure can be negative (when disagreements exceed what can be expected by chance) but this is not useful for our reflective purposes, so we show participants 0% accuracy at minimum. \mathcal{A}_i is displayed to team member i as a percentage for easier interpretation. By controlling for agreements by random chance, this measure emphasizes discrepancies between an individual’s

evaluation of others' emotions and their actual emotions, with the goal of drawing attention to the valence and arousal levels of others' emotions.

4 METHOD

To investigate the impacts of *Empathosphere* on teams' expression and handling of conflicting opinions and diverse perspectives, we conducted a between-subjects study. This study was approved by our university's Institutional Review Board.

4.1 Participants

Participants were recruited from Amazon Mechanical Turk (AMT). We restricted participation to workers located in the United States who had completed at least 100 tasks and had an approval rate greater than 95%. All participants were compensated at a rate of \$15/hr with a bonus that was adjusted based on the time they spent on the experiment. Prior work on team viability has studied teams ranging in size from four to eight [92] members so we aimed to form teams of similar size in our study. We wanted to ensure that teams studied had at least four members so we split participants into groups of six to account for subsequent reduction in group size due to disconnections and dropouts. We only analyzed data from teams that had at least four members till the end of the task. If at any point the number of team members dropped below four, the task was terminated and participants were compensated for their time. In total, 38 participants were unable to complete the full study either because they dropped out themselves or because their teammates dropped out and the team strength dropped below four.

4.2 Experimental setup

We extended the Meteor.js collaboration application that houses *Empathosphere*, with the Turk-Server [57] framework. This allowed us to recruit participants on AMT and draw them into a lobby where they waited for other participants to arrive. Once there were a sufficient number of participants, the application automatically assigned teams to the different experimental conditions and created multiple parallel worlds where different teams worked simultaneously. Before they were added to chatrooms with their teams, participants were asked to set a pseudonym for themselves to preserve privacy, allowing them to control what aspects of their identity they wanted to reveal. Teams were randomly assigned to either the *Empathosphere* or control condition:

Empathosphere Condition: For teams in this condition, *Empathosphere* was triggered at the midpoint of the task and they were prompted to carry out the perspective-taking exercise.

Control Condition: Teams in the control condition were asked to take a two-minute pause and reflect on their teamwork experience individually. The specific prompt we used was: "*The experiment will proceed after a brief two-minute pause. Use this time to revisit the messages exchanged in the conversation so far and reflect on how the experience of working with this group has been.*"

We chose this silent, individual reflection activity to design a conservative control condition. Prior work suggests that a break in work can have some benefits for team members' ability to navigate conflict, irrespective of the activity they engage in during the break [17, 51, 86]. This research suggests that time delays can be effective at de-escalating negative emotions and encourage mindful decision making and so our control condition consists of a break with a silent individual activity. This control thus seeks to isolate the effects of processes engaged by specific activities in a counter-normative space, rather than outcomes that could be attributed to the presence of a break.

Through our pilots, we observed that the perspective-taking exercise took around two minutes on average and so, the control group's break duration was set to this time; so that the average time spent on the task would be similar for both the control and the *Empathosphere* group. Preserving

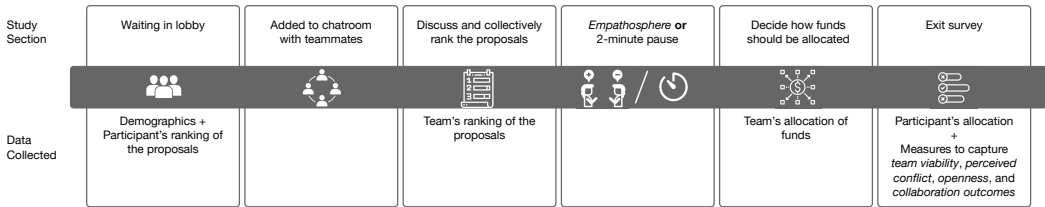


Fig. 3. The study workflow and data collected in the different stages of the study.

the task time between conditions seeks to further eliminate differences in familiarity between team members as a potential confound.

4.3 Task

Similar to previous studies of group work [91, 92] and group decision making [30], we asked teams in both conditions to work on a negotiation task. The task required teams to decide how to allocate funds across competing project proposals. This task was an instantiation of a “cognitive conflict” task in the McGrath’s Task Circumplex Model [59] and as such required teams to work interdependently and resolve conflicting opinions and perspectives to arrive at a solution.

We use the well-established “foundation task” [30, 90] to create our specific task: groups were asked to allocate \$500,000 across five competing project proposals, each in need of \$500,000. The specific proposals were: 1) To establish a community arts program featuring art, music, and dance programs for children and adults 2) To create a tourist bureau to develop advertising and other methods of attracting tourism into the community 3) To purchase additional volumes for the community’s library system 4) To establish an additional shelter for the homeless in the community 5) To purchase art for display in the community’s art gallery. While the original “foundation task” [90] had a single phase where participants decided allocations, our experiment adapts this task to use two phases (discuss and decide) as below, to better measure the effects of perspective-taking.

4.4 Procedure

Figure 3 illustrates the study procedure. The workflow of the study was as follows:

4.4.1 Waiting in the lobby. On joining the study, participants were added to a waiting room till there were a sufficient number of participants for the study. Before they were assigned to groups and the task began, each participant was presented the proposals and asked to rank them in terms of their relative merit. They were asked to evaluate the projects based on their own beliefs and values. During this time, we also collected demographic information of the participants.

Data collected: Demographic Information, each participant’s ranking of the proposals.

4.4.2 Added to chatroom with teammates. As soon as there were a sufficient number of participants, they were split into teams of six participants each and were added to a chatroom with their team members. Participants were asked to pick a pseudonym before they began to interact through the chat-based interface. The task was split into two phases: *discuss* and *decide*.

4.4.3 Discuss and collectively rank proposals. In the *discuss* phase, the teams were asked to weigh the pros and cons of each proposal and rank them in terms of their relative merit collaboratively. Participants were asked to advocate for projects that aligned with their personal values. At the end

of this phase, we asked teams if they were able to agree on a ranking of the proposals and if so, we ask them to submit their collective ranking. Each team had nine minutes for this phase.

Data collected: Each team's ranking of the proposals, chatlogs.

4.4.4 Empathosphere or 2-minute reflective pause. Following this, teams in the *Empathosphere* condition carried out the perspective-taking exercise while teams in the control condition were asked to pause for two minutes and reflect on how their experience of working in the team had been thus far.

4.4.5 Decide how funds should be allocated. After this, the teams moved on to the *decide* phase. Teams were informed that they had to allocate \$500,000 across the different projects. Each project was in need of \$500,000, and the more money a project got, the more likely it was to succeed. The rankings decided in the *discuss* phase were meant to guide the *decide* phase, but teams were told that their final allocations need not reflect their rankings from the *discuss* phase. At the end of the task, we asked each team to enter their decision on how the \$500,000 should be allocated following which the participants were directed to an exit survey. Each team had nine minutes for this phase.

Data collected: Each team's allocation of funds, chatlogs.

4.4.6 Exit survey. The exit survey included questions corresponding to the measures described in 4.5. Finally, the exit survey also asked each participant how they would have allocated the \$500,000 themselves.

Data collected: Each participant's allocation of the funds, responses to likert-type and open-ended questions.

4.5 Measures

To capture different aspects of teams, their collaboration experiences, and their collaboration outcomes, we included measures for (1) baseline disagreement, (2) team viability, (3) perceived conflict, (4) openness, (5) collaboration outcomes, and (6) conversational behavior. Forms of data captured included proposal rankings provided by participants while in the lobby, chatlogs, teams' allocation of funds, and open-ended and Likert-based questions in the exit survey. Items for quantitative measures included in the exit-survey were scored on a 5-point Likert scale.

4.5.1 Baseline disagreement. We computed the amount of disagreement between members in a team using each participant's initial ranking of proposals as entered by them while in the lobby. For every team, we computed the Spearman footrule distance [77] between the rank vectors of team members in a pairwise fashion. For every pair, the Spearman footrule distance provides a proxy for disagreement between the two members. Averaging this pairwise disagreement across all possible pairs in a team, we obtained a measure of team-level disagreement.

4.5.2 Team viability. To measure participants' desire to continue collaborating with their teammates, we measured team viability using a three-item scale ($\alpha = 0.76$): "Most of the members of this team would welcome the opportunity to work as a group again in the future" "As a team this work group shows signs of falling apart." "The members of this team could work together for a long time." The items were selected from an item pool developed to measure viability [16] and have been used in other studies to measure team viability [91, 92]. To elicit honest responses, we told participants that we might use their responses to decide whether to team them up with the same or different people if we ran subsequent experiments.

4.5.3 Measures to capture perceived conflict. To understand difference in perceptions of conflict across conditions, we measured perceived conflict on task-related issues- *task conflict* as well as perceived interpersonal tension- *relationship conflict*.

- *Task conflict* was measured using a two-item scale [39] ($\alpha = 0.79$) where items were: “*there was a lot of conflict of ideas in our group*”, and “*my team had frequent disagreements relating to the task we were assigned.*”
- *Relationship conflict* was also a two-item scale [39] ($\alpha = 0.86$). Items included: “*people in my team often got angry while working together.*”, and “*there was a lot of relationship tension in my group.*”

4.5.4 *Measures to capture openness.* Openness is characterized as frank communication of issues and feelings about both task-related and personal matters [50]. To further understand differences in openness in teams across the two conditions, we measure team members’ desire to give each other feedback and their receptiveness to feedback. We also include an open-ended question to probe for openness.

Quantitative measures:

- To measure participants’ *willingness to give feedback*, we ask them if they would be willing to give feedback to other members of this group on their teamwork practices. To elicit more honest responses, we added that giving feedback was optional, and that if they were willing, we might follow up later to collect their feedback.
- To measure participants’ *willingness to receive feedback*, we ask if they would be willing to receive feedback from other members of their team on their teamwork practices.

Open-ended question:

- We included an open-ended question asking participants if they would characterize the conversation in their group as open or guarded and asked them to explain their characterization.

4.5.5 *Measures to capture collaboration outcomes.* As there is no clear performance measure for cognitive conflict tasks [91], we measure collaborative outcomes in terms of *satisfaction with solution* and the degree of *compromise*.

- *satisfaction with solution* captures the extent to which individual team members are satisfied with the team’s final allocation of funds (survey item- “*I am satisfied with my team’s final solution*”).
- To measure *compromise* in a team, we measured the differences between participants’ individual allocation of funds as provided by them in the exit survey, and the allocations that were decided by the team. For each team, we calculated the compromise measure as the mean divergence of individuals’ allocations from the group’s allocation. (This is calculated as the arithmetic mean of root-mean-square-of-differences between each members’ allocation vector and the team’s allocation vector.) This measure captures the extent to which the group’s final decision aligns with individual members opinions¹. A higher average divergence suggests that members did not agree with the final consensus of the group— the final decision did not involve a fair compromise.

4.5.6 *Conversational Behavior.* To capture changes in the conversational behavior induced by *Empathosphere*, we compared shifts in LIWC indicators across the two conditions. We also included two open-ended questions in the exit survey.

Quantitative measures:

- *Changes in LIWC indicators in the two conditions* To analyze differences in teams’ chatlog in the two conditions, we use the popular linguistic dictionary Linguistic Inquiry and Word Count, known as LIWC [66]. LIWC contains words bucketed into 125 psychometric categories

¹We also tried other measures such as the mean absolute differences between individual and group allocations. We use our current measure because it is less susceptible to outlier inputs. Results are consistent with other measures we tried.

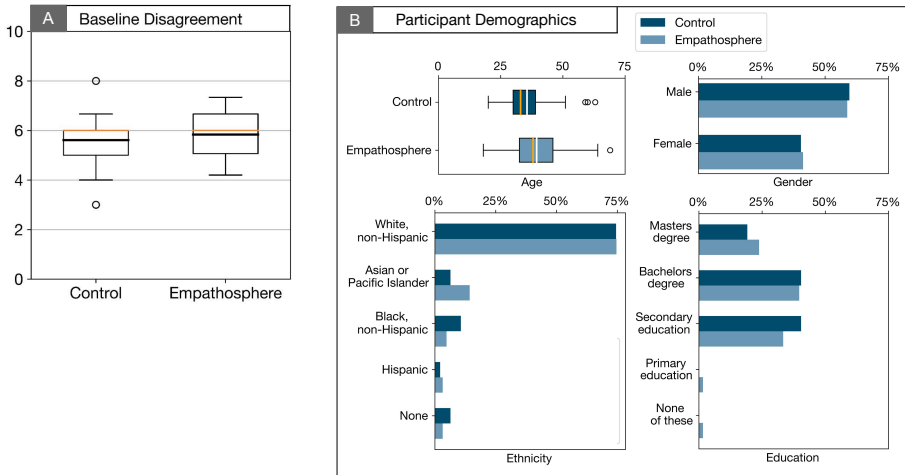


Fig. 4. A) The baseline disagreement in teams in the two conditions where disagreement in a team is measured as the average of the Spearman footrule distance between the ranking of proposals across all possible pairs of members in that team. Median disagreement in both conditions is indicated by the orange marker and the mean in black. B) Demographic information of participants in our study across the control group (■) and the intervention group (▨). Median age is indicated by the orange marker and mean in white.

including categories such as ‘first-person pronouns’, ‘second-person pronouns’, ‘positive emotion’, ‘negative emotion’, ‘informality’ etc. This allows us to analyze a given text along these 125 dimensions. For every message, we compute normalized frequency counts for the LIWC categories, i.e. the number of times words from the category were present in the message divided by the total number of words in the message. To obtain psychometric measures that are independent of the length of the chatlog, we then average these normalized frequency counts across all messages in a chatlog. Therefore, the mean LIWC indicator for *positive emotion* for a chatlog measures *on average* how positive messages in the conversation were.

Open-ended question:

- We included an open-ended question to capture perceived changes in conversational behavior: “How did you engage with the group in the second stage?”

5 RESULTS

A total of $N = 110$ participants completed the experiment across 24 teams with 4-6 members each. We conducted our analyses on these 24 teams which included 11 teams in the control condition and 13 teams in the intervention condition. 59% of the participants were male and the average age of participants was 38 years ($\sigma = 10.4$). 74.5% participants identified as White, non-Hispanic, 10.9% identified as Asian or Pacific Islander, 7.3% as Black, and 2.7% as Hispanic. 21.8% reported having a masters degree, 40% a bachelors degree, and 36.4% a secondary education. Figure 4B shows a breakdown of demographics by experiment condition.

5.1 Baseline disagreement

To check whether team composition varied significantly across the two conditions, we compared the baseline disagreement in teams in the two conditions (Figure 4A). We found no significant

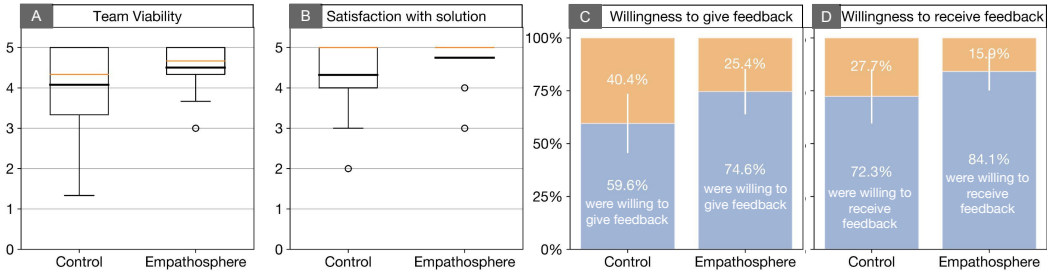


Fig. 5. A) Compared to the control condition, *Empathosphere* led to significantly higher team viability. Median score is indicated by the orange marker and the mean by the black marker. B) Participants in the *Empathosphere* condition expressed significantly higher satisfaction with their teams' solution than participants in the control condition. Median score is indicated by the orange marker and the mean by the black marker. C) Participants in the *Empathosphere* condition were more likely to give their teammates feedback. The chart shows the proportion of participants that were willing and unwilling to give feedback to other team members, with 95% CI at the boundary. D) Participants in the *Empathosphere* condition were also more open to receiving feedback from their teammates. The chart shows the proportion of participants that were willing and unwilling to receive feedback from other team members, with 95% CI at the boundary.

	Outcome Variables							
	Team Viability				Satisfaction with solution			
	Coeff.	SE	z	p	Coeff.	SE	z	p
<i>Fixed Effects</i>								
(Intercept)	4.02***	0.17	23.43	<0.001	4.29***	0.14	30.48	<0.001
Condition	0.49*	0.23	2.13	0.033	0.45*	0.19	2.35	0.018
Disagreement	-0.35*	0.15	-2.34	0.019	-0.11	0.13	-0.89	0.372
ConditionXDisagreement	0.29	0.23	1.31	0.191	0.08	0.58	0.41	0.684
<i>Random Effects</i>	Var.	SE			Var.	SE		
Team	0.21	0.17			0.09	0.10		

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 1. Results of mixed effect linear regression analyzing the impact of experiment condition and disagreement within the group on measures of *team viability* and *satisfaction with solution*. The condition had a significant effect on both measures with participants in the *Empathosphere* condition expressing higher viability and satisfaction with solution.

difference between disagreement in the control condition teams ($\mu = 5.6, \sigma = 1.27$) and intervention condition groups ($\mu = 5.83, \sigma = 1.01$) via a t-test: $t(109) = -0.45, p = 0.65$. Hedges' $g = 0.197$.

5.2 Team Viability

Empathosphere improved team viability. Participants in the *Empathosphere* condition scored team viability to be higher ($\mu = 4.50, \sigma = 0.63$) than in the control condition ($\mu = 4.08, \sigma = 0.98$) (Figure 5A). We fit a mixed effects linear regression model with the experiment condition and team disagreement as fixed effects, their interaction, team grouping as a random effect, and team viability scores as the outcome variable. Results in Table 1. We observe a significant effect of both the condition ($\beta = 0.49$; 95% CI = 0.06, 0.93; $p < 0.05$) and disagreement ($\beta = -0.35$; 95% CI = $-0.63, -0.07$; $p < 0.05$) on team viability, with no significant interaction effects. Teams with lower initial disagreement had higher viability, suggesting that teams with lower opinion diversity, and therefore inherently lower potential for conflict, tend to exhibit higher viability. Meanwhile, the

	Outcome Variables							
	Task Conflict				Relationship Conflict			
Fixed Effects	Coeff.	SE	z	p	Coeff.	SE	z	p
(Intercept)	2.19***	0.21	10.64	<0.001	1.55***	0.17	9.28	<0.001
Condition	-0.36	0.27	-1.28	0.199	-0.15	0.23	-0.66	0.512
Disagreement	0.31	0.18	1.73	0.084	0.11	0.15	0.77	0.438
ConditionXDisagreement	-0.04	0.27	-0.13	0.896	0.10	0.22	0.46	0.644
Random Effects	Var.	SE			Var.	SE		
Team	0.29	0.19			0.19	0.15		

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 2. Results of mixed effect linear regression analyzing the impact of experiment condition and disagreement on *perceived task conflict*, and *perceived relationship conflict* showing the absence of a significant relationship between the condition and either variables.

	Outcome Variables							
	Willingness to give feedback				Willingness to receive feedback			
Fixed Effects	Coeff.	SE	z	p	Coeff.	SE	z	p
(Intercept)	0.27	0.31	0.90	0.366	0.93*	0.38	2.45	0.014
Condition	0.78	0.43	1.83	0.067	0.95	0.56	1.68	0.092
Disagreement	-0.33	0.29	-1.15	0.252	-0.62	0.37	-1.65	0.098
ConditionXDisagreement	0.55	0.44	1.26	0.209	0.15	0.58	0.26	0.793
Random Effects	Var.	SE			Var.	SE		
Team	0	0			0.13	0.37		

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 3. Results of mixed effect logistic regression analyzing the impact of experiment condition and disagreement on participants' *willingness to give feedback* to teammates and their *willingness to receive feedback* from teammates. The condition had a marginally significant effect on both measures with participants in the *Empathosphere* condition expressing higher willingness to give and receive feedback.

effect of *Empathosphere* on team viability shows that prompting perspective-taking is a promising approach to improving viability of all teams, regardless of the degree of diversity in team members opinions.

5.3 Perceived Conflict

We did not observe differences in *perceived task or relationship conflict* across the conditions. Following the same analysis strategy as before, we fit mixed effects linear regression models with perceived task conflict and perceived relationship conflict as the outcome variables and found no significant effects of condition, disagreement, and their interaction on either outcome variable (Table 2). There was a marginally significant effect of initial disagreement on the perceived task conflict ($\beta = 0.31$; 95% CI = $-0.02, 0.66$; $p = 0.08$) with higher initial disagreement leading to higher perceptions of task conflict, which is in line with expectations. However, the *Empathosphere* condition did not yield significant differences in perceived task or relationship conflict suggesting that prompted perspective-taking might not change perceptions of conflict or trigger negative emotional responses to conflict.

5.4 Openness

5.4.1 *Empathosphere improved willingness to give and receive feedback.* Since *willingness to give feedback* and *willingness to receive feedback* had binary responses, for each of them as outcome

variables, we fit a mixed effects logistic regression model with the experiment condition and team disagreement as fixed effects, their interaction, team grouping as a random effect (Table 3). We found a marginally significant effect of condition on *willingness to give feedback* ($\beta = 0.78$; 95% CI = $-0.05, 1.76$; $p = 0.067$). Of the participants in the *Empathosphere* condition, 74.6% were willing to give their teammates feedback while only 59.6% of the participants in the control condition were willing to do the same (Figure 5C). Similarly, we found a marginally significant effect of the experiment condition on *willingness to receive feedback* ($\beta = 0.95$; 95% CI = $-0.14, 2.24$; $p = 0.092$). 84.1% of the participants in the *Empathosphere* condition were willing to receive feedback from their teammates while 72.3% of the participants in the control condition were willing to do the same (Figure 5D). We also saw a marginally significant effect of disagreement ($\beta = -0.62$; 95% CI = $-1.47, 0.09$; $p = 0.098$), such that participants in teams with higher disagreement were less keen on receiving feedback from their teammates.

5.4.2 *Empathosphere encouraged participants to voice disagreement while also making participants more perceptive to other team members' behaviors.* We compared participants' responses to the open-ended question on whether they thought the conversation in their group as open or guarded. Participants in the control condition mentioned how some team members chose to stay silent: "Some people had nothing to say at all while two others were very open." (P5) Some other participants mentioned how there was very little opposition or open disagreement: "It was not really as engaging as I hoped. I had to get the ball rolling and didn't really get any conflicting opinions" (P11), and "It did not appear that anyone wanted to "dominate" the conversation/debate and therefore potentially yielded quicker than they would in person or make real decisions." (P41)

In contrast to this, participants in the intervention group reported higher levels of open disagreement: "We never strongly attacked an idea and views were able to change. People were allowed to suggest their ideal solution and did not seem bothered by challenges." (P94), "everyone brought something to the table and it was a great group" (P108), "I felt like everyone could voice their opinions, and no one was shot down unfairly." (P106) One participant noted a distinct shift in their behavior after the *Empathosphere* exercise: "I was guarded on the first stage and wanted to recommend tourism as the first priority but when others said homeless I agreed. I knew I would have to speak up quickly and I was less guard[ed] on round two because they all seemed like nice people would [who] wouldn't be rude." (P71)

Even though we did not explicitly probe for it, participants in the *Empathosphere* condition also made more specific observations about their teammates suggesting that they were perceptive to how their teammates were behaving. One participant noted: "I knew I should suggest amounts quickly because William would have a proposal and would be more particular, I think. He seems like a leader or someone who wants to be in charge and doesn't ask anyone else what they want probably. However he did compromise." (P71) Another participant mentioned: "Kate seemed to be the one that had the most ideas that differed from the group. The other 2 people seemed to be the most in line with me." (P96)

5.5 Collaboration outcomes

5.5.1 *Empathosphere improved participants' satisfaction with their teams' solutions.* Participants in the *Empathosphere* condition reported higher satisfaction with their teams' solutions ($\mu = 4.74$, $\sigma = 0.50$) than participants in the control condition ($\mu = 4.31$, $\sigma = 0.94$) (Figure 5B). We fit a mixed effects linear regression model with the experiment condition and team disagreement as fixed effects, their interaction, team grouping as a random effect, and satisfaction with solution as the outcome variable. Results in Table 3. We observed a significant effect of the experiment condition on

satisfaction with solution ($\beta = 0.45$; 95% CI = 0.09,0.80; $p < 0.05$) while there was no significant effect of disagreement. There was also no interaction effect.

5.5.2 We did not observe significant differences in compromise across teams in the two conditions. The difference between *compromise* in teams in the control condition ($\mu = 0.075, \sigma = 0.022$) and teams in the *Empathosphere* condition ($\mu = 0.070, \sigma = 0.036$) was not statistically significant, possibly due to small sample size (*compromise* was measured at the group level, and so had fewer observations than some of the individual level measures above).

5.6 Conversational behavior

5.6.1 Teams in the Empathosphere condition used more second-person pronouns and exchanged more informal messages. We compare the changes in LIWC indicators from the *discuss* to *decide* phase in the *Empathosphere* and control conditions. We find that *Empathosphere* was followed by an increase in use of second-person pronouns (you, you've y'all, u). Teams used 89% more second-person pronouns in the *decide* phase, after *Empathosphere* ($p < 0.05$) but there was no significant difference in the usage of second person pronouns between the *discuss* and *decide* phase in the control condition. This indicates that *Empathosphere* potentially shifted the conversation to be more other-focused, engaging with what other team members are saying and drawing them into the conversation.

The intervention was also followed by an increase in informality (okay, yaas) and netspeak. Usage of informal words was 27% higher in the *decide* phase than the *discuss* phase in the intervention condition ($p < 0.05$). Similarly, usage of netspeak was 281% higher in the *decide* phase than the *discuss* phase for the intervention condition ($p < 0.001$). Meanwhile, in the control condition, the use of informal language decreased in the *decide* phase compared to the *discuss* phase by 24%, however, this difference was only marginally significant ($p = 0.09$). Taken together, this also suggests that *Empathosphere* has the potential to improve social bonds in teams [93].

5.6.2 Empathosphere helped foster higher comfort levels and led to team members respectfully engaging with each other. We analyzed participants' responses to the open-ended question asking how they engaged with the group in the second stage. Participants in the intervention condition noted higher comfort levels in their team: "It felt like coming back to a group of coworkers that I know well." (P91) They mentioned being able to voice their opinions with their team: "People threw out ideas while others were more intent on keeping the group focused on coming up with an actual answer to the task. I chimed in when I wanted something addressed or wanted to broach a specific idea that mattered to me" (P85), "I gave my thoughts and everyone listened and gave me constructive feedback" (P107), and "I suggested an alternative allocation of funds at one point and the group reached an amicable decision taking in everyone's vote" (P67).

They also took efforts to accommodate each others' perspectives: "I tried to take everyone's ideas and formulate them into a single plan. I took the numbers and tried to make them work, so as to make the plan easier to visualize. I think this helped the group to come to an agreement more quickly, because they could see the plan." (P81) They tried to make their teammates feel heard: "I knew Williams' personality and he would have a suggestion and would want to be heard." (P71) Several participants mentioned how they attempted to strike a balance between voicing their perspectives and listening to others' opinions: "I tried to take in their perspectives but wanted to make sure they understood me too" (P94), "I tried to summarize everyone's ideas as well as contribute my own" (P96), and "I made suggestions but also was open to what they thought as well." (P93)

5.6.3 *Teams in the control condition had more polarized experiences, with some reporting either too little or too much conflict.* Analyzing responses to the open-ended question by participants in the control condition, we found that participants hinted at loafing: “I engaged with caution, trying to let some of the other members bounce ideas off of one another, but no one was really into it.” (P11) Some tried to avoid conflict altogether: “I gave my proposal, but it seemed like the group wanted to finish quickly so I didn’t push my ideas that much.” (P36) On the other hand, some participants in the control condition noted how their opinions were disregarded: “I expressed my thoughts and ideas about how to distribute money but was getting into some arguments about the merits of some programs versus others. I was getting frustrated because it felt like my group was ignoring my suggestions” (P24). Such dismissal of ideas also triggered negative emotions: “I tried to keep the group focused. However, one person was not respectful of my ideas and made snide remarks about me being insecure. That definitely hampered our progress.” (P25)

6 DISCUSSION

With *Empathosphere*, we have demonstrated that focusing on the barriers to effective communication patterns - namely, insufficient or poorly calibrated perspective-taking - is an effective lever for providing technological support to ad-hoc virtual teams, as it provides the scaffolding for behavior change while still providing teams with the agency to regulate their own communication. We show how *Empathosphere*, by prompting reflection about team members’ emotional states, improves work satisfaction and encourages more open communication and feedback within the team, and boosts team viability. As such, *Empathosphere* expands on and brings together several ideas that have been previously explored in HCI research. We situate our work alongside existing research focused on harnessing diverse perspectives of team members, supporting constructive deliberation, inducing perspective-taking, and leveraging technology to support team development. We then discuss the limitations of our work and suggest possible directions for future exploration.

6.1 Harnessing disagreement to improve teamwork

Empathosphere attempts to help teams surface the diverse opinions of teams members and manage the resulting disagreement, towards the goal of improving collaborative outcomes. Prior work has established that well-handled conflict can help improve team creativity, identify and resolve issues, and strengthen common values [34, 46]. Previous systems have sought to manifest constructive conflict in different ways. Some have attempted to assemble teams by maximizing diversity in member perspectives [36, 55]. Hive [71] brings new members into teams, at later stages, in an attempt to introduce fresh perspectives and challenge a teams’ ideas. In doing so, it brings in conflicting opinions from sources external to the team. However, prior work has shown that mere existence of diverse opinions does not translate to disagreement and the positive benefits that arise from disagreement [34, 71], especially in the presence of a challenging group climate [62]. We, instead, shift the focus from the *source* of diverse perspectives towards the *expression* of diverse perspectives. We attempt to illuminate the diverse perspectives and experiences of team members so that teams may benefit from their differences in perspectives. As such, this work serves to complement existing work: using established team assembly systems in concert with interventions to foster constructive communication has the potential for greater success than does either approach by itself.

6.2 Supporting deliberation while enforcing minimal structure

Research aiming to improve group deliberation, has largely focused on the design of platforms that provide effective *structure* for deliberation. For instance, ConsiderIt structures deliberations around user-curated pros/cons lists [47]. OpinionSpace structures deliberation by plotting opinions

in a two-dimensional space [21]. Lead Line structures discussion in accordance to a pre-authored script [22]. StarryThoughts supports visual exploration of diverse opinions while emphasizing the relationship between peoples' identities and opinions [42]. Structure is powerful but constraining—by reducing flexibility in how people can interact, it can encourage people to elaborate their reasoning [74], encourage people to listen and consider others' opinions [21, 42, 47], and allow people to scope out and consider the broader context of the deliberation [47, 95]. However, in practice, virtual teams rarely move to these dedicated platforms to make use of their *structure* for deliberations; they deliberate within the *flexible* communication channels they already occupy (for instance Discord [1] or Slack [71, 91]). This suggests the need to embed deliberation support in the unstructured, synchronous communication channels that teams choose to use while maintaining the *flexibility* that drew teams to those channels in the first place. Researchers have explored two strategies for this: (1) allowing teams to invoke structure temporarily [43, 44], and (2) designing behavior change systems that promote positive communication behaviors through heightened social awareness [29, 52, 53]. *Empathosphere* explores the potential of systems at the confluence of these two strategies: when triggered, *Empathosphere* creates structure for two minutes during which it attempts to heighten social awareness and facilitate behavior change. Thus, *Empathosphere* shows the promise of invoking structure for micromoments and using those structured micromoments to foster behavior change.

6.3 The tradeoff between control and authenticity in mechanisms to induce perspective-taking

Empathosphere contributes an operationalization of strategies to induce perspective-taking, that relies on self-regulated emotional exchanges. A prominent line of prior work that explicitly attempts to foster empathy or perspective-taking, via emotional expressiveness, has explored the use of expressive biosignals as cues to cognitive or affective states [29, 54]. The key difference between these approaches and ours is how emotional signals are sourced, represented, and interpreted. Unlike *Empathosphere* that asked people to self-report their emotions, biosignals are directly sensed from a person's body and they convey personal and intimate information. So, biosignals make it harder for people to conceal or feign their emotions when compared to our approach. However, the fact that people can only regulate whether or not they share biosignals but not the extent of expression [18], can make them hesitant to share their biosignals altogether [54]. To reduce the hesitation and to promote greater safety in the sharing of emotions, we consciously choose to give people control over the extent to which they want to express their emotions. While this agency may, in theory, come at the potential cost of authenticity, we also aimed to increase psychological safety by not revealing individual team members' own emotional self-reports, instead providing accuracy metrics at the aggregate level. Additionally, since biosignals require physiological sensing, they require users to possess these sensing systems— a constraint *Empathosphere* doesn't have. In addition, biosignals as cues to emotions are inherently ambiguous, often providing more information about the extremity rather than the valence of an emotion and, thus, requiring subjective interpretation [18]. In contrast, we rely on a more straightforward reporting of emotional states, and provide an accuracy measure to help support users as they attempt to interpret others' emotions.

6.4 Designing interventions for different stages in a team's development process

Truckman identifies four stages teams go through: forming, storming, norming, and performing—from when they are convened to when they realize their full potential [83]. Forming is the initiation of collaboration where team members first come together and get to know each other. This is followed by the storming stage, where team members push against established boundaries, challenge each other, and sometimes even clash with each other. Norming is marked by team members moving

past their differences, accepting each other, and this is when team members feel comfortable and safe in the team. Ultimately teams arrive at the performing stage where they have learned to work together and perform at their fullest potential. Yet, unlike Truckman's linear model, real-world teams are increasingly fluid [8], and membership changes frequently [71, 85]. For such high-turnover teams, forming and storming stages are frequent: developing effective communication and accommodating new team members is an ongoing challenge. Each of these stages provides unique opportunities for technology-mediated protocols to support team development and adaptability.

In the forming stages, teams can benefit from protocols such as team dating which help teams identify good collaborators (and prune undesirable ones) by allowing them to experience working with different collaborators [56]. The storming stage is especially tricky to navigate. Several teams never move past it: escalating conflict and rising tensions can cause teams to fracture [91]. Part of the problem is that unregulated team interactions can lead to teams unintentionally establishing unhealthy interaction norms that are hard to alter later [92]. To help teams navigate the storming phase, previous work has explored the benefits of allowing teams to "start over" [92], by masking teammates' identities and repeatedly convening the same team till they adopt positive interaction norms. This works well when teams are assembled all at once, but masking identities and restarting is harder for a fluid, partially formed team. While this current study did not specifically study storming processes, *Empathosphere* may empower teams to collectively navigate the storming phase by equipping team members to effectively surface and manage disagreements and conflict. *Empathosphere* may provide teams with the opportunity to alter problematic norms they might have previously established and, as suggested by results in Section 5.6, may also help establish healthy conversational norms, such as respectfully engaging with each other. Finally, perspective-taking interventions, if used in conjunction with protocols like Hive [71] could help help periodically expose teams to new ideas and fresh perspectives during the norming and performing stages.

6.5 Limitations

6.5.1 Limitations of task design and collaboration duration. Consistent with prior studies on group work [30, 91], we use the negotiation task - an instantiation of a cognitive conflict task - to effectively simulate situations where team's must resolve conflicting viewpoints. Specifically, teams in our study worked on a cognitive conflict task within McGrath's Circumplex Model [59]. While our results may be relevant to other kinds of tasks such as creative generative tasks that also surface, and benefit from, diverse perspectives, we did not study those tasks directly.

Furthermore, teams in our experiment only worked together for a short duration of about 20 minutes. This has two implications for our findings. First, we can not comment on the long term effects of our intervention and whether the changes in behavior persist beyond the short time frame. Secondly, prior work has shown that ad-hoc teams that expect to work for a short duration- as was the case with Amazon Mechanical Turk workers in our study- are less likely to be positively influenced by interpersonal interventions. The short time frame and anticipated future separation causes team members to focus on the completion of the task without attempting to form cohesive team norms that would only benefit the members if they were going to remain together for the performance of future task [10]. As such, our study setup could have potentially muted the effects of *Empathosphere* because team members did not have an incentive to invest effort in forming group norms, which would have benefited from perspective-taking.

6.5.2 We do not study teams with pre-existing relationships. Participants in our study were recruited from Amazon Mechanical Turk and assigned to ad-hoc teams after they picked pseudonyms. Members in our study teams had not worked together before and even if they had, the pseudonyms made it hard for them to recognize this. While this served to mimic real world scenarios where

teams might be meeting for the first time [32, 91], we caution against generalizing our findings to teams where members might be familiar with each other and have preexisting relationships.

6.5.3 Our experimental design cannot quantify the effect of Empathosphere on social perceptiveness. Our qualitative analysis of groups' conversational data, provides evidence of the impact of our intervention on team members' displayed social perceptiveness in deliberation. However, since we did not include a direct measure of social perceptiveness in our study, we were unable to quantify the effect. Measures of changes to social perceptiveness that we considered- such as a measure that involves having participants draw a capital letter E on their foreheads and assessing whether it was drawn from a self-oriented or other-oriented perspective [28]- were ultimately deemed to be either too unnatural or obtrusive to fit within the flow of the study. Future work can explore the design and use of appropriate measures of social perceptiveness to test the effects of interventions like *Empathosphere*.

6.5.4 Our experimental design cannot measure the impact of team members' individual characteristics. To protect privacy, we asked participants in our study to pick pseudonyms before they were added to chatrooms with their teams. However, doing so may have reduced the salience of their individual characteristics, such as gender and race in our study. This may have in turn muted impression formation and stereotyping as interfering phenomena. At the same time, several participants chose gendered names, and so these effects were not entirely absent. Future work can explore the effects of team members' impressions of each other on their behaviors and how social capital, impression formation and stereotyping might interact with the effects of interventions designed to improve group norms.

6.6 Future Work

6.6.1 Interventions to boost minority voices in teams. *Empathosphere* demonstrates the promise of creating spaces that provide a layer of safety and anonymity for team members to express and reflect on their experiences of group climate. Such mechanisms, which encourage members to voice their emotions, issues, and disagreements, are critical for creating inclusive team dynamics. Psychological safety, for example, is critical for boosting minority opinions and voices in groups and centering the perspectives of individuals from marginalized groups. Prior work has shown that experiences with implicit bias and other forms of micro-inequities threaten the confidence and self-esteem of minority group members [23, 26, 73]. Even when members of the group might care about fairness and equitable participation, a lack of intentionality in enforcing inclusive norms can create a "chilly climate" for marginalized or underrepresented group members [37]. This problem is amplified by the fact that minority group members might feel hesitant and powerless to express their issues openly for fear of judgment and further alienation [31, 81]. Ultimately, this can cause minority members to leave such groups [7, 35]. Future work can explore the design of interventions that build on *Empathosphere* and have particularly powerful benefits for increasing equity of participation and inclusion for diverse teams.

6.6.2 Expanding the notion of "experimental spaces" and exploring the role of virtual spaces in transforming team norms. Researchers using "experimental spaces" to transform team norms have primarily done so by moving team interaction to different social settings- demarcated spaces (contributing to their sense of *space*) with well defined rules of interaction (contributing to their sense of *place*). However, most of this work has explored the role of "experimental spaces", situated in the real world [97]. *Empathosphere* explores the alternate possibility of creating spaces in the *virtual* world, where teams convene increasingly often. However, exciting opportunities might also lie at the intersection of these two approaches in hybrid spaces- virtual spaces overlaid on

physical realities. Pokemon GO [2] showed the exciting possibilities of overlaying a virtual space over a real one and how both the virtual and real space can support and enable entirely different interaction norms. Can we use such hybrid spaces to motivate counternormative behaviors that facilitate change? Two examples of prior work explore this direction. Inneract [70] was an app that allowed users to create spaces with their own rules for interaction. For instance: “we trade secrets in this space”. It allowed other users within a certain physical radius to “enter” spaces they were close to by opting into the interaction norms of those spaces. Similarly, Situationist [48] was an app that allowed people to pick counternormative experiences they would like to be a part of (such as “giving a 5-second hug”) and when they were in physical proximity to another user of the app, they were prompted to engage in the experience together. While both these projects were short-lived social experiments, future work can further explore the design of experiences in alternate realities that foster social connection in teams and improve team climate.

6.6.3 Generalizing to long-lived teams: identifying opportune moments to intervene. While our study focused on short-lived teams, it illuminated important questions on designing interventions for well-being of long-lived teams. Our short-term study did not allow us to measure the persistence of *Empathosphere*'s effects on perspective-taking but by investigating how effects of such team interventions dampen over time, future work can formulate strategies for periodic deployment of interventions that maintain social perceptiveness in teams. In our setting, *Empathosphere* was triggered automatically and teams might well need such hand-holding in initial interactions [92, 96]. However, once teams move past initial interactions and learn to work and engage with each other, should the intervention switch to being passive? This would allow team members to invoke the system when necessary (e.g. on arrival of a new member) as opposed to the system deciding the right time to interrupt teams. Future work can investigate stages in a team's interaction after which the system can step back and shift control to the users.

7 CONCLUSION

With *Empathosphere*, we have demonstrated the promise of introducing spaces for reflection and perspective-taking as a means to foster open communication, encourage expression of diverse and conflicting viewpoints, and, ultimately, improve team satisfaction and viability among members of ad-hoc teams. This work provides a foundation for designing systems that leverage social perceptiveness to foster greater safety and efficacy within virtual teams. More broadly, we hope that, going forward, this work signals a greater focus on well-being and group climate as greater priorities for the design of technological supports for virtual groups.

ACKNOWLEDGMENTS

We are grateful to Mark Whiting and Andrew Mao for their valuable inputs and guidance on conducting collaborative real-time experiments on Amazon Mechanical Turk. Our experimental platform was built atop TurkServer- a package Andrew helped develop and maintain- and we thank him for helping us understand and work with TurkServer. We also thank Yasmine Kotturi, Julia Cambre, Sayan Chaudhry, Franklin Li, Jaemarie Solyst, and our anonymous reviewers for their valuable comments and feedback on drafts of this work. We thank the participants in our pilots and the Amazon Mechanical Turk workers who participated in our study. This work was supported by the Office of Naval Research but this article solely reflects the opinions and conclusions of the authors and not our funder.

REFERENCES

- [1] [n.d.]. Discord Loves Open Source. <https://discord.com/open-source>.

- [2] [n.d.]. Pokemon Go. https://en.wikipedia.org/wiki/Pok%C3%A9mon_Go.
- [3] [n.d.]. Wikipedia:IRC. <https://en.wikipedia.org/wiki/Wikipedia:IRC>.
- [4] Chris Argyris. 1991. Teaching smart people how to learn. *Harvard business review* 69, 3 (1991).
- [5] Kathleen T Ashenfelder. 2007. *Simultaneous analysis of verbal and nonverbal data during conversation: symmetry and turn-taking*. University of Notre Dame.
- [6] Julia B Bear and Anita Williams Woolley. 2011. The role of gender in team collaboration and performance. *Interdisciplinary science reviews* 36, 2 (2011), 146–153.
- [7] Maya A Beasley and Mary J Fischer. 2012. Why they leave: The impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors. *Social Psychology of Education* 15, 4 (2012), 427–448.
- [8] Suzanne T Bell and Brian J Marentette. 2011. Team viability for long-term and ongoing organizational teams. *Organizational Psychology Review* 1, 4 (Nov. 2011), 275–292.
- [9] Osvald M Bjelland and Robert Chapman Wood. 2008. An inside view of IBM's 'Innovation Jam'. *MIT Sloan management review* 50, 1 (2008), 32.
- [10] John Bradley, Barbara Jo White, and Brian E Mennecke. 2003. Teams and Tasks: A Temporal Framework for the Effects of Interpersonal Interventions on Team Performance. *Small Group Research* 34, 3 (June 2003), 353–387.
- [11] Susan L Bryant, Andrea Forte, and Amy Bruckman. 2005. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. 1–10.
- [12] Silke Bucher and Ann Langley. 2016. The interplay of reflective and experimental spaces in interrupting and reorienting routine dynamics. *Organization Science* 27, 3 (2016), 594–613.
- [13] Daniel Buschek, Alexander De Luca, and Florian Alt. 2015. There is more to typing than speed: Expressive mobile touch keyboards via dynamic font personalisation. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 125–130.
- [14] Hancheng Cao, Vivian Yang, Victor Chen, Yu Jin Lee, Lydia Stone, N'godjigui Junior Diarrassouba, Mark E Whiting, and Michael S Bernstein. 2021. My Team Will Go On: Differentiating High and Low Viability Teams through Team Interaction. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–27.
- [15] Abraham Carmeli, Jane E Dutton, and Ashley E Hardin. 2015. Respect as an engine for new ideas: Linking respectful engagement, relational information processing and creativity among employees and teams. *Hum. Relat.* 68, 6 (June 2015), 1021–1047.
- [16] Jessica Nicole Cooperstein. 2017. *Initial Development of a Team Viability Measure*. Ph.D. Dissertation. DePaul University.
- [17] Anna L Cox, Sandy JJ Gould, Marta E Cecchinato, Ioanna Iacovides, and Ian Renfree. 2016. Design frictions for mindful interactions: The case for microboundaries. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 1389–1397.
- [18] Max T Curran, Jeremy Raboff Gordon, Lily Lin, Priyashri Kamlesh Sridhar, and John Chuang. 2019. Understanding digitally-mediated empathy: An exploration of visual, narrative, and biosensory informational cues. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [19] James E Driskell, Eduardo Salas, and Joan Johnston. 1999. Does stress lead to a loss of team perspective? *Group dynamics: Theory, research, and practice* 3, 4 (1999), 291.
- [20] Stale Einarsen, Helge Hoel, and Cary Cooper. 2002. *Bullying and emotional abuse in the workplace: International perspectives in research and practice*. CRC press.
- [21] Siamak Faridani, Ephrat Bitton, Kimiko Ryokai, and Ken Goldberg. 2010. Opinion space: a scalable tool for browsing online comments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1175–1184.
- [22] Shelly Farnham, Harry R Chesley, Debbie E McGhee, Reena Kawal, and Jennifer Landau. 2000. Structured online interactions: improving the decision-making of small discussion groups. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. 299–308.
- [23] Faye Flam. 1991. Still a "chilly climate" for women? *Science* 252, 5013 (1991), 1604.
- [24] Vivian Franco, R Piirto, H-Y Hu, Bruce V Lewenstein, Russ Underwood, and Noni Korf Vidal. 1995. Anatomy of a flame: conflict and community building on the Internet. *IEEE Technology and Society Magazine* 14, 2 (1995), 12–21.
- [25] Santi Furnari. 2014. Interstitial spaces: Microinteraction settings and the genesis of new practices between institutional fields. *Academy of management review* 39, 4 (2014), 439–462.
- [26] Roberta M Hall and Bernice R Sandler. 1982. The Classroom Climate: A Chilly One for Women?. (1982).
- [27] Andrew B Hargadon and Beth A Bechky. 2006. When Collections of Creatives Become Creative Collectives: A Field Study of Problem Solving at Work. *Organization Science* 17, 4 (Aug. 2006), 484–500.
- [28] R Glen Hass. 1984. Perspective taking and self-awareness: Drawing an E on your forehead. *Journal of personality and social psychology* 46, 4 (1984), 788.
- [29] Mariam Hassib, Daniel Buschek, Paweł W Wozniak, and Florian Alt. 2017. Heartchat: Heart rate augmented mobile chat to support empathy and awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing*

Systems. 2239–2251.

- [30] Helen Ai He, Naomi Yamashita, Chat Wacharamanatham, Andrea B Horn, Jenny Schmid, and Elaine M Huang. 2017. Two Sides to Every Story: Mitigating Intercultural Conflict through Automated Feedback and Shared Self-Reflections in Global Virtual Teams. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW (Dec. 2017), 1–21.
- [31] Catherine Hill, Christianne Corbett, and Andresse St Rose. 2010. *Why so few? Women in science, technology, engineering, and mathematics*. ERIC.
- [32] Pamela J Hinds and Diane E Bailey. 2003. Out of sight, out of sync: Understanding conflict in distributed teams. *Organization science* 14, 6 (2003), 615–632.
- [33] Pamela J Hinds and Diane E Bailey. 2003. Out of Sight, Out of Sync: Understanding Conflict in Distributed Teams. *Organization Science* 14, 6 (Dec. 2003), 615–632.
- [34] Inga J Hoefer, Daan van Knippenberg, Wendy P van Ginkel, and Harry G Barkema. 2012. Fostering team creativity: perspective taking as key to unlocking diversity’s potential. *J. Appl. Psychol.* 97, 5 (Sept. 2012), 982–996.
- [35] Shannon E Holleran, Jessica Whitehead, Toni Schmader, and Matthias R Mehl. 2011. Talking shop and shooting the breeze: A study of workplace conversation and job disengagement among STEM faculty. *Social Psychological and Personality Science* 2, 1 (2011), 65–71.
- [36] Sujin K Horwitz and Irwin B Horwitz. 2007. The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of management* 33, 6 (2007), 987–1015.
- [37] Lynne A Isbell, Truman P Young, and Alexander H Harcourt. 2012. Stag parties linger: continued gender bias in a female-rich scientific discipline. *PLoS One* 7, 11 (2012), e49682.
- [38] Ryo Ishii, Kazuhiro Otsuka, Shiro Kumano, Ryuichiro Higashinaka, and Junji Tomita. 2018. Analyzing Gaze Behavior and Dialogue Act during Turn-taking for Estimating Empathy Skill Level. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction (Boulder, CO, USA) (ICMI ’18)*. Association for Computing Machinery, New York, NY, USA, 31–39.
- [39] Karen A Jehn and Elizabeth A Mannix. 2001. The dynamic nature of conflict: A longitudinal study of intragroup conflict and group performance. *Academy of management journal* 44, 2 (2001), 238–251.
- [40] Kristiina Jokinen, Hirohisa Furukawa, Masafumi Nishida, and Seiichi Yamamoto. 2013. Gaze and turn-taking behavior in casual conversational interactions. *ACM Trans. Interact. Intell. Syst.* 3, 2 (Aug. 2013), 1–30.
- [41] Malte F Jung, Nikolas Martelaro, and Pamela J Hinds. 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 (Portland, Oregon, USA) (HRI ’15)*. Association for Computing Machinery, New York, NY, USA, 229–236.
- [42] Hyunwoo Kim, Haesoo Kim, Kyung Je Jo, and Juho Kim. 2021. StarryThoughts: Facilitating Diverse Opinion Exploration on Social Issues. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–29.
- [43] Soomin Kim, Jinsu Eun, Changhoon Oh, Bongwon Suh, and Joonhwan Lee. 2020. Bot in the bunch: Facilitating group chat discussion by improving efficiency and participation with a chatbot. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [44] Soomin Kim, Jinsu Eun, Joseph Seering, and Joonhwan Lee. 2021. Moderator Chatbot for Deliberative Discussion: Effects of Discussion Structure and Discussant Facilitation. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–26.
- [45] Taemie Kim, Pamela Hinds, and Alex Pentland. 2012. Awareness as an antidote to distance: making distributed groups cooperative and consistent. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (Seattle, Washington, USA) (CSCW ’12)*. Association for Computing Machinery, New York, NY, USA, 1237–1246.
- [46] Aniket Kittur, Bongwon Suh, Bryan A Pendleton, and Ed H Chi. 2007. He says, she says: conflict and coordination in Wikipedia. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 453–462.
- [47] Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. 2012. Supporting reflective public thought with considerit. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. 265–274.
- [48] Amy Lee. [n.d.]. Situationist App Lets Strangers Do Random Things With You. https://www.huffpost.com/entry/situationist-app_n_835594.
- [49] Joonhwan Lee, Soojin Jun, Jodi Forlizzi, and Scott E Hudson. 2006. Using kinetic typography to convey emotion in text-based interpersonal communication. In *Proceedings of the 6th conference on Designing Interactive systems*. 41–49.
- [50] Michael Y Lee, Melissa Mazmanian, and Leslie Perlow. 2020. Fostering positive relational dynamics: The power of spaces and interaction scripts. *Academy of Management Journal* 63, 1 (2020), 96–123.
- [51] Jennifer S Lerner, Ye Li, Piercarlo Valdesolo, and Karim S Kassam. 2015. Emotion and decision making. *Annual review of psychology* 66 (2015).
- [52] Gilly Leshed, Jeffrey T Hancock, Dan Cosley, Poppy L McLeod, and Geri Gay. 2007. Feedback for guiding reflection on teamwork practices. In *Proceedings of the 2007 international ACM conference on Supporting group work (Sanibel Island, Florida, USA) (GROUP ’07)*. Association for Computing Machinery, New York, NY, USA, 217–220.

- [53] Gilly Leshed, Diego Perez, Jeffrey T Hancock, Dan Cosley, Jeremy Birnholtz, Soyoung Lee, Poppy L McLeod, and Geri Gay. 2009. Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) (CHI '09). Association for Computing Machinery, New York, NY, USA, 537–546.
- [54] Fannie Liu, Laura Dabbish, and Geoff Kaufman. 2017. Supporting social interactions with an expressive heart rate sharing application. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–26.
- [55] Ioanna Lykourantzou, Angeliki Antoniou, Yannick Naudet, and Steven P Dow. 2016. Personality matters: Balancing for personality types leads to better outcomes for crowd teams. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 260–273.
- [56] Ioanna Lykourantzou, Robert E Kraut, and Steven P Dow. 2017. Team Dating Leads to Better Online Ad Hoc Collaborations. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) (CSCW '17). Association for Computing Machinery, New York, NY, USA, 2330–2343.
- [57] Andrew Mao, Yiling Chen, Krzysztof Z Gajos, David C Parkes, Ariel D Procaccia, and Haoqi Zhang. 2012. Turkserver: Enabling synchronous and longitudinal online experiments. In *Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence*.
- [58] Michelle A Marks, John E Mathieu, and Stephen J Zaccaro. 2001. A temporally based framework and taxonomy of team processes. *Academy of management review* 26, 3 (2001), 356–376.
- [59] Joseph Edward McGrath. 1984. *Groups: Interaction and performance*. Vol. 14. Prentice-Hall Englewood Cliffs, NJ.
- [60] Jessica Menold and Kathryn Jablowski. 2019. Exploring the effects of cognitive style diversity and self-efficacy beliefs on final design attributes in student design teams. *Design Studies* 60 (Jan. 2019), 71–102.
- [61] Elizabeth Wolef Morrison and Frances J Milliken. 2000. Organizational Silence: A Barrier to Change and Development in a Pluralistic World. *AMRO* 25, 4 (Oct. 2000), 706–725.
- [62] Elizabeth Wolfe Morrison, Sara L Wheeler-Smith, and Dishan Kamdar. 2011. Speaking up in groups: a cross-level study of group voice climate and voice. *Journal of Applied Psychology* 96, 1 (2011), 183.
- [63] Philipp Müller, Michael Xuelin Huang, and Andreas Bulling. 2018. Detecting Low Rapport During Natural Interactions in Small Groups from Non-Verbal Behaviour. In *23rd International Conference on Intelligent User Interfaces* (Tokyo, Japan) (IUI '18). Association for Computing Machinery, New York, NY, USA, 153–164.
- [64] Gary M Olson and Judith S Olson. 2000. Distance matters. *Human-computer interaction* 15, 2-3 (2000), 139–178.
- [65] Soya Park, April Yi Wang, Ban Kawas, Q Vera Liao, David Piorkowski, and Marina Danilevsky. 2021. Facilitating knowledge sharing from domain experts to data scientists for building nlp models. In *26th International Conference on Intelligent User Interfaces*. 585–596.
- [66] James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates* 71, 2001 (2001), 2001.
- [67] Leslie Perlow and Stephanie Williams. 2003. Is silence killing your company? *Ieee Engineering Management Review* 31, 4 (2003), 18–23.
- [68] Sajjadur Rahman, Pao Siangliulue, and Adam Marcus. 2020. MixTAPE: Mixed-initiative Team Action Plan Creation Through Semi-structured Notes, Automatic Task Generation, and Task Classification. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–26.
- [69] Daniela Retelny, Sébastien Robaszekiewicz, Alexandra To, Walter S Lasecki, Jay Patel, Negar Rahmati, Tulsee Doshi, Melissa Valentine, and Michael S Bernstein. 2014. Expert crowdsourcing with flash teams. In *Proceedings of the 27th annual ACM symposium on User interface software and technology* (Honolulu, Hawaii, USA) (UIST '14). Association for Computing Machinery, New York, NY, USA, 75–85.
- [70] Mike Rose. [n.d.]. No Comment: Inneract provides antidote to restrictive social norms. <https://www.engadget.com/2011-05-05-no-comment-inneract-provides-antidote-to-restrictive-social-nor.html>.
- [71] Niloufar Salehi and Michael S Bernstein. 2018. Hive: Collective Design Through Network Rotation. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (Nov. 2018), 1–26.
- [72] Niloufar Salehi, Andrew McCabe, Melissa Valentine, and Michael Bernstein. 2017. Huddler: Convening Stable and Familiar Crowd Teams Despite Unpredictable Availability. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) (CSCW '17). Association for Computing Machinery, New York, NY, USA, 1700–1713.
- [73] BR Sandler, LA Silverberg, and RM Hall. 1996. Pedagogy and the classroom. *The Chilly Classroom Climate: A Guide to Improve the Education of Women*. The National Association for Women in Education, Washington, DC (1996), 37–47.
- [74] Mike Schaeckermann, Joslin Goh, Kate Larson, and Edith Law. 2018. Resolvable vs. irresolvable disagreement: A study on worker deliberation in crowd work. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–19.

- [75] Valerie I Sessa. 1996. Using perspective taking to manage conflict and affect in teams. *The Journal of applied behavioral science* 32, 1 (1996), 101–115.
- [76] Dan Sholler, Igor Steinmacher, Denae Ford, Mara Averick, Mike Hoye, and Greg Wilson. 2019. Ten simple rules for helping newcomers become contributors to open projects. *PLoS computational biology* 15, 9 (2019), e1007296.
- [77] Charles Spearman. 1906. Footrule for measuring correlation. *British Journal of Psychology* 2, 1 (1906), 89.
- [78] Angela E B Stewart, Hana Vrzakova, Chen Sun, Jade Yonehiro, Cathlyn Adele Stone, Nicholas D Duran, Valerie Shute, and Sidney K D’Mello. 2019. I Say, You Say, We Say: Using Spoken Language to Model Socio-Cognitive Processes during Computer-Supported Collaborative Problem Solving. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (Nov. 2019), 1–19.
- [79] Annie Tat and M Sheelagh T Carpendale. 2002. Visualising human dialog. In *Proceedings Sixth International Conference on Information Visualisation*. IEEE, 16–21.
- [80] Annie Tat and Sheelagh Carpendale. 2006. CrystalChat: Visualizing personal chat history. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS’06)*, Vol. 3. IEEE, 58c–58c.
- [81] Holly E Tatum, Beth M Schwartz, Peggy A Schimmoeller, and Nicole Perry. 2013. Classroom participation and student-faculty interactions: does gender matter? *The Journal of Higher Education* 84, 6 (2013), 745–768.
- [82] Yla R Tausczik and James W Pennebaker. 2013. Improving teamwork using real-time language feedback. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 459–468.
- [83] Bruce W Tuckman. 1965. Developmental sequence in small groups. *Psychological bulletin* 63, 6 (1965), 384.
- [84] Rajan Vaish, Snehal Kumar (Neil) S Gaikwad, Geza Kovacs, Andreas Veit, Ranjay Krishna, Imanol Arrieta Ibarra, Camelia Simoiu, Michael Wilber, Serge Belongie, Sharad Goel, et al. 2017. Crowd research: Open and scalable university laboratories. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. 829–843.
- [85] Melissa A Valentine, Daniela Retelny, Alexandra To, Negar Rahmati, Tulsee Doshi, and Michael S Bernstein. 2017. Flash Organizations: Crowdsourcing Complex Work by Structuring Crowds As Organizations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI ’17). Association for Computing Machinery, New York, NY, USA, 3523–3537.
- [86] Kathleen Van Royen, Karolien Poels, Heidi Vandebosch, and Philippe Adam. 2017. “Thinking before posting?” Reducing cyber harassment on social networking sites through a reflective message. *Computers in Human Behavior* 66 (2017), 345–352.
- [87] Fernanda B Viegas and Judith S Donath. 1999. Chat circles. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 9–16.
- [88] Hana Vrzakova, Mary Jean Amon, Angela E B Stewart, and Sidney K D’Mello. 2019. Dynamics of Visual Attention in Multiparty Collaborative Problem Solving using Multidimensional Recurrence Quantification Analysis. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI ’19, Paper 342). Association for Computing Machinery, New York, NY, USA, 1–14.
- [89] Hua Wang, Helmut Prendinger, and Takeo Igarashi. 2004. Communicating emotions in online chat using physiological sensors and animated text. In *CHI’04 extended abstracts on Human factors in computing systems*. 1171–1174.
- [90] Richard T Watson, Gerardine DeSanctis, and Marshall Scott Poole. 1988. Using a GDSS to facilitate group consensus: Some intended and unintended consequences. *Mis Quarterly* (1988), 463–478.
- [91] Mark E Whiting, Allie Blaising, Chloe Barreau, Laura Fiuza, Nik Marda, Melissa Valentine, and Michael S Bernstein. 2019. Did It Have To End This Way? Understanding The Consistency of Team Fracture. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (Nov. 2019), 1–23.
- [92] Mark E Whiting, Irena Gao, Michelle Xing, N’godjigui Junior Diarrassouba, Tonya Nguyen, and Michael S Bernstein. 2020. Parallel worlds: Repeated initializations of the same team to improve team viability. *proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–22.
- [93] Chien Wen Yuan, Leslie D Setlock, Dan Cosley, and Susan R Fussell. 2013. Understanding informal communication in multilingual contexts. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 909–922.
- [94] Jamil Zaki. 2014. Empathy: a motivated account. *Psychological bulletin* 140, 6 (2014), 1608.
- [95] Amy X Zhang, Lea Verou, and David Karger. 2017. Wikum: Bridging discussion forums and wikis using recursive summarization. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 2082–2096.
- [96] Sharon Zhou, Melissa Valentine, and Michael S Bernstein. 2018. In search of the dream team: Temporally constrained multi-armed bandits for identifying effective team structures. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.
- [97] Charlene Zietsma and Thomas B Lawrence. 2010. Institutional work in the transformation of an organizational field: The interplay of boundary work and practice work. *Administrative science quarterly* 55, 2 (2010), 189–221.

Received July 2021; revised November 2021; accepted November 2021