Does the transition to online meetings exacerbate or alleviate gender and racial inequities in student participation in project teams? Machine-learning-based analysis of online collaborative work during the pandemic

In contrast to popular beliefs that online meetings during the pandemic provided equal opportunities for participation to otherwise marginalized populations, whether these meetings exacerbated or alleviated gender and racial inequities or not remains unknown. The lack of intelligent tools that enable unobtrusive observations of complex project teams in diverse fields is considered as a major obstacle to the investigations into the impact of online transition on collaborative work. To fill this important gap in the literature, we examined, during the transition to online meetings over 14 months from March 2020 to April 2021, whether an online meeting format altered students' participation in 17 project teams in various disciplines, and whether this effect differed for historically marginalized groups, such as female and racial minorities. Second, we analyzed whether and how an individual's contributions to teamwork, as perceived by other team members, changed during the transition, and whether this change affected historically marginalized groups disproportionally. To do so, we built a machine learning pipeline for speech diarization, speaker recognition, and transcription, as well as supervised learning classification of members' roles (i.e., giving information, asking information, and others). In addition to the results produced by this pipeline, we also collected 101 students' responses from 34 surveys. Our results indicated that both female students and minority students were affected by the transition to online meetings due to the COVID-19 pandemic. However, these groups were not all affected alike. Female students have participated equally as they negotiated their roles with peers at the beginning of the pandemic, but their participation rapidly decreased as the new norms settled. In contrast, racial minority groups were heavily affected in the beginning of the transition, but these negative impacts decreased as they became acclimated to new norms and settled with new roles. As such, future solutions should continue to be sensitive to the diversity of historically marginalized populations and should integrate the idea that a multitude of interventions might be necessary to increase equity amongst culturally distinct subpopulations. This study contributes to advancing generalizable scientific knowledge about how to support diverse groups of individuals to work in complex social systems and create a more inclusive environment to reduce inequities. In addition, this research demonstrates how future researchers can utilize machine learning algorithm to automate the diagnoses of problems in project teams.

CCS Concepts: • **Applied Computing**→ Law, social and behavioral sciences; Education; • **Human-centered computing** → Human computer interaction; Collaborative and social computing

KEYWORDS: Project teams, online meetings, virtual meetings, video conference, natural language processing, student participation, females, minorities, COVID

1 INTRODUCTION

As complex, large-scale projects are becoming more prevalent, it is increasingly important for team members to make high-quality contributions to complex, interdependent, and collaborative work so that the group can achieve optimal decisions, performance, and intended goals. However, making meaningful contributions to groupwork has become more challenging due to the abrupt transition to online meetings during the COVID-19 pandemic, which upended the norms for collaborative work. This imposed transition to online meeting spaces required team members to negotiate their roles with peers to establish new norms for collaboration. Consequently, researchers need to identify how group dynamics have changed in the transition to online meetings. This diagnosis of what has changed will be essential for identifying individuals who need further support for online cooperative work and aid in the development of technical interventions to increase the equity of participation in such projects.

Current research has yet to fully explicate how the new norms for collaborative work have changed, or perhaps not changed, group dynamics for individuals, including those with diverse backgrounds. This lack of research is largely attributed to the absence of intelligent tools for unobtrusively collecting participation data (e.g., speech). Specifically, many extant studies have relied solely on self-reported measures (Alharbi et al., 2021; Kaptelinin et al., 2021; Karl et al., 2021; Bayhan et al., 2022; Standaert and Thunus, 2022), which are known to suffer from response biases. Self-response bias hinders the fair observations of individuals from historically marginalized groups. In addition, extant studies have examined a small number of teams in only one field or discipline, thus hindering the generalization of findings to diverse types of collaborations in various areas. Finally, extant studies have observed group dynamics in a short period of time, often ranging from weeks to a month or two, which is not enough to capture the changes in group dynamics over the span of the pandemic.

To fill this important gap, we examine, during the transition to online meetings (over the 14 months from March 2020 to April 2021), whether an online meeting format alters students' participation in project teams in various disciplines, and whether this effect differs for historically marginalized groups. Second, we analyze whether and how an individual's contributions to teamwork, as *perceived* by other team members, have changed during the transition, and whether this change affects historically marginalized groups disproportionally.

To achieve these goals, we collected speech from online meetings and accompanied them with the surveys of 17 undergraduate and graduate student project teams in social sciences and STEM majors over four semesters. These data were collected within an organic setting as a part of their class projects—covering the onset, peaks, and troughs of changes due to COVID-19. Team meetings were video or audio recordings, which were then analyzed by machine-learning methods. Our machine-learning-based analytical pipeline includes, but is not limited to, (i) speech diarization, speaker recognition, and transcription, and (ii) classification of roles team members played in groupwork (i.e., giving information, asking information and others). This analytical pipeline streamlines data flows to improve the speed and quality of data analyses and, thus, facilitates the translation of results for the attainment of our research objectives. In addition, for every project team in each semester, we conducted two surveys—(i) an entry survey to collect the students' demographic factors and (ii) a final survey to evaluate their perceived contributions to project teams. As such, we conducted a total of 34 surveys over 14 months.

This study's examination of group dynamics over 14 months, observed with intelligent tools as opposed to self-reported measures, demonstrates how the transition to online meetings has affected the existing inequities of participation in projects. In so doing, we lay groundwork to support diverse groups of individuals to work in complex social systems (Cross et al., 2010; Bayhan et al., 2022; Karl et al., 2022) and to create a more inclusive environment to reduce inequities (Wu et al., 2022; Bagmar et al., 2022, Do et al., 2022, Kim et al., 2021). In addition, this research demonstrates how future researchers can utilize machine learning methods to automate the diagnoses of problems in project teams (Mikolov, 2013), and demonstrate social implications of equity in group work (Garicano and Rossi-Hansberg, 2015; Deming, 2017; Garicano, 2020).

2 LITERATURE REVIEW

Computer-Supported Cooperative Work (CSCW) research focuses on how peers and teams collaborate and how technology facilitates that collaboration. Included in the CSCW focus is a strand of literature that examines the effectiveness and affordances of online meetings compared to offline meetings. Prior to the transformation of work and school to online modalities due to the Covid-19 pandemic, online meetings had largely been studied in comparison to offline meetings. This was often in groups that collaborated through a hybrid format of both online and offline interactions. Pre-pandemic studies focused on collaboration via online and technology-enabled meetings in education (Kinnula et al., 2018; Willermark and Pareto, 2020), healthcare (Islind et al., 2019; Constantinides, 2011), and co-authored writing (Larsen-Ledet and Korsgarrd, 2019), among many other settings. In the context of educational collaboration, seven positive affordances of online meetings previously identified are that participants (1) engage in a joint task, (2) communicate, (3) share resources, (4) engage in productive collaborative learning processes, (5) engage in co-construction, (6) monitor and regulate collaborative learning, and (7) find and build groups and communities" (Jeong and Hmelo-Silver, 2016).

However, other studies found that online meetings exacerbated biases that already occur during in-person meetings. A recent study that proposed video conferencing etiquette to promote gender equity (Dhawan et al.,

2021) found that pre-pandemic, many women faced difficulty being heard in in-person meetings with male peers and struggled to say as much or more than male peers during these meetings. Dhawan et al. (2021) went on to discuss how many of those gender inequities were likely to transition into digital meetings as well. Other studies have found similar inequities for both racial and LGBTQ-identified minorities (Houtti et al., 2022). Likewise, during online meetings, there are significant mismatches in native English speakers' attributions of non-native speakers' behaviors, but no significant mismatch exists in non-native speakers' attributions of native speakers' behaviors (He et al., 2017). It was also found that non-native speakers were only able to engage in "compromised" impression management during the task (He et al., 2017). In a similar vein, intercultural conflict is found in global virtual teams and is shown to negatively affect communication and project results (He et al., 2017), while team diversity is known to facilitate creativity (Ye and Robert Jr, 2017).

Another strand of CSCW literature centers on the learning effects between two meeting modalities during the pandemic. This strand discusses how the transition to an online learning paradigm affects students' participation in learning processes and, thus outcomes (Chen et al., 2021; Chen et al., 2021; Ravi et al., 2021; Lacy et al., 2022; Manshaei et al., 2022; Gui et al., 2022). Most of these studies have shown sustained disadvantages of online meetings in higher education compared to offline, co-located meetings (Kisworo et al. 2022), as shown pre-pandemic. For instance, online meetings do not effectively support hands-on activities (Labrie et al., 2022) and the experience of "we-ness" (Kaptelinin et al., 2021) because online meetings make it harder to engage in private chats (Guo et al., 2022), while the lack of structure in private chats is not conducive to reaching consensus (Kim et al., 2021). In addition, using text-based applications (e.g., Slack) for distributed teams decreases the ability to perceive, understand, and regulate emotions between team members (Benke et al., 2021).

Additionally, these studies have shown that the disadvantages are exacerbated for historically marginalized sub-populations such as females and cultural minorities. For instance, female learners perform better than males in preparatory, performance, and appraisal phases in online self-regulated learning (Liu et al., 2021), but there is a lack of inclusion for potential female participants and wider community support in STEM to solve the lack of female students in STEM majors (MacArthur et al., 2021). Also, language and cultural barriers negatively affect minorities in terms of their ability to participate in communication and group tasks (He et al., 2017a, 2017b).

Summary of the existing literature and the identified gap: Both before and during the pandemic, prior research has shown that online meetings have more disadvantages than offline, co-located meetings. These disadvantages are often exacerbated for women and racial minority groups. And while these studies expose where some progress needs to be made to increase inclusion of historically marginalized groups, these studies are limited by data and in perspective, as they do not investigate how the changes in meeting norms during the pandemic interact with the participation of historically marginalized groups, *over time*. However, it is likely that the abrupt transition from offline to online modality during the early pandemic in 2020—2021 has created new norms for online meetings. What is not yet understood is in what ways this changing norm interacts with gender and race in their negotiations of roles in groupwork, and this insight is an important topic that has not been done. Especially, deeply ingrained social roles for women and racial minorities may remain intact regardless of the modality changes or could have overcome as minority members become familiarized with the new meeting norms and established their new roles, but existing studies have not yet answered this important question. To fill these important gaps in the current literature, we propose the following research questions.

[**RQ1**] Does female participation change over time during the transition to online meetings? Do they participate more or less? What roles do females play in online cooperative work? Do these roles change as new norms arise as online meetings become the sole meeting modality?

[**RQ2**] Does racial minority groups' participation change during the transition to online meetings? Do they participate more or less? What roles do racial minority participants play in online cooperative work? Do these roles change as new norms are negotiated as online meetings become the sole meeting modality?

[**RQ3**] Do other members' perceptions of women and racial minority members' contributions to project teams change during the transition to online meetings? If so, is this change heterogenous by gender and race, reflecting the existing stereotypes and biases?

Alleviating subjectivity is important when examining the above research questions as the researchers and participants alike can be subject to pervasive stereotypes and implicit biases. For instance, a prior study has shown that teachers have continuously evaluated white male students' participation more positively than other male students and female students, especially in STEM-related subjects (Copur-Genturk et al., 2020). Thus, using a more objective tool that can minimize the researchers' and the respondents' biases and stereotypes is essential, but prior studies have not yet done so. Therefore, we propose our exploratory question.

[**RQ4**] Does unobstructive data collection in organic settings facilitate the observations of members' participation and roles that change dynamically over time?

3 MACHINE LEARNING METHODS

We employed an innovative machine learning pipeline to transform original video and audio files to intuitive tabulated data, which can aid in the early detection of communication and decision problems in teams (Arum and Roska, 2011). Specifically, our method conducts speech diarization, speaker recognition, transcription, and classification of participation (giving information, asking information and others). Next, we introduce the corresponding components in detail.

3.1 Speech Diarization, Speaker Recognition, and Transcription

The speech diarization module splits audio signals based on speaker identity. It first partitions the audio into small segments, each of which ideally contains a variable-length utterance from a unique speaker. Then, by comparing the audio signals in each segment, the module determines the number of speakers in a conversation and judges which speaker each segment belongs to. To perform this step, we applied a deep-learning-based speaker diarization model released in the Kaldi speech recognition toolkit (Povey et al., 2011). The model first encodes utterances to fixed-dimensional embeddings called "x-vectors" (Snyder et al., 2018), then performs agglomerative hierarchical clustering on x-vectors to produce an initial diarization output. Finally, a variational Bayes Hidden Markov model is applied over x-vectors to improve the diarization results. In our implementation, we use the pre-trained model from the Kaldi project (2021) and run it in ONNXRuntime (ONNXRuntime, 2021).

Following the speech diarization, the speech recognition module then transcribes speakers' utterances in audio into text. After a meeting, an audio file is split by the speaker diarization module into segments, each segment is passed through the speech recognition module to get the utterance in text format. This speech recognition module is then followed by the transcription module. In this module, the utterance is then transcribed into text over three steps: (i) acoustic feature extraction, where an acoustic model is applied to convert raw audio signals into acoustic features; (ii) word selection, which chooses candidate words according to the acoustic features from the system's dictionary; and (iii) sentence-level matching, where a language model is applied to determine the final words based on their contexts. In our implementation, we use the pre-trained ASpIRE chain model (Povey, 2021) from the Kaldi project.

3.2 Classification of Participants' Roles

Following the transcription module, the classification module groups the transcribed text from speakers into three categories: giving information (G), asking information (A), and others (O). This module aims to determine whether the transcribed text is giving information to others, asking for information from others, or neither (called "others"), to help us understand members' roles in project teams. This categorization has been frequently used in prior studies in CSCW that examined how members exchange information and collaborate to obtain a shared goal (Liang et al., 2021).

For this classification task, we trained a deep-text classifier. Specifically, a 1-layer recurrent neural network (RNN) with gated recurrent units (GRU) is used as the text classification model. The size of word embeddings is set as 50, and the hidden size of the RNN is 128. The model is trained on 932 training instances manually labeled by human coders with a batch size of 64 for 20 epochs. As shown in Table 1, our classifier had an acceptable accuracy of 84.12, F1-macro of 80.70 and F1-micro for 84.12.

Table 1: The Confusion Matrix of RNN

		Predicted Labels		
		G	A	O
True Labels	G	119	9	13
	A	2	51	0
	O	10	3	26

4 DATA

4.1 Data Collection

We collected data at a large Big Ten public university during four semesters: Spring 2020, Summer 2020, Fall 2020, and Spring 2021, all of which were required to be all or partially online by the University. Table 2 summarizes course modalities across different semesters. The sample covers 101 students assigned to 17 teams in nine sections of seven courses. Table 3 summarizes the semester and course for each team, the number of members in each team, the number of videos and audios recorded for each team, the average length of these recordings, and the time frame of the project. Social science classes encompass economics or human resources and labor relations; technical classes include computer science or engineering classes; hybrid classes consist of both social and technical components.

Table 2: The University's course modality policies

Norm negotiation phase	Spring 2020	January 2020 to May 2020 On March 11, 2020, the university switched all classes to online meetings.
	Summer 2020	May 2020 to August 2020 All classes were online.
Post new-norm phase	Fall 2020	August 2020 to December 2020 All classes were online (announced before the beginning of Fall semester).
	Spring 2021	January 2021 to May 2021 Most classes were online, with some in-person classes (not in our sample).

Table 3: Summary of project teams

Semester	Team ID	Course type	Team size	number of videos	Average length (minutes)	Video and audio recording date range
2020 spring	A	Social	6	3	25	03/27/2020-04/24/2020
2020 spring	C	Social	8	2	33	04/16/2020-04/27/2020
2020 spring	D	Social	8	4	21	03/27/2020-04/24/2020
2020 spring	F	Technical	9	3	39	04/09/2020-04/20/2020
2020 summer	H	Hybrid	5	5	46	07/14/2020-08/01/2020
2020 fall	I	Hybrid	4	13	43	10/02/2020-11/27/2020
2020 fall	J	Social	5	3	42	11/03/2020-11/19/2020
2020 fall	K	Social	5	7	53	10/30/2020-11/24/2020
2020 fall	L	Social	5	3	44	11/03/2020-11/17/2020
2020 fall	M	Social	5	5	26	10/21/2020-11/20/2020
2020 fall	N	Social	4	3	29	11/17/2020-11/29/2020
2020 fall	O	Social	4	3	12	11/18/2020-12/04/2020

2020 fall	P	Technical	6	7	17	11/06/2020-12/11/2020
2020 fall	Q	Technical	5	2	37	10/23/2020-11/17/2020
2020 fall	R	Technical	6	7	22	11/03/2020-12/02/2020
2021 spring	S	Social	5	8	22	03/04/2021-04/07/2021
2021 spring	T	Social	5	3	23	03/10/2021-03/29/2021
Total	17		101	81	32	03/27/2020-04/07/2021

4.2 Data Processing via Machine Learning Pipeline

The raw dataset contains video and audio recordings of meetings and two surveys for each of 101 participants.[1] We then utilized the machine-learning pipeline to process the raw audio and video.

4.2.1 Variables

Table 4 presents the definitions of variables included in this study. The main independent variable of interest is the meeting norm (henceforth, "norm"). It is a dummy variable that equals 0 if the sample is from Spring 2020 and Summer 2020 semesters, or the "pre-norm negotiation phase," and 1 is for Fall and Spring 2021, or the post -norm phase. Empirical evidence showed that most of the participants taking the Summer 2020 course in the sample were visiting scholars and exposed to online modality for course interactions for the first time in this semester. Thus, Spring and Summer 2020 participants in the dataset were included in the norm negotiation phase. Many control variables were selected to control for demographic characteristics and course information. Female and minority are dummy variables representing gender and race, and they are the main independent variables of interest. The average number of members in each team is 5.64. The average cumulative grade point average (i.e., GPA) in the sample is 3.59 out of 4.00, equivalent to an A in the alphabetic grading system. The sample contains roughly equal numbers of undergraduate courses and graduate courses.

Table 4: Summary Statistics

	definition	mean	Standard Deviations (SDs)	min	max
pdur	Percent of speak time duration	18.71	20.06	0	99
dur	Minutes of speak time duration	5.91	8.73	0	88
givepdur	Percent of time giving information	15.45	17.03	0	99
askpdur	Percent of time asking for information	2.86	4.60	0	33
norm	1 if the post new-norm phase	0.73	0.45	0	1
female	1 if female	0.50	0.50	0	1
minority	1 if ethnicity is Asian, Black, or Hispanic	0.44	0.50	0	1
size	Size of the team	5.64	1.41	4	9
GPA	Grade Point Average	3.59	0.61	0	4
technical	1 if project contains technical materials	0.46	0.50	0	1
social	1 if project contains social science materials	0.72	0.45	0	1
hybrid	1 if project is both technical and social	0.18	0.38	0	1
undergraduate	1 if undergraduate student	0.42	0.49	0	1

Notes: The unit of observation is meeting-individual, and the sample consists of 433 such meeting-individual data points.

4.2.2 Participation Measure

To measure students' participation, we employed *pdur,* which is the percentage of speaking time (out of 100) in a meeting. Each student speaks 18.71% of the total meeting time on average.

4.2.3 Measure for the Perceived Contribution to Teams

Students were surveyed twice throughout their project: once in the beginning of the project (called the "entry

[1] 9 out of 101 students did not complete the exit survey.

survey") and the other at the end of the project (called "final survey"). In each survey, students were asked to evaluate their own and teammates' contributions to teamwork. This perceived contribution measure focuses on twelve aspects that are essential to the success of group projects, and they are listed in Table 5. Each measure employs a 4-point Likert scale, with 1 being strongly disagree, and 4 being strongly agree. To eliminate the potential bias of self-evaluation, we used the average of evaluation scores from the other team members as the individual's final score. For example, for a team containing J total members, the average perceived contribution of question $l21$ for the first team member will be

$$l21 = \frac{\sum_{j \neq 1} l21_{\mathrm{j}}}{J-1}.$$

The average of these 12 measures is then generated the overall contribution score.

Table 5: Perceived Contributions to Teams

Notation	Description
Proficiency (adopted from Kaufman et al. 1999, Ohland et al. 2012 and Mentzer et al. 2017)	
$l21$	Completed his/her tasks with the expected quality
$l22$	Completed his/her tasks to achieve the overall project goals
$l23$	Collaborated with team members from other disciplines to achieve the project goals
$l24$	Provided information to team members from other disciplines when needed
Adaptivity (adopted from Raelin et al. 2011 and Mentzer et al. 2017)	
$l31$	Adapted well to changes in the project that altered how he/she completed his/her task
$l32$	Coped with unforeseen demands placed on him/her
$l33$	Understood well how to deal with changes in the project to achieve the project goals
$l34$	Coped with changes to the project that altered how members from different disciplines collaborated on project goals
Proactivity (adopted from Raelin et al. 2011 and Mentzer et al. 2017)	
$l41$	Contributed new, improved ways to develop his/her tasks
$l42$	Initiated changes to the ways in which his/her tasks were done that helped accomplish project goals
$l43$	Created innovative solutions to improve the project quality
$l44$	Developed alternative solutions to achieve the project goals ahead of time

5. RESULTS

5.1 Descriptive Statistics

Figure 1 demonstrates total meeting lengths of all teams in different semesters.

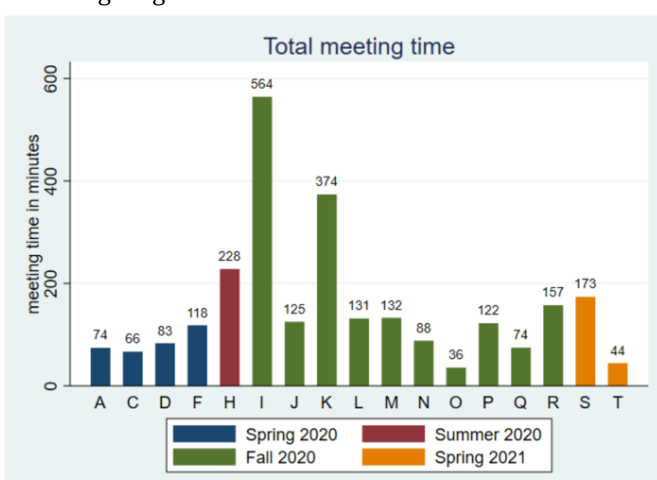

Figure 1: Total meeting lengths of all teams in different semesters

Since total meeting length is affected by meeting occurrence, a better metric to compare meeting lengths is the average duration per meeting (Figure 2). Figure 2 however shows no evidence that average meeting time differs in the norm-negotiation phase and the post new-norm phase.

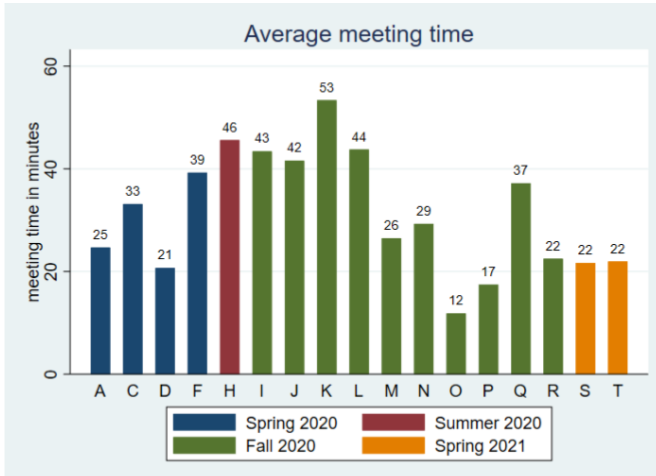

Figure 2: Average meeting time of all teams in different semesters

The next question is whether there are clear differences in average meeting times in different types of classes. As noted earlier, all teams are separated into one of three categories: technical, social science, or both, each of which has varying numbers of teams. While it is important to note that the percentage of course grades assigned to the team projects and in-class time allocated to team projects are likely to vary, we have one preliminary conclusion from Figure 3: Technical and social (hybrid) teams have longer average meeting time, compared to the other two categories. There is no clear pattern between technical and social teams. In addition, the meeting length for technical teams decreases sharply from the norm-negotiation phase to the post new-norm phase.

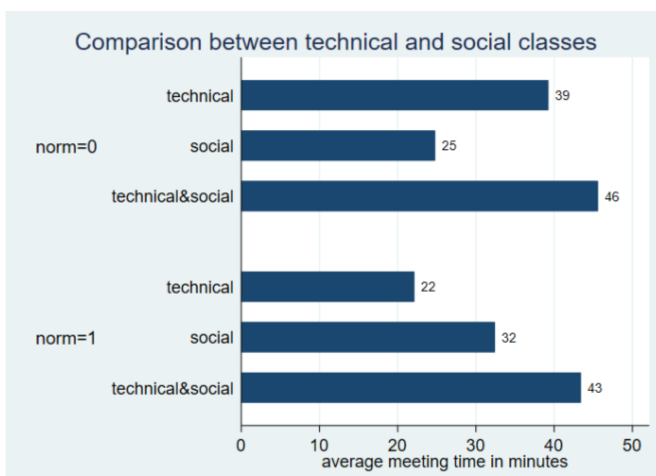

Figure 3: Comparison of average meeting time across disciplines

Notes: The top three bars represent the norm-negotiation phase, while the bottom three bars show the post-new-norm phase

5.2 Statistical analysis of online meetings

In this section, we apply ordinary least squares (OLS) regressions to answer the first, second, and third research questions: how the changes in meeting norms during the pandemic affected the participation and perceived contributions of females and racial minorities. The linear models to be estimated are as follows:

$$y_i = \beta_0 + \beta_1 * norm + \beta_2 * female + \beta_3 * female * norm + \beta_4 * norm * size + \beta_5 * minority + \beta_6 * size + \beta_7 * technical + \beta_8 * social + \beta_9 * undergraduate + \beta_{10} * minority * norm + \epsilon_i \quad (1)$$

The descriptions and descriptive statistics of the variables can be found in Table 3.

5.2.1 Student participation

For research question 1 (**RQ1**), Table 6 shows that the coefficient of *female* is 2.608 (p = 0.537), so the participation gap between female and male students does not exist in the norm-negotiation phase. The coefficient of the interaction between the variables *female* and *norm* is -8.772 (p=0.087, <0.1) in Models (1). This indicates that female students' participation demonstrates a statistically significant 8.772 percentage points decrease than their male counterparts as online meetings become the norm.

The second and third column in Table 6 presents the estimation results for the percent of time giving information and the percent of time asking information. The coefficient on the interaction of *female* and *norm* is - 7.254 (p = 0.091, <0.1) for giving information, and - 1.794 (p = 0.093, <0.1) for asking information, which are both negative and statistically different from zero. This implies that the decrease in female participation in the post new-norm phase can be attributed to both a decline in giving and asking information, with the decline in giving information explaining most of the change. Thus, female students change their roles in team meetings as online meetings become the sole meeting modality.

Table 6: Regression Results on Student Participation

	Percent of speaking time duration (pdur)	Percent of time giving information	Percent of time asking for information
norm=1	-0.912	3.276	-1.216
	(0.957)	(0.825)	(0.732)
female=1	2.608	2.277	0.790
	(0.537)	(0.514)	(0.306)
norm=1 * female=1	-8.772*	-7.254*	-1.794*
	(0.087)	(0.091)	(0.093)
norm=1 * size	-0.353	-0.985	0.302
	(0.903)	(0.704)	(0.642)
minority=1	-9.830**	-7.072**	-2.018***
	(0.015)	(0.035)	(0.003)
minority=1 * norm=1	8.827*	6.445*	1.557*
	(0.060)	(0.099)	(0.092)
size	-3.296*	-2.685*	-0.350
	(0.088)	(0.093)	(0.331)
technical	-1.687	-1.981	0.329
	(0.717)	(0.621)	(0.747)
social	0.400	-0.946	1.238
	(0.958)	(0.886)	(0.410)
undergraduate	-2.047	-1.912	-0.146
	(0.456)	(0.419)	(0.782)
Constant	44.55**	37.20**	4.390
	(0.031)	(0.029)	(0.265)
Observations	433	433	433

Notes: p-values in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

For research question 2 (**RQ2**), we find that the participation of minority students (Asian, Black, and Hispanic/Latino students[2]) is significantly lower than that for White students at the beginning of the COVID-19 pandemic (-9.830, $p = 0.015$, < 0.05). Table 6 shows that compared with White students, the percent of speaking time for minority students is 9.83 percentage points less. However, as the online meeting format becomes the norm over time, the gap between minority and White students shrinks by 8.827 percentage points ($p = 0.06$, <0.1). This indicates that online meetings might help to eliminate racial inequality of participation. Tests on giving and asking information show that minority students' time of giving information (6.445, $p = 0.099$, <0.1) and asking information (1.557, $p=0.092$, <0.1) both recovered and nearly eliminate the racial gap seen during the transition period of online meetings.

5.2.2 How does the norm changes affect students' perceived contribution to teams?

Next, we examined research question 3 (**RQ3**): whether students' participation was translated into students' perceived contributions to teams. By using Model (1), we find that most measures in Table 5 increased in the post new-norm phase, which in sum increased the average of the 12 ratings by 0.6 (out of 4, $p = 0.026$, < 0.05). The detailed regression results for each contribution measure are listed in Appendix A1. No measure had a significantly negative coefficient for *norm*. This indicates that individual student's perceived contribution of other members increased from the norm-negotiation phase to the post new-norm phase.

Female students' contribution ratings did not significantly differ from their male counterparts in the norm-negotiation phase (0.0256, not significant at $p = 0.851$) nor in the post new-norm phase (-0.0796, not significant at $p = 0.768$). In contrast, the average perceived contributions of minority students were lower by 0.389 (out of 4, $p = 0.026$, < 0.05) in the norm-negotiation phase, and this gap closed by 0.136 (not significant at p=0.513) in the post new-norm phase.

Table 7: Changes in Students' Perceived Contributions to Team

	Perceived contributions to team
norm=1	0.600**
	(0.026)
female=1	0.0256
	(0.851)
norm=1 * female=1	-0.0796
	(0.768)
minority=1	-0.389**
	(0.026)
norm=1 * minority=1	0.136
	(0.513)
GPA	0.282
	(0.136)
size	0.182*
	(0.054)
technical	-0.0160
	(0.945)
social	0.486*
	(0.076)
undergraduate	0.182
	(0.257)
Constant	0.906
	(0.523)
Observations	93

Notes: p-values in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[2] To deal with a small sample size for Hispanic and Black students, we used only two categories, namely white and minority students, where minority students contain all non-white students. While acknowledging the subgroup differences, we opted for the binary grouping to obtain an adequate level of power for the quantitative analysis.

5.2.3 Robustness check

Even though the factors that could affect *pdur* across teams (e.g. team size) are already controlled, team size may still pose systematic effect on participation. To address this concern, we use two normalized versions of *pdur* and ran the above analysis again. The measure *pdur1* is normalized by the entire sample, which transforms the distribution of *pdur* to resemble a standard normal distribution:

$$\mathrm{pdur}_1 = \frac{\mathrm{pdur} - \mathrm{mean(pdur)}}{\mathrm{sd(pdur)}}.$$

The measure *pdur2* considers that group size has an inevitable influence on the value of speak time percent. It thus subtracts the group mean from *pdur* (100/team size) and divides the normalized standard deviation from the team mean:

$$\mathrm{pdur}_2 = \frac{(\mathrm{pdur} - 100\ /\mathrm{team\ size})}{\mathrm{sd\ (pdur\ within\ team\ within\ meeting)}}.$$

Both *pdur1* and *pdur2* have a normalized mean of zero. We do not observe any noticeable differences in regression results compared with the main results in section 5.2.1 and 5.2.2, which demonstrate the robustness of our findings. The detailed regression tables can be found in Appendix A2.

Another concern is that our findings on group dynamics may be attributed to changes in meeting length as online meetings gradually become the norm. However, section 5.1 shows that there is no evidence that total or average meeting time changes from the norm-negotiation phase to the post new-norm phase, so we can rule out this potential confounding variable.

6 DISCUSSIONS AND CONCLUSION

From a unique dataset collected at a public Big Ten university spanning 14 months from March 2020 to April 2021, we examined how the COVID-19 pandemic and its imposed transition to online meetings have affected students' participation in, and contributions to, project teams. In this analysis, we divided the dataset into two phases: the norm-negotiation phase in which the initial transition to online meetings occurred (Spring 2020 and Summer 2020), and the post new-norm phase in which students had acclimated to online meetings during the Fall 2020 and the Spring 2021 semesters. Our results indicated that both female students and minority students were affected by the transition to online meetings due to the COVID-19 pandemic and transition to virtual workspaces. However, these groups were not all affected alike. Female students have participated equally in the norm-negotiation phase, but their participation rapidly decreased in the post new-norm phase. In contrast, racial minority groups were heavily affected in the beginning of the transition while they negotiated roles with their peers, but these negative impacts decreased as they become acclimated to new norms and settled with new roles.

Overall, in the first two semesters of online meetings there was no significant gap between male and female student participation. After the initial months passed, and the transition to online learning had become normalized, female student participation decreased compared with their male counterparts. The decrease in giving information during online communication contributes more to the negative effect than that in asking for information. The decrease in female students' time for giving information may reflect the deeply engrained gender roles whereby female students hesitated to play leadership roles when the groups are once established.

Paradoxically, in the case of racial minority students, there was a participation gap between them and their white peers during the norm-negotiation phase. Initially, minority students participated less compared to white peers. However, after new norms were established, minority students closed the gap in participation with their white peers. During the norm-negotiation phase, the average percent of speaking time for minority students are 10.22 percentage points lower than that for white students. The catch-up of minority students can be attributed to an increase in both giving information and asking information.

We also look at how online meetings affect students' perceived contributions to their teams. It is shown that the average of the 12 cross-evaluation measures rose by 0.600 (out of 4) after online meetings became the norm. There is no gender heterogeneous effect on perceived contributions. Minority students' perceived contributions were negatively affected in the beginning of the transition and recovered partially during the post new-norm phase.

Finally, given the above results, we demonstrate the promise of the machine-learning pipeline to generate speaker diarization for analyzing team members' participation. The generated G/A/O information classification enables early detection of member roles and thus identifications of potential communication problems and inequity in participation. A multi-method approach combining these machine-generated results with survey results can construct progress loops for teams and detect issues in team coordination and communication.

6.1 Contributions

In this study, we showed how changing norms from March 2020 until April 2021 affected both female student and the minority student participation in project teams. This result provides a springboard to future studies which will examine why this participation level varied and whether any interventions might create a more equal online meeting environment. Towards this end, this paper contributes to theory by identifying and helping to understand how to overcome the shortcomings of online meetings, and diagnoses inequity in online meeting participation. It also contributes to ongoing work within STEM that stresses the ingrained nature of inequality in professional environments, and that more work must be done to close the gap for historically marginalized groups. Especially, this study shows that the monolithic snapshot of inequities among minority students should be avoided because the patterns of inequity in participation may change through members' negotiations of their roles with their peers. Female students were less affected initially, but their roles changed and thus their participation decreased significantly in the later phase. While racial minority students were heavily affected while they negotiated their roles with their peers during the transition, they recovered in the post new-norm phase, and their peers' evaluations have improved accordingly. These results show that researchers must avoid the uniform understanding of inequities among various subpopulations of the historically marginalized and instead consider the intersections between these subpopulations and other socioeconomic factors.

6.2 Limitations and Future Research

As with any research, there are some potential limitations to this study which can be addressed in future research. Future researchers may consider recruiting more underrepresented minorities to be able to have more analyses about individual minority populations that are sufficiently powered since the authors are aware of the important cultural differences and perceptions of different racial minority populations.

Furthermore, this study's results indicate that women began to participate less as time passed as students entered the post new-norm phase, while racial minority students participated less at the beginning of the transition but recovered in the post new-norm phase. This contradiction between the timing of the gaps in participation might indicate that minority students faced difficult home living situations during the transition which negatively affected their participation, while their white female peers might have been able to transition to more affluent home situations which allowed for comfort in online participation as the new norms were being settled. However, for female students, a reversion to old 'pre-pandemic' norms (such as women talking less in professional and scholastic settings) began to take hold as online meetings continued into the post new-norm phase of the study. To further investigate these possibilities, future studies should include more demographic information pertaining to participants' home life, level of income, and technology access, as these data points could illuminate whether racialized gaps in income and technology access affected participation level. If true, these findings might indicate that future solutions should continue to be sensitive to the diversity of historically marginalized populations and should integrate the idea that a multitude of interventions might be necessary to increase equity amongst culturally distinct subpopulations.

Given our results showing inequities in participation, future research may examine and develop a set of interventions, which include (i) team meetings to set milestones at the beginning of the project, and (ii) individual-level nudges to improve individual—and ultimately team—performance, among others. These milestones and nudges can depend on demographics and other performance-related characteristics.

Appendix

Appendix A1: Regression Results for Each Contribution Measure

Table A1: How Online Learning Format Becoming the Norm Affects Student Contributions

	(1)	(2)	(3)	(4)
	I21 Completed his/her tasks with the expected quality	I22 Completed his/her tasks to achieve the overall project goals	I23 Collaborated with team members from other disciplines to achieve the project goals	I24 Provided information to team members from other disciplines when needed
norm=1	0.502**	0.230	0.749**	0.708***
	(0.044)	(0.365)	(0.038)	(0.003)
female=1	0.0459	0.128	-0.0638	0.0314
	(0.731)	(0.255)	(0.628)	(0.766)
norm=1 * female=1	-0.167	-0.214	-0.00570	-0.0825
	(0.546)	(0.417)	(0.984)	(0.753)
minority=1	-0.320**	-0.257*	-0.324*	-0.381***
	(0.032)	(0.081)	(0.063)	(0.004)
norm=1 * minority=1	0.155	0.116	0.150	0.163
	(0.391)	(0.533)	(0.489)	(0.356)
GPA	0.330*	0.309**	0.182	0.258*
	(0.059)	(0.044)	(0.243)	(0.087)
size	0.0861	-0.00827	0.157	0.168**
	(0.315)	(0.919)	(0.127)	(0.032)
technical	-0.382	-0.0482	-0.0424	0.100
	(0.107)	(0.822)	(0.875)	(0.603)
social	-0.0350	0.189	0.475	0.585**
	(0.894)	(0.448)	(0.149)	(0.010)
undergraduate	0.146	0.155	0.269*	0.259*
	(0.285)	(0.253)	(0.062)	(0.054)
Constant	2.049*	2.631**	1.459	1.068
	(0.097)	(0.027)	(0.309)	(0.318)
Observations	93	93	93	93

Notes: p-values in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Table A2: How Online Learning Format Becoming the Norm Affects Student Contributions (continued)

	(1) l31 Adapted well to changes in the project that altered how he/she completed his/her task	(2) l32 Coped with unforeseen demands placed on him/her	(3) l33 Understood well how to deal with changes in the project to achieve the project goals	(4) l34 Coped with changes to the project that altered how members from different disciplines collaborated on project goals
norm=1	0.849***	0.550*	0.452	-0.240
	(0.006)	(0.088)	(0.162)	(0.147)
female=1	0.0435	0.0922	0.0418	0.0511
	(0.815)	(0.622)	(0.827)	(0.451)
norm=1 * female=1	-0.205	-0.107	-0.0942	-0.124
	(0.495)	(0.732)	(0.764)	(0.616)
minority=1	-0.472**	-0.512**	-0.459**	-0.180**
	(0.021)	(0.013)	(0.031)	(0.033)
norm=1 * minority=1	0.244	0.273	0.172	-0.107
	(0.294)	(0.248)	(0.479)	(0.460)
GPA	0.290	0.370*	0.302	0.115
	(0.176)	(0.070)	(0.148)	(0.284)
size	0.215*	0.181*	0.138	-0.0952
	(0.051)	(0.097)	(0.216)	(0.161)
technical	-0.246	0.0201	-0.0834	-0.365**
	(0.353)	(0.945)	(0.776)	(0.050)
social	0.322	0.521	0.307	-0.142
	(0.312)	(0.140)	(0.379)	(0.512)
undergraduate	0.163	0.212	0.126	-0.0312
	(0.298)	(0.180)	(0.433)	(0.786)
Constant	0.880	0.660	1.436	4.504***
	(0.598)	(0.689)	(0.401)	(0.000)
Observations	93	93	93	89

Notes: p-values in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Table A3: How Online Learning Format Becoming the Norm Affects Student Contributions (continued)

	(1) l41 Contributed new, improved ways to develop his/her tasks	(2) l42 Initiated changes to the ways in which his/her tasks were done that helped accomplish project goals	(3) l43 Created innovative solutions to improve the project quality	(4) l44 Developed alternative solutions to achieve the project goals ahead of time
norm=1	0.964***	1.029***	0.721***	1.530***
	(0.003)	(0.005)	(0.008)	(0.005)
female=1	-0.0540	-0.0426	-0.131	0.0699
	(0.746)	(0.785)	(0.266)	(0.776)
norm=1 * female=1	-0.0532	0.0528	0.0495	-0.0969
	(0.860)	(0.859)	(0.856)	(0.789)
minority=1	-0.353*	-0.359*	-0.271*	-0.466*
	(0.075)	(0.061)	(0.071)	(0.098)
norm=1 * minority=1	0.170	0.0458	-0.102	0.108
	(0.486)	(0.853)	(0.631)	(0.744)
GPA	0.272	0.202	0.114	0.205
	(0.200)	(0.311)	(0.505)	(0.435)
size	0.234**	0.297***	0.144*	0.425***
	(0.029)	(0.009)	(0.085)	(0.008)
technical	-0.0760	0.309	0.121	0.332
	(0.757)	(0.273)	(0.545)	(0.386)
social	0.400	0.821**	0.523**	0.908*
	(0.183)	(0.018)	(0.039)	(0.058)
undergraduate	0.338*	0.302**	0.344**	0.409**
	(0.051)	(0.044)	(0.029)	(0.037)
Constant	0.572	-0.0526	1.574	-1.340
	(0.723)	(0.973)	(0.178)	(0.541)
Observations	93	93	93	93

Notes: p-values in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Appendix A2: Robustness Check using alternative measures of participation

The measure *pdur1* is normalized by the entire sample, which transforms the distribution of *pdur* to resemble a standard normal distribution:

$$\text{pdur}_1 = \frac{\text{pdur} - \text{mean(pdur)}}{\text{sd(pdur)}}.$$

The measure *pdur2* considers that group size has an inevitable influence on the value of speak time percent. It thus subtracts the group mean from *pdur* (100/team size) and divides the normalized standard deviation from the team mean:

$$\text{pdur}_2 = \frac{(\text{pdur} - 100\ /\text{team size})}{\text{sd (pdur within team within meeting)}}.$$

Table A4: Examine gender and racial participation gap using alternative measures of participation

	(1) pdur1	(2) pdur2
norm=1	-0.0455	-0.475
	(0.957)	(0.538)
female=1	0.130	0.130
	(0.537)	(0.493)
norm=1 * female=1	-0.437*	-0.357
	(0.087)	(0.128)
norm=1 * size	-0.0176	0.0689
	(0.903)	(0.594)
minority	-0.490**	-0.436**
	(0.015)	(0.029)
norm=1 * minority=1	0.440*	0.335
	(0.060)	(0.140)
size	-0.164*	-0.0383
	(0.088)	(0.644)
technical	-0.0841	-0.0562
	(0.717)	(0.789)
social	0.0200	0.0621
	(0.958)	(0.856)
undergraduate	-0.102	-0.0706
	(0.456)	(0.581)
Constant	1.288	0.491
	(0.209)	(0.580)
Observations	433	433

Notes: p-values in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

REFERENCES

[1] Alharbi, R. A., Altayyari, F. T., Alamri, F. S., & Alharthi, S. A. (2021, October). Pandemic-Driven Technology During COVID-19: Experiences of Older Adults. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing* (pp. 5-9).

[2] Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses.* University of Chicago Press.

[3] Ayobi, A., Eardley, R., Soubutts, E., Gooberman-Hill, R., Craddock, I., & O'Kane, A. A. (2022). Digital Mental Health and Social Connectedness: Experiences of Women from Refugee Backgrounds. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), 1-27.

[4] Bagmar, A., Hogan, K., Shalaby, D., & Purtilo, J. (2022). Analyzing the Effectiveness of an Extensible Virtual Moderator. *Proceedings of the ACM on Human-Computer Interaction*, 6(GROUP), 1-16.

[5] Bayhan, H. G., Mollaoglu, S., Zhang, H., & Frank, K. A. (2022). Project Team Collaborations during Time of Disruptions: Transaction Costs, Knowledge Flows, and Social Network Theory Perspective. In *Construction Research Congress (CRC) on Project Management and Delivery, Contracts, and Design and Materials* (pp. 1012-1023).

[6] Benke, I., Knierim, M. T., & Maedche, A. (2020). Chatbot-based emotion management for distributed teams: A participatory design study. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2), 1-30.

[7] Bird, S. (2006). NLTK: the natural language toolkit. In *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*. 69–72.

[8] Chen, X., Chen, S., Wang, X., & Huang, Y. (2021). " I was afraid, but now I enjoy being a streamer!" Understanding the Challenges and Prospects of Using Live Streaming for Online Education. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3), 1-32.

[9] Chen, Z., Cao, H., Deng, Y., Gao, X., Piao, J., Xu, F., ... & Li, Y. (2021, May). Learning from home: A mixed-methods analysis of live streaming based remote education experience in Chinese colleges during the COVID-19 pandemic. In *Proceedings of the 2021 CHI Conference on human factors in computing systems* (pp. 1-16).

[10] Copur-Gencturk, Y., Cimpian, J. R., Lubienski, S. T., & Thacker, I. (2020). Teachers' Bias Against the Mathematical Ability of Female, Black, and Hispanic Students. Educational Researcher, 49(1), 30–43.

[11] Cross, R., Gray, P., Cunningham, S., Showers, M., & Thomas, R. J. (2010). How to make employee networks really work. *MIT Sloan management review*, 52(1), 83-90.

[12] Deming, D. J. (2017). The growing importance of social skills in the labor market. The Quarterly Journal of Economics, 132(4), 1593-1640.

[13] Dhawan, N., Carnes, M., Byars-Winston, A., & Duma, N. (2021). Videoconferencing Etiquette: Promoting Gender Equity During Virtual Meetings. *Journal of Women's Health*, 30(4), 460–465.

[14] Do, H. J., Kong, H. K., Lee, J., & Bailey, B. P. (2022). How Should the Agent Communicate to the Group? Communication Strategies of a Conversational Agent in Group Chat Discussions. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), 1-23.

[15] Garcia, A. J., Mollaoglu-Korkmaz, S., & Miller, V. D. (2014). Progress loops in interorganizational project teams: An IPD case. In *Construction Research Congress 2014: Construction in a Global Network* (pp. 2011-2020).

[16] Garicano, L. (2000). Hierarchies and the Organization of Knowledge in Production. *Journal of political economy*, 108(5), 874-904.

[17] Garicano, L., & Rossi-Hansberg, E. (2015). Knowledge-based hierarchies: Using organizations to understand the economy. *Annual Review of Economics*, 7(1), 1-30.

[18] Gui, X., Li, Y., & Wu, Y. (2021). Teacher-Guardian Collaboration for Emergency Remote Learning in the COVID-19 Crisis. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1-26.

[19] Guo, J., & Fussell, S. R. (2022). " It's Great to Exercise Together on Zoom!": Understanding the Practices and Challenges of Live Stream Group Fitness Classes. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW1), 1-28.

[20] Grønbæk, J. E., Rasmussen, M. K., Halskov, K., & Petersen, M. G. (2020, April). Kirigami Table: Designing for proxemic transitions with a shape-changing tabletop. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-15).

[21] He, H. A., Yamashita, N., Hautasaari, A., Cao, X., & Huang, E. M. (2017, February). Why did they do that? Exploring attribution mismatches between native and non-native speakers using videoconferencing. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 297-309).

[22] He, H. A., Yamashita, N., Wacharamanotham, C., Horn, A. B., Schmid, J., & Huang, E. M. (2017). Two Sides to Every story: Mitigating intercultural conflict through automated feedback and shared self-reflections in global virtual teams. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW), 1-21.

[23] Houtti, M., Zhou, M., Terveen, L., & Chancellor, S. (2022). "All of the White People Went First": How Video Conferencing Consolidates Control and Exacerbates Workplace Bias. *arXiv preprint arXiv:2212.00849*.

[24] Hutto, C., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media* (Vol. 8, No. 1, pp. 216-225).

[25] Islind, A. S., Snis, U. L., Lindroth, T., Lundin, J., Cerna, K., & Steineck, G. (2019). The virtual clinic: two-sided affordances in consultation practice. *Computer Supported Cooperative Work (CSCW)*, *28*(3), 435-468.

[26] Jeong, H., & Hmelo-Silver, C. E. (2016). Seven affordances of computer-supported collaborative learning: How to support collaborative learning? How can technologies help? *Educational Psychologist*, 51(2), 247-265.

[27] Kaldi Project. 2021. Kaldi Project. https://github.com/kaldi-asr/kaldi

[28] Kaptelinin, V., Danielsson, K., Kaiser, N., Kuenen, C., & Nordin, M. (2021, October). Understanding the Interpersonal Space of Online Meetings: An Exploratory Study of "We-ness". In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing* (pp. 79-83).

[29] Karl, K. A., Peluchette, J. V., & Aghakhani, N. (2022). Virtual work meetings during the COVID-19 pandemic: The good, bad, and ugly. *Small Group Research*, 53(3), 343-365.

[30] Kaufman, D.B., Felder, R.M. and Fuller, H. (1999). Peer ratings in cooperative learning teams. *Annual American Society for Engineering Education Meeting Proceedings of the 1999 Annual ASEE Meeting*.

[31] Kim, S., Eun, J., Seering, J., & Lee, J. (2021). Moderator chatbot for deliberative discussion: Effects of discussion structure and discussant facilitation. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 1-26.

[32] King, M. M., & Frederickson, M. E. (2021). The pandemic penalty: the gendered effects of COVID-19 on scientific productivity. *Socius*, 7, 23780231211006977.

[33] Kinnula, M., Iivari, N., Isomursu, M., & Laari-Salmela, S. (2018). 'Worksome but Rewarding'–Stakeholder Perceptions on Value in Collaborative Design Work. *Computer Supported Cooperative Work (CSCW)*, *27*(3), 463-494.

[34] Kirshenbaum, N., Davidson, K., Harden, J., North, C., Kobayashi, D., Theriot, R., ... & Leigh, J. (2021). Traces of Time through Space: Advantages of Creating Complex Canvases in Collaborative Meetings. *Proceedings of the ACM on Human-Computer Interaction*, 5(ISS), 1-20.

[35] Kisworo, A. Y., Wardhanie, A. P., & Erdiana, S. P. (2021, September). One Year Reflection of Online Learning: The Students' Perceptions of the Effectiveness of Online Learning amidst the COVID-19 Pandemic in a University. In *Proceedings of the 5th International Conference on Learning Innovation and Quality Education* (pp. 1-7).

[36] Labrie, A., Mok, T., Tang, A., Lui, M., Oehlberg, L., & Poretski, L. (2022). Toward Video-Conferencing Tools for Hands-On Activities in Online Teaching. *Proceedings of the ACM on Human-Computer Interaction*, 6(GROUP), 1-22.

[37] Lacy, A., Polsley, S., Ray, S., & Hammond, T. (2022). A seat at the virtual table: Emergent inclusion in remote meetings. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), 1-20.

[38] Larsen-Ledet, I., & Korsgaard, H. (2019). Territorial functioning in collaborative writing. *Computer Supported Cooperative Work (CSCW)*, *28*(3), 391-433.

[39] Liang, Y. (2017, February). Knowledge sharing in online discussion threads: What predicts the ratings? In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 146-154).

[40] Liu, X., He, W., Zhao, L., & Hong, J. C. (2021). Gender differences in self-regulated online learning during the COVID-19 lockdown. *Frontiers in Psychology*, 12.

[41] MacArthur, C., Wong, C., & Hancock, M. (2019). Makers and quilters: Investigating opportunities for improving gender-imbalanced maker groups. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1-24.

[42] Manshaei, R., Mayat, U., Imtiaz, S. A., Andric, V., Aliar, K., Abu Hantash, N., ... & Mazalek, A. (2022). Tangible Chromatin: Tangible and Multi-surface Interactions for Exploring Datasets from High-Content Microscopy Experiments. *Proceedings of the ACM on Human-Computer Interaction*, 6(ISS), 51-72.

[43] McInnis, B., Xu, X., & Dow, S. P. (2018). How features of a civic design competition influence the collective understanding of a problem. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW), 1-25.

[44] Mentzer, N., Laux, D., Zissimopoulos, A., & Richards, K. A. R. (2016). Peer evaluation of team member effectiveness as a formative educational intervention. *Journal of Technology Education*, 28, 53-82.

[45] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26.

[46] Morris, M. R., Morris, D., & Winograd, T. (2004, November). Individual audio channels with single display groupware: effects on communication and task strategy. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (pp. 242-251).

[47] Nepal, S., Wang, W., Vojdanovski, V., Huckins, J. F., DaSilva, A., Meyer, M., & Campbell, A. (2022, April). COVID Student Study: A Year in the Life of College Students during the COVID-19 Pandemic Through the Lens of Mobile Phone Sensing. In *CHI Conference on Human Factors in Computing Systems* (pp. 1-19).

[48] ONNXRuntime. (201) Optimize and accelerate machine learning inferencing and training. https://github.com/Microsoft/onnxruntime

[49] Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C., Layton, R., Pomeranz, H., & Schmucker, D. G. (2012). The comprehensive assessment of team member effectiveness: Development of a behaviorally anchored rating scale for self-and peer evaluation. *Academy of Management Learning & Education,* 11(4), 609-630.

[50] Povey, D. (2021). ASpIRE Chain Model. http://kaldi-asr.org/models/m1

[51] Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., Hannemann, M., Motlicek, P., Qian, Y., Schwarz, P., et al. 2011. The Kaldi speech recognition toolkit. In *IEEE 2011 workshop on automatic speech recognition and understanding*. IEEE Signal Processing Society.

[52] Raelin, J. A., Bailey, M., Hamann, J., Pendleton, L., Raelin, J., Reisberg, R., & Whitman, D. (2011). The effect of cooperative education on change in self-efficacy among undergraduate students: Introducing work self-efficacy. *Journal of Cooperative Education and Internships*, 45(2), 17-35.

[53] Ravi, P., Ismail, A., & Kumar, N. (2021). The pandemic shift to remote learning under resource constraints. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1-28.

[54] Riedl, C., Kim, Y. J., Gupta, P., Malone, T. W., & Woolley, A. W. (2021). Quantifying collective intelligence in human groups. *Proceedings of the National Academy of Sciences*, 118(21), e2005737118.

[55] Standaert, W., & Thunus, S. (2022). Virtual Meetings during the Pandemic: Boon or Bane for Gender Inequality.

[56] Staniscuaski, F., Kmetzsch, L., Soletti, R. C., Reichert, F., Zandonà, E., Ludwig, Z. M., ... & de Oliveira, L. (2021). Gender, race, and parenthood impact academic productivity during the COVID-19 pandemic: from survey to action. *Frontiers in Psychology*, 12, 663252.

[57] Shi, W. W., Jagannadharao, A., Lee, J., & Bailey, B. P. (2021). Challenges and Opportunities for Data-Centric Peer Evaluation Tools for Teamwork. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW2), 1-20.

[58] Snyder, D., Garcia-Romero, D., Sell, G., Povey, D., & Khudanpur, S. (2018, April). X-vectors: Robust dnn embeddings for speaker recognition. In *2018 IEEE international conference on acoustics, speech, and signal processing* (ICASSP) (pp. 5329-5333). IEEE.

[59] Shrivastava, A., & Pandey, A. (2022). Understanding intra-group response biases in culturally diverse teams. *International Journal of Intercultural Relations*, *88*, 189-199.

[60] Tomprou, M., Kim, Y. J., Chikersal, P., Woolley, A. W., & Dabbish, L. A. (2021). Speaking out of turn: How video conferencing reduces vocal synchrony and collective intelligence. *PLoS One*, 16(3), e0247655.

[61] Tripathi, P., & Burleson, W. (2012, February). Predicting creativity in the wild: Experience sample and sociometric modeling of teams. In *Proceedings of the ACM 2012 conference on computer supported cooperative work* (pp. 1203-1212).

[62] Valente, Thomas. 2012. Network Interventions. *Science* 337, 6090 (2012), 49–53.

[63] Willermark, S., & Pareto, L. (2020). Unpacking the role of boundaries in computer-supported collaborative teaching. *Computer Supported Cooperative Work (CSCW)*, 29(6), 743-767.

[64] Wu, F., and Teets, T. S. (2021). Effects of the Covid-19 pandemic on student engagement in a general chemistry course. *Journal of Chemical Education*, 98, 12: 3633-3642.

[65] Wu, Y., Sun, Y., and Sundar, S. S. (2022, November). What Do You Get from Turning on Your Video? Effects of Videoconferencing Affordances on Remote Class Experience During COVID-19. In *Proceedings of the ACM on Human-Computer Interaction* 6, no. CSCW2.

[66] Xue, F., Merrill, M., Housefield, J., & McNeil, T. (2021, October). Beyond isolation: Benefits and challenges as perceived by students throughout online collaboration during the covid-19 pandemic. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing* (pp. 195-198).

[67] Ye, T., & Robert Jr, L. P. (2017, February). Does collectivism inhibit individual creativity? The effects of collectivism and perceived diversity on individual creativity and satisfaction in virtual ideation teams. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (pp. 2344-2358).

Received January 2023