ProtoTeams: Supporting Team Dating in Co-Located Settings

GUSTAVO UMBELINO, Northwestern University, USA
MATIN YARMAND, University of California, San Diego, USA
SAMUEL BLAKE, University of California, San Diego, USA
VIVIAN TA, University of California, San Diego, USA
AMY LUO, Carnegie Mellon University, USA
STEVEN P. DOW, University of California, San Diego, USA

Team dating, or small-group interactions, can expose people to diverse perspectives and inform the potential for longer-term collaboration. However, rapidly configuring groups and facilitating interactions among strangers can be difficult, especially in co-located settings. We present ProtoTeams, a system that leverages personal mobile devices to support rapid group formation, to facilitate group activities, and to collect data about the potential for future collaboration. We report on a field study where 406 students in eight different project-based classes used ProtoTeams to interact with classmates through multiple rounds of brief discussion activities before selecting teammates for a term project. We found that the system enables groups to form in about one minute, allows for meaningful interactions with a diverse range of peers, and can significantly influence subsequent teammate selection. We discuss design implications and challenges for in-person team dating in classrooms and other contexts.

CCS Concepts: • **Human-centered computing → Collaborative and social computing systems and tools**; *Interactive systems and tools*; **Field studies**.

Additional Key Words and Phrases: co-located collaboration, team formation, team dating, groupware

ACM Reference Format:
Gustavo Umbelino, Matin Yarmand, Samuel Blake, Vivian Ta, Amy Luo, and Steven P. Dow. 2020. ProtoTeams: Supporting Team Dating in Co-Located Settings. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3, Article 273 (December 2020), 27 pages. https://doi.org/10.1145/3434182

1 INTRODUCTION

Effective small-group interactions not only build and strengthen the bonds between people, but can also create a socio-emotional environment conducive for interactive learning [15] and teamwork [22]. In educational settings, especially those that have adopted the studio model for design education [90] or the flipped classroom [74], instructors often divide their students into small groups of various sizes to share personal perspectives and debate open-ended issues [12, 65], or to participate in group activities, such as jigsaw learning, where group members specialize and teach each other a different part of a topic [5, 8].

Authors' addresses: Gustavo Umbelino, Northwestern University, USA, gustavo@u.northwestern.edu; Matin Yarmand, University of California, San Diego, USA, myarmand@eng.ucsd.edu; Samuel Blake, University of California, San Diego, USA, sblake@ucsd.edu; Vivian Ta, University of California, San Diego, USA, vivian.vta@gmail.com; Amy Luo, Carnegie Mellon University, USA, awluo@andrew.cmu.edu; Steven P. Dow, University of California, San Diego, USA, spdow@ucsd.edu.

2573-0142/2020/12-ART273 $15.00
https://doi.org/10.1145/3434182

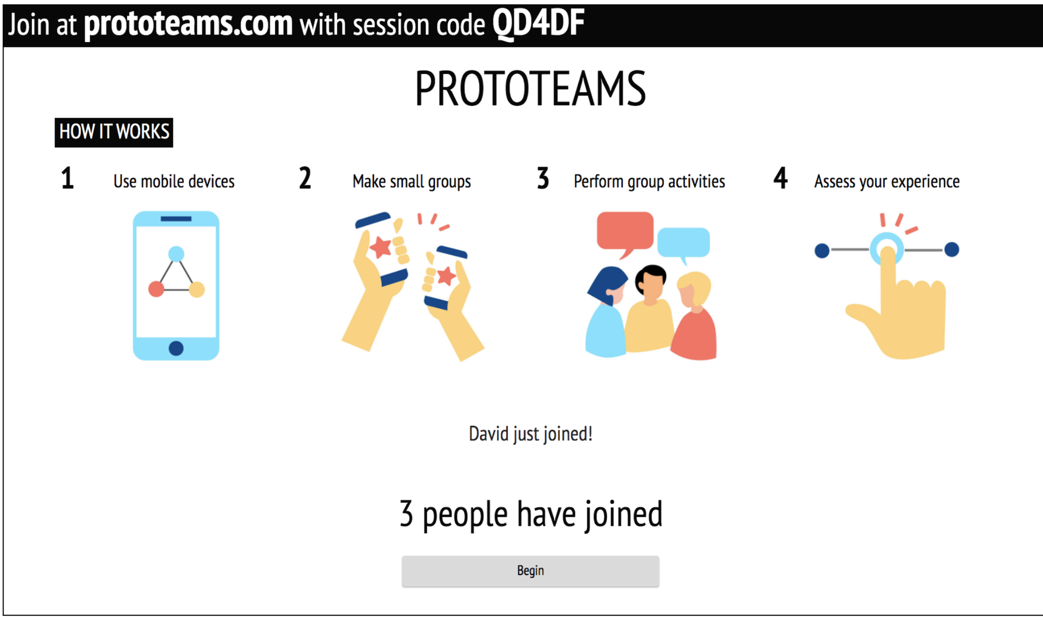

Fig. 1. Participants see this overview screen on the shared display when joining ProtoTeams (www.prototeams.com). The system leverages personal mobile devices to facilitate co-located *team dating* by quickly assembling small groups, scaffolding group activities, and supporting peer assessment for potential long-term collaboration.

The computer-supported cooperative work (CSCW) community has long focused on understanding what makes teams successful and how to design technologies to support formation [35], coordination [76], and collaboration [71]. The literature typically recommends forming diverse groups, based on individual characteristics, like gender [94], background [62], and personality [31, 59]. Several software tools support the creation of criteria-based teams [56], however, the rigidity of this approach often leaves students desiring more agency [45] and instructors seeking more flexibility for team composition.

Other researchers and educators have explored an approach to team formation that relies on the intuitions that form during short small-group interactions. This *team dating* method involves creating a series of short interactions with potential teammates before forming into final teams [19, 60, 61]. The concept builds on the idea of *thin slicing* [3, 4, 10] where individuals gather information, assess interpersonal dynamics, and make a judgement, in a short amount of time. While researchers have demonstrated the efficacy of this approach in online settings [60, 61], it remains a challenge to setup these kinds of rapid interactions in co-located settings, like classrooms. Instructors may verbally instruct students to quickly form small groups, but this does not assure that students interact with a diverse set of peers and makes it difficult to collect data about the potential of future teams.

To support co-located team dating, we designed and developed a web-based application called *ProtoTeams* to enable rapid and interactive group activities facilitated through personal mobile devices. As depicted in Figure 1, the system works as follows: an audience gathers in a physical space and then a facilitator asks people to take out their mobile devices and browse to a short URL. Participants enter their preferred name and then their phone displays a shape-color symbol (e.g., blue circle, red triangle, etc.) that visually connects them with other people in the room (Figure

Fig. 2. Group formation stage where participants find their group members using their assigned symbol. All participants included in this study provided consent to be videotaped for the purposes of analysis and presentation.

2). After participants gather into small groups using these visual cues, instructions appear on each person's mobile device to lead each group through short activities, including introductions, icebreakers, creative thinking, and skill sharing. Each "date" ends by asking peers to do a short assessment about the potential of working together in the future. ProtoTeams allows facilitators to repeat this process of rapid group formation, interaction, and assessment for as many rounds as needed to help people mix with each other or as long as time allows.

This paper describes our design process and goals, including how we tested low-fidelity prototypes to explore different mechanisms for group formation, interactive activities, and assessment methods. To evaluate ProtoTeams, we conducted nine deployments in eight different project-based classes where 406 students participated in multiple rounds of team dating with groups of two to six people in each date. After using ProtoTeams, students filled out a short survey to report on their experience and then formed teams on their own without any technology. We found that ProtoTeams enables small groups to form rapidly (~60 seconds on average) and crowds got faster at forming groups each round. Students reported that they enjoyed the formation exercise where they find others based on colored shapes and they appreciated the opportunity to meaningfully interact with peers they would not have otherwise. We deployed a pre-survey to capture existing social ties and other personal factors, and we found that co-located team dating can affect teammate selection, especially for audiences where there are fewer existing social ties. This paper offers the following contributions:

- We create a novel system, ProtoTeams, for facilitating co-located team dating through personal mobile devices and describe our process to design and prototype these social interactions;
- We perform a mixed-method analysis of nine deployments across eight different project-based classrooms with 406 students. We found that ProtoTeams enables groups to form quickly, allows for meaningful interactions with diverse peers, and can significantly influence teammate selection;
- We discuss key considerations for supporting co-located team-dating activities, including how we might leverage this approach for a variety of small-group activities in different contexts.

2 RELATED WORK

To situate our work, we review prior research on the benefits of group work in classroom settings, strategies for forming groups and teams, and tools to support co-located team dating.

2.1 Group Work in Classrooms

Many instructors and researchers recognize the value of small-group, collaborative learning activities. Some of the benefits of cooperative learning include improved class participation [80], enhanced socialization among peers [47, 49], academic gains [38], and greater psychological health, such as social competence and self-esteem [48]. To kick-start group interactions, instructors often deploy low-stakes activities, such as icebreakers and team-building exercises. These exercises are shown to be important for improving emotional resilience to stress during collaborative tasks [68] and for overcoming initial unfamiliarity between members of a newly formed team [37, 45]. While there are numerous strategies for supporting peer interactions, such as peer feedback exchange [55, 77, 83], this paper focuses on how to initially create safe opportunities for peers to mix diversely, towards the eventual goal of forming teams for projects or collaborative learning.

Research on small-group classroom activities shows that factors such as group size, composition, task type, and time on task can influence the outcomes and member satisfaction [6, 13, 66, 81, 92]. With respect to group size, some evidence indicates that larger (5-person) groups achieve better outcomes than smaller (2-person) groups, since more people means more overall effort [43]. Other research indicates that individuals working in smaller groups are more productive and that larger groups exhibit more social loafing [69], where one or more members exerts less effort [50]. Our research provides an opportunity to collect a range of empirical data on key factors, such as group size, to explore how they might influence groups' experiences during small-group classroom activities.

2.2 Strategies for Forming Groups and Teams

Deciding how to form small groups for collaboration is challenging because of the many trade-offs to consider. In classrooms, instructors employ a variety of approaches [27], including self-selection, random assignment, and criteria-based algorithms [45]. These approaches vary in terms of their ease of deployment, student satisfaction, and overall balance of familiarity and diversity of groups. Self-selection gives students agency to choose group members, which generally increases satisfaction with the resulting groups [45]. However, students tend to lean on their existing social ties to form groups [28]. More familiar groups generally feel safer with each other and can begin performing faster [85]. However, assembling groups only based on familiarity often results in group members having very similar skills and life experiences [30, 46], which can hinder creative output [84].

Other strategies are intended to foster diversity by infusing a wider range of skills and experiences into groups. For instance, randomly assigning students might lead to less familiar and more diverse groups. However, this strategy can result in less cohesion, higher conflict rates [20], and a longer period to cultivate a high-performing team [85]. Random assignment is an efficient way for instructors to form diverse teams, but it also lacks a rational basis, which is valued by students [45].

Many researchers and instructors have explored criteria-based approaches to group formation, including personality traits [31, 59], gender and language preference [2], interest and background [63], and other individual characteristics. However, a recent review of team formation algorithms revealed that there is still an ongoing debate about the importance of different criteria for automatically forming effective teams [70]. Moreover, this approach requires information about individuals that can be difficult or time-consuming to collect in advance.

Towards balancing for familiarity, diversity, and student agency, Curşeu and colleagues [19] introduced the concept of *team dating* where students repeatedly participate in short, small-group activities. After each interaction, students assess their preferred peers for a longer-term task. During team dating interactions, subjects usually perform short, low-stakes activities with the goal of getting to know potential teammates. The idea is to allow participants to interact with and to assess their fit with many peers, something that is difficult to measure or predict solely through individual characteristics. The researchers found that teams performed better when formed based on reciprocal preferences rather than external criteria (e.g., skill balance) [19]. Their findings have also been replicated in an online setting by Lykourentzou et al. [60, 61] where the authors used an online system to form groups, coordinate interactions, and support peer assessment.

In classroom settings, the team dating approach not only can help students create more informed teams, it can also help instructors facilitate a wide range of small-group learning activities. However, while technology exists to support team dating in online settings [60, 61], there are no tools, to our knowledge, to coordinate team dating in co-located settings.

2.3 Tools to Support Co-Located Team Dating

The CSCW community has a long history of creating tools to facilitate group formation [35], but these tools do not support student agency in co-located settings. Criteria-based tools for team formation, such as *CATME* and *MyDreamTeam*, attempt to balance for multiple students characteristics, but generally do not collect and account for student preferences for teammates [17, 56]. A recent study indicates that students want to have more agency over team formation [45], but very few systems support this feature [21, 36]. Moreover, most group formation tools are primarily designed for online settings where algorithms organize which people are grouped together. For example, the *Hive* system repeatedly reorganizes groups by rotating group membership over time [76]. Similarly to the group formation stage in team dating, rotation strategies could be difficult to perform in co-located settings where individuals need to physically move around.

A number of research systems have explored how to facilitate co-located group interactions, beyond the goal of team formation. For instance, researchers have created tools to support co-located peer feedback exchange [34, 77, 83], brainstorming activities [14], shared note-taking [73], and in-class discussions [18]. Our work on ProtoTeams builds on these approaches by considering how mobile technology can support student interaction in a physical classroom. In particular, we strive to create activities that help students build familiarity with a diverse set of peers.

Similarly, a number of tools facilitate peer assessment in different contexts [58, 72]. However, these tools are generally not designed for rapid, on-the-spot evaluations, the kind needed in a co-located team-dating session. Performing assessments can surface a range of socio-emotional factors, such psychological safety and trust [88], which might bias the assessments. These biases may be even more problematic when peers assess each other in a face-to-face setting [72]. A key goal for ProtoTeams is to allow peers to quickly and comfortably perform on-the-spot assessments to inform the potential for future teams.

3 PROTOTEAMS SYSTEM

We created ProtoTeams to help facilitate team dating in co-located settings. The system leverages personal mobile devices to quickly and repeatedly form small groups within a large co-located crowd and then provides prompts for short group interactions and data collection that can inform long-term teams. First, we describe our prototyping process and insights from a series of informal evaluations within our research lab. Next, we articulate design goals and principles for three key aspects of team dates: formation, interaction, and assessment. Finally we share the final system design, including the technology stack, interface design, and user experience.

3.1 Prototyping Process

As noted in prior research, developing social-interaction technologies, also known as groupware, can be challenging due to the difficulties of recruiting groups of people, accounting for existing social ties, and creating authentic emergent behavior between participants [32, 33]. To get a sense for how this social computing application could work, we first prototyped the system on paper. Then, inspired by small-group party games [24], we built a simple digital prototype that we could test informally before building a more robust experience for larger groups of people.

We conducted several rounds of informal paper prototyping with 15 people per session as part of our lab's weekly research meetings and collected qualitative data through observations and interviews. Our early prototypes divided the sessions into multiple rounds of two stages: group formation and group activity. To coordinate these stages, we created paper cards with individual instructions and input fields as well as shapes (e.g., square, circle, star, etc.). The instructions/input card recorded individual responses during the group activity and the shapes helped group members find each during the formation phase. During the group activity, participants were instructed to verbally share their responses within their small groups. Each participant received one shape-instructions pair of cards per round in the session. We also created a companion slide deck with general instructions for each stage and used a timer to help guide the sessions.

Paper prototyping revealed a number of issues that we addressed through our low-fidelity digital prototype. We learned through practice that the system would need to be resilient to people coming in and out of the session. Late-arrivers often disrupted the onboarding and grouping process, and the facilitator would need to re-iterate instructions. In the digital prototype, we chose to show the participation URL and a unique session code on a shared screen at all times. Late-arrivers would be able to join after observing a round of group interaction.

We also learned that facilitators found it difficult to know when all groups have formed, especially with large numbers of participants. This is important to keep the groups synchronized and pace the interactions accordingly. To address this, our digital prototype asked participants to confirm when their group forms. Anecdotally, during our paper prototyping sessions, we noticed that, perhaps out of convenience, participants would often form groups among the nearest people, even if they had interacted with the same people in prior rounds. To instigate more diverse interactions, our digital prototype included a group formation algorithm that randomly grouped people and avoided putting people together multiple times across rounds.

Finally, in order to assess whether groups had potential for longer-term interactions, the digital prototype introduced an assessment interface at the end of each round of group activity. Building on the prior literature on team dating [19, 60, 61], participants were asked a short question about group member preference: *"Rate how likely you are to work with these members again"*. For each group member, they were provided with a slider on a 5-point Likert scale (1: "Extremely unlikely" to 5: "Extremely likely"). The purpose of this question was to assess the future potential of interacting groups.

We tested the digital prototype in an undergraduate project-based design course through a series of three pilot studies [87]. Sessions took place during three different class periods over a 4-week period and varied in size between 50 to 67 students. The research team collected qualitative data through debrief surveys at the end of each session, observations, and interviews.

Our observations from these sessions showed that participants found their group members, but did not always tap their names to confirm, as instructed. This was a challenge because the system relied on this confirmation to deliver group-specific instructions and to collect group formation times. Moreover, the assessment interface created awkwardness among group members, because they were asked to answer a personal question regarding their group members while standing next to them. To address these issues and to clarify our overall design goals for the system, we compiled a set of design principles for supporting in-person team dating activities.

3.2 Design Principles for Supporting Co-Located Team Dating

Based on insights from paper and digital prototyping, we present the key goals and design principles, organized around the three main stages of ProtoTeams: group formation, interactions, and assessment. During the formation stage, ProtoTeams should help facilitators organize an audience into smaller groups within co-located spaces. Audience members who choose to participate will move around the space such that they are face-to-face with their assigned group members. Key design principles for group *formation* include:

- **Minimal**: The system should be able to form groups with minimal or no information about the audience or the space;
- **Rapid**: The system should help people efficiently find group members in a large crowd;
- **Flexible**: The system should be adaptable to a large range of environments and user behaviors, including supporting participants who join late or drop out;
- **Accessible**: The system should be very easy to use and work on personal mobile devices.

For the interaction stage, ProtoTeams should motivate and scaffold interactions between group members so that participants get to know and learn from each other. Key design principles for group *interaction* include:

- **Social**: The system should present activities to help group members get to know each other as potential collaborators;
- **Brief**: Interactions should be short and rich, so that participants have time to meet more people within a dedicated time period;
- **Equitable**: The system should support activities that are balanced and inclusive, so all members have equal opportunity to participate.

For the assessment stage, ProtoTeams should provide a simple and quick way for participants to privately evaluate their "fit" with other group members to reflect the potential for longer-term collaboration. Key design principles for group *assessment* include:

- **Informative**: The system should collect data that meaningfully inform the future potential of interacting group members;
- **Socially Acceptable**: The system should avoid creating feelings of awkwardness among participants, which can bias the results;
- **Timely**: The system should collect data as soon as the group interaction ends so that participants recall details they might otherwise forget.

3.3 System Design and User Experience

The ProtoTeams system comprises of three key components: a mobile user interface that appears on individuals' personal devices, a shared screen that displays information for the entire audience, and a dashboard for facilitators. We implemented ProtoTeams using Meteor, an open-source JavaScript framework for rapid web application development. We used MongoDB to store participant and interaction data, and React for the front-end.

ProtoTeams supports two main types of users: *hosts* who facilitate the sessions and *participants* who participate in team-dating activities. The *shared display* provides participants and hosts an overview of the session status. For instance, as shown in Figure 3, as participants form groups, the shared display shows the system's URL and session code, as well as all the assigned shapes and their status. Filled shapes represent confirmed groups while outlined shapes represent groups that are still being formed. The shared display is also used by the host to manually advance between activity stages.

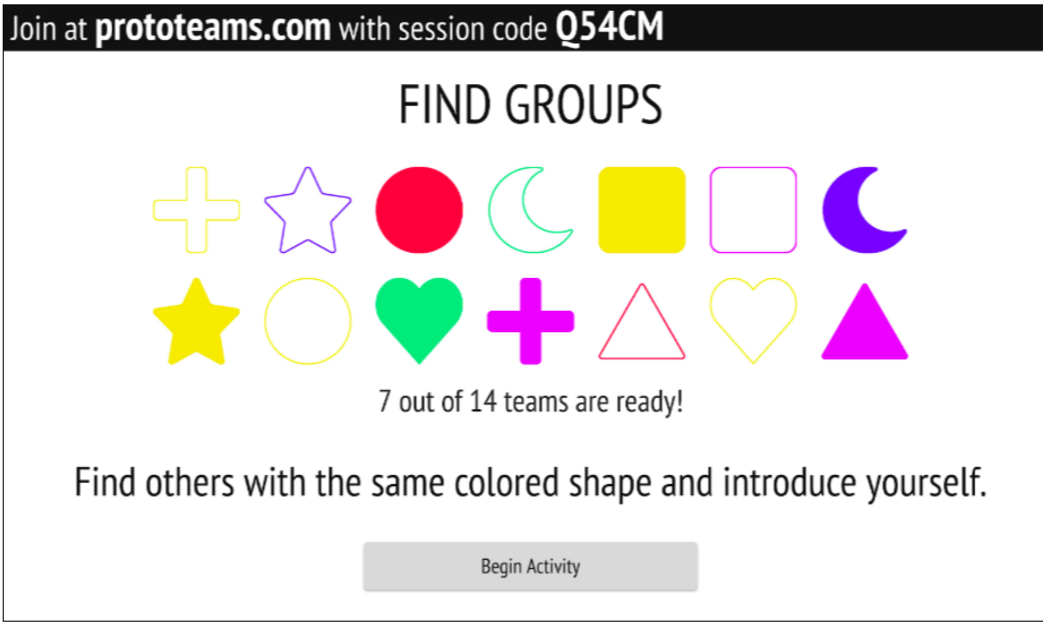

Fig. 3. Shared display during the group formation phase. Shapes become filled when groups solve the collaborative puzzle correctly. Unconfirmed groups appear as an outlined shape. The group activity phase begins automatically once all groups confirm. Alternatively, instructors can manually begin the activity by clicking the “Begin Activity” button.

ProtoTeams primarily structures interactions through a *mobile interface* accessed via web browser on participants’ personal mobile devices (see Figure 4). We take advantage of personal mobile devices due to their popularity, accessibility, and privacy. Finally, session facilitators can use a private *dashboard* to create and customize aspects like the number of activities, duration of each activity, group size, number of rounds, and discussion prompts.

3.3.1 Joining a session. Each ProtoTeams session can support multiple rounds with three stages in each round: formation, interaction, and assessment. The system selects reasonable defaults for group size (three), the number of rounds (six), and the duration of each round (three minutes). Hosts can also configure these settings in the dashboard.

Once the host and participants are located at the same physical space, the host can present the shared display where a welcome screen greets participants and provides instructions to join. Participants navigate to the short URL on their personal mobile devices. The welcome screen also provides a randomly-generated 5-character session code that ensures only people in the room who can see the shared display can participate. Next, participants enter a name to use during the session and wait for the host to begin the activities. The shared display includes the instructions throughout the entire session, so participants can join at any time. The process when joining late works the same (URL, session code, and name), but if someone joins mid-activity, they wait in a virtual queue until the next round of group formation.

3.3.2 Forming groups. During the group formation stage, the shared display shows empty outlines of each colored shape, as well as the instructions: "Find others with the same colored shape and introduce yourself." On the mobile devices, each individual sees a colored shape and the names

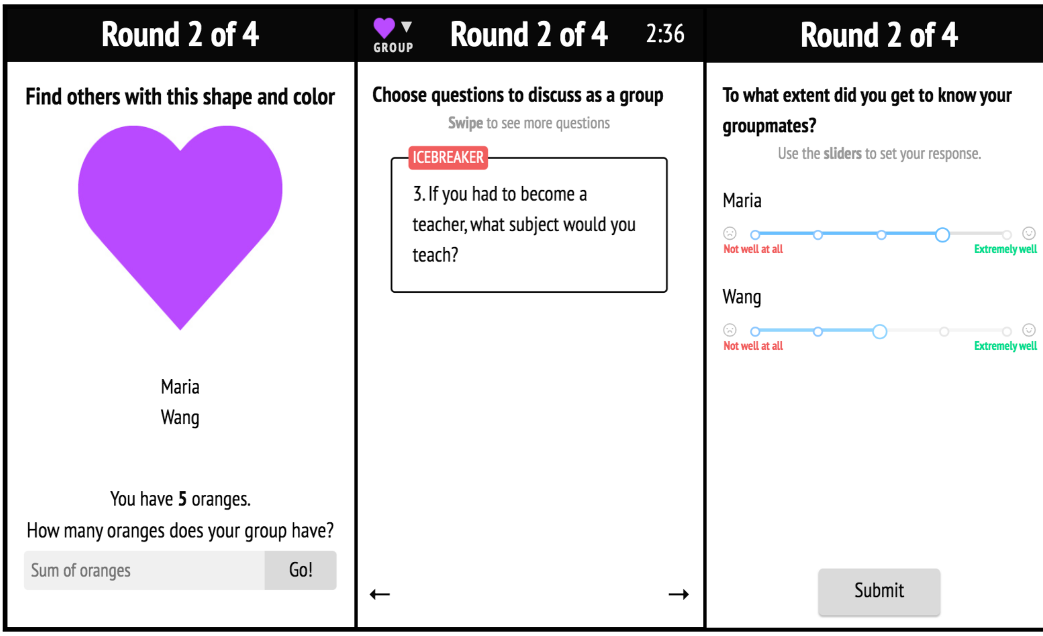

Fig. 4. Examples of ProtoTeams' mobile interface screens: (left) during group *formation*, the interface displays the participant's assigned colored shape, a random number of oranges, and an input field for the confirmation question (i.e., total number of oranges for the group); (middle) during *interaction*, ProtoTeams presents discussion prompts which can be accessed by swiping left or right; (right) during *assessment*, the interface shows group members' names and participants rate how much they got to know each member.

of their group members for that round, as shown in Figure 4 (left). ProtoTeams also includes a short collaborative puzzle to confirm that groups find each other. The puzzle shows each member a random number of oranges and then asks *"How many oranges does your group have?"* As soon as one group member solves the puzzle correctly, the system confirms the group and records the elapsed time since groups were assigned. The shared display is updated by filling in the colored shape for that group. While waiting for other groups to form, confirmed groups are encouraged to introduce themselves to each other.

3.3.3 Interacting in groups. The interaction stage begins when all groups confirm or the host manually advances the session through the shared display. The host may choose to advance manually rather than wait for all groups to solve their puzzles, in order to keep the session moving forward in a timely manner. At this stage, the system presents discussion prompts designed to induce a deeper level of familiarity between the group members. The research team compiled three types of discussion prompts: *icebreakers*, *team-building*, and *ideation* prompts. Each type of prompt was labeled and color-coded on the mobile interface for easy distinction. Participants can swipe left or right through prompts on their own mobile devices, and it does not affect what other group members see on their screens. Examples of each type of prompt are shown in Table 1.

ProtoTeams includes over 75 different discussion prompts, which can be further customized and supplemented by the host. For each round, the system randomly selects a mix of all three prompt types. Participants receive different icebreaker and ideation prompts to allow the small-group interactions to remain unique and interesting. The team-building prompts are held consistent each round to help individuals discern potential teammates across rounds. We set the default length of

Types	Goals	Example prompts
Icebreaker	Offset the cold start of interactions and to create familiarity between members of the newly formed group	"What is the last movie you watched and would you recommend it?" "Share one of your favorite childhood memories."
Team-building	Elicit conversation about role preferences and past teamwork experiences	"What was your role in previous group projects that have you been a part of?" "What do you believe an important factor for a successful team?"
Ideation	Encourage the small group to quickly think of solutions to a subject-specific problem and help group members learn about each other's approaches to problem-solving	"Someone you know often forgets where they last placed things. What can you design to help them?" "An adult struggles with overeating. What solution can you design to help them develop better eating habits?"

Table 1. Types of discussion prompts, goals, and examples used during ProtoTeams' interactive activities.

the interaction stage to be three minutes. Prior research suggests that short interactions can be sufficient for evaluating potential teammates [3, 4, 10]. Also, by keeping each interaction stage short, it leaves time for more rounds and more opportunities to foster new connections between participants.

3.3.4 Assessing group members. During the assessment stage of each round, participants assess how much they got to know their group members. As shown in Figure 4 (right), participants answer the question *"To what extent did you get to know your groupmates?"* using a 5-point Likert scale with values "Not well at all", "Slightly well", "Moderately well', "Very well", and "Extremely well", respectively. Informed by prior literature on social tie strength [26, 64], the question estimates the strength of existing social ties by measuring closeness, since closeness is a strong predictor of tie strength [64]. Once all participants respond to this question, or the host decides to manually advance the activity, the next round begins. If there are no more rounds left, the session ends and participants are asked to fill out a debriefing survey. The survey requested input and perspectives about the ProtoTeams system, and would be excluded from a production version of the user experience.

4 FIELD STUDY

To evaluate ProtoTeams, we conducted a field study in classroom settings where students form teams for a course project. To identify deployment opportunities, we scanned syllabi published on our university registration system and reached out to instructors who taught project-based courses. This led to nine deployment sessions (with a total of 406 unique participants) across a range of course topics, including cognitive science, engineering, design, and business (see Table 2). One instructor invited our research team back for a second day. Sessions S2 and S8 are the same population of students, although the second session had fewer students that day. We worked with each instructor to select team-dating parameters. Table 2 shows how the group size varied between two to five people, time per round varied between three to five minutes, and the number

Session	Type	Rounds	Minutes per Round	Group Size+	N
S1	Design	6	3	3	65
S2*	Design	5	5	5	57
S3	Design	5	4	4	52
S4	Design	5	4	3	50
S5	Design	4	4	2	56
S6	Cognitive Science	5	5	3	30
S7	Engineering	5	5	3	21
S8*	Design	3	4	3	39
S9	Business	5	4	3	73

Table 2. Deployment sessions presented in chronological order. *S2 and S8 happened in the same class separated by a few weeks; the second session had fewer students in class that day. +Group size varied (plus or minus one) when the number of participants was not evenly divisible by the default group size.

of rounds ranged between three and six. We also noted differences in the physical spaces where the deployments took place. Sessions S2, S3, S4, S5, and S8 were held in the same classroom with movable tables and chairs that were pushed aside to create a large open space for participants to move around. Sessions S1, S6, and S7 did not have movable furniture but had space in the front for participants to gather. Sessions S9 took place in a lecture hall with multiple rows of immovable chairs and tables. In this deployment, students could only interact on the stairs and in a flat area between the podium and the first row of seats.

The primary goal of the field deployments was to understand emergent behaviors and reactions around the ProtoTeams system for structuring co-located team-dating activities. However, the natural variance between the different deployments also gave us an opportunity to investigate more specific research questions. In particular, instructors had different preferences for the number of students in each group, allowing us to examine how *group size* affects outcomes such as group formation times. Group size also varied within each session since the number of participants was not always evenly divisible by the desired group size. For instance, in S2 where we had 57 students, most groups resulted in the desired size of 5, but two groups had 6 members. We report on student opinions about the issue of group size.

In all classes, we deployed the system early in the term such that students would have a chance to interact through ProtoTeams before they formally created teams. For two sessions, S1 and S9, students formed teams immediately after the ProtoTeams session. We captured the final team composition for these two sessions to analyze the effects of team-dating activities on teammate selection in comparison to other factors, such as existing social ties. Based on the the differences between each classroom deployment, we asked three research questions:

- **RQ1.** What behaviors and reactions emerge during in-class deployments of ProtoTeams?
- **RQ2.** (How) does group size affect co-located team dating?
- **RQ3.** (How) do interactions during co-located team dating influence teammate selection?

To explore these questions, we collected quantitative data through pre-study surveys and system logs, including group formation time and group member assessment scores. We also gathered qualitative data through post-study surveys, semi-structured interviews, and observational data during ProtoTeams deployments.

4.1 Participants

A total of 406 students participated in at least one of nine different deployment sessions in undergraduate courses at a large public university in the United States. 286 participants filled out a pre-study survey with demographic information and prior social ties (64% response rate). The majority of students were American (33.2%) or Chinese (29.4%). 57.3% self-identified as Male, 37.4% as Female, and 1.1% as Non-binary, while 4.2% preferred not to disclose. Students were encouraged, but not required, to participate in the team-dating activity. Nearly everyone brought personal mobile phones to class on the deployment dates. We provided spare phones for those who did not have a mobile phone.

4.2 Procedure

Before each ProtoTeams session, we worked with instructors to email students the pre-study survey. At the start of each session, our research team briefly explained the team-dating activity, while the shared display showed the URL and session code so that students could join the session. During each session, the instructors helped facilitate and guide the students. Once a session ended, participants filled out the post-study survey.

4.3 Data Collection & Analysis

We collected a mix of quantitative and qualitative data before, during, and after ProtoTeams sessions. The research team also took notes to identify behaviors that could have not been captured by the system, such as observable group dynamics. The sources of data were combined and analyzed to reveal insights towards our research questions.

4.3.1 System log data. The ProtoTeams system collected a range of quantitative data, including timing (e.g., how long each participant took to join the session, how long for groups to confirm each round, how long each individual hovered on each discussion prompt), rating data (i.e., what assessment ratings each participant assigned to their group members), and group composition for each round (i.e., which students were randomly grouped together). To understand the effects of group size on the amount time participants need to find group members, we used an ANOVA and post-hoc t-tests. To analyze the potential effects of team dating interactions on teammate selection, we used a logistic regression.

4.3.2 Pre-study survey. Before each session, we sent out a pre-study survey to obtain written consent and to collect information about individual demographics (e.g., age, gender, nationality) and existing relationships (i.e., previous social connections with peers). We built on Gilbert and colleagues' approach for predicting existing tie strength [26] by asking students to specify their *closeness* to each classmate. We asked students *"How well do you know this person?"* for each peer in the class in a 5-point Likert scale (1: "I don't know them", 5: "We are very close"). Combined with the final composition of teams, these data were used in a logistic regression analysis to determine what factors—such as existing social ties, nationality, or gender—affected students' teammate choices.

4.3.3 Post-study survey. To gain insight on the ProtoTeams experience, we asked students to fill out a short, anonymous post-study survey. The survey collected students' responses immediately after each session through short-answer and multiple-choice questions. Participants provided overall impressions, such as what they did or did not like about the session. They also answered more specific questions, such as their preferred number of rounds and favorite types of discussion prompts.

4.3.4 Semi-structured interviews. To supplement the post-survey, after each session, we sent out an email to participants inviting them to take part in a 25-minute semi-structured interview with the research team. 18 students were interviewed within a week of their session and were compensated with a $5 Amazon gift card for their time. The interviews captured participants' overall thoughts, as well as specific insights about the three phases of ProtoTeams (i.e., formation, interaction, and assessment). We performed an inductive thematic analysis on the interview data, following Braun and Clarke's approach [11]. We first transcribed all interviews and categorized snippets based on their associated phase of ProtoTeams. We began open-coding the data by identifying explicit topics mentioned by the participants. Initial codes were combined into preliminary themes, which were discussed among the research team. Finally, after iteratively revising the preliminary themes, we agreed on the final themes described in detail below.

5 RESULTS

Overall, students enjoyed team dating with ProtoTeams, providing an average rating of 4.3 on a 5-point Likert scale (1: "Did not enjoy at all" to 5: "Extremely enjoyed") in the post-study survey. Students joined the system within an average of 35 seconds of presenting the URL on the shared screen. 6.5% of students joined their session late, due to being late to class. Once the group formation activity started, groups formed within an average of 54 seconds. At a high-level, our findings suggest that ProtoTeams supports our design goal of creating a rapid, flexible, and social way for students to interact and assess potential teammates. Here we present more detailed findings, organized around our three research questions.

5.1 What Behaviors and Reactions Emerge During In-Class Deployments of ProtoTeams?

To understand participants' experiences using ProtoTeams, we analyzed our observations, interview data, and log data. Figure 2 illustrates the group formation process where peers find group members (same color-shape) by holding up their mobile devices. Post-study interviews revealed that while it can take up to a minute for groups to form, students generally enjoyed going around the room and searching for their assigned group members. As one participant said, "The fact that you had to find people who had the same symbol made it very interactive and it was fun" (P4, S3). Another student described the clustering of people that happened during that formation stage:

> As chaotic as it was and it felt a little disorderly, I think that was part of its appeal. When you're told, okay everybody stand up, the first thing is that rush of adrenaline that you're like, 'Oh no, I've gotta go meet new people.' (P15, S9)

Many students described how the process felt chaotic, but seemed to help disarm people and make them more open to meeting new people. Moreover, some mentioned how the randomness of the group assembly forced them to interact with lesser known peers:

> The fact that it is random, and then you basically have to talk to them, it brings up new conversations that you wouldn't have had before. [...] I am meeting all these people and it is fun. (P4, S3)

By forming groups randomly, and offering interactive activities, ProtoTeams enabled students to meet people they normally would not interact with during a typical class.

5.1.1 Groups form faster each round. To understand how group formation behavior evolved throughout a session, we analyzed the formation times in each round (see Figure 5). For consistency, we excluded S5 and S8 since these sessions contained less than 5 rounds. We also removed S9 from this analysis, since we observed that the physical structure of this classroom hindered people from

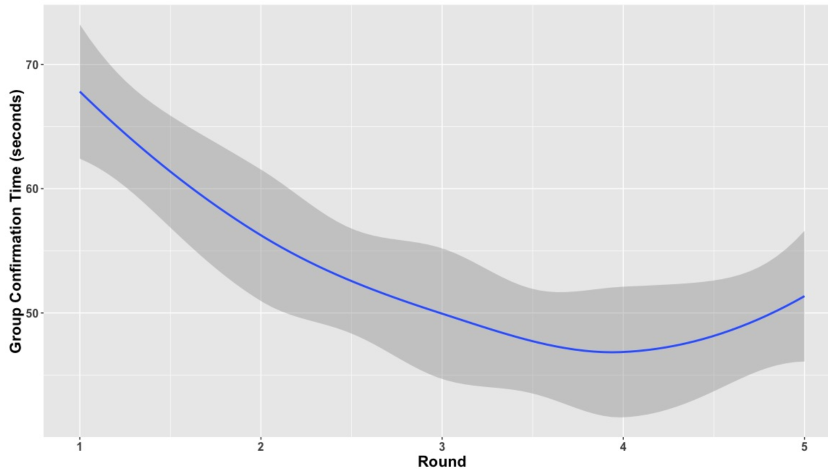

Fig. 5. Average group confirmation time (seconds) in each round (from sessions S1-S4, S6-S7, and S8) with standard error indicators. Group formation goes faster with practice.

moving freely. The average group confirmation time in S9 (~86 seconds) was considerably longer than all other sessions (~52 seconds).

Figure 5 shows how group formation time decreases by round. It took an average 68 seconds in the first round, and reduced to 46 seconds by round four. This result also aligns with qualitative evidence, as one participant mentioned:

> I think towards the end it seemed like it was getting easier to find people. At the beginning, it was a little weird because we were just understanding it. (P1, S3)

Notably, the average confirmation time increased slightly (~5 seconds) from round four to round five. In our anonymous post-study survey, some participants indicated that "five rounds can get exhausting to start conversations and remember people's names", and "it got a little tiring by the 5th round". While we did not directly observe fatigue, the nature of a team dating session, including searching for group mates and initiating multiple conversations with new people, can be quite demanding.

5.1.2 Participants appreciated the interaction prompts, but wanted more constructive activities. To scaffold peer interactions through ProtoTeams, we provided three types of discussion prompts: icebreakers, ideation tasks, and team-oriented questions. The log data revealed how much time groups spent discussing each type of prompt. This analysis assumes that groups discussed the prompts that appeared on their screens, otherwise they would have skipped by swiping to the next prompt. After normalizing by the number of participants, total activity time, and number of rounds, the top viewed prompt across all sessions was a team-building prompt (*"How is your course load this quarter?"*). On average, participants spent 7.8 seconds per minute of allocated time on this prompt every round. The most popular icebreaker and ideation prompts were *"What is the best TV show/movie that you watched recently?"* (6.0 seconds/minute) and *"Give an elevator pitch of your project idea."* (5.9 seconds/minute), respectively.

During interviews, we asked participants how they felt about the different types of discussion prompts. Participants described how the icebreaker prompts helped groups "build [their] own unique conversation” (P2, S4) and “make everybody feel comfortable” (P9, S7). However, the icebreaker prompts seemed to have only marginal value for assessing potential teammates.

> On the one hand, [the icebreaker prompts] are a great way for me to get to know very straightforward, basic details about somebody. But, that doesn’t give me insight as to what kind of person they are in the professional or academic environment. (P7, S4)

Overall, participants gave the most positive feedback for the team-building prompts. Participants liked how these prompts were relevant to factors that they found important in a successful team:

> [The team-building prompts] were the ones where we actually talked and discussed the most, because we finally found something in common. We are all students. We can all relate to these problems. (P7, S4)

While these team prompts helped people gain more information about the personality and skills of other group members, some participants mentioned that they would have preferred to engage more through cooperative activities:

> I feel like the ideal kind of questions would be one that forces us to work together, something like a puzzle where we have to solve it cooperatively. That way we bond naturally without having to ask forced questions about each other, like how is your course load or how did your summer go. (P2, S4)

Similarly, some participants talked about how they wanted more time and methods for gathering information to select a teammate. P8 said they "would be interested to learn a bit more in some unstructured way" such as the ability to "click a name and get their information." (P8, S4) In general, the ProtoTeams discussion prompts were effective in terms of initiating conversations, but could be improved by helping users simulate the experience of working with a potential teammate through constructive and cooperative activities.

5.1.3 Most participants failed to discern between group members when assessing them. To understand the informational value of the assessment rating, we measured to what extend participants discerned their group members in each round. To quantify this, we calculated the standard deviation of ratings given by each participant in each round. The standard deviation equaled zero in 69.8% of all assessment opportunities. That is, participants generally did not discern between group members. This analysis excludes groups of two, since they only have one other member to evaluate.

In the interviews, participants discussed how they often gave everyone in their group the same rating due to lack of time: "I basically gave everybody a three, because we only have five minutes" (P9, S7). Other participants noted how difficult it was to "remember some of the students’ responses and how I reacted after so many rounds” (P6, S6). Many participants met up to 20 new people during a session, making it hard to recall group members and conversations:

> We’ve been exposed to a lot of new faces and then it was really hard to remember and match the name to the face. (P5, S4)

Whether due to time constraints, memory, or discomfort with doing on-the-spot ratings, the group assessment did not yield discerning data about the potential for longer-term collaboration. More thoughtful design will need to go into the assessment stage before an algorithm would be able to recommend potential teams.

5.2 How Does Group Size Affect Team Dating?

To understand the effects of group size on team-dating activities, we analyzed group formation time, assessment patterns, and interview data.

5.2.1 Smaller groups form faster. We explored the relationship between group size and the amount of time to form groups (i.e., confirmation time). To control for potential confounding variables, we limited our analysis to three sessions that had the same instructor with the same course structure. Sessions S3, S4, and S5 were composed of groups of 4, 2, and 3 students, respectively. By comparing these three sessions, we can hold constant variables such as the instructor, course topic, and classroom that could influence the analysis of group formation times.

We find a significant effect of group size on formation time, as 2-person groups form in an average of 32 seconds (SD = 14), 3-person groups form in 46 seconds (SD = 21), and 4-person groups form in 66 seconds (SD = 12). A one-way ANOVA shows a statistically significant difference in the mean confirmation times between different group sizes ($F(2,245) = 101$, $p < 0.001$). This quantitative result aligns with our observations during the deployments. We observed that once the formation phase began, all participants actively moved around the classroom and searched for the other group members. As soon as two members found each other, they stood together and began to passively scan around the room for the missing members. With groups of four, we often observed two group members pairing up in different areas of the room and then they had to find the other pair. Groups of two likely formed fastest because they did not need to scan for other members.

5.2.2 Participants preferred three-person groups. We explored how group size can impact the interactions within groups, such as the engagement level of members. Since some students participated in groups of different sizes (either within a session or across sessions), they could provide valuable points of comparison.

Overall, three key insights emerged about interaction with respect to group size. First, most users in larger groups had difficulty with the time constraint in each round. As one participant mentioned: "Three [people per group] was really comfortable, [because] we each got time to share. Whereas in four, it was a little too big" (P9, S7). This is intuitive since the larger the group, the less time for each person to share. Second, students brought up the compatibility of different personality traits. Some participants observed that the more introverted group members can feel intimidated in larger groups:

> I would have preferred a three-person group over a four-person, because if there were more people and one of them is shy, they don't tend to speak as much. (P10, S7)

Third, we observed a general consensus in favor of smaller groups, because it helped people engage in conversations with others. One student liked groups of three because "you can hear everyone and everyone can be involved in the conversation" (P2, S4). Overall, based on the pacing of interactions, individual traits, and level of engagement, participants seem to prefer three-person groups.

5.2.3 Smaller groups rated their members higher. To explore whether group size affected ratings, we plotted the mean assessment ratings against the group size for all sessions. Figure 6 shows an inverse relationship between group size and the assessment ratings. Groups of six participants resulted in the lowest mean ratings (2.6), groups of five, four, and three had mean ratings of 3.0, 3.5, and 3.7, respectively, and two-member groups had the highest mean ratings (4.0). A one-way analysis of variance (ANOVA) shows a significant effect of group size on the mean assessment ratings ($F(1,1853) = 182.5$, $p < 0.001$). We ran post-hoc, paired t-test comparisons with Bonferroni corrections to test each pairwise comparison. As seen on Figure 6, the pairwise comparisons between groups of two and three and between groups of four and five were statistically significant. Overall, this result aligns with the intuition that smaller groups have more time to express opinions and to get to know each other better given a time constraint.

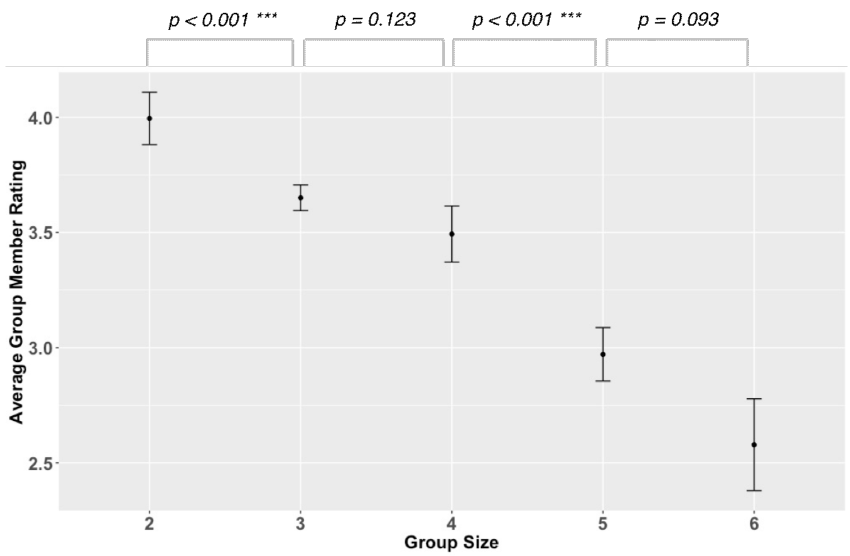

Fig. 6. The mean assessment ratings (ranging from 1 to 5) for each group member (with confidence intervals) by participants of different sized groups. At the top, the p-values for pair-wise comparisons show significant differences group member ratings by group size (*** = p < 0.001)

5.3 (How) do interactions during team dating influence teammate selection?

To determine if interactions in ProtoTeams affected how people form long-term teams, we performed an in-depth analysis on two deployment sessions, S1 and S9, where students were given a chance to form project teams immediately following the ProtoTeams session. In all other sessions, teams formed days or weeks later, and so we excluded these session since we could not account for the additional external factors (e.g., meeting peers outside of class or other group activities). S1 happened on the first day of the quarter, while S9 happened mid-way through the quarter.

5.3.1 Participants used team dating insights when forming groups. Participants reported how ProtoTeams played a role in team formation decisions by exposing them to character traits:

> [After the ProtoTeams session] myself and another one of my classmates were approached by a couple of students because we were very vocal and the students were like, I want to be on their team because they're going to get a good grade. (P14, S9)

These participants observed some desired personal traits in others, such as "being vocal". Not only did ProtoTeams help peers form these impressions, it warmed students up for further interaction, allowing them to approach each other after team dating and make more informed teaming decisions.

5.3.2 Team-dating interactions provided most value for audiences with fewer initial social ties. To quantify whether meeting someone on a team date improves the likelihood of forming teams together, we performed a dyadic regression analysis [53, 54]. Before each session, students filled out a pre-study survey with demographic information (e.g., age, gender, and nationality) and ratings on existing social ties to peers in the class. As described above, the social tie question asked, *"How well do you know this person?"* on a 5-point Likert scale for all other students in class. We also tracked whether or not the dyads interacted with each other during the ProtoTeams sessions. We excluded 13.7% of assessment scores because students did not assess their group members, which sometimes occurred when the instructors advanced to the next round before students had a chance to input ratings. Finally, a binary variable coded whether any two participants ended up on a project team together.

We compiled data for all possible dyads, or pairs of students, where we had complete information (filled out the pre-survey and participated in the ProtoTeams deployment). Dyads were only counted once (i.e., person A and person B were not double counted as dyad AB and dyad BA). In S1, 45 students filled out the survey and participated in ProtoTeams, resulting in 990 dyads. In S9, 53 students filled out the survey and participated in ProtoTeams for a total of 1378 dyads.

	S1	**S9**
Descriptive measures	Counts	Counts
Total possible dyads	990	1378
Dyads that interacted in ProtoTeams	202	339
Dyads with existing social ties	36	190
Regression variables	Mean (SD)	Mean (SD)
ProtoTeams Interaction (binary indicator)	0.20 (0.40)	0.18 (0.38)
Social Tie (mean pairwise rating)	1.03 (0.18)	1.12 (0.49)
Age Difference (in years)	2.82 (2.82)	2.90 (3.69)
Same Gender (binary indicator)	0.46 (0.50)	0.51 (0.50)
Same Nationality (binary indicator)	0.33 (0.47)	0.28 (0.45)

Table 3. Differences between sessions S1 and S9. Top: Raw counts for the total number of dyads, the dyads that interacted in ProtoTeams, and the dyads that had existing social ties. Bottom: Mean values (and st. dev.) for key variables in the regression analysis. Note that participants in S9 reported more prior social ties compared to S1.

Table 3 shows the mean and standard deviation for key variables in both sessions in this analysis. Building on existing research on factors that affect team formation [30], we created metrics for *Social Tie*, *Age Difference*, *Same Gender*, and *Same Nationality* which we collected in the pre-study survey. Social Tie refers to the mean pair-wise ratings dyads gave each other during the pre-study survey. Age Difference encodes the absolute difference in age between the two members. Gender and Nationality are binary indicators of whether members of a dyad reported the same gender and same nationality, respectively. For these binary indicators, a 0.5 mean suggests about half the dyads have the same gender/nationality. ProtoTeams Interaction is also a binary indicator of whether or not members of the dyad were grouped together during ProtoTeams activities. We use the binary indicator rather than the mean pair-wise ratings because about 70% of participants gave their group members the same rating when assessing them.

Of the 990 possible dyads in S1, 202 dyads (20%) interacted through ProtoTeams and 36 dyads (4%) reported existing social ties. In contrast, out of the 1378 possible dyads in the S9 session, 339 dyads (24%) interacted through ProtoTeams and 190 dyads (14%) had existing social ties. Since the S9 session took place halfway through the term, these students likely had more time to create social ties with each other. In both sessions, about a third of existing social ties also interacted with each other in ProtoTeams.

Looking at the final composition of teams, we see that 24% of S1 dyads compared to 15% of S9 teams met for the first time during the ProtoTeams session. ProtoTeams appears to factor into how teams form, especially in the session that had fewer existing social ties. To explore this, we ran a

	S1	S9
	Estimates (SE)	Estimates (SE)
ProtoTeams Interaction	0.61 (0.25)*	0.45 (0.15)
Social Tie	2.57 (0.53)***	1.20 (0.15)***
Age Difference	-0.13 (0.05)*	0.01 (0.03)
Same Gender	0.67 (0.24)**	0.34 (0.23)
Same Nationality	0.28 (0.24)	0.44 (0.25)

Table 4. Results from logistic regressions on dyad-level data from S1 and S9, where the dependent variable is whether or not a dyad ends up on the same team. The "Estimates" reflect the model coefficients for each variable in the logistic regression (with standard error). Only S1 shows that team dating significantly factored into team selection. A negative estimate value indicates an inverse relationship. (* = $p < 0.05$, ** $p < 0.01$, *** = $p < 0.001$.)

logistic regression using the dyad features in Table 3 as the independent variables. Whether or not a dyad ended up in a final project team together served as the dependent variable. Analysis of the correlation coefficients for the dyad features found no evidence of multicollinearity, meaning the variables were not linearly related. We ran a separate logistic regressions for S1 and S9 since the two deployments had different participants and were run at different times in the school year.

The logistic regression analysis for deployment S1, shown in Table 4, suggests that Social Tie, ProtoTeams Interaction, Same Gender, and Age Similarity all significantly factor into whether dyads end up on the same team. Among the variables included in the model, Social Tie showed the greatest effect size on teammate selection ($\beta = 2.57$, $z = 4.81$, $p < 0.001$). ProtoTeams Interaction also proved to be a significant variable ($\beta = 0.61$, $z = 2.35$, $p < 0.05$) which means the team dating interaction significantly affected teammate selection in S1. Additionally, Same Gender and Age Difference are also strong predictors of final teams ($\beta = 0.67$, $z = 2.78$, $p < 0.01$, and $\beta = -0.13$, $z = -2.42$, $p < 0.05$, respectively).

In contrast, the logistic regression for S9, seen in Table 4, shows that only Social Tie significantly predicted final teammate selection ($\beta = 1.20$, $z = 7.77$, $p < 0.001$). In both sessions, as found in prior work [30], students often gravitate towards people they already know when forming teams. However, we see in S1, where only 4% of the dyads report existing social ties, the team dating activity appears to provide students an opportunity form preliminary social ties that carry over into team selection.

6 DISCUSSION

ProtoTeams supports team dating activities by leveraging personal mobile devices to enable co-located audiences to quickly form into small groups, interact, and assess the potential for future collaboration. An evaluation with 406 students across 8 project-based classes shows how ProtoTeams supports the design goals of enabling rapid and diverse interactions, creating equitable and engaging discussions, and gathering timely and informative assessments. Participants gave ProtoTeams an average rating of 4.3 out of 5 in terms of how much they enjoyed their experience using the system. While high ratings could indicate a novelty effect, some prior work suggests that novel classroom systems can have lasting benefits [89].

A number of behaviors and reactions emerged that yielded key insights about formation, interaction, and assessment. The symbol-based group formation method was perceived as chaotic, yet enjoyable. They talked about how ProtoTeams randomly exposed them to peers they would not have met otherwise. Groups formed faster each round, as participants learned how to quickly find peers in a crowd. In terms of supporting interactions, participants appreciated seeing discussion prompts that probed at team behaviors. They saw the potential for ProtoTeams to go further to scaffold cooperative activities that truly simulate the experience of working together. In terms of gathering assessments, in nearly 70% of cases, participants did not discern (provide different ratings for) their group members, whether due to time constraints, forgetfulness, or the discomfort of performing assessments right in front of peers. While peer assessments are not necessary for the core team dating experience, accurate and discerning data would open up the possibility for recommending teams after team dating, or for more intentionally forming groups during team dating, rather than randomly.

Differences in each session allowed us to investigate the effects of group size. Smaller groups formed significantly faster than larger ones. This was not obvious *a priori*, as large clusters of peers, in theory, could have made them easier to find. Participants also preferred smaller groups because they support better discussions. Students reported they were less likely to share and get to know everyone in groups of four or larger. We observed this in the peer assessments where participants rated group members higher in smaller groups. While the general insight seems to be "smaller is better", groups of two have a key disadvantage: peers only get to know one person per round which means they get exposed to fewer peers in the population. Groups of *three* seems to be optimal in terms of forming groups quickly, allowing for many students to mix, supporting equitable interactive discussions, and being able to discern between peers for potential future collaboration.

Our dyad analysis of S1 and S9—the only two sessions where teams formed directly after the ProtoTeams deployment—shows that team-dating interactions can be a significant predictor of teammate selection, along with known factors, such as existing social ties and same gender [29, 30]. Team dating increases the chance that two participants end up on a team together, especially in session S1 where peers had fewer existing ties. This shows students value the exposure to new peers and are willing to commit to longer-term teams based only on a short interaction.

6.1 Limitations

The deployment study provided insights on emergent behaviors and reactions that formed during co-located team dating. A key limitation is that we do not know how ProtoTeams compares to a low-tech approach, like paper cards. Hosts could, for example, predetermine groups for each round and provide paper handouts with discussion prompts and assessment requests. While a paper-based approach might suffice, ProtoTeams still makes it easier to dynamically assign groups and to collect assessment data with no prior knowledge of the audience. Further research would be needed to validate ProtoTeams compared to other team-dating approaches, and to understand the utility of team dating compared to other methods for team formation, such as criteria-based approaches [56].

Our low-fidelity prototypes and deployments surfaced a number of socio-technical challenges that are common when deploying groupware [32, 33]. For example, students connected to the Internet using the private university network, which occasionally created bandwidth issues. If even just one participant must wait for pages to load, it could negatively impact the session for all participants. Likewise, individuals or groups may simply move at a slower pace than other groups. To mitigate pacing issues, ProtoTeams allows groups to advance to the interaction phase without waiting for other groups. The host can also manually advance the experience on all devices, which essentially allows slower individuals/groups to sync up with the rest of the audience. More research is needed to understand emergent socio-technical challenges that could arise in diverse settings and with different audiences.

Finally, this paper opportunistically explored research questions related to group size and teammate selection. We performed these analyses on a subset of the overall data and we did not tightly control for all potential confounding factors. For example, our regression model uncovers factors that predict two members will end up on the same team together, building on prior literature [53, 54]. However, teammate selection is not just a dyadic process [54]. Choosing to team-up with another person may not be not independent of decisions to team-up with other participants. There could also be any number of hidden structural patterns [39, 40], that also predict teammate selection. While our dyad analysis provides a crude first-pass indicator of significant factors, future work could explore more sophisticated network analysis methods [41, 44]. Furthermore, our analysis focused on the immediate affect on team formation, not team performance [85, 86]. Future work could focus on whether co-located team dating leads to better collaborations in the longer term.

6.2 Future Work

This paper offers a preliminary study of how technology can mediate team dating in co-located settings. Further studies are needed to carefully vary and measure the effects of group size, interaction length, frequency of interaction, activity types, and different formation techniques (i.e., random or deliberate) to uncover effective conditions for team dating.

Another key direction is to explore team dating in different contexts, such as hackathons, civic workshops, or orientation meetings where participants need to meet and interact with new people. This raises the issue of investigating the effects of different audience demographics—such as the number of existing social ties. Beyond these specific science directions, we offer a number of angles for future design and development.

6.2.1 Minimize friction to support on-the-spot participation. ProtoTeams supports co-located group activities with minimal preparation from hosts and participants. By leveraging personal mobile phones, classroom displays, and wireless networks, ProtoTeams side steps some issues around adopting new classroom technologies. We did not require students to obtain additional hardware, such as clickers [51], but we did have to loan out spare phones to about 1% of study participants. While the "digital divide" affects each community differently [52, 91], hosts should not assume that everyone has a device. Future work on technology-mediated team dating might explore how to assure everyone can participate, even with a shortage of devices.

Not only does ProtoTeams build on existing infrastructure, it requires no prior information information about participants. This means that hosts can launch ProtoTeams and create a rich social experience for an audience with little to no upfront planning. Participants do not need to fill out a survey ahead of time (students did fill out a pre-survey for this paper, but only for research purposes). While audience details are often available in classroom settings, we imagine a wide range of use cases where hosts might want to run ProtoTeams with no information about who or how many people are present in the audience.

Despite this ability to walk up and participate, we anticipate a number of issues that could hinder equitable participation. For example, in our pilot testing, we observed how participants, for one reason or another, would often arrive late or leave a session early. We designed flexibility into ProtoTeams by always showing the join URL on the shared display and allowing newcomers to jump in mid-activity. Similarly, when and if participants leave a session early, they can be removed from the session (either by hosts, other participants, or the individual can indicate their status) so they do not get assigned to groups in subsequent rounds.

Before using ProtoTeams, hosts need to consider the cultural practices and social norms of their audience. It may not be appropriate for all settings and audiences to ask strangers to mix with each other. Even when most audience members seem keen, there may be individual participants who experience socio-emotional stress, whether due to peer pressure [72], social anxiety [75], physical impairments [7], or any number of accessibly issues.

6.2.2 Simulate team cooperation scenarios. ProtoTeams quickly organizes participants into small groups and offers different types of discussion prompts (i.e., icebreaker, ideation, and team-building). While this was helpful for understanding other group members' personalities and preferences, our study participants pined for group activities that realistically simulate team conditions. Simple discussion prompts represent only the tip of the iceberg for the kinds of co-located group activities that could be supported through mobile devices. Imagine activities that ask participants to cooperatively solve a puzzle [93], make sense of data [67], improvise [25], or create an artifact [57]. In theory, these more constructive activities promote deeper learning [16, 78, 79] and lead to better insights on potential teammates [19].

Future work might explore designs for new kinds of cooperative activities that build on the affordances of mobile devices [23, 42, 82], while also balancing for time constraints. For the sake of both learning and forming teams, hosts will need to consider the trade-off between providing more time each round to foster deeper relationships and offering more (but shorter) rounds so that participants meet more people.

6.2.3 Provide real-time session data to support facilitation. Managing a large in-person audience can be daunting, especially when coordinating team dating. The audience moves around, and the host often has to shout verbal instructions so that participants can hear amidst the chatter. While ProtoTeams supports certain aspects of facilitation, we see many opportunities to provide better tools for hosts. During a live session, the facilitator interface could provide real-time data to inform decisions, such as when to advance the action for all groups, whether to intervene for problematic groups or individuals, or how to match people in subsequent rounds of team dating.

The interface could also provide post-session insights, such as understanding which prompts triggered the most fruitful discussion among participants, or assessing which groups effectively balance participation among members [1, 9]. Several participants mentioned the desire to take notes on group members after each round. Such information could help support team formation algorithms that try to optimize for sentiment, pair-wise ratings, skill balance, and personal characteristics. Since students prefer agency over team formation [45], this information could also be useful for self-forming teams where individuals need to recall their experiences during team dating.

7 CONCLUSION

Researchers have proposed *team dating*, a series of rapid small-group interactions to help individuals gather tacit knowledge about the potential of working with others in the future. This approach can be difficult to manage for large, physically co-located audiences. This paper presents ProtoTeams, a system that leverages personal mobile devices, and other existing infrastructure, to organize people to quickly form groups, engage in discussion, and assess peers as potential future teammates. We deployed ProtoTeams in eight different project-based classes with 406 participants to understand the dynamics of co-located team dating, including whether and how it influences teammate selection. We found that ProtoTeams enables small groups to form rapidly (60 seconds on average), and faster with each round of practice. Participants enjoyed the group forming aspect where they find others based on colored shapes and interacted with peers they would not have otherwise. Our analysis provides preliminary evidence that three-person groups are most effective for co-located team dating. We also found that co-located team dating can affect teammate selection for longer-term projects, especially for audiences where there are fewer existing social ties.

8 ACKNOWLEDGEMENTS

We thank the National Science Foundation for funding on grants #1821618 and #1122206. We thank Marlon Twyman II and Diego Gómez-Zará for their feedback on early drafts of this paper, as well as Eric Truong for help with early-stage design and prototyping. We also appreciate Amaya Mali, Enyu Alison Chen, and Justin Chen for helping us refine the system and website.

REFERENCES

[1] Hiroyuki Adachi, Seiko Myojin, and Nobutaka Shimada. 2015. ScoringTalk: A Tablet System Scoring and Visualizing Conversation for Balancing of Participation *(SA '15)*. Association for Computing Machinery, New York, NY, USA, Article 9, 5 pages. https://doi.org/10.1145/2818427.2818454

[2] I. Amarasinghe, D. Hernández-Leo, and A. Jonsson. 2017. Intelligent Group Formation in Computer Supported Collaborative Learning Scripts. In *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*. 201–203.

[3] Nalini Ambady and Robert Rosenthal. 1992. Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin* 111, 2 (1992), 256–274. https://doi.org/10.1037/0033-2909.111.2.256

[4] Nalini Ambady and Robert Rosenthal. 1993. Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology* 64, 3 (1993), 431–441. https://doi.org/10.1037/0022-3514.64.3.431

[5] Elliot Aronson and et al. 1978. *The jigsaw classroom.* Sage, Oxford, England.

[6] Nikolaos Avouris, Meletis Margaritis, and Vassilis Komis. 2004. The effect of group size in synchronous collaborative problem solving activities. In *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE), 4303–4308.

[7] Jayne R Beilke and Nina Yssel. 1999. The chilly climate for students with disabilities in higher education. *College Student Journal* 33, 3 (1999), 364–364.

[8] Roland Berger and Martin Hänze. 2015. Impact of Expert Teaching Quality on Novice Academic Performance in the Jigsaw Cooperative Learning Method. *International Journal of Science Education* 37, 2 (Jan. 2015), 294–320. https://doi.org/10.1080/09500693.2014.985757

[9] Tony Bergstrom and Karrie Karahalios. 2007. Seeing More: Visualizing Audio Cues. In *Human-Computer Interaction – INTERACT 2007*, Cécilia Baranauskas, Philippe Palanque, Julio Abascal, and Simone Diniz Junqueira Barbosa (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 29–42.

[10] Peter Borkenau, Nadine Mauer, Rainer Riemann, Frank M. Spinath, and Alois Angleitner. 2004. Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology* 86, 4 (April 2004), 599–614. https://doi.org/10.1037/0022-3514.86.4.599

[11] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[12] Spencer Carlson, Kristine Lu, Evey Huang, Elizabeth Gerber, and Matthew Easterday. 2020. Designing a Model for Deliberation-Based Learning. *The Interdisciplinarity of the Learning Sciences, 14th International Conference of the Learning Sciences (ICLS)* 3 (2020), 1553–1556.

[13] Avner Caspi, Paul Gorsky, and Eran Chajut. 2003. The influence of group size on nonmandatory asynchronous instructional discussion groups. *The Internet and Higher Education* 6, 3 (2003), 227–240.

[14] Joel Chan, Steven Dang, and Steven P Dow. 2016. IdeaGens: enabling expert facilitation of crowd brainstorming. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*. 13–16.

[15] Michelene T. H. Chi and Ruth Wylie. 2014. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist* 49, 4 (Oct. 2014), 219–243. https://doi.org/10.1080/00461520.2014.965823

[16] Elizabeth G Cohen. 1994. Restructuring the classroom: Conditions for productive small groups. *Review of educational research* 64, 1 (1994), 1–35.

[17] Noshir Contractor, Leslie Ann DeChurch, Anup Sawant, and Xiang Li. 2013. My dream team assembler. (2013).

[18] Marguerite Cronk. 2012. Using gamification to increase student engagement and participation in class discussion. In *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE), 311–315.

[19] Petru L. Curşeu, Patrick Kenis, Jörg Raab, and Ulrik Brandes. 2010. Composing Effective Teams through Team Dating. *Organization Studies* 31, 7 (2010), 873–894. https://doi.org/10.1177/0170840610373195

[20] Petru L Curşeu, Patrick Kenis, and Jörg Raab. 2009. Reciprocated relational preferences and intra-team conflict. *Team Performance Management: An International Journal* (2009).

[21] Elena del Val, Juan Miguel Alberola, Victor Sanchez-Anguix, Alberto Palomares, and Ma Dolores Teruel. 2014. A Team Formation Tool for Educational Environments. In *Trends in Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection*, Javier Bajo Perez, Juan M. Corchado Rodríguez, Philippe Mathieu, Andrew Campbell, Alfonso Ortega, Emmanuel Adam, Elena M. Navarro, Sebastian Ahrndt, María N. Moreno, and Vicente Julián (Eds.). Springer International Publishing, Cham, 173–181.

[22] Piet Van den Bossche, Wim H. Gijselaers, Mien Segers, and Paul A. Kirschner. 2006. Social and Cognitive Factors Driving Teamwork in Collaborative Learning Environments: Team Learning Beliefs and Behaviors. *Small Group Research* 37, 5 (2006), 490–521. https://doi.org/10.1177/1046496406292938 arXiv:https://doi.org/10.1177/1046496406292938

[23] Qing-Ke Fu and Gwo-Jen Hwang. 2018. Trends in mobile technology-supported collaborative learning: A systematic review of journal publications from 2007 to 2016. *Computers & Education* 119 (2018), 129–143.

[24] Jackbox Games. 2013. *You Don't Know Jack*. Retrieved August 19, 2020 from https://www.jackboxgames.com/

[25] Elizabeth Gerber. 2009. Using improvisation to enhance the effectiveness of brainstorming. *Conference on Human Factors in Computing Systems - Proceedings*, 97–104. https://doi.org/10.1145/1518701.1518718

[26] Eric Gilbert and Karrie Karahalios. 2009. Predicting Tie Strength with Social Media *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 211–220. https://doi.org/10.1145/1518701.1518736

[27] Diego Gómez-Zará, Leslie A Dechurch, and Noshir S Contractor. 2020. A Taxonomy of Team-Assembly Systems: Understanding How People Use Technologies to Form Teams. *Proceedings of the 2020 ACM Conference on Computer Supported Cooperative Work and Social Computing* (2020).

[28] Diego Gómez-Zará, Mengzi Guo, Leslie A. DeChurch, and Noshir Contractor. 2020. The Impact of Displaying Diversity Information on the Formation of Self-Assembling Teams *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376654

[29] Diego Gómez-Zará, Mengzi Guo, Leslie A. DeChurch, and Noshir Contractor. 2020. The Impact of Displaying Diversity Information on the Formation of Self-Assembling Teams *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376654

[30] Diego Gómez-Zará, Matthew Paras, Marlon Twyman, Jacqueline N. Lane, Leslie A. DeChurch, and Noshir S. Contractor. 2019. Who Would You Like to Work With? *(CHI '19)*. ACM, New York, NY, USA, Article 659, 15 pages. https://doi.org/10.1145/3290605.3300889

[31] Sabine Graf and Rahel Bekele. 2006. Forming Heterogeneous Groups for Intelligent Collaborative Learning Systems with Ant Colony Optimization. In *Intelligent Tutoring Systems*, Mitsuru Ikeda, Kevin D. Ashley, and Tak-Wai Chan (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 217–226.

[32] Catherine Grevet and Eric Gilbert. 2015. Piggyback Prototyping: Using Existing, Large-Scale Social Computing Systems to Prototype New Ones *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 4047–4056. https://doi.org/10.1145/2702123.2702395

[33] Jonathan Grudin. 1994. Groupware and Social Dynamics: Eight Challenges for Developers. *Commun. ACM* 37, 1 (Jan. 1994), 92–105. https://doi.org/10.1145/175222.175230

[34] Emily Harburg, Daniel Rees Lewis, Matthew Easterday, and Elizabeth M. Gerber. 2018. CheerOn: Facilitating Online Social Support for Novice Project-Based Learning Teams. *ACM Trans. Comput.-Hum. Interact.* 25, 6, Article 32 (Dec. 2018), 46 pages. https://doi.org/10.1145/3241043

[35] Alexa M. Harris, Diego Gómez-Zará, Leslie A. DeChurch, and Noshir S. Contractor. 2019. Joining Together Online: The Trajectory of CSCW Scholarship on Group Formation. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article Article 148 (Nov. 2019), 27 pages. https://doi.org/10.1145/3359250

[36] Emily M. Hastings, Albatool Alamri, Andrew Kuznetsov, Christine Pisarczyk, Karrie Karahalios, Darko Marinov, and Brian P. Bailey. 2020. LIFT: Integrating Stakeholder Voices into Algorithmic Team Formation *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376797

[37] Emily M. Hastings, Farnaz Jahanbakhsh, Karrie Karahalios, Darko Marinov, and Brian P. Bailey. 2018. Structure or Nurture?: The Effects of Team-Building Activities and Team Composition on Team Outcomes. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 68 (Nov. 2018), 21 pages. https://doi.org/10.1145/3274337

[38] Gayle W. Hill. 1982. Group versus individual performance: Are N+1 heads better than one? *Psychological Bulletin* 91, 3 (1982), 517–539. https://doi.org/10.1037/0033-2909.91.3.517

[39] Paul W Holland and Samuel Leinhardt. 1971. Transitivity in structural models of small groups. *Comparative group studies* 2, 2 (1971), 107–124.

[40] Paul W Holland and Samuel Leinhardt. 1977. A method for detecting structure in sociometric data. In *Social Networks*. Elsevier, 411–432.

[41] Paul W. Holland and Samuel Leinhardt. 1981. An Exponential Family of Probability Distributions for Directed Graphs. *J. Amer. Statist. Assoc.* 76, 373 (1981), 33–50. http://www.jstor.org/stable/2287037

[42] Yu-Jen Hsu and Ju-Ling Shih. 2013. Developing computer adventure education games on mobile devices for conducting cooperative problem-solving activities. *International Journal of Mobile Learning and Organisation 4* 7, 2 (2013), 81–98.

[43] Chia En Hsueh, Hsinju Lee, Yu-Han Lu, Nien-Hsin Wu, Yu-Wei Wu, Shao En Lin, and Pei-Yi Kuo. 2020. Exploring the Effect of Group Size on Goal Setting Sharing to Reduce Procrastination: Lessons Learned in a Field Study *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3383020

[44] Rustom Ichhaporia, Diego Gómez-Zará, Leslie DeChurch, and Noshir Contractor. 2020. A Network Approach to the Formation of Self-assembled Teams. In *Complex Networks and Their Applications VIII*, Hocine Cherifi, Sabrina Gaito, José Fernendo Mendes, Esteban Moro, and Luis Mateus Rocha (Eds.). Springer International Publishing, Cham, 969–981.

[45] Farnaz Jahanbakhsh, Wai-Tat Fu, Karrie Karahalios, Darko Marinov, and Brian Bailey. 2017. You Want Me to Work with Who?: Stakeholder Perceptions of Automated Team Formation in Project-based Courses *(CHI '17)*. ACM, New York, NY, USA, 3201–3212. https://doi.org/10.1145/3025453.3026011

[46] David S Jalajas and Robert I Sutton. 1984. Feuds in student groups: Coping with whiners, martyrs, saboteurs, bullies, and deadbeats. *Organizational Behavior Teaching Review* 9, 4 (1984), 94–102.

[47] David W. Johnson and Roger T. Johnson. 1989. *Cooperation and competition: Theory and research.* Interaction Book Company, Edina, MN, US.

[48] David W Johnson, Roger T Johnson, and Karl A Smith. 1998. Cooperative learning returns to college what evidence is there that it works? *Change: the magazine of higher learning* 30, 4 (1998), 26–35.

[49] Don W. Jordan and Joanna Le Métais. 1997. Social skilling through cooperative learning. *Educational Research* 39, 1 (1997), 3–21. https://doi.org/10.1080/0013188970390101 arXiv:https://doi.org/10.1080/0013188970390101

[50] Steven J Karau and Kipling D Williams. 1993. Social loafing: A meta-analytic review and theoretical integration. *Journal of personality and social psychology* 65, 4 (1993), 681.

[51] Robin H Kay and Ann LeSage. 2009. Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education* 53, 3 (2009), 819–827.

[52] Jared Keengwe and Malini Bhargava. 2014. Mobile learning and integration of mobile technologies in education. *Education and Information Technologies* 19, 4 (2014), 737–746.

[53] Adam M. Kleinbaum. 2018. Reorganization and Tie Decay Choices. *Management Science* 64, 5 (2018), 2219–2237. https://doi.org/10.1287/mnsc.2016.2705 arXiv:https://doi.org/10.1287/mnsc.2016.2705

[54] Adam M. Kleinbaum, Alexander H. Jordan, and Pino G. Audia. 2015. An Altercentric Perspective on the Origins of Brokerage in Social Networks: How Perceived Empathy Moderates the Self-Monitoring Effect. *Organization Science* 26, 4 (2015), 1226–1242. https://doi.org/10.1287/orsc.2014.0961 arXiv:https://doi.org/10.1287/orsc.2014.0961

[55] Chinmay E. Kulkarni, Michael S. Bernstein, and Scott R. Klemmer. 2015. PeerStudio: Rapid Peer Feedback Emphasizes Revision and Improves Performance *(L@S '15)*. Association for Computing Machinery, New York, NY, USA, 75–84. https://doi.org/10.1145/2724660.2724670

[56] Richard A. Layton, Misty L. Loughry, Matthew W. Ohland, and George D. Ricco. 2010. Design and Validation of a Web-Based System for Assigning Members to Teams Using Instructor-Specified Criteria. *Advances in Engineering Education* 2, 1 (2010). https://eric.ed.gov/?id=EJ1076132

[57] Henri Lipmanowicz and Keith McCandless. 2013. *The surprising power of liberating structures: Simple rules to unleash a culture of innovation.* Liberating Structures Press Seattle, WA.

[58] Andrew Luxton-Reilly. 2009. A systematic review of tools that support peer assessment. *Computer Science Education* 19, 4 (2009), 209–232.

[59] Ioanna Lykourentzou, Angeliki Antoniou, Yannick Naudet, and Steven P. Dow. 2016. Personality Matters: Balancing for Personality Types Leads to Better Outcomes for Crowd Teams *(CSCW '16)*. Association for Computing Machinery, New York, NY, USA, 260–273. https://doi.org/10.1145/2818048.2819979

[60] Ioanna Lykourentzou, Robert E. Kraut, and Steven P. Dow. 2017. Team Dating Leads to Better Online Ad Hoc Collaborations *(CSCW '17)*. ACM, New York, NY, USA, 2330–2343. https://doi.org/10.1145/2998181.2998322

[61] Ioanna Lykourentzou, Shannon Wang, Robert E. Kraut, and Steven P. Dow. 2016. Team Dating: A Self-Organized Team Formation Strategy for Collaborative Crowdsourcing *(CHI EA '16)*. ACM, New York, NY, USA, 1243–1249. https://doi.org/10.1145/2851581.2892421

[62] D. Maria-Iuliana, B. Constanta-Nicoleta, and B. Alexandru. 2013. Platform for Creating Collaborative E-Learning Communities Based on Automated Composition of Learning Groups. In *2013 3rd Eastern European Regional Conference on the Engineering of Computer Based Systems*. 103–112. https://doi.org/10.1109/ECBS-EERC.2013.21

[63] D. Maria-Iuliana, B. Constanta-Nicoleta, and B. Alexandru. 2013. Platform for Creating Collaborative E-Learning Communities Based on Automated Composition of Learning Groups. In *2013 3rd Eastern European Regional Conference on the Engineering of Computer Based Systems (ECBS-EERC)*. IEEE Computer Society, Los Alamitos, CA, USA, 103–112. https://doi.org/10.1109/ECBS-EERC.2013.21

[64] Peter V. Marsden and Karen E. Campbell. 1984. Measuring Tie Strength. *Social Forces* 63, 2 (1984), 482–501. http://www.jstor.org/stable/2579058

[65] Paula McAvoy and Diana Hess. 2013. Classroom Deliberation in an Era of Political Polarization. *Curriculum Inquiry* 43, 1 (Jan. 2013), 14–47. https://doi.org/10.1111/curi.12000

[66] Richard L Moreland, JM Levine, and ML Wingert. 2013. Creating the ideal group: Composition effects at work. *Understanding group behavior* 2 (2013), 11–35.

[67] Meredith Ringel Morris, Jarrod Lombardo, and Daniel Wigdor. 2010. WeSearch: Supporting Collaborative Search and Sensemaking on a Tabletop Display *(CSCW '10)*. Association for Computing Machinery, New York, NY, USA, 401–410. https://doi.org/10.1145/1718918.1718987

[68] Catherine Neubauer, Joshua Woolley, Peter Khooshabeh, and Stefan Scherer. 2016. Getting to Know You: A Multimodal Investigation of Team Behavior and Resilience to Stress *(ICMI '16)*. Association for Computing Machinery, New York, NY, USA, 193–200. https://doi.org/10.1145/2993148.2993195

[69] Adrian C North, P Alex Linley, and David J Hargreaves. 2000. Social loafing in a co-operative classroom task. *Educational Psychology* 20, 4 (2000), 389–392.

[70] Chinasa Odo, Judith Masthoff, and Nigel Beacham. 2019. Group Formation for Collaborative Learning. In *Artificial Intelligence in Education*, Seiji Isotani, Eva Millán, Amy Ogan, Peter Hastings, Bruce McLaren, and Rose Luckin (Eds.). Springer International Publishing, Cham, 206–212.

[71] Lora Oehlberg, Jasmine Jones, Alice Agogino, and Björn Hartmann. 2012. Dazzle: Supporting Framing in Co-located Design Teams Through Remote Collaboration Tool *(CSCW '12)*. ACM, New York, NY, USA, 183–186. https://doi.org/10.1145/2141512.2141573

[72] Annelies Raes, Ellen Vanderhoven, and Tammy Schellens. 2015. Increasing anonymity in peer assessment by using classroom response technology within face-to-face higher education. *Studies in Higher Education* 40, 1 (2015), 178–193.

[73] Mark Reilly and Haifeng Shen. 2011. Shared note-taking: a smartphone-based approach to increased student engagement in lectures. In *The 11th International Workshop on Collaborative Editing Systems in Conjunction with ACM Conference on Computer Supported Cooperative Work*.

[74] Amy Roehl, Shweta Linga Reddy, and Gayla Jett Shannon. 2013. The Flipped Classroom: An Opportunity to Engage Millennial Students through Active Learning Strategies. *Journal of Family and Consumer Sciences* 105, 2 (2013), 44–49.

[75] Graham Russell and Phil Topham. 2012. The impact of social anxiety on student learning and well-being in higher education. *Journal of Mental Health* 21, 4 (2012), 375–385. https://doi.org/10.3109/09638237.2012.694505 arXiv:https://doi.org/10.3109/09638237.2012.694505 PMID: 22823093.

[76] Niloufar Salehi and Michael S. Bernstein. 2018. Hive: Collective Design Through Network Rotation. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 151 (Nov. 2018), 26 pages. https://doi.org/10.1145/3274420

[77] Amy Shannon, Jessica Hammer, Hassler Thurston, Natalie Diehl, and Steven Dow. 2016. PeerPresents: A Web-Based System for In-Class Peer Feedback during Student Presentations *(DIS '16)*. Association for Computing Machinery, New York, NY, USA, 447–458. https://doi.org/10.1145/2901790.2901816

[78] Shlomo Sharan. 1980. Cooperative Learning in Small Groups: Recent Methods and Effects on Achievement, Attitudes, and Ethnic Relations. *Review of Educational Research* 50, 2 (1980), 241–271. https://doi.org/10.3102/00346543050002241 arXiv:https://doi.org/10.3102/00346543050002241

[79] Leonard Springer, Mary Elizabeth Stanne, and Samuel S Donovan. 1999. Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of educational research* 69, 1 (1999), 21–51.

[80] Robert J. Stevens and Robert E. Slavin. 1995. The Cooperative Elementary School: Effects on Students' Achievement, Attitudes, and Social Relations. *American Educational Research Journal* 32, 2 (1995), 321–351. https://doi.org/10.3102/00028312032002321 arXiv:https://doi.org/10.3102/00028312032002321

[81] Susan G Straus and Joseph E McGrath. 1994. Does the medium matter? The interaction of task type and technology on group performance and member reactions. *Journal of applied psychology* 79, 1 (1994), 87.

[82] Siobhán Thomas. 2006. Pervasive learning games: Explorations of hybrid educational gamescapes. *Simulation & Gaming* 37, 1 (2006), 41–55. https://doi.org/10.1177/1046878105282274 arXiv:https://doi.org/10.1177/1046878105282274

[83] David Tinapple, Loren Olson, and John Sadauskas. 2013. CritViz: Web-based software supporting peer critique in large creative classrooms. *Bulletin of the IEEE Technical Committee on Learning Technology* 15, 1 (2013), 29.

[84] Harry C Triandis, Eleanor R Hall, and Robert B Ewen. 1965. Member heterogeneity and dyadic creativity. *Human relations* 18, 1 (1965), 33–55.

[85] Bruce W Tuckman. 1965. Developmental sequence in small groups. *Psychological bulletin* 63, 6 (1965), 384.

[86] Bruce W Tuckman and Mary Ann C Jensen. 1977. Stages of small-group development revisited. *Group & Organization Studies* 2, 4 (1977), 419–427.

[87] Gustavo Umbelino, Vivian Ta, Samuel Blake, Eric Truong, Amy Luo, and Steven Dow. 2019. ProtoTeams: Supporting Small Group Interactions in Co-Located Crowds *(CSCW '19)*. Association for Computing Machinery, New York, NY, USA, 392–397. https://doi.org/10.1145/3311957.3359505

[88] Nanine AE Van Gennip, Mien SR Segers, and Harm H Tillema. 2010. Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions. *Learning and Instruction* 20, 4 (2010), 280–290.
[89] Alf Inge Wang. 2015. The wear out effect of a game-based student response system. *Computers & Education* 82 (2015), 217–227.
[90] Tsungjuang Wang. 2010. A New Paradigm for Design Studio Education. *International Journal of Art & Design Education* 29, 2 (2010), 173–183. https://doi.org/10.1111/j.1476-8070.2010.01647.x
[91] Mark Warschauer. 2004. *Technology and social inclusion: Rethinking the digital divide.* MIT press.
[92] Noreen M Webb. 1982. Group composition, group interaction, and achievement in cooperative small groups. *Journal of Educational Psychology* 74, 4 (1982), 475.
[93] Patrick Williams. 2018. Using escape room-like puzzles to teach undergraduate students effective and efficient group process skills. In *2018 IEEE Integrated STEM Education Conference (ISEC).* 254–257.
[94] Zehui Zhan, Patrick S. W. Fong, Hu Mei, and Ting Liang. 2015. Effects of gender grouping on students' group performance, individual achievements and attitudes in computer-supported collaborative learning. *Computers in Human Behavior* 48 (July 2015), 587–596. https://doi.org/10.1016/j.chb.2015.02.038

Received January 2020; revised June 2020; accepted July 2020