Joining Together Online: The Trajectory of CSCW Scholarship on Group Formation

ALEXA M. HARRIS*, Northwestern University, USA
DIEGO GÓMEZ-ZARÁ*†, Northwestern University, USA
LESLIE A. DECHURCH, Northwestern University, USA
NOSHIR S. CONTRACTOR, Northwestern University, USA

The field of Computer Supported Cooperative Work (CSCW) has an enduring interest in studying and designing technologies that bring people together in partnerships, teams, crowds, communities, and other collectives. As the technologies enabling group formation have evolved, so too have the guiding questions pursued by CSCW scholars. This review outlines the trajectory of scholarship on group formation with an eye towards the most pressing future questions in this area. To understand how CSCW researchers have studied technology-enabled group formation, we systematically review articles published at CSCW from 1992 to 2018. Exploring more than 2,000 potentially relevant works, we identified 35 focused on technologies and group formation. Content coding and thematic analysis revealed four periods and six themes in the study of online group formation. These themes include: group composition, self-presentation, assembly mechanisms, recruitment, organizing structures, and group culture. *Quo vadis?* Based on our review, we offer recommendations for the next generation of CSCW scholarship seeking to understand and enable collectives joining together online.

CCS Concepts: • **Human-centered computing → Collaborative and social computing theory, concepts and paradigms**.

Additional Key Words and Phrases: group formation; group assembly; groups; teams; crowds; communities; CSCW models; systematic literature review

ACM Reference Format:
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 148 (November 2019), 27 pages. https://doi.org/10.1145/3359250

1 INTRODUCTION

Research on how groups emerge through collaborative technologies has been evolving in CSCW over the past 25 years. Since the creation of CSCW in 1984, scholars have been investigating how technologies affect groups of different scales: from small bounded teams sharing an office in the 1990s, to modern-day crowds who contribute to the creation of a collective product while working

*Both authors contributed equally to this research.
†Also with Pontificia Universidad Católica de Chile, Facultad de Comunicaciones.

Authors' addresses: Alexa M. Harris, alexaharris2021@u.northwestern.edu, Northwestern University, School of Communication, Evanston, USA; Diego Gómez-Zará, dgomezara@u.northwestern.edu, Northwestern University, McCormick School of Engineering and School of Communication, Evanston, USA; Leslie A. DeChurch, dechurch@northwestern.edu, Northwestern University, School of Communication, Evanston, USA; Noshir S. Contractor, nosh@northwestern.edu, McCormick School of Engineering, Kellogg School of Management, and the School of Communication, Northwestern University, Evanston, USA.

2573-0142/2019/11-ART148 $15.00
https://doi.org/10.1145/3359250

from across the globe. Online social platforms enabled the emergence of online communities in the 2000s, and ultimately, large-scale collaborative networks of crowds ten years later. Historically, CSCW scholars have drawn on multiple disciplines to understand the nature of work performed by technologically-enabled groups [114].

One central aspect of groups is their *formation*, which refers to the mechanisms and processes explaining how and why groups form. As an example, understanding how users search for and evaluate potential partners to work with, or which online communities to join, are questions of group formation [33, 124]. Whether groups involve relatively small well-bounded teams or large-scale crowd worker communities, there is a basic need to understand how technologies and organizing intersect in the formation of groups. The expression, "joining together online," has evolved considerably since 1984, brought on by rapid developments in computing. While there is a vibrant body of CSCW literature reviews and conceptual models that focus on groups interacting with technologies [68, 93, 113, 114, 119], studies on group formation and the role of technologies have not been synthesized comprehensively. That is, what aspects of group formation have been examined by CSCW scholars over the past 25 years? Given the importance of the topic of group formation and the rapid evolution of technologies, a comprehensive review of CSCW scholarship on group formation is needed. Such a review synthesizes current work aimed at understanding how group formation is enabled by collaborative technologies and traces the trajectory forward [48, 63, 109].

In this paper, we systematically review the CSCW literature on group formation to take stock of what we know and identify opportunities for future research. Our systematic review of the ACM conference proceedings of Computer Supported Cooperative Work (CSCW) spans 1992-2018. We screened more than 2,000 articles returned in response to search queries using the terms *group*, *team*, *assembly*, and *formation*. After closely reviewing more than 150 articles, we identified 35 focusing on technologies, mechanisms, processes, and user behaviors associated with group formation. We first classified the nature of this research, and then identified four periods in the nature of CSCW scholarship on group formation. Lastly, we conducted a thematic analysis of these papers, identifying six core themes in the study of group formation.

The kinds of questions posed about group formation have evolved over the past 25 years. In the 1990s, CSCW articles on group formation focused on how small intact groups could be connected. These papers focused mostly on technology and less on social processes. This technological focus gave way to online groups and communities in the mid to late 2000s, and then on to crowds around 2013. The next shift studied group formation in crowd-sourced groups, and the consequences of how groups form online for their performance. Our thematic analysis reveals six enduring themes about group formation: (1) how the composition of the group affects group outcomes; (2) how users and groups self-present on collaborative platforms; (3) group recruitment, or the factors that influence individuals' decision to join pre-existing groups or to start a new group; (4) the assembly mechanisms by which groups are created; (5) the organizing structures in which groups form and the ways existing structures affect the formation of groups; and (6) the culture of groups.

Based on our results and inspired by previous CSCW models, we map the different contexts in which groups form, exemplary supporting technologies in each context, and an intersecting dimension characterizing a core distinction in the group formation literature: personal agency. CSCW work on group formation differs in the conceptualization of personal agency. For example, whereas some group formation research explores technologies designed to increase personal agency by allowing individuals to form groups organically with wider sets of prospective members, other papers explore technologies that reduce agency, deploying algorithms to assign membership. Taking stock of where we have been, and where we are, is useful in suggesting where we should go next. For instance, despite the increasing number of technologies bringing people together, CSCW

scholarship has placed more emphasis on describing what happens *after* groups form online, with less attention paid to how, why, and when they form.

This review is organized as follows. The next section provides a definition of group formation and a historical overview of the major lines of inquiry about groups posed by CSCW scholars. We next explain the shifts in research over time and themes explored by these works. Finally, we build on these shifts and themes to encourage the next wave of CSCW scholarship to consider how, why, and when groups form, as well as the outcomes associated with group formation processes.

2 THEORETICAL BACKGROUND

The topic of group formation has been an important one in multiple adjacent fields. For example, scholars in organizational behavior, social and organizational psychology, and communication have studied the human tendencies that shape and constrain group formation. Group formation research within CSCW explores human tendencies to form groups as they are intertwined with technology. As the field has matured, the scope of group formation research has expanded. Whereas the focus was initially on technologies that support existing teams, it has evolved into technologies that assemble groups online, that could not otherwise exist. We begin by reviewing existing definitions of group formation and perspectives on groups at CSCW.

2.1 Defining Group Formation

Group formation describes the process of identifying prospective members and bringing them together to collaborate [38]. Informed by the group's purpose, location, network structure, hierarchy, norms, and level of identity [108], formation mechanisms allow us to understand who works together, how well they work together, and how productive they are [71, 116]. Group formation mechanisms are affected by technologies that channel how potential members can be searched, recognized, selected, invited, or otherwise incorporated into a group. Furthermore, groups often form differently depending on their goals. For example, a team could search for specific members with the skills needed to solve a particular problem [6], or an online community might invite users who share similar characteristics to join [95]. Research on group formation varies in the degree to which technologies enable or bypass personal agency. Some research studies the ways technologies enable groups to form themselves; others study the ways groups can be formed top-down by a manager or other gate-keeper, and still others, how they can be formed by algorithmic assignment [10]. Furthermore, group formation depends on the information available about the members and their objectives [115]. Groups can be configured to achieve the right combination of skills, demographic attributes, knowledge, and social networks.

In studying group formation, scholars have used related terms, such as *assembly*, *configuration*, *design*, and *development* as descriptors [10]. These terms are often associated with group *composition*, a related topic, concerned with how the combination of members' characteristics and attributes affect group processes and outcomes [15, 70]. Group formation differs from group composition in that the former describes the organizing processes through which groups come together, whereas the latter describes the attributes of group members [117]. Group formation often sets in motion the composition of online groups.

Group formation has been examined from multiple disciplinary perspectives. Psychologists have long been interested in how groups socialize newcomers to their goals, norms, and status hierarchies. For example, Tuckman's [107] oft-cited linear progression model characterizes the forming, storming, norming, and performing stages groups follow in order to regularize member interactions. Moreland [82] elaborates group formation as the process of "social integration," defined by creating or strengthening the bonds among people. Kozlowski and Bell describe the critical role of group formation in establishing a meaningful boundary [60]. A group does not gain legitimacy

and meaning to its members until all members have accepted their respective roles. Sociologists study the "social processes" involved in individual and organizational decision-making behaviors [10]. Organizational scholars [43] have linked individuals' basic need for predictability to their choices of whom to work with.

From a computational perspective, scholars explore how the social-technological infrastructures influence user decisions and experiences as they assemble into groups. For example, a human-computer interaction article analyzed how different group formation criteria affected users' expectations [46]. *Flash teams* and *flash organizations* are two examples of grouping technologies that take advantage of expert users and consider the organizational hierarchical structures required to coordinate projects [96, 110]. Another work explored how computational systems can automatize group formation processes by considering the users, tasks, resources, and requirements needed to assemble groups [5]. More sophisticated techniques leverage machine learning and computational augmentation to shift group norms, hierarchies, composition, and member assessment [111, 122]. Furthermore, *team formation systems* have gained ground by considering each individuals' skills, or human capital, and also their social capital which consists of previous relationships and previous experiences working together [4, 33]. A group assembly approach considers individuals as component parts, and the challenge, to find the best combination of parts needed to optimize group output. This mechanistic approach is efficient, but fails to incorporate the role of motivational and social processes [115].

2.2 Research of Groups in CSCW

CSCW was conceived in 1984 as an "effort by technologists to learn from economists, social psychologists, anthropologists, organizational theorists, educators, and anyone else who could shed light on group activity" [35, p. 19]. Personal computers and workstations started showing up in work environments and eventually became ubiquitous. The widespread adoption of computer technologies raised new opportunities and challenges for the workforce. To this end, CSCW, as a community, has aimed to design and understand collaborative technologies for supporting groups. The major premise underlying the study and development of these technologies was the coordination of activities and people across time and space [47], and, at the same time, the goal of ensuring that prospective users would have an appropriate understanding of their use as a collective tool [90].

It was not until after some CSCW conferences were held that scholars started questioning the differences between analyzing technological components, such as the devices, interfaces, and ontologies used, and the necessary elements for supporting cooperative work through computational systems [30]. In 1994, Grudin introduced a taxonomy for linking work-oriented technologies with user populations [35]. Using a concentric model comprised of four rings, Grudin places the individual at the center and, moving from the inner ring to the outermost ring, increases in scale and organizational structure of small groups, projects, and, finally, to organizations. In doing so, this model distinguishes CSCW and related disciplines, such as human-computer interaction and information studies. This conceptualization not only enabled a better understanding of technologies for small groups, projects or organizations, but it also invited CSCW scholars to consider a broader range of coordinated activities among systems with a varying number of users and organizational structures.

Despite interdisciplinary efforts, early empirical work examining technology and collaboration in groups was criticized as being "meager, and, unsystematic in coverage of all of the relevant factors and their interplay" [78]. Most of the CSCW research targeted group processes from a technological vantage point rather than comprehending the social and learning aspects behind these processes, such as group cohesion, social identification, leadership, or trust [103]. The technology formed

the image, whereas the group was often in the shadow and slightly out of focus. Focusing on the technology advanced our understanding of how systems mediate user behaviors such as the awareness of other members [29], multi-tasking [89], and communication [12, 31]. Other CSCW scholars brought forth new conceptual models articulating the social dimensions in conjunction with the technological aspects, such as the kinds of contextual information available for users in systems, structural, social, or organizational, [30], the support and access mechanisms necessary for allowing users to work together [17], and the degree of collaboration enabled through these systems [51].

As technologies were advancing and providing large-scale and long-term infrastructures, the goals of CSCW expanded beyond a focus on collaboration in small groups. New studies examined the use of systems to support group work carried out by departments, business units, and organizations. Yet, scholars still were mostly focused on the technical dimensions, group effectiveness, and tangible outcomes of these systems. Pinelle and Grudin [93] discovered in their 1990-98 CSCW literature review that almost one-third of groupware systems were not formally evaluated and that only about 25% of those were evaluated in naturalistic settings affording a richer perspective on motivational and social processes commencing among groups "in the wild." It became evident that the design of computational models and systems may be missing important social dimensions [101]. In 1999, Ackerman made the distinction of what process can be *socially* supported and what *technically* is possible to support, where several of the assumptions and mechanisms provided by those collaborative systems lacked the flexibility of social life [2].

2.2.1 Moving to Social Systems. After the attention devoted to small groups and groupware systems, the CSCW community embraced other kinds of collectives that were adopting technologies for supporting their activities. Around 1992, people and organizations began using the Internet and the World Wide Web. By taking advantage of this unprecedented level of digital connectivity, systems and software started to become more *social*. In a special issue of *Communications* of the ACM, Schuler's study [102] described how software was establishing social relations at a distance and modifying existing social patterns. He defined *social computing* as "any type of computing application in which software serves as an intermediary or a focus for a social relation." Essentially, these discussions brought social processes more clearly into focus inspiring what would be the next generation of software for groups. Digital platforms were able to connect people and enable them to interact in ways that improve their group awareness and support their decision-making processes. CSCW researchers started conducting observational and empirical studies in social networking sites [28, 64], virtual offices [37, 104], peer-to-peer systems [91], online communities [49], collaborative and people recommender systems [41, 77], blogs [3, 87], wikis [57], collaborative environments [26], and online gaming platforms [19, 27, 112]. Although these studies and technologies explored more social aspects of groups than did those published previously, the unit of analysis was rather narrow. The focus was mostly on users' interactions at the individual level with their respective groups, rather than focusing on those groups and their participation within online spaces [103, p. 201:1].

2.2.2 The Model of Coordinated Action. In 2015, a new CSCW framework emerged considering the new technological affordances introduced by these large-scale groups. Lee and Paine proposed the Model of Coordinated Action (MoCA), which classifies collaborative systems according to seven dimensions built upon previous CSCW models [68]. Inspired by the Johansen's matrix [47] and Grudin's taxonomy [35], Lee and Paine adapted the dimensions of time to *synchronicity*, place *physical distribution*, and work-level *scale*. Additionally, they introduced the number of communities of practice as an important dimension to measure a platform's diverse user base. This dimension gathers individuals, groups, or networks that share specific cultures, backgrounds, perspectives.

Then, *planned permanence* (from short-term to long-term) and *turnover* (from low to high) crystallize the arrival, continuity, and departure of users. These last two dimensions not only evaluate the local and temporal presence of group members, but also consider how cooperative systems manage inclusion, maintenance, and removal of individuals. For example, systems must consider how to manage the planning of member coordination, and how this will affect the performance of the group. This framework offers a systematic method to classify systems according to the social processes involved in group formation. However, as [114] pointed out, operationalizing these dimensions as continuous metrics can be problematic. Group sizes, cultures, and membership attributes are highly contextual, specific to each case, and users can adapt these systems according to their group needs. For example, users on a crowdsourcing system can be assembled into expert-novice dyads or scale up to large groups consisting of thousands of novices and experts working together [80].

2.2.3 Crowdsourcing Systems. In the early 2010s, CSCW scholars started paying attention to crowdsourcing systems and how large-scale tasks were performed by crowds [58]. These socio-technical infrastructures decomposed tasks into workflows which are executed by other users, called crowd workers, representing a mix of novices, experts, and algorithms. Inspired by Taylorism ideas and distributed computing, individuals are assigned to different roles according to their expertise, availability, and incentives. These advanced systems are highly supported by algorithms and digital tools that enable requesters to find workers to perform their tasks, and the mechanisms for generating incentives, payments, and deliverables. With collocated workgroups, tasks are usually assigned to workers via a manager, this framework hands over the management of tasks, interactions between workers, resources, and deliverables to the computing platform. Because of these architectural decisions, the collectives assembled by crowdsourcing platforms tend to be short-term, project-driven, and their communication constrained by platform design. Users are recruited for specific tasks, and then upon completion of the task, the group is dissolved. Studies have reported issues of poor coordination, lack of support, and lack of social identification with the group, which may diminish the group's productivity [34, 65, 83]. In response to these shortcomings, CSCW studies were shifting the perspective on these platforms to the whole being greater than the sum of its parts [11, 123]. New crowdsourcing techniques have been implemented to nurture and support constructive social dynamics [56, 100]. For example, *crowd guilds* enable interdependencies between workers [118] by replacing individual's pay rates with peer assessments, emulating reputation and social evaluation processes in these socio-technical platforms.

2.2.4 Focus on Large Social Systems. A more recent line of research systematized the CSCW literature of the past 25 years. Wallace, Oji, and Anslow [114] reviewed articles published at CSCW from 1990 to 2015, and analyzed publications across four classification schemes to characterize CSCW work. They found the scope of CSCW investigation has expanded from small groups to large social systems during the past two decades. Their main results also show that CSCW research has increased in descriptive studies that analyze existing users or tasks, decreased in explanatory and systems research that develops theories or proposes novel systems, and a lack of bibliographic research that synthesizes CSCW findings. Similarly, Seering and colleagues [103] found that most CSCW theoretical and methodological research has focused on individuals as the unit of analysis, rather than focusing specifically on groups and their participation within online spaces. Based on Social Identity perspectives, which explains how people organize themselves into and within groups, they provide a guide for future CSCW research. They identify five potential research domains: self-presentation, social support, collaboration, misbehavior, and leadership in online communities. However, group formation is not discussed in this systematic review. Moreover, these two systematic reviews reveal how group formation mechanisms behind these collectives have not been fully explored; that is, who, how, and why people are coming together online [92].

3 RESEARCH QUESTION

Though the CSCW community has certainly delved into complex social and cultural barriers in the use of technology by groups [114], there is still much to be learned about the earliest stage of a group's life cycle and how online social collectives emerge. This has two main implications. First, technologies are shaping, moderating, and influencing the way people are brought together. With the increasing adoption of online systems, there is a desire to assemble new kinds of groups, spontaneous or long-standing, with a shared purpose or simply shared identity [53, 94, 97]. Group formation may be even more important to nascent groups forming online, and lacking deep prior relationships, and so this topic is especially critical to the CSCW community.

Second, the factors that explain users' decisions during the group assembly stage are not particularly well understood. The work on groups conducted in the past decade has focused mostly on group composition (i.e., the mix of people in the group), and less so on formation. Composition research reveals factors that lead to success, but how do these factors come about during the formation stage, and how do technologies shape behavior when forming groups [23, 36]? Organizational studies find group formation affects performance by determining group composition and setting early norms in motion [15, 74, 82]. However, designing collaborative systems for group formation requires us to bring both the technology and social processes clearly into focus. Given this motivation, we conduct a scoping review to answer the following research question:

RQ. *How has group formation been addressed by CSCW researchers?*

4 METHOD

To answer this research question, we followed a scoping review methodology, a highly referenced method in the health science field [7, 69, 88, 106], and used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to report the methods and results [81]. Scoping reviews are used to map literature in a specific area of research or field to identify gaps in the research may lie [7, p. 2]. Scoping reviews have a number of advantages: they are transparent, comprehensive, less prone to bias, and easier to reproduce the detailed information reported about each step of the review and how it is conducted. PRISMA commences in four stages: Identification, Screening, Eligibility, and Inclusion. We tabulated articles in a shared spreadsheet, capturing metadata such as publication year, ACM ID, abstract, and keywords. Once the metadata was recorded, the authorship information was hidden to avoid any potential bias during the coding phase.

4.1 Eligibility Criteria: Inclusion and Exclusion of Articles

Our review includes all CSCW articles that refer to the formation of groups. Articles were eligible for inclusion if they described, analyzed, or elaborated technologies, constructs, or mechanisms involved in searching for, selecting, and incorporating members into a group. We included articles that experimented with different group compositions, or where the research questions considered the processes or consequences of adding, modifying, or removing team members. We excluded articles that did not study some aspect of group formation. Articles that developed or analyzed technologies to support *existing* groups did not meet the inclusion criteria because they did not study the formation of groups or technologies that support group formation.

4.2 Search Strategy and Data Sources

We conducted the first step "Identification" by using the ACM Digital Library. We searched for articles on group formation published in the CSCW conference. Since CSCW switched format from publishing conference proceedings to a journal in 2017, we searched in the CSCW proceedings from

1990 to 2017 and the Proceedings of the ACM on Human-Computer Interaction (PACMHCI) from 2017 to 2018 [52]. We first performed several searches to assess the volume of potentially relevant studies according to the eligibility criteria. We built and identified keywords and search terms from our research question, theoretical background, and discussions [106, p. 215]. After several iterations, our final set of searches included combinations of the word "group" with "assembly" or "formation", and the word "team" with "assembly" or "formation". This search strategy also allowed us to identify articles with phrases such as "forming groups," "assembling teams," and related terms such as "teamwork". After defining these four search queries, we conducted each query on the ACM Digital Library and exported the results to a CSV file. The four files were merged into a single CSV file and duplicates were removed using the article IDs.

Following [114]'s search strategy, we excluded the Companion and Extended Abstracts publications to maintain consistency between literature reviews and to ensure that our review included work in advanced research stages. To ensure reproducibility, a member of our research group, who was not a co-author, peer-reviewed the search strategy using another computer and following the Peer Review of Electronic Search Strategies (PRESS) checklist [78].

4.3 Article Selection

As preparation for the "Screening" phase, two of the researchers calibrated the screening form independently with a random sample of 50 articles. We quantified the agreement between these researchers using the Kappa statistic [66]. Once the researchers achieved a consistent article selection ($\kappa \geq 0.8$), they proceeded to screen articles for inclusion through a three-stage process. First, one researcher reviewed the retrieved articles' titles and keywords (i.e., level-one screening). Articles whose titles or keywords met the eligibility criteria were retained. Next, in the level-two screening, the two researchers independently performed a second review including the articles' titles and keywords, and also the abstracts. If the researcher considered the article eligible, he/she coded it as 1. Otherwise, as 0. The two researchers then compared their lists and resolved any disagreements with discussion and consensus. Once rater agreement on inclusion was high ($\kappa \geq 0.8$), the two researchers started the full-text article review (i.e., "Eligibility" phase). In a first-cycle of revisions, the two researchers classified each article independently as "include," "unclear," or "exclude" to determine whether they met the inclusion criteria. Articles that were classified in the first two categories were reviewed independently in a second-cycle by each author. Final eligibility disagreements were resolved by consensus or arbitration. Once consistent article selection was achieved ($\kappa \geq 0.8$), the two researchers moved onto the data extraction and synthesis stage.

4.4 Data Extraction and Synthesis

The two reviewers independently extracted data from all included citations using a pre-designed electronic form that was pilot tested using a random sample of 10 citations. Once data were consistently abstracted ($\kappa \geq 0.8$), reviewers proceeded with full data extraction. Based on the categories provided by [35, 114], we extracted data pertained to: (1) research type (e.g., design and evaluation, descriptive, explanatory, bibliographic, not empirical); (2) evaluation type (e.g., field experiment, field and case studies, laboratory experiments); (3) evaluation methodology (e.g., quantitative, qualitative, mixed-methods); and (4) focus level (e.g., individual, small group, project, organization). The researchers also included terms used to refer to group formation, open notes, and memos. We managed and analyzed these data using Google Forms and Google Spreadsheets.

After the two researchers completed the full data extraction, they met and discussed their main results. To illustrate the evolution of group formation research at the conference over time, the two researchers plotted each article and a one-sentence summary by year on a timeline. The researchers then created categories based on the main shifts that emerged from plotting articles and their

respective summary sentences. To summarize and integrate the articles' findings across the corpus, we use thematic analysis [18]. One of the researchers developed an open and iterative coding scheme, and then classified articles into this scheme. This process continued iteratively until a final thematic codeset was established. From these multiple rounds of coding, the researcher identified two theme levels, where the main level's themes encapsulate concepts of the second level. Since codes were not mutually exclusive (i.e., articles could contain multiple distinct themes), reliability was measured using Krippendorff's alpha for each subtheme [62]. To ensure that the themes coded by the first researcher were reliable and exhaustive, the second researcher independently coded the selected papers following researcher 1's coding scheme until they achieved satisfactory levels of agreement for each theme ($\alpha \geq 0.5$) [8]. Discrepancies were resolved by discussion and the researchers reached a consensus on the final classification of papers.

5 RESULTS

The procedures searches identified 2,292 unique articles that were screened for inclusion, from which 127 full-text articles were retrieved for further assessment. From these articles, 35 were included in the final review (Figure 1). These articles addressed technologies, mechanisms, processes, and users behaviors behind group formation. The most common reason articles were excluded in the eligibility stage was that they did not focus on group formation and/or group assembly. 92 articles passed through the screening stage but were not included as they did not focus on group formation. Many of these that did not focus on group formation developed technologies to support group interaction, observed existing groups performing an experimental task, or examined factors (other than group formation) affecting group outcomes.

5.1 Description of the Included Articles

Table 1 provides a description of the included articles. The majority of the articles were published from 2010 onwards (68.57%), with very few papers published prior to 2000 (5.71%). Most articles report exploratory research (40%), followed by descriptive research (22.86%). Excluding bibliographic articles, most articles reported field and case studies (67.86%), followed in prevalence by reports of laboratory (25%) and field experiments (7.14%). From the papers in which hypotheses or systems were evaluated, 35.71% were quantitative, 32.14% were qualitative, and 32.14% used mixed-methods. Regarding the focal level, most of these papers examined small groups (42.86%) or projects (28.57%).

5.2 Group Formation Periods

We discovered meaningful shifts in the nature of CSCW scholarship on group formation over time. Reviewing these papers in order of publication suggested four periods, noted in Figure 2 that depicts a timeline from 1990 and 2018.

5.2.1 Period #1: Groups Interacting with Technologies. The first period is highlighted by three articles from the years 1992 and 2004. As indicated by the small number of papers, group formation was not yet a central topic at CSCW. Groupware gained a lot of attention from CSCW scholars. In this work, groups were seen as fixed entities "interacting with technologies" [30]. Two panels held between 1992 and 2004 advocated for future work that explores the social processes involved in group formation [14, 78]. This represented a call to advance understanding of the social processes needed by groups in order to join together and work together effectively. We labeled this period "groups interacting with technologies," borrowing the phrase and framing from McGrath's early CSCW panel.

5.2.2 Period #2: Enabling Online Groups and Communities. The second period started in the 2000s. As was described in the background section, socio-technological infrastructure started becoming

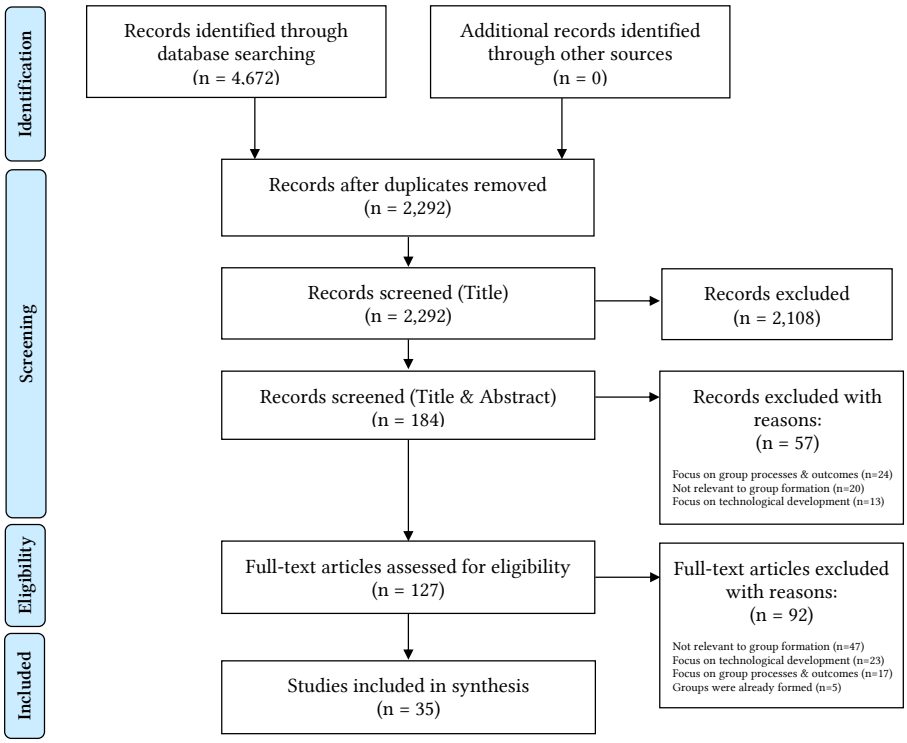

Fig. 1. PRISMA Flow Diagram for this study. It presents the details of the article selection process.

more social enabling users to find dynamic ways to build online groups and communities. Some papers described different aspects and mechanisms that could be used to assemble groups. For example, one paper introduced a recommender system in order to find experts inside the company [77], another article tested how group structures affected their coordination [42], and another tested pairing individuals for specific tasks [80]. Social networking sites [61, 94], wikis [76, 105], software development platforms [75], virtual worlds [16], and online games [13, 86] are some examples of collaboration spaces that enabled users to meet, interact, and share with others. CSCW scholarship moved from the strong focus on how intact groups "interact with technologies" to a focus on how technologies could enable the formation and nucleation of groups. The papers tended how group formation was driven by users' characteristics and behaviors (e.g., gaining social capital, improving self-presentation, and impression), as well as the emergence of larger and virtual social structures; the latter being represented in Figure 2.

5.2.3 Period #3: Enabling Crowds. One of the most cited CSCW articles in our corpus belongs to this period. "The Future of Crowd Work" [58] details a new socio-technical infrastructure where large-scale tasks are requested by users on a crowd work platform. The task is then decomposed into workflows which are executed by other users, called crowd workers, comprised of novices, experts, and algorithms. In crowds, group formation is simplified to assigning skilled workers to micro-tasks. Therefore, several of the papers in this review explore the relationships between these virtual workers and their work environments and describe the social obstacles and barriers that

Characteristic	Number	Percentage
Years of Publication		
1990-1999	2	5.71%
2000-2010	9	25.71%
2010-current	24	68.57%
Research Type		
Bibliographic	6	17.14%
Descriptive	8	22.86%
Design and Evaluation	6	17.14%
Explanatory	14	40.00%
Not Empirical	1	2.86%
Evaluation Type		
Field and Case Studies	19	67.86%
Field Experiment	2	7.14%
Laboratory Experiments	7	25.00%
Evaluation Methodology		
Mixed-Methods	9	32.14%
Qualitative	9	32.14%
Quantitative	10	35.71%
Focal Level		
Individual	4	11.43%
Organization	6	17.14%
Project	10	28.57%
Small Group	15	42.86%

Table 1. Characteristics of included articles

users face when collaborating [34, 85]. In this period, there was a shift in how group formation was viewed. Whereas previous studies focused on how to make support or fit small groups with technologies, and then how to assemble small groups into larger communities. In this period, the focus was on how groups could be nucleated into crowds from larger communities.

5.2.4 Period #4: The Renaissance of Small Groups. The most recent period is marked by a return to forming small interacting teams, though, now from a larger collective. As the socialization problem and lack of interaction in collaborative platforms became evident, the CSCW community started discussing how collaborative platforms could support group performance and effectiveness matching social needs with technological affordances. One exemplary article emphasized the relevance of group formation and its consequences on the group performance [39]. Two papers studied how collective intelligence is able to predict team performance and group satisfaction [22, 55]. Other articles experimented with different strategies for assembling groups: based on team dating [73], familiarity [99], community deliberation processes [117], merging existing groups [54], and member rotation among teams to increase creativity [98].

5.3 Thematic Analysis

We identified six main themes present in the articles reviewed in this study, namely, *group composition, self-presentation of users and groups, recruitment mechanisms, assembly mechanisms, organizing structures, and group culture.* In the following subsections, we detail the main distinctions between these topics and basic descriptive statistics of their prevalence. A summary of these results can be found in Table 2.

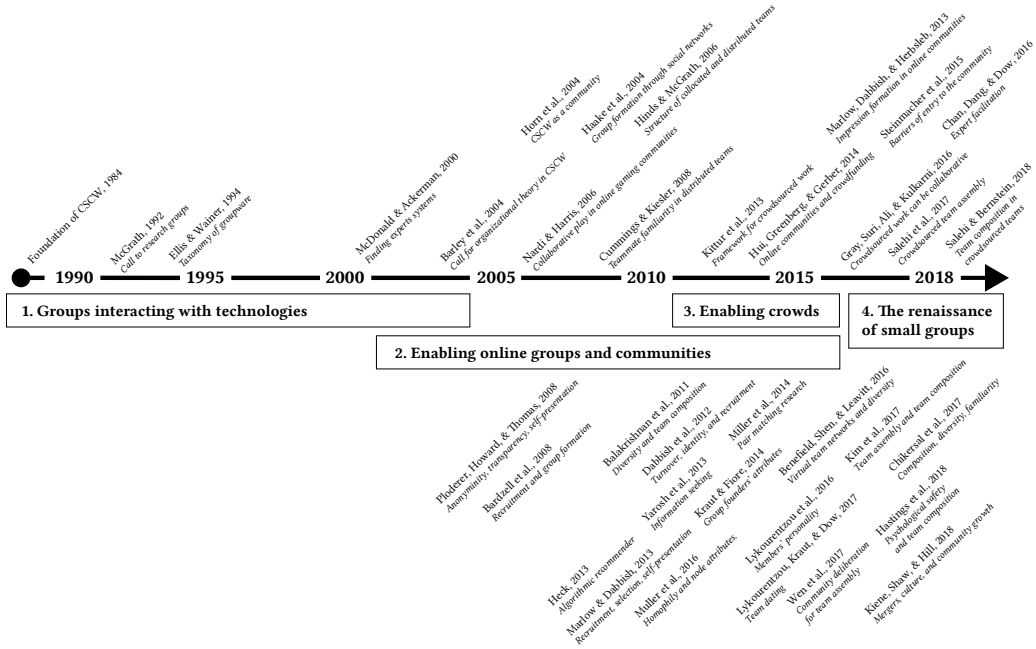

Fig. 2. The evolution of group formation research in CSCW.

5.3.1 Theme #1: Group Composition. The first theme describes how the combination of members' personal characteristics and interpersonal relationships affect group processes and outcomes [22, 70]. Papers on *group composition* have explored the effects of group members' *similarity* (10 articles), prior *familiarity* (9 articles), *networks* (7 articles), *personality traits* (2 articles), and *group size* (2 articles).

The first sub-theme is *similarity*, including papers that apply homophily theory to group formation [21, 79]. Across a variety of contexts, CSCW papers have demonstrated that "like attracts like." Individuals are more likely to work with those who are similar to themselves. This has been found in papers exploring academic collaborations [40, 44], World of Warcraft teams [13], and other studies on video game platforms that confirm users can both benefit from and prefer similarity between group members [16, 55].

The second sub-theme is *familiarity*, including papers that explore the role of prior relationships on group formation. These studies find that while technologies open up the possibility of collaborating with those who are near and far, people often use technologies to form groups with those whom they already know. An exemplar finding is that member familiarity mitigates the otherwise negative effects of distance on performance [24]. Another study on familiarity is [99], which describes the design of Huddler, a system to match people based on a familiarity metric.

The third sub-theme is *networks*. These papers explore the role of network positions and structures in group formation. For example, [42] examined flat organizational hierarchies in collocated and distributed teams, finding flatter organizational hierarchies provided smoother collaboration in collocated teams, but not for distributed teams. Additionally, [98] Hive system supports better creativity in teams by rotating users through different groups at appropriate times using an algorithm.

The fourth sub-theme, *personality traits*, includes articles that are concerned with the effects of assembling groups with conflicting personality traits as well as trying to find combinations of traits that benefit outcomes. An exemplary article for this subcategory is [72] in which the authors examine the effects of personality compatibility in crowd teams on performance and individual perceptions. The other paper in this thematic category is by [86] on how players collaborate in World of Warcraft. This study reports issues of work style and personality play a role in successful collaborations. The authors suggest World of Warcraft gives people a chance to engage in lightweight collaborations and then assess whether they wish to continue the relationship based on early interactions.

The fifth sub-theme, *group size*, includes articles investigating the effects of group size and membership rules as they relate to group formation. For example, in a study on access rights to group membership, [38] found that users should be able to form groups and adapt access rights when they deem necessary; thus, allowing the founding user to preemptively decide group size.

5.3.2 Theme #2: Self Presentation. The second thematic category is *self-presentation*. These papers describe how users decide and are able to present themselves to others through these social platforms. This concept is based on Goffman's framework [32]. The reviewed articles analyzed three different aspects that tell how users present themselves in order to reduce uncertainty during group formation: *identity* (13 articles), *information seeking* (7 articles), and *information availability* (11 articles).

The first sub-theme, *identity*, refer to studies that explore the role of social identity processes in group formation. Social identity describes the set of categorizations that an individual feels a sense of belongingness to. CSCW work on identity and group formation has explored users' view of themselves as belonging to one or more salient social groups [61], how profiles are constructed on platforms [76], group culture [54], and both the outward- and inward-facing presentation of that group, and how these serve as a driving force in group formation [39]. For example, [26] found that the influence of turnover on participation is mediated by common social identity among members such that the effect was strongest when there was a shared social identity among group members.

The second sub-theme, *information seeking*, pertains to a bidirectional search individuals make when looking for potential collaborators or new groups [25, 77, 120] or for a group to invite new members [13, 75]. In essence, these papers detail how was the process of searching for new individuals and reflect that this is a critical first step in group formation. This theme is exemplified by [75] as they found that information seeking plays an integral role in online impression formation, which, in turn, plays an important role in the decision to accept pull requests on GitHub.

The third sub-theme, *information availability*, captures the degree of information provided by users to the systems and to other users [34, 77]. Articles containing this theme focused primarily on the effect the amount of identifying information present has on group formation. [94] found that in an online community of bodybuilders that post identifiable progress pictures of themselves users reported a higher degree of trust in advice received by other community members as compared to online forums with high degrees of anonymity.

5.3.3 Theme #3: Group Recruitment. This third theme is *group recruitment*. These studies explore the factors that influence individuals' decision to join pre-existing groups or to begin a new group with someone. This theme is defined by three sub-themes, *visibility* (8 articles), *member attributes* (6 articles), *motivations and goals* (10 articles).

The first sub-theme, *visibility*, emerged from research on groups or individuals attempting to attract others, promote themselves [75], and raise the awareness of their needs. We see this most clearly in [61] article on how important group founder attributes are on the formation, survival, and longevity of groups. Visibility is exemplified, then, as the founder's social capital during the

early stages of formation; higher social capital means greater visibility of the group the founder is beginning, and this equates to faster growth during a critical phase of the group's life cycle.

The second sub-theme, *member attributes*, captures the appeal of the attributes held by current members of the group or of the individual looking to start a group. [61] also exemplifies member attributes and how they drive formation and recruitment. This article found that member attributes are not limited to social capital and that the founder's resources, skills, and behavior also heavily influence the survival of a new group. Another exemplar paper that contains member attributes as a theme is [13] on the formation of World of Warcraft instance groups; the authors find that users wanted to group up with others they knew from prior experience were competent and those with skills that matched specific roles that needed to be filled.

The third sub-theme, *motivations and goals* includes articles that focused on how a group or individual broadcasts their goals and motivations when looking for others [45, 75], and how that broadcast attracts others to join when there are goal and motivation alignment [13, 76]. For example, [45] found that crowdfunded entrepreneurs expressed the difficulty of fostering interest in their products; to do so, they relied on social media to broadcast their values and motivations for their projects. Additionally, these crowdfunded entrepreneurs fostered a community comprised of other crowdfunded entrepreneurs and functioned as a guide and mentor to new members as they go through the stages of developing and fostering their own communities around their products; demonstrating a goal and motivation alignment between the mentee and the mentor, and between the user and the community.

5.3.4 Theme #4: Assembly Mechanisms. This theme captures the various processes in which groups are created. Second-order themes include *random assignment* (2 articles), *criteria-based assignment* (6 articles), *self-assembly* (14 articles), *open enrollment* (4 articles), and *mergers* (1 article).

The first sub-theme, *random assignment* includes studies that draw comparisons to randomly formed groups and teams. This assembly mechanism was mentioned as a formation alternative and as a control in [39]'s experiment on team formation which experimentally tested differences between random assignment without socialization, random assignment with socialization, criteria-based assignment without socialization, and criteria-based assignment with socialization.

The second sub-theme, *criteria-based assignment*, is an assembly mechanism that is based on a specific attribute that either a figure in a role of power or an algorithm uses to assign people to groups; most often these are work teams and the goal is to optimize performance outcomes by correctly fitting individual attributes together. Huddler system [99], for example, forms crowdsourced teams based on two criteria, familiarity and availability. They found that these criteria-based teams doubled the performance outcomes of ad-hoc teams.

The third sub-theme, *self-assembly*, encompasses articles investigating the formation mechanism for when individuals have the agency to choose their group association. In that sense, self-assembly differs from the criteria-based assignment. Self-assembly includes an element of exclusivity to the group; that is, members of the group must extend an invite to someone before they are accepted into the group. Self-assembly is one of the most common thematic categories we found present in our corpus. Articles that focused on methods of self-assembly ranged from work on how people form teams for collaborative play in World of Warcraft [86], to a team dating method that quickly fosters familiarity in potential working relationships [73].

The fourth sub-theme of assembly mechanisms, *open enrollment*, encompasses mechanisms of formation that are not exclusive and joining a group, mainly a community, is elective. An exemplar article here is on open source software communities; the authors found that while the common conception for open source software communities is often perceived as having few barriers of entry, they found 58 barriers of entry [105].

The fifth sub-theme includes a recent article investigating the handling of organizational culture in group *mergers*, when two or more groups form a new group, via World of Warcraft guilds [54]. This was the only article investigating mergers as a mechanism for group assembly.

5.3.5 Theme #5: Organizing Structures. This theme essentially represents an extension of Grudin's work-level taxonomy; that is, clear themes emerged from the data regarding the specificity of terms used to describe both the scale and the structure of what was focal entity was being formed. We found five second-order themes: *dyads* (7 articles), *teams* (14 articles), *groups* (5 articles), *crowds* (8 articles), and *communities* (12 articles).

The first sub-theme, *dyads*, are comprised of two people with a purpose to complete a goal or task. For example, [22] found that psychophysiological correlates, such as facial expression, were predictive of collective intelligence measures in dyads, which has implications for assembling dyads outside of a lab setting. Additionally, [80] applied a method of assembling dyads using what they call a paired research approach, in which individuals are paired to work on each other's research for one week before rotating membership.

The second sub-theme, *teams*, are formed for the purpose of completing tasks and achieving goals for which members are mutually reliant on one another. Teams contain from 3 to 8 people and involve exclusive membership. [73] looked at creating dyads through the team dating process. In the end, individuals rated whom they would like to form a team with most.

The third sub-theme, *groups*, is a catch-all term in CSCW. Some articles use the term group interchangeably with other organizing structures, and this is somewhat problematic. For example, [16] used group and team interchangeably in their article on team performance in a massively multiplayer online game, "Dragon Nest." The most prevalent use of group was when describing a social collective that requires exclusive membership, but does not have a distinct purpose or task to perform. For example, [54] examine the organizational culture during two guilds merging in World of Warcraft; the guild is the group. The guild lacks a clear purpose and members' report differing perceptions on the reason for forming the guild.

The fourth sub-theme, *crowds*, are large in size and have a purpose of completing tasks independently when called on to do so. Crowds are comprised of a wide array of individuals with skills and abilities, ranging from novice to expert. For example, [58] examines the future of crowdsourced work by weighing the pros and cons of this organizing structure.

The fifth sub-theme, *communities* describe an often large, loosely formed organizing structure that is not exclusive, has little to no requirements for membership and has no purpose other than a social identity and discourse. An exemplary article is [45] on the community behind crowdsourced work. This work describes the difference between a crowd and a community: crowd workers build a community outside of the crowd work platform that fosters social identity and support.

5.3.6 Theme #6: Group Culture. This last theme emerged containing one second-order thematic category, *cultural practices* (4 articles). Cultural practices encompass processes that become institutionalized within an organizing structure. For example, [76] examined cultural practices they coined activity traces and signals in which members of the GitHub community would signal their competency to one another using signals only those in the community would understand and value. Additionally, one article investigated various team assembly mechanisms and their effects on the adoption of a psychologically safe culture where members value and encourage one another to speak up and voice concerns openly [39]. The authors note, "for instructors deploying a team formation tool, creating an expectation among team members that their team can perform well is as important as tuning the criteria in the tool." [39, p. 68:2], which has noteworthy implications for future assembly mechanism design.

Theme	Definition	N	Articles
Group Composition	The combination of members' personal characteristics and interpersonal relationships effect on group processes and outcomes.		
Similarity	Articles pertaining to the effects he similarity, or likeness, of a group's members.	10	[11, 13, 16, 22, 24, 40, 44, 55, 85, 117]
Familiarity	Articles pertaining to the effects of members having, or not having, prior experience with one another.	9	[24, 44, 54, 73, 76, 78, 98, 99, 117]
Networks	Articles pertaining to or directly manipulating network structures or organizational hierarchies of groups.	7	[16, 22, 42, 55, 58, 78, 98]
Personality	Articles pertaining to group assembly based on combinations of personality traits to improve outcomes.	2	[72, 86]
Group Size	Articles pertaining to the effects of group size and member requirements for group formation.	2	[38, 72]
Self-Presentation	How users decide and are able to present themselves to others through these social platforms.		
Identity	Articles pertaining to user identification with a group, how profiles are constructed on platforms, identifying with its culture, and both the outward and inward facing presentation of that group.	13	[11, 13, 26, 30, 34, 39, 54, 61, 75, 76, 85, 86, 94]
Information Seeking	Articles regarding the bidirectional search individuals make when looking for potential collaborators or for new groups, or for a group to invite new members.	7	[13, 26, 34, 75, 76, 120]
Information Availability	The degree of information provided by users to the systems and to other users.	11	[13, 24, 26, 30, 34, 75–77, 85, 94, 120]
Recruitment	The factors that influence individuals' decision to join pre-existing groups or to begin a new group with someone.		
Visibility	Articles pertaining to groups or individuals attempting to attract others, promote themselves, and or raise the visibility of their needs.	8	[13, 24, 26, 30, 44, 75–77]
Motivation & Goals	Articles pertaining to how a group or individual broadcast their goals and motivations to attract others to join.	10	[11, 13, 14, 45, 54, 58, 75, 76, 85, 105]
Member's Attributes	Articles pertaining to the appeal of the attributes held by current members to potential new members.	6	[13, 14, 54, 61, 78, 85]
Assembly Mechanisms	The various processes and mechanisms in which groups are created.		
Random	Articles pertaining to randomly selecting individuals and placing them into groups.	2	[39, 117]
Criteria-Based Assignment	Articles pertaining to group assignment based on a specific attribute to optimize performance outcomes by fitting individual attributes together methodically.	6	[39, 40, 80, 98, 99, 117]
Self-Assembly	Articles pertaining to when individuals have the agency to choose their group association with closed membership.	14	[11, 24, 26, 38, 40, 55, 58, 73, 75, 77, 85, 86, 99, 120]
Open Enrollment	Articles pertaining to when individuals have the agency to choose their group association with open membership.	4	[34, 38, 85, 105]
Merger	Articles pertaining to group formation occurring between two or more groups.	1	[54]
Organizing Structures	The specificity of terms used to describe both the scale and the structure of what was being formed, simultaneously.		
Dyad	Articles pertaining to a group comprised of two people with a purpose to complete a goal or task.	7	[20, 22, 38, 77, 80, 86, 120]
Team	Articles pertaining to groups that are formed for the purpose of completing tasks and achieving goals, and is comprised of 3-8 people.	14	[11, 13, 16, 24, 38–40, 72, 73, 80, 86, 98, 99, 117]
Group	Articles regarding group as a social collective that requires exclusive membership, but does not have a distinct purpose or task to perform.	5	[38, 61, 85, 86, 94]
Crowd	Articles regarding crowds as being large in size with a latent purpose of completing tasks independently when called on to do so.	8	[20, 45, 58, 72, 73, 85, 98, 99]
Community	Articles regarding communities as large, loosely formed organizing structure that is not exclusive, has little to no requirements for membership and has no purpose other than a social identity and discourse.	12	[11, 16, 26, 34, 54, 75–77, 86, 94, 105, 117]
Group Culture			
Cultural Practices	Articles pertaining to processes that become institutionalized within an organizing structure.	4	[26, 39, 54, 75]

Table 2. Typology and prevalence of themes

6 DISCUSSION

In order to answer our research question –*How has group formation been addressed by CSCW researchers?*– we systematically reviewed the CSCW literature on group formation. Group formation, the mechanisms, and processes explaining how and why groups form has been a consistent topic of interest in CSCW since the 1990s. The results of this scoping review provide a comprehensive answer to this question. We now elaborate on some of the key findings and their future implications.

6.1 Rigor and Relevance

The first aspect of CSCW scholarship on group formation we considered is the nature of the research, as it bears directly on the kinds of inferences that can ultimately be drawn from this work. Rigor is often used to describe the ability of a study to enable firm causal conclusions to be drawn. The gold standard for causal inferences being experimentation. In the case of group formation, rigor allows precision in understanding which mechanisms are and are not responsible for group formation, and which aspects of group formation lead to which collaborative outcomes. Relevance, on the other hand, is often used to describe the degree to which findings bear on the phenomenon of interest. Comparing the relative proportion of studies using field versus laboratory settings bodes well for the relevance of group formation research. Examining the kinds of investigations, we see a literature rich in relevance, but perhaps lagging in causal rigor, as true experiments on group formation have been somewhat rare (25% of the sample). However, the strength of this work is that the vast majority of studies have been conducted in naturalistic settings (more than 40%). Dovetailing a trend in the larger CSCW literature, we saw an increase in mixed-method research in the most recent time period, suggesting work on group formation has benefited from multiple approaches that triangulate rich description and qualitative insight with the merits of quantitative deductive analysis and systematic evaluation.

6.2 Evolution of Group Formation at CSCW

Examining these papers over time shows four periods of shifting questions about group formation within CSCW. A first period shows how CSCW scholars started studying groups that interacted with technologies and called for incorporating the study of social aspects of group formations for those technologies, which were mostly developed to support small group interaction and decision making. A second period examined the formation of larger groups, taking on questions of group formation within communities formed in online social networking sites, wikis, and virtual worlds, to name a few. A third period examined the formation of crowds, a new kind of collective enabled mostly by automated systems where workers are assigned to micro-tasks. Most recently, CSCW work on group formation shows a renewed interest in understanding many of the social aspects of small interacting groups who leverage collaborative technologies. The evolution of these periods is consistent with Wallace et al. [114] study, which emphasized that the scope of CSCW literature has expanded from small groups to large social systems. We extend these results by considering CSCW articles published after 2016 and their renewed focus on small groups, which examine how group formation and the social characteristics of groups affect their performance and effectiveness.

It is noteworthy that Ackerman's call [2] continues echoing in CSCW research 20 years later. The study of group formation has evolved in accordance with the main shifts addressed in CSCW. Fundamentally, the CSCW research goals have moved from assembling groups using primarily technological criteria for designing functional groups supported by social processes augmented by technologies. Even though the majority of articles on group formation, 63%, have been published in the past five years (more than the number published in any preceding 5-year period), the study of

group formation still appears to be at an early stage at CSCW. Next, we consider the implications of our review for future work on group formation.

6.3 The Boundaries of the Collective

Results from our thematic analysis show that "organizing structures" are a highly researched topic in CSCW. Groups assembled online depend heavily on the socio-technical infrastructure. However, as this infrastructure has been researched and developed to support more individuals collaborating or participating collectively, three notable changes happen as a result. First, since 1994, new organizing structures have been created, such as crowds and communities. Second, we draw a distinction between groups, teams, and dyads as our review of the CSCW literature on group formation indicates these are distinct organizing structures that require varying assembly mechanisms. Finally, the third notable change is the blurred lines of group membership and the increasing ease of mobility between organizing structures that is afforded to users by the development in socio-technical infrastructures. Our thematic analysis categories are not mutually exclusive as many of the articles examine what we call *group membership mobility*. For example, papers investigating crowds are examining the formation process of pulling from a crowd and assembling or forming those selected into a dyad or team. A representative example of this proposed concept is [117] as their organizing structure begins as a crowd, then one experimental condition is assigned to a team and the other condition is encouraged to engage in a community-wide discussion before they are assigned a team. Upon completion of their task, both conditions dissolve back into the crowd. While the prior example is an experiment, group formation occurs similarly in non-experimental settings. Another example is [20], in which dyads are assembled by selecting members from a crowd, and each dyad is then matched with an expert facilitator creating a small group.

6.4 Group Formation Taxonomy

Using concepts introduced by Grudin's taxonomy (1994) and the MoCA framework [68], we paired the themes found in our thematic analysis within the CSCW literature. Based upon that, we provide a conceptual framework to describe the CSCW research on group formation along with their supporting technology types in Figure 3. We provide a spectrum of assembly mechanisms arranged by agency, horizontally; beginning on the left, high user agency entails the user has complete control over who they group with and the cost of membership is low, to low agency, on the far right, where users have no choice in who they are assigned to work with and the cost of membership is high.

After surveying the CSCW literature produced over the past 25 years, it appears that the boundaries that once starkly defined the difference between varying organizing structures have blurred with the advancement of group formation technology and infrastructures. Scholars believe this is a result of the internet, and, consequently, the ease of connecting with others [50, 121]. However, we believe this may also be a result of the ease of group membership mobility that is provided by the current state of socio-technical systems, where a user can switch between being an individual, part of a group, or part of a community. Group membership mobility presents a novel and interesting space to focus on in future research and system design [84].

6.5 Theoretical Implications

We now consider the implications of these findings for future CSCW scholarship. Our first recommendation is to explicitly theorize the interplay of technologies and group dynamics during the process of group formation. Our review shows CSCW scholarship has attended more to intact groups and their interaction processes than to understanding and enabling the formation processes

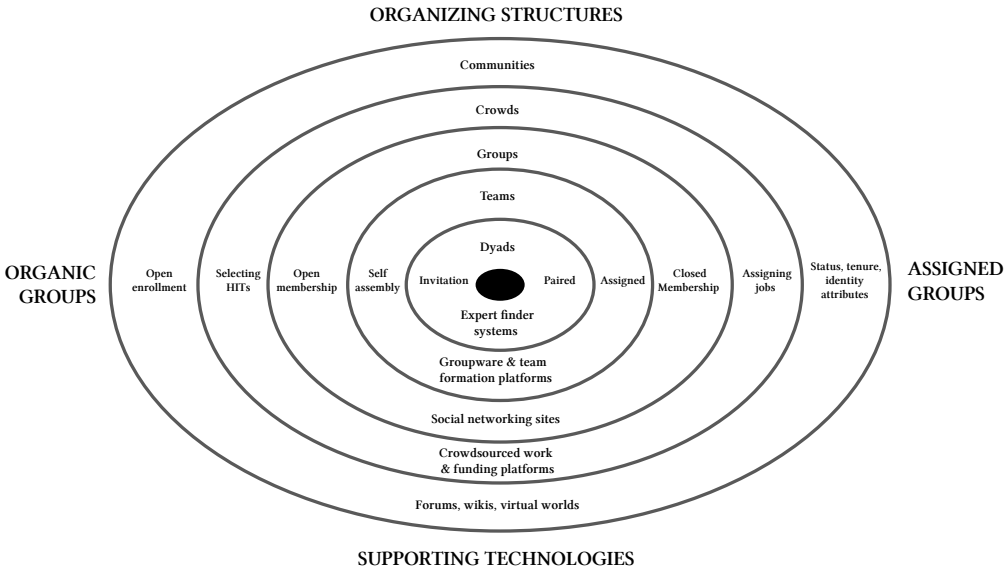

Fig. 3. Group formation taxonomy based upon our review. The levels "Work-level" and "Corresponding-systems" consider the themes found in the thematic analysis. The degree of agency present in group formation mechanisms are represented on the horizontal axis, from "Organic Groups" with complete agency to "Assigned Groups" with little or no agency.

involved in how groups come together online. [103] emphasized that CSCW must acknowledge that group identities are by no means an "inevitable" outcome in every online social context. Designers and platform builders must consider in-depth that forming groups is not only about putting users together, but that it is also a social process that users must experience through contact with others, identifying with the group, and understanding the group's nature and context. Future studies in CSCW should consider how technologies are mediating platform mechanisms for assembling groups, how newcomer obstacles are overcome by social support, and how to deal with the imbalance of experts and novice users on these platforms.

McGrath's call [78] to understand group formation is even more relevant now as social computing provides myriad opportunities for groups to form in different ways. The study of group formation has been less programmatic. Particular studies work with their own technology factors, group types, task types, research strategies, and use different subsets of dependent variables. Our second recommendation is for future work to explicitly consider the effects of technologies on member attributes, group structures, task characteristics, and the context. These four sets of factors were identified by McGrath, and attending to them in primary studies will expedite knowledge growth on when, how, and why technologies affect group formation. Consider the effects of group formation algorithms that use prior relations as a factor for making group recommendations. Over time, the effect will be that the "rich get richer" [1]. Those with established relations are more likely to get even more opportunities through these recommendation systems [9, 33]. Further, consider the consequences to resulting group diversity [67]. The point is not to value one outcome or another, but to consider a fuller range of consequences when exploring new technologies that support group formation.

Finally, considering the vast literature on group processes [59, 103], future work on technologies for group formation would benefit by incorporating social theories more centrally in the work. The thematic analysis shows a fair amount of breadth in the types of formation factors examined, but much less breadth in terms of social consequences. For example, how do technologies affect status hierarchies, leadership emergence, psychological safety, and/or long term group viability? Our third recommendation is for future work to leverage social theories at the early stages of technology design and in evaluating technologies. Social theories can be usefully applied to technology evaluation by considering how features shape the formation of group processes and properties, in addition to users' reactions to and experiences with the technology. As a starting point, we suggest the robust literature on group processes can be used to guide technology design and evaluation. The groups literature supports four essential processes needed for group performance and viability: affective states (e.g., trust, cohesion, identity), motivational states (e.g., collective efficacy), cognitive states (e.g., transactive memory systems, shared mental models), and behavioral processes (e.g., planning, backup behavior, coordinating) [59]. These four provide a useful criterion domain for work that explores group formation technologies. For example, how can technology features representing members' expertise and social networks fast track, or not, the formation of team transactive memory systems? Existing studies tend to examine a single outcome. Social theories on groups are rich with additional possibilities. Exploring the ways technologies can augment and enable groups to develop these four essential properties represents an exciting new vista for CSCW scholarship on group formation.

6.6 Limitations

There are two main limitations in our review that need to be considered. First, our search strategy relied on keywords, and thus, it is possible that we may have missed relevant articles due to the lack of indexing terminology specific to group formation. Second, our review was scoped to CSCW, and thus we did not include articles from similar proceedings (e.g., ACM GROUP, IEEE, ACM CHI) and adjacent fields (e.g., Academy of Management, Organizational Science). Incorporating knowledge on group formation from other adjacent fields would be beneficial for CSCW, and so we reference some of that work in the introduction and discussion sections as a linking mechanism between CSCW and adjacent areas.

7 CONCLUSION

Reviewing CSCW scholarship on group formation reveals key shifts and themes in how CSCW scholars have investigated group formation since 1990. In this period, the scope of research has expanded from initially studying existing groups interacting with technologies, to more recently, enabling large social systems. The early work studied groupware technologies and their "fit" with existing group practices. Subsequent work studied larger groups, like online communities and crowd workers, whose collaboration was not possible without the technology. The most recent work shifts again to focus on small interacting groups. This work explores how small, purposive groups can form out of larger collectives. Thematic analysis reveals six key aspects of group formation that have been the focus of this work: group composition, the way that users and groups self-presented in these platforms, recruitment mechanisms, assembly mechanisms, the organizing structures inherent to technologies, and the cultural practices that reside in these groups.

Together, these insights provide a better understanding of the social mechanisms behind group formation, which are necessary for technologies to support the nucleation of successful groups. *Quo Vadis?*, the current socio-technological ecosystem provides several opportunities to expand the CSCW focus on the nascence of online groups. From designing new group formation systems to evaluating assembly mechanisms, there are many opportunities for future work on technology and

team formation. Despite recent questions surrounding the impact of literature reviews of similar kin, our hope is that this review will raise awareness of the importance of online group formation and the impact it can have on group outcomes, draw the interest of researchers and practitioners to the topic, and to supply those interested with a thorough overview and useful point of reference for future research.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under grants IIS-1514427 and SES-1738297, the U. S. Army Research Laboratory under grant W911NF-09-2-0053, the National Aeronautics and Space Administration under grant NNX15AM32G, and the National Institutes of Health under award number R01GM112938-01. We thank the anonymous reviewers for their highly constructive feedback and suggestions.

REFERENCES

[1] Andrés Abeliuk, Gerardo Berbeglia, Pascal Van Hentenryck, Tad Hogg, and Kristina Lerman. 2017. Taming the Unpredictability of Cultural Markets with Social Influence. In *Proceedings of the 26th International Conference on World Wide Web (WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 745–754. https://doi.org/10.1145/3038912.3052680

[2] Mark S. Ackerman. 2000. The Intellectual Challenge of CSCW: The Gap Between Social Requirements and Technical Feasibility. *HumanâĂŞComputer Interaction* 15, 2-3 (2000), 179–203. https://doi.org/10.1207/S15327051HCI1523_5 arXiv:https://doi.org/10.1207/S15327051HCI1523_5

[3] Ban Al-Ani, Gloria Mark, Justin Chung, and Jennifer Jones. 2012. The Egyptian Blogosphere: A Counter-narrative of the Revolution. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 17–26. https://doi.org/10.1145/2145204.2145213

[4] Oznur Alkan, Elizabeth M. Daly, and Inge Vejsbjerg. 2018. Opportunity Team Builder for Sales Teams. In *23rd International Conference on Intelligent User Interfaces (IUI '18)*. ACM, New York, NY, USA, 251–261. https://doi.org/10.1145/3172944.3172968

[5] Aris Anagnostopoulos, Luca Becchetti, Carlos Castillo, Aristides Gionis, and Stefano Leonardi. 2010. Power in Unity: Forming Teams in Large-scale Community Systems. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM '10)*. ACM, New York, NY, USA, 599–608. https://doi.org/10.1145/1871437.1871515

[6] Aris Anagnostopoulos, Luca Becchetti, Carlos Castillo, Aristides Gionis, and Stefano Leonardi. 2012. Online Team Formation in Social Networks. In *Proceedings of the 21st International Conference on World Wide Web (WWW '12)*. ACM, New York, NY, USA, 839–848. https://doi.org/10.1145/2187836.2187950

[7] Hilary Arksey and Lisa O'Malley. 2005. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology* 8, 1 (2005), 19–32. https://doi.org/10.1080/1364557032000119616 arXiv:https://doi.org/10.1080/1364557032000119616

[8] Ron Artstein and Massimo Poesio. 2008. Inter-coder agreement for computational linguistics. *Computational Linguistics* 34, 4 (2008), 555–596.

[9] Ricardo Baeza-Yates. 2016. Data and Algorithmic Bias in the Web. In *Proceedings of the 8th ACM Conference on Web Science (WebSci '16)*. ACM, New York, NY, USA, 1–1. https://doi.org/10.1145/2908131.2908135

[10] Jasmón L. Bailey and John Skvoretz. 2017. The Social-psychological Aspects of Team Formation: New Avenues for Research. *Sociology Compass* 11, 6 (2017), e12487. https://doi.org/10.1111/soc4.12487 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/soc4.12487 e12487 SOCO-1220.R2.

[11] Aruna D. Balakrishnan, Sara Kiesler, Jonathon N. Cummings, and Reza Zadeh. 2011. Research Team Integration: What It is and Why It Matters. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW '11)*. ACM, New York, NY, USA, 523–532. https://doi.org/10.1145/1958824.1958905

[12] Jakob Bardram. 1998. Designing for the Dynamics of Cooperative Work Activities. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work (CSCW '98)*. ACM, New York, NY, USA, 89–98. https://doi.org/10.1145/289444.289483

[13] Shaowen Bardzell, Jeffrey Bardzell, Tyler Pace, and Kayce Reed. 2008. Blissfully Productive: Grouping and Cooperation in World of Warcraft Instance Runs. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW '08)*. ACM, New York, NY, USA, 357–360. https://doi.org/10.1145/1460563.1460621

[14] Stephen R. Barley, William H. Dutton, Sara Kiesler, Paul Resnick, Robert E. Kraut, and JoAnne Yates. 2004. Does CSCW Need Organization Theory?. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (CSCW '04)*. ACM, New York, NY, USA, 122–124. https://doi.org/10.1145/1031607.1031628

[15] Suzanne T Bell. 2007. Deep-level composition variables as predictors of team performance: a meta-analysis. *Journal of applied psychology* 92, 3 (2007), 595.

[16] Grace A. Benefield, Cuihua Shen, and Alex Leavitt. 2016. Virtual Team Networks: How Group Social Capital Affects Team Success in a Massively Multiplayer Online Game. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 679–690. https://doi.org/10.1145/2818048.2819935

[17] John Bowers. 1994. The Work to Make a Network Work: Studying CSCW in Action. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work (CSCW '94)*. ACM, New York, NY, USA, 287–298. https://doi.org/10.1145/192844.193030

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

[19] Barry Brown and Marek Bell. 2004. CSCW at Play: 'There' As a Collaborative Virtual Environment. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (CSCW '04)*. ACM, New York, NY, USA, 350–359. https://doi.org/10.1145/1031607.1031666

[20] Joel Chan, Steven Dang, and Steven P. Dow. 2016. Improving Crowd Innovation with Expert Facilitation. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 1223–1235. https://doi.org/10.1145/2818048.2820023

[21] Shuo Chang, Vikas Kumar, Eric Gilbert, and Loren G. Terveen. 2014. Specialization, Homophily, and Gender in a Social Curation Site: Findings from Pinterest. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 674–686. https://doi.org/10.1145/2531602.2531660

[22] Prerna Chikersal, Maria Tomprou, Young Ji Kim, Anita Williams Woolley, and Laura Dabbish. 2017. Deep Structures of Collaboration: Physiological Correlates of Collective Intelligence and Group Satisfaction. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 873–888. https://doi.org/10.1145/2998181.2998250

[23] Noshir Contractor. 2013. Some assembly required: leveraging Web science to understand and enable team assembly. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371, 1987 (2013), 20120385. https://doi.org/10.1098/rsta.2012.0385 arXiv:https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2012.0385

[24] Jonathon N. Cummings and Sara Kiesler. 2008. Who Collaborates Successfully?: Prior Experience Reduces Collaboration Barriers in Distributed Interdisciplinary Research. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW '08)*. ACM, New York, NY, USA, 437–446. https://doi.org/10.1145/1460563.1460633

[25] Laura Dabbish, Rosta Farzan, Robert Kraut, and Tom Postmes. 2012. Fresh faces in the crowd: turnover, identity, and commitment in online groups. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 245–248.

[26] Laura Dabbish, Colleen Stuart, Jason Tsay, and Jim Herbsleb. 2012. Social Coding in GitHub: Transparency and Collaboration in an Open Software Repository. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 1277–1286. https://doi.org/10.1145/2145204.2145396

[27] Laura A. Dabbish. 2008. Jumpstarting Relationships with Online Games: Evidence from a Laboratory Investigation. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW '08)*. ACM, New York, NY, USA, 353–356. https://doi.org/10.1145/1460563.1460620

[28] Joan DiMicco, David R. Millen, Werner Geyer, Casey Dugan, Beth Brownholtz, and Michael Muller. 2008. Motivations for Social Networking at Work. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW '08)*. ACM, New York, NY, USA, 711–720. https://doi.org/10.1145/1460563.1460674

[29] Paul Dourish and Victoria Bellotti. 1992. Awareness and Coordination in Shared Workspaces. In *Proceedings of the 1992 ACM Conference on Computer-supported Cooperative Work (CSCW '92)*. ACM, New York, NY, USA, 107–114. https://doi.org/10.1145/143457.143468

[30] Clarence Ellis and Jacques Wainer. 1994. A Conceptual Model of Groupware. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work (CSCW '94)*. ACM, New York, NY, USA, 79–88. https://doi.org/10.1145/192844.192878

[31] Susan R. Fussell, Robert E. Kraut, F. Javier Lerch, William L. Scherlis, Matthew M. McNally, and Jonathan J. Cadiz. 1998. Coordination, Overload and Team Performance: Effects of Team Communication Strategies. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work (CSCW '98)*. ACM, New York, NY, USA, 275–284. https://doi.org/10.1145/289444.289502

[32] Erving Goffman et al. 1978. *The presentation of self in everyday life.* Harmondsworth London.
[33] 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?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 659, 15 pages. https://doi.org/10.1145/3290605.3300889
[34] Mary L. Gray, Siddharth Suri, Syed Shoaib Ali, and Deepti Kulkarni. 2016. The Crowd is a Collaborative Network. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 134–147. https://doi.org/10.1145/2818048.2819942
[35] Jonathan Grudin. 1994. Computer-supported cooperative work: History and focus. *Computer* 27, 5 (1994), 19–26.
[36] Roger Guimerà, Brian Uzzi, Jarrett Spiro, and Luís A. Nunes Amaral. 2005. Team Assembly Mechanisms Determine Collaboration Network Structure and Team Performance. *Science* 308, 5722 (2005), 697–702. https://doi.org/10.1126/science.1106340 arXiv:http://science.sciencemag.org/content/308/5722/697.full.pdf
[37] Carl Gutwin and Saul Greenberg. 1998. Design for Individuals, Design for Groups: Tradeoffs Between Power and Workspace Awareness. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work (CSCW '98)*. ACM, New York, NY, USA, 207–216. https://doi.org/10.1145/289444.289495
[38] Joerg M. Haake, Anja Haake, Till Schümmer, Mohamed Bourimi, and Britta Landgraf. 2004. End-user Controlled Group Formation and Access Rights Management in a Shared Workspace System. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (CSCW '04)*. ACM, New York, NY, USA, 554–563. https://doi.org/10.1145/1031607.1031702
[39] 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
[40] Tamara Heck. 2013. Combining Social Information for Academic Networking. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. ACM, New York, NY, USA, 1387–1398. https://doi.org/10.1145/2441776.2441932
[41] Jonathan L. Herlocker, Joseph A. Konstan, and John Riedl. 2000. Explaining Collaborative Filtering Recommendations. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work (CSCW '00)*. ACM, New York, NY, USA, 241–250. https://doi.org/10.1145/358916.358995
[42] Pamela Hinds and Cathleen McGrath. 2006. Structures That Work: Social Structure, Work Structure and Coordination Ease in Geographically Distributed Teams. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW '06)*. ACM, New York, NY, USA, 343–352. https://doi.org/10.1145/1180875.1180928
[43] Pamela J Hinds, Kathleen M Carley, David Krackhardt, and Doug Wholey. 2000. Choosing work group members: Balancing similarity, competence, and familiarity. *Organizational behavior and human decision processes* 81, 2 (2000), 226–251.
[44] Daniel B. Horn, Thomas A. Finholt, Jeremy P. Birnholtz, Dheeraj Motwani, and Swapnaa Jayaraman. 2004. Six Degrees of Jonathan Grudin: A Social Network Analysis of the Evolution and Impact of CSCW Research. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (CSCW '04)*. ACM, New York, NY, USA, 582–591. https://doi.org/10.1145/1031607.1031707
[45] Julie S. Hui, Michael D. Greenberg, and Elizabeth M. Gerber. 2014. Understanding the Role of Community in Crowdfunding Work. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 62–74. https://doi.org/10.1145/2531602.2531715
[46] 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. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3201–3212. https://doi.org/10.1145/3025453.3026011
[47] Robert Johansen. 1991. Groupware: Future directions and wild cards. *Journal of Organizational Computing* 1, 2 (1991), 219–227. https://doi.org/10.1080/10919399109540160 arXiv:https://doi.org/10.1080/10919399109540160
[48] Benjamin F Jones, Stefan Wuchty, and Brian Uzzi. 2008. Multi-university research teams: Shifting impact, geography, and stratification in science. *science* 322, 5905 (2008), 1259–1262.
[49] Quentin Jones, Sukeshini A. Grandhi, Steve Whittaker, Keerti Chivakula, and Loren Terveen. 2004. Putting Systems into Place: A Qualitative Study of Design Requirements for Location-aware Community Systems. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (CSCW '04)*. ACM, New York, NY, USA, 202–211. https://doi.org/10.1145/1031607.1031640
[50] Brigitte Jordan. 2009. Blurring boundaries: The" real" and the" virtual" in hybrid spaces. *Human Organization* (2009), 181–193.
[51] Simon M. Kaplan, William J. Tolone, Douglas P. Bogia, and Celsina Bignoli. 1992. Flexible, Active Support for Collaborative Work with ConversationBuilder. In *Proceedings of the 1992 ACM Conference on Computer-supported Cooperative Work (CSCW '92)*. ACM, New York, NY, USA, 378–385. https://doi.org/10.1145/143457.143552

[52] Karrie Karahalios, Geraldine Fitzpatrick, and Andrés Monroy-Hernández. 2017. Editor's Note/Chairs' Welcome. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 16 (Dec. 2017), 1 pages. https://doi.org/10.1145/3134651

[53] Andrea Kavanaugh, John M Carroll, Mary Beth Rosson, Than Than Zin, and Debbie Denise Reese. 2005. Community networks: Where offline communities meet online. *Journal of Computer-Mediated Communication* 10, 4 (2005), JCMC10417.

[54] Charles Kiene, Aaron Shaw, and Benjamin Mako Hill. 2018. Managing Organizational Culture in Online Group Mergers. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 89 (Nov. 2018), 21 pages. https://doi.org/10.1145/3274358

[55] Young Ji Kim, David Engel, Anita Williams Woolley, Jeffrey Yu-Ting Lin, Naomi McArthur, and Thomas W. Malone. 2017. What Makes a Strong Team?: Using Collective Intelligence to Predict Team Performance in League of Legends. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 2316–2329. https://doi.org/10.1145/2998181.2998185

[56] Aniket Kittur, Susheel Khamkar, Paul André, and Robert Kraut. 2012. CrowdWeaver: Visually Managing Complex Crowd Work. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 1033–1036. https://doi.org/10.1145/2145204.2145357

[57] Aniket Kittur and Robert E. Kraut. 2008. Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW '08)*. ACM, New York, NY, USA, 37–46. https://doi.org/10.1145/1460563.1460572

[58] Aniket Kittur, Jeffrey V. Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The Future of Crowd Work. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. ACM, New York, NY, USA, 1301–1318. https://doi.org/10.1145/2441776.2441923

[59] Steve WJ Kozlowski and Daniel R Ilgen. 2006. Enhancing the effectiveness of work groups and teams. *Psychological science in the public interest* 7, 3 (2006), 77–124.

[60] Steve W. J. Kozlowski and Bradford S. Bell. 2003. *Work Groups and Teams in Organizations*. American Cancer Society, Chapter 14, 333–375. https://doi.org/10.1002/0471264385.wei1214 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/0471264385.wei1214

[61] Robert E. Kraut and Andrew T. Fiore. 2014. The Role of Founders in Building Online Groups. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 722–732. https://doi.org/10.1145/2531602.2531648

[62] Klaus Krippendorff. 2011. Computing Krippendorff's alpha-reliability. (2011).

[63] Shyong K. Lam, Jawed Karim, and John Riedl. 2010. The Effects of Group Composition on Decision Quality in a Social Production Community. In *Proceedings of the 16th ACM International Conference on Supporting Group Work (GROUP '10)*. ACM, New York, NY, USA, 55–64. https://doi.org/10.1145/1880071.1880083

[64] Cliff Lampe, Nicole Ellison, and Charles Steinfield. 2006. A Face(Book) in the Crowd: Social Searching vs. Social Browsing. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW '06)*. ACM, New York, NY, USA, 167–170. https://doi.org/10.1145/1180875.1180901

[65] Airi Lampinen and Barry Brown. 2017. Market Design for HCI: Successes and Failures of Peer-to-Peer Exchange Platforms. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4331–4343. https://doi.org/10.1145/3025453.3025515

[66] J Richard Landis and Gary G Koch. 1977. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* (1977), 363–374.

[67] Susan Leavy. 2018. Gender Bias in Artificial Intelligence: The Need for Diversity and Gender Theory in Machine Learning. In *Proceedings of the 1st International Workshop on Gender Equality in Software Engineering (GE '18)*. ACM, New York, NY, USA, 14–16. https://doi.org/10.1145/3195570.3195580

[68] Charlotte P. Lee and Drew Paine. 2015. From The Matrix to a Model of Coordinated Action (MoCA): A Conceptual Framework of and for CSCW. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM, New York, NY, USA, 179–194. https://doi.org/10.1145/2675133.2675161

[69] Danielle Levac, Heather Colquhoun, and Kelly K. O'Brien. 2010. Scoping studies: advancing the methodology. *Implementation Science* 5, 1 (20 Sep 2010), 69. https://doi.org/10.1186/1748-5908-5-69

[70] John M Levine and Richard L Moreland. 1998. Small groups. (1998).

[71] Alina Lungeanu, Dorothy R. Carter, Leslie A. DeChurch, and Noshir S. Contractor. 2018. How Team Interlock Ecosystems Shape the Assembly of Scientific Teams: A Hypergraph Approach. *Communication Methods and Measures* 12, 2-3 (2018), 174–198. https://doi.org/10.1080/19312458.2018.1430756 arXiv:https://doi.org/10.1080/19312458.2018.1430756

[72] Ioanna Lykourentzou, Angeliki Antoniou, Yannick Naudet, and Steven P. Dow. 2016. Personality Matters: Balancing for Personality Types Leads to Better Outcomes for Crowd Teams. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 260–273. https://doi.org/10.1145/2818048.2819979

[73] Ioanna Lykourentzou, Robert E. Kraut, and Steven P. Dow. 2017. Team Dating Leads to Better Online Ad Hoc Collaborations. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 2330–2343. https://doi.org/10.1145/2998181.2998322

[74] Elizabeth A Mannix. 1993. Organizations as resource dilemmas: The effects of power balance on coalition formation in small groups. *Organizational Behavior and Human Decision Processes* 55, 1 (1993), 1–22.

[75] Jennifer Marlow and Laura Dabbish. 2013. Activity Traces and Signals in Software Developer Recruitment and Hiring. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. ACM, New York, NY, USA, 145–156. https://doi.org/10.1145/2441776.2441794

[76] Jennifer Marlow, Laura Dabbish, and Jim Herbsleb. 2013. Impression Formation in Online Peer Production: Activity Traces and Personal Profiles in Github. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. ACM, New York, NY, USA, 117–128. https://doi.org/10.1145/2441776.2441792

[77] David W. McDonald and Mark S. Ackerman. 2000. Expertise Recommender: A Flexible Recommendation System and Architecture. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work (CSCW '00)*. ACM, New York, NY, USA, 231–240. https://doi.org/10.1145/358916.358994

[78] Joseph E. McGrath. 1992. Groups Interacting with Technology: The Complex and Dynamic Fit of Group, Task, Technology, and Time. In *Proceedings of the 1992 ACM Conference on Computer-supported Cooperative Work (CSCW '92)*. ACM, New York, NY, USA, 4–. https://doi.org/10.1145/143457.371588

[79] Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 27, 1 (2001), 415–444.

[80] Robert C. Miller, Haoqi Zhang, Eric Gilbert, and Elizabeth Gerber. 2014. Pair Research: Matching People for Collaboration, Learning, and Productivity. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 1043–1048. https://doi.org/10.1145/2531602.2531703

[81] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, Prisma Group, et al. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine* 6, 7 (2009), e1000097.

[82] Richard L Moreland. 1987. The formation of small groups. (1987).

[83] Meredith Ringel Morris, Jeffrey P. Bigham, Robin Brewer, Jonathan Bragg, Anand Kulkarni, Jessie Li, and Saiph Savage. 2017. Subcontracting Microwork. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1867–1876. https://doi.org/10.1145/3025453.3025687

[84] Mark Mortensen and Martine R. Haas. 2018. Perspective–Rethinking Teams: From Bounded Membership to Dynamic Participation. *Organization Science* 29, 2 (2018), 341–355. https://doi.org/10.1287/orsc.2017.1198 arXiv:https://doi.org/10.1287/orsc.2017.1198

[85] Michael Muller, Mary Keough, John Wafer, Werner Geyer, Alberto Alvarez Saez, David Leip, and Cara Viktorov. 2016. Social Ties in Organizational Crowdfunding: Benefits of Team-Authored Proposals. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 1246–1259. https://doi.org/10.1145/2818048.2819955

[86] Bonnie Nardi and Justin Harris. 2006. Strangers and Friends: Collaborative Play in World of Warcraft. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW '06)*. ACM, New York, NY, USA, 149–158. https://doi.org/10.1145/1180875.1180898

[87] Bonnie A. Nardi, Diane J. Schiano, and Michelle Gumbrecht. 2004. Blogging As Social Activity, or, Would You Let 900 Million People Read Your Diary?. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (CSCW '04)*. ACM, New York, NY, USA, 222–231. https://doi.org/10.1145/1031607.1031643

[88] Daniel J. Niven, Kelly J. Mrklas, Jessalyn K. Holodinsky, Sharon E. Straus, Brenda R. Hemmelgarn, Lianne P. Jeffs, and Henry Thomas Stelfox. 2015. Towards understanding the de-adoption of low-value clinical practices: a scoping review. *BMC Medicine* 13, 1 (06 Oct 2015), 255. https://doi.org/10.1186/s12916-015-0488-z

[89] Judith S. Olson and Stephanie Teasley. 1996. Groupware in the Wild: Lessons Learned from a Year of Virtual Collocation. In *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work (CSCW '96)*. ACM, New York, NY, USA, 419–427. https://doi.org/10.1145/240080.240353

[90] Wanda J. Orlikowski. 1993. Learning from notes: Organizational issues in groupware implementation. *The Information Society* 9, 3 (1993), 237–250. https://doi.org/10.1080/01972243.1993.9960143 arXiv:https://www.tandfonline.com/doi/pdf/10.1080/01972243.1993.9960143

[91] Gérald Oster, Pascal Urso, Pascal Molli, and Abdessamad Imine. 2006. Data Consistency for P2P Collaborative Editing. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW '06)*. ACM, New York, NY, USA, 259–268. https://doi.org/10.1145/1180875.1180916

[92] Manoj Parameswaran and Andrew B Whinston. 2007. Social computing: An overview. *Communications of the Association for Information Systems* 19, 1 (2007), 37.

[93] D. Pinelle and C. Gutwin. 2000. A review of groupware evaluations. In *Proceedings IEEE 9th International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE 2000)*. 86–91. https://doi.org/10.1109/

ENABL.2000.883709

[94] Bernd Ploderer, Steve Howard, and Peter Thomas. 2008. Being Online, Living Offline: The Influence of Social Ties over the Appropriation of Social Network Sites. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW '08)*. ACM, New York, NY, USA, 333–342. https://doi.org/10.1145/1460563.1460618

[95] Yuqing Ren, Robert Kraut, and Sara Kiesler. 2007. Applying common identity and bond theory to design of online communities. *Organization studies* 28, 3 (2007), 377–408.

[96] Daniela Retelny, Sébastien Robaszkiewicz, Alexandra To, Walter S. Lasecki, Jay Patel, Negar Rahmati, Tulsee Doshi, Melissa Valentine, and Michael S. Bernstein. 2014. Expert Crowdsourcing with Flash Teams. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 75–85. https://doi.org/10.1145/2642918.2647409

[97] Catherine M Ridings and David Gefen. 2004. Virtual community attraction: Why people hang out online. *Journal of Computer-mediated communication* 10, 1 (2004), JCMC10110.

[98] 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

[99] Niloufar Salehi, Andrew McCabe, Melissa Valentine, and Michael Bernstein. 2017. Huddler: Convening Stable and Familiar Crowd Teams Despite Unpredictable Availability. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 1700–1713. https://doi.org/10.1145/2998181.2998300

[100] Niloufar Salehi, Jaime Teevan, Shamsi Iqbal, and Ece Kamar. 2017. Communicating Context to the Crowd for Complex Writing Tasks. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 1890–1901. https://doi.org/10.1145/2998181.2998332

[101] Kjeld Schmidt. 1994. The Organization of Cooperative Work: Beyond the "Leviathan" Conception of the Organization of Cooperative Work. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work (CSCW '94)*. ACM, New York, NY, USA, 101–112. https://doi.org/10.1145/192844.192883

[102] Doug Schuler. 1994. Social computing. *Commun. ACM* 37, 1 (1994), 28–29.

[103] Joseph Seering, Felicia Ng, Zheng Yao, and Geoff Kaufman. 2018. Applications of Social Identity Theory to Research and Design in Computer-Supported Cooperative Work. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 201 (Nov. 2018), 34 pages. https://doi.org/10.1145/3274771

[104] Markus Sohlenkamp and Greg Chwelos. 1994. Integrating Communication, Cooperation, and Awareness: The DIVA Virtual Office Environment. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work (CSCW '94)*. ACM, New York, NY, USA, 331–343. https://doi.org/10.1145/192844.193041

[105] Igor Steinmacher, Tayana Conte, Marco Aurélio Gerosa, and David Redmiles. 2015. Social Barriers Faced by Newcomers Placing Their First Contribution in Open Source Software Projects. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM, New York, NY, USA, 1379–1392. https://doi.org/10.1145/2675133.2675215

[106] David Tranfield, David Denyer, and Palminder Smart. 2003. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management* 14, 3 (2003), 207–222. https://doi.org/10.1111/1467-8551.00375 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-8551.00375

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

[108] John C Turner, Michael A Hogg, Penelope J Oakes, Stephen D Reicher, and Margaret S Wetherell. 1987. *Rediscovering the social group: A self-categorization theory.* Basil Blackwell.

[109] Brian Uzzi, Satyam Mukherjee, Michael Stringer, and Ben Jones. 2013. Atypical combinations and scientific impact. *Science* 342, 6157 (2013), 468–472.

[110] Melissa A. Valentine, Daniela Retelny, Alexandra To, Negar Rahmati, Tulsee Doshi, and Michael S. Bernstein. 2017. Flash Organizations: Crowdsourcing Complex Work by Structuring Crowds As Organizations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3523–3537. https://doi.org/10.1145/3025453.3025811

[111] Jennifer Wortman Vaughan. 2017. Making Better Use of the Crowd: How Crowdsourcing Can Advance Machine Learning Research. *J. Mach. Learn. Res.* 18, 1 (Jan. 2017), 7026–7071. http://dl.acm.org/citation.cfm?id=3122009.3242050

[112] Amy Voida, Sheelagh Carpendale, and Saul Greenberg. 2010. The Individual and the Group in Console Gaming. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10)*. ACM, New York, NY, USA, 371–380. https://doi.org/10.1145/1718918.1718983

[113] Jacques Wainer and Claudia Barsottini. 2007. Empirical research in CSCW — a review of the ACM/CSCW conferences from 1998 to 2004. *Journal of the Brazilian Computer Society* 13, 3 (01 Sep 2007), 27–35. https://doi.org/10.1007/BF03192543

[114] James R. Wallace, Saba Oji, and Craig Anslow. 2017. Technologies, Methods, and Values: Changes in Empirical Research at CSCW 1990 - 2015. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 106 (Dec. 2017), 18 pages.

https://doi.org/10.1145/3134741

[115] X. Wang, Z. Zhao, and W. Ng. 2016. USTF: A Unified System of Team Formation. *IEEE Transactions on Big Data* 2, 1 (March 2016), 70–84. https://doi.org/10.1109/TBDATA.2016.2546303

[116] Amy Wax, Leslie A. DeChurch, and Noshir S. Contractor. 2017. Self-Organizing Into Winning Teams: Understanding the Mechanisms That Drive Successful Collaborations. *Small Group Research* 48, 6 (2017), 665–718. https://doi.org/10.1177/1046496417724209 arXiv:https://doi.org/10.1177/1046496417724209

[117] Miaomiao Wen, Keith Maki, Steven Dow, James D. Herbsleb, and Carolyn Rose. 2017. Supporting Virtual Team Formation Through Community-Wide Deliberation. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 109 (Dec. 2017), 19 pages. https://doi.org/10.1145/3134744

[118] Mark E. Whiting, Dilrukshi Gamage, Snehalkumar (Neil) S. Gaikwad, Aaron Gilbee, Shirish Goyal, Alipta Ballav, Dinesh Majeti, Nalin Chhibber, Angela Richmond-Fuller, Freddie Vargus, Tejas Seshadri Sarma, Varshine Chandrakanthan, Teogenes Moura, Mohamed Hashim Salih, Gabriel Bayomi Tinoco Kalejaiye, Adam Ginzberg, Catherine A. Mullings, Yoni Dayan, Kristy Milland, Henrique Orefice, Jeff Regino, Sayna Parsi, Kunz Mainali, Vibhor Sehgal, Sekandar Matin, Akshansh Sinha, Rajan Vaish, and Michael S. Bernstein. 2017. Crowd Guilds: Worker-led Reputation and Feedback on Crowdsourcing Platforms. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 1902–1913. https://doi.org/10.1145/2998181.2998234

[119] Anita Williams Woolley, Christopher F. Chabris, Alex Pentland, Nada Hashmi, and Thomas W. Malone. 2010. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science* 330, 6004 (2010), 686–688. https://doi.org/10.1126/science.1193147 arXiv:http://science.sciencemag.org/content/330/6004/686.full.pdf

[120] Svetlana Yarosh, Tara Matthews, Michelle Zhou, and Kate Ehrlich. 2013. I Need Someone to Help!: A Taxonomy of Helper-finding Activities in the Enterprise. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. ACM, New York, NY, USA, 1375–1386. https://doi.org/10.1145/2441776.2441931

[121] Stephen J Zaccaro and Paige Bader. 2003. E-leadership and the challenges of leading e-teams: Minimizing the bad and maximizing the good. *Organizational dynamics* (2003).

[122] Sharon Zhou, Melissa Valentine, and Michael S. Bernstein. 2018. In Search of the Dream Team: Temporally Constrained Multi-Armed Bandits for Identifying Effective Team Structures. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 108, 13 pages. https://doi.org/10.1145/3173574.3173682

[123] Haiyi Zhu, Steven P. Dow, Robert E. Kraut, and Aniket Kittur. 2014. Reviewing Versus Doing: Learning and Performance in Crowd Assessment. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 1445–1455. https://doi.org/10.1145/2531602.2531718

[124] Doug Zytko, Guo Freeman, Sukeshini A. Grandhi, Susan C. Herring, and Quentin (Gad) Jones. 2015. Enhancing Evaluation of Potential Dates Online Through Paired Collaborative Activities. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM, New York, NY, USA, 1849–1859. https://doi.org/10.1145/2675133.2675184

Received April 2019; revised June 2019; accepted August 2019