Board 286: Formative Assessment of Equity and Inclusion in Student Teams

Andrew Moffat, University of Michigan

Andrew Moffat is a Postdoctoral Research Fellow at the University of Michigan, working with the Engineering Education Research Unit and Center for Academic Innovation on an NSF-funded project to assess the effectiveness of Tandem, an in-house software platform designed to support and nurture teamwork skills in undergraduate engineering students. Andrew has a background in education research and evaluation, having previously worked on a project at the University of Leeds, UK, evaluating an institution-wide curriculum transformation initiative. He holds a PhD in Applied Linguistics from the University of Nottingham, UK, prior to the undertaking of which he spent a decade teaching English as a foreign language.

Dr. Robin Fowler, University of Michigan

Robin Fowler is a Technical Communication lecturer and a Engineering Education researcher at the University of Michigan. Her teaching is primarily in team-based engineering courses, and her research focuses on equity in communication and collaboration as well as in group design decision making (judgment) under uncertainty. She is especially interested in how power relationships and rhetorical strategies affect group judgment in engineering design; one goal of this work is to to understand factors that inhibit full participation of students who identify with historically marginalized groups and investigate evidence-based strategies for mitigating these inequities. In addition, she is interested in technology and how specific affordances can change the ways we collaborate, learn, read, and write. Teaching engineering communication allows her to apply this work as she coaches students through collaboration, design thinking, and design communication. She is part of a team of faculty innovators who originated Tandem (tandem.ai.umich.edu), a tool designed to help facilitate equitable and inclusive teamwork environments.

Rebecca L Matz, University of Michigan

Becky Matz is a Research Scientist on the Research and Analytics team at the Center for Academic Innovation at the University of Michigan. She directs and supports research projects across Academic Innovation's portfolio of educational technologies. Her research expertise is in assessing the efficacy of software tools that support student learning and success, analyzing quantitative equity disparities in STEM courses across institutions, and developing interdisciplinary activities for introductory chemistry and biology courses. Matz holds a B.S. in Chemistry from the University of Illinois and an M.S. in Educational Studies and a Ph.D. in Chemistry from the University of Michigan.

Miss Xiaping Li, University of Michigan

Xiaping Li is a Ph.D. candidate in Engineering Education Research at the University of Michigan. Her research interests encompass faculty development and change, neurodiverse college student learning experiences and outcomes, international students in engineering, and cognitive sciences. She holds a B.S. in Hydrology and Water Resources Engineering and an M.S. in Geological Sciences.

Spencer JaQuay, University of California, Irvine

I study how individual differences, teamwork, and language intersect to enhance collaboration and communication. My aim is to equip students with the skills they need to thrive in the workforce by leveraging the strength of diversity.

Madison Jeffrey, University of Michigan

Higher Education Graduate Intern

Mark Mills, University of Michigan

Mark Mills (he/him) is a Data Scientist on the Research & Analytics team at University of Michigan's Center for Academic Innovation. He directs and supports analytics across CAI's portfolio of educational technologies. His experience is in prediction and classification of longitudinal and hierarchically cross-classified data structures such as students in courses measured over time.

Formative Assessment of Equity and Inclusion in Student Teams

Abstract
Teamwork is both widely employed as a pedagogical tool and expected as an important learning outcome in engineering education. However, it cannot be assumed that students' interactions within teams will always be constructive and positive experiences. Inequitable patterns of interaction can exclude individuals from participation, and reproduce existing structures and systems of race- and gender-based marginalization that exist in wider society. Educational institutions should provide appropriate support to foster equitable and inclusive teamwork environments in order to maximize learning and affective outcomes for all students. Here, the authors present on a team support software tool designed to detect and respond to team behaviors and surface patterns of inequities, with interfaces for both students and faculty. The tool centers questions of equity and inclusion and provides formative feedback to students in the form of tailored messages and instructional content, including graphs of data situating team ratings. The tool asks students to reflect on the messages and patterns that they see in their team, as well as to describe behaviors they might try next using strategies from motivational interviewing.

The National Science Foundation program for Improving Undergraduate STEM Education (IUSE) awarded the authors a grant to support evaluating the effectiveness of this tool, both in terms of its ability to detect inequity and exclusion and in terms of its interventions. In this short paper and associated poster we summarize some of this work. Specifically, we will present how we have operationalized "diverse" and "effective" teams, as well as how statistical measures of these variables are related to student outcomes, student identities, and team behaviors. We will highlight patterns in student responses showing, for example, relationships between lesson interventions and student ratings and how patterns in team ratings change over time. We will also present the results of a scoping review synthesizing academic discourse around the notion of team equity. Forthcoming research projects will be described, including an initiative to explore instructors' experiences with the software tool and how it assists their efforts to foster equitable teamwork.

Introduction

Teamwork is both widely employed as a pedagogical tool and expected as an important learning outcome in engineering education. However, research has shown that it cannot be assumed that students' interactions within teams will always be constructive and positive experiences [1], [2]. Inequitable patterns of interaction can exclude individuals from participation and reproduce existing structures and systems of race- and gender-based marginalization that exist in wider society [3], [4]. Educational institutions should provide appropriate support to foster equitable and inclusive teamwork environments in order to maximize learning and affective outcomes for all students. Tandem is a software platform designed for that purpose, offering support to instructors in the formation and monitoring of student teams, and to students in providing feedback on their team experiences and flagging any concerns they may have [5]. This paper and accompanying poster present an overview of research undertaken to date aimed at assessing Tandem's effectiveness as part of the NSF IUSE-funded project titled 'Testing the Effectiveness of Tandem in Assessing and Supporting Inclusive and Equitable Teamwork in Engineering' (grant number 2120252).

Tandem
Tandem is a software platform developed in collaboration between the Center for Academic Innovation and faculty innovators in the College of Engineering at the University of Michigan. The platform incorporates both student-facing and instructor-facing components. Students initially complete an onboarding survey at the beginning of the term, during which they are asked several questions about their experiences of teamwork, their preferences in this regard, and their sense of themselves as engineering students (Tandem is used in other disciplines as well, but in this project and paper we limit the scope to engineering). Responses to these questions, together with demographic variables including race/ethnicity and gender (where students choose to report these), form input for algorithms that suggest team formations to instructors. By default, the algorithms seek to promote diversity while also avoiding 'stranding' students with historically marginalized identities in engineering, especially students of color and female students. The result is intentionally designed to be provisional, and instructors are then presented with an interface in which they can make adjustments based on their more nuanced understanding of the context and the students: instructors are better able to anticipate and account for students' intersectional identities than the algorithm alone.

As the term continues, students complete weekly team checks in which they give feedback about their team experiences. Responses to these questions are used to monitor teams and flag potential problems for the instructor's attention. In addition, these responses can trigger formative lesson content to be presented to the teams, regarding aspects of teamwork such as cognitive diversity, conflict, and task distribution. An end-of-term survey is completed by students to support reflection and peer evaluation, and a mid-term survey may also be administered at instructors' discretion.

Research Activities
The following sections provide an overview of four research activities undertaken as part of the overarching project. A brief description and key findings are presented here; full details can be found in cited publications.

Project 1: Mapping Academic Discourse on Team Equity

Rationale
In exploring notions of equity and equitable interaction in related literature we found a lack of focus in defining and conceptualizing equity at the level of team interactions. We therefore undertook a scoping literature review to explore how the term equity is used and conceptualized in recent academic discourse around teamwork, and to situate the term in a landscape of related concepts.

Methods
The scoping review was structured following the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for scoping reviews (PRISMA-ScR, [6]). Search terms were iteratively developed using the Scopus advanced search interface, then replicated in three other academic databases: Web of Science, ERIC, and PsycInfo. Through a process of filtering and exclusions detailed in [7] and [8] we arrived at a set of 42 publications

from the years 2017 to 2021. The publications were coded for explicit uses of the term equity as well as for adjacent concepts related to fairness and equality. The resulting codes were analyzed and grouped for common thematic content.

Findings

The result of the process was a set of seven underlying themes that we characterize as "facets" of team equity: alignment, dialogism, heterophily, participation, power, ownership, and risk (figure 1). These facets are interconnected and overlapping, together describing team equity as a function of team environments in which team members acknowledge the value of difference in those around them, achieving common alignment in both social-relational-affective and cognitive dimensions through a process of mutual, dialogic perspective-taking. Externally-derived power dynamics must be overcome or set aside to create a space of safety for all team members to take interpersonal risks and feel able and empowered to contribute without fear of negative consequences, promoting feelings of attachment to, and investment in, the team and the task, and motivation to participate fully in expectation of equitable returns on this investment. Full methodological details, along with a more detailed analysis of the seven facets and a discussion of the implications of the framework can be found in [8].

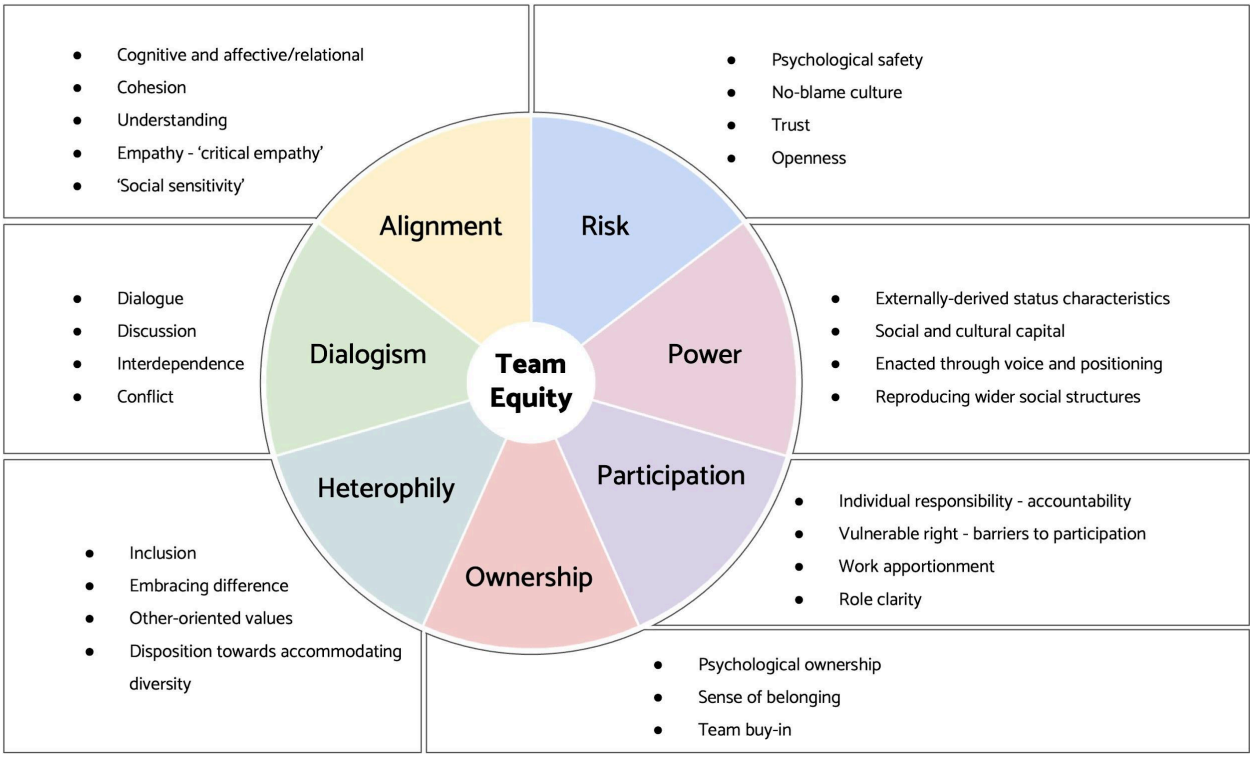

Figure 1 - The seven facets of team equity identified in our scoping review

Project 2: Assessing Change in Halo Effects Following Peer Evaluation Item Format Change

Rationale

A change was made to the format of peer evaluation items in Tandem from 1) asking students to rate each teammate on all items before moving to the next teammate to 2) asking students to rate

all teammates on one item before moving to the next item. This change was made based on hiring decision literature [9] about minimizing halo effects, that is, when an overall positive impression of someone influences ratings on specific behaviors. An analytical project was undertaken to evidence the effect of the change.

Methods

We employed a multilevel linear modeling approach to examine the impact of the peer evaluation format change by selected demographic characteristics, including race, gender, and nationality, among a sample of over 5,000 college students. For each of the characteristics, we established a four-level linear model where responses are nested in the crossing of students and peer evaluation items, which in turn are nested in teams within courses. Peer evaluation served as the dependent variable and the main factors were the change (i.e., before and after change), identity of the rater (i.e., the student rating their teammates), and identity of the target (i.e., the student being rated).

Findings

Our findings reveal statistically significant identity-based effects in peer evaluations and suggest that the implemented format change may have reduced identity-based effects for some groups. Overall, our analysis shows that female, white, and domestic students were more likely to receive higher ratings from their teammates. Before the format change, female students tended to assign lower ratings to male teammates compared to male raters; however, this trend evened out after the change, with female and male raters rating male teammates similarly (figure 2). In terms of race-based effects, white students consistently rated their white peers more favorably, although their Asian peers had the second-highest average rating. Notably, students from minoritized groups tended to receive lower ratings from their white teammates even though they rated their white teammates higher than others. Importantly, white and Asian students' ratings of teammates from minoritized groups improved as the format changed (figure 3). Furthermore, our results highlight that international students were consistently rated lower than domestic students, particularly white domestic teammates, a pattern that seemed to be exacerbated by the format change (figure 4). This concerning observation warrants closer examination in future research.

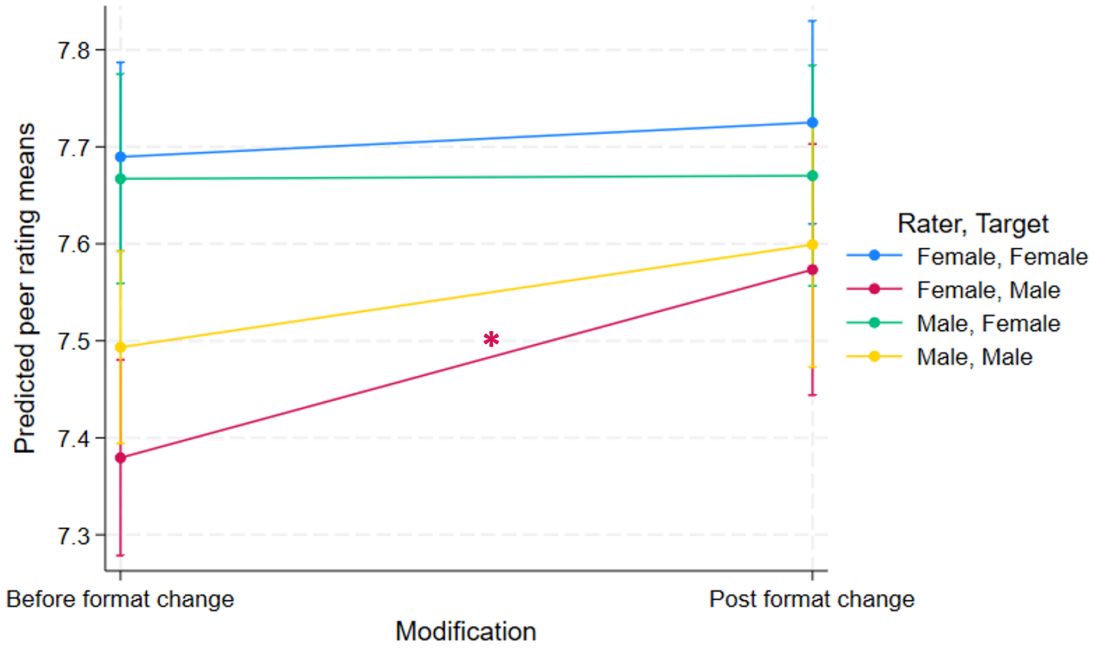

Figure 2 - The impact of format change on gender-based effects in peer rating means
Note: The error bars denote 95% confidence intervals. An asterisk (*) indicates a statistically significant shift in average peer ratings associated with the format change.

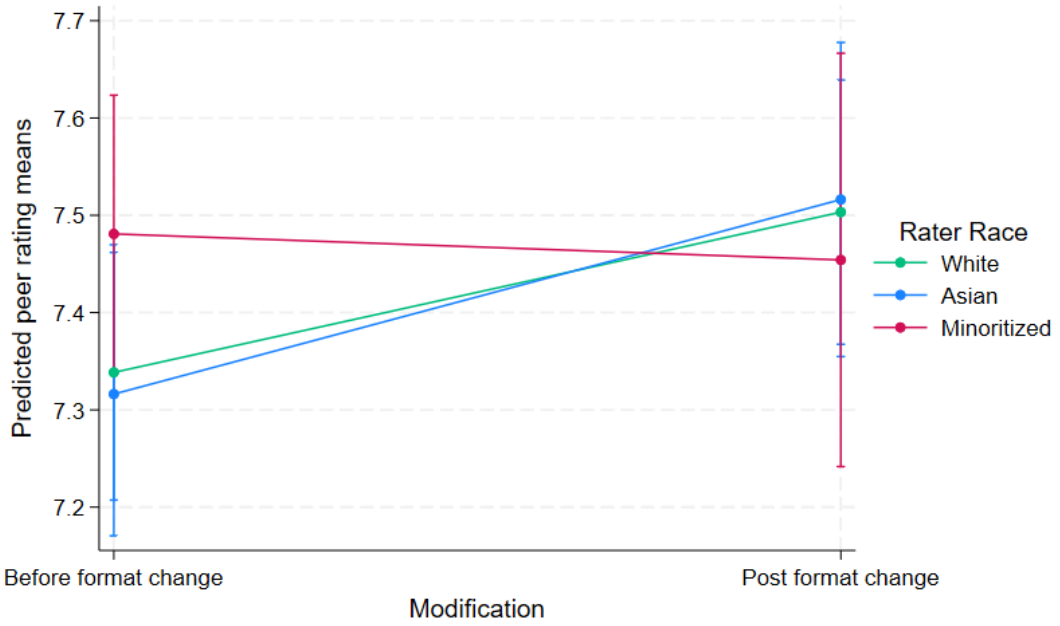

Figure 3 - The impact of format change on race-based effects in peer rating means ***for targets from minoritized groups***
Note: The error bars denote 95% confidence intervals. The pre-format-change discrepancy in average peer ratings between the red and green points (0.133, $p < 0.05$) shifted to a non-significant difference post-change (0.03, $p > 0.05$).

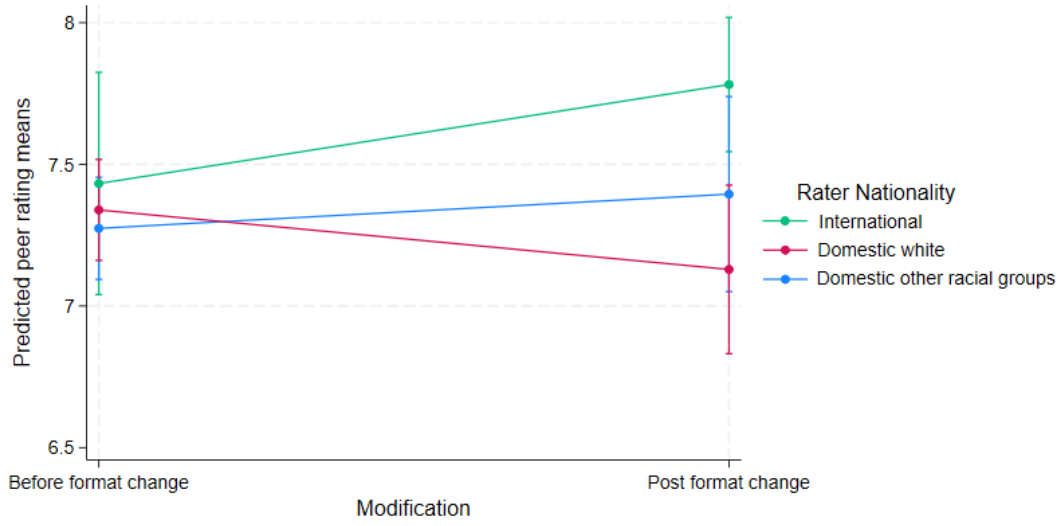

Figure 4 - The impact of format change on nationality-based effects in peer rating means ***for international targets***
Note: The error bars denote 95% confidence intervals. The pre-format-change discrepancy in average peer ratings between the red and green points (0.08, $p > 0.05$) shifted to a significant difference post-change (0.69, $p < 0.05$).

Project 3: Exploring Possible Measures of Equity in Teams

Rationale
Quantitative measures of equity in team interactions are often reductive, using equality of some measure such as talking time as a "rough proxy" for equity [10]. However, it has been demonstrated that such appearances of equal status can mask fundamental inequities [11]. As part of our aim to monitor equity in team interactions using Tandem, we wanted to explore the possibilities of detecting equitable patterns of interaction in more sophisticated ways.

Methods
We undertook a four-part analysis of transcript data created from audio recordings of three engineering course teams making design decisions with the intention of creating a path from human interpretation to quantitative data. Two members of the research team created a qualitative description of the data, providing their interpretations of the team dynamics. Next, a deductive coding process was undertaken using a coding scheme derived from positioning theory [12] to understand how the team members positioned themselves and one another in relation to the team and the task. Table 1 presents the coding scheme with examples from the data. Then two forms of computational analysis were performed using Linguistic Inquiry and Word Count (LIWC, [13]) and Group Communication Analysis (GCA, [14]). LIWC provides measures of psychometric constructs for the team and teammates using pre-existing dictionaries of words indexing those constructs. GCA uses an algorithmic approach to score teammates on six constructs, of which we used three: social impact, the degree to which an individual's contributions are taken up by the team; responsiveness, the degree to which an individual picks

up and develops the contributions of others; and participation, measured as the number of utterances above or below the team average.

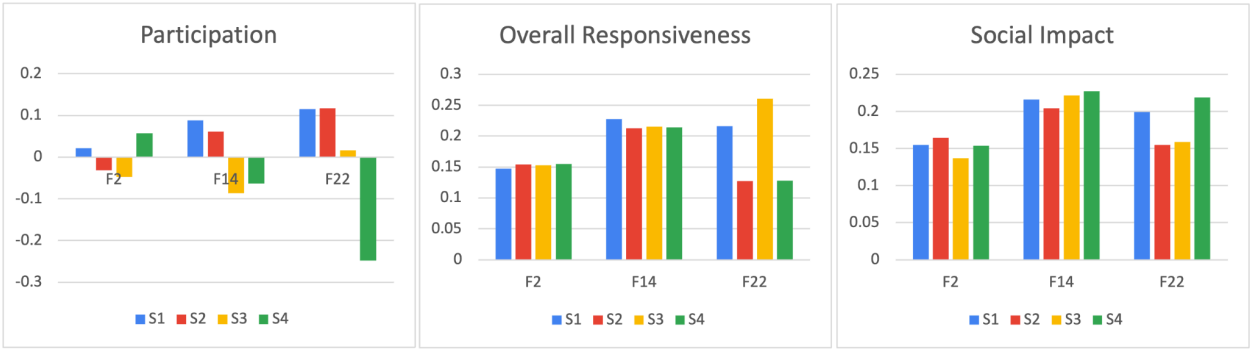

Figure 5 - Scores for each member (S1-S4) of each team for each of the three GCA constructs. The results for team F22 are skewed by S4's very small number of utterances.

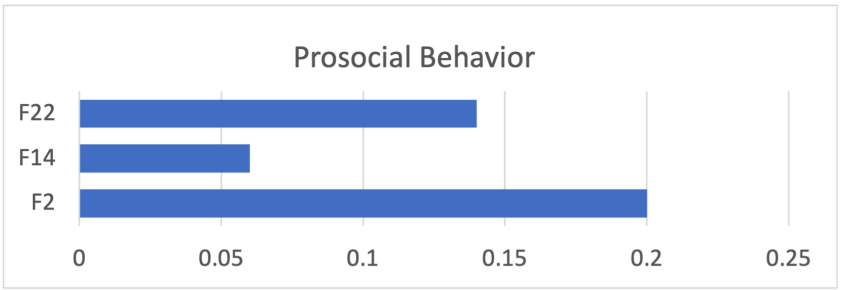

Figure 6 - LIWC team-level prosocial behavior score

Table 1 - Descriptions and examples of interactional positioning codes, taken from [10].

Positional move (code)	Description	Example from data
Expert (C1)	Firm statements of fact or firm or strong disagreement	"The least amount of time is gonna be the kid [shoveling]".
Intermediate expert (C2)	Softened statements or softened disagreement, using hedging, question tags etc.	"Safety should probably be first"
Intermediate novice (C3)	Questions that demonstrate understanding and make constructive contributions to the discussion	"Do we need 'effective' on the list?
Novice (C4)	Questions or statements that convey helplessness or general confusion	"Yeah I've never actually lived with snow"
Facilitator (C5)	Metalevel statements or questions that facilitate the discussion	"Should we move on to the matrix thing?"

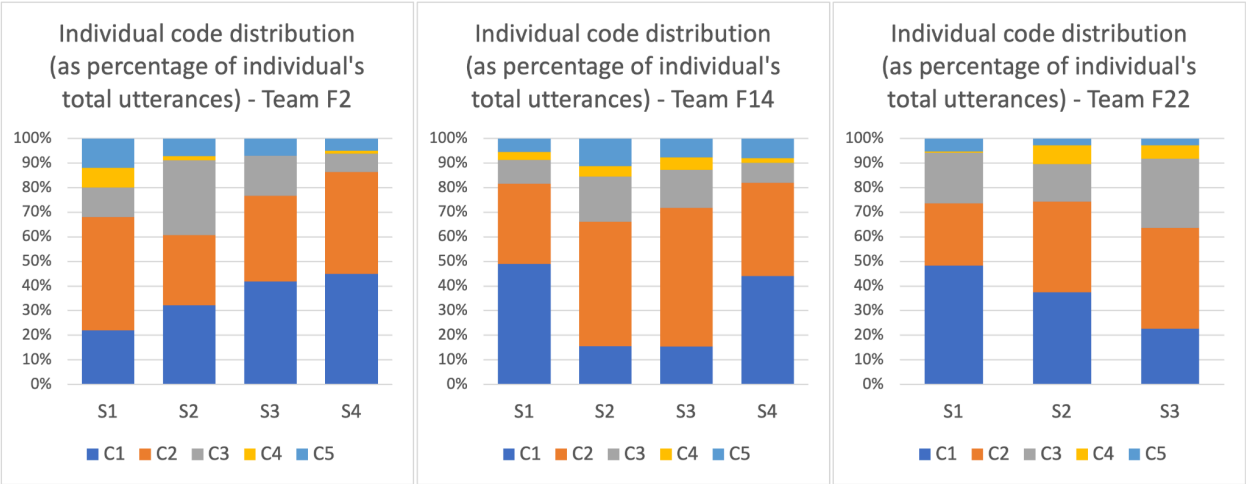

Figure 7 - Distribution of positioning moves (C1-C5; see Table 1) among members (S1-S4) in each team from manual coding. S4 in team F22 had no coded utterances.

Findings

Congruence among the different analyses suggested that the GCA measure of social impact (figure 5) and the group-level LIWC measure of prosocial behaviors (figure 6) most closely paralleled the researchers' understandings of the dynamics at play within the teams. The deductive coding (figure 7) showed promise in identifying team members who positioned themselves as experts with the use of bald declarative statements and others who used more tentative language and thus positioned themselves as less authoritative. However, imprecisions in the transcription process made rigorous coding difficult: further work in this area would approach transcription with greater precision following conventions from conversation analysis. Full details of this project and its findings can be found in [15].

Project 4 (Ongoing): Equity Support Strategies among Instructors and the Role of Tandem

Rationale

This project, in the data transcription phase at the time of writing, aims to explore the experiences of instructors in facilitating student teamwork, with a focus on strategies to promote equity in team interactions. The research question in particular is: What supports instructors in facilitating equitable teamwork in their courses? With the goal of providing appropriate support to instructors in Tandem, we sought to understand how instructors conceptualize challenges in implementing equitable teamwork and the degree to which Tandem's functionality addresses (and doesn't address) these challenges.

Methods

Semi-structured interviews were conducted with 31 instructors, both those using (N = 16) and not using (N = 15) Tandem. When the interview data have been fully transcribed they will be analyzed for how the instructors talk about equity in team formation, team dynamics, and peer evaluation. We will particularly focus on how instructors using Tandem talk about the tool's functionality in their teaching and areas of concern voiced by instructors not using Tandem that users consider the tool to provide solutions for (or could potentially be built into Tandem).

Broader Impacts and Future Work

The main broader impacts of this work lie in the improvements to Tandem effected through the project. Our research team is constantly interacting with the software developers, user experience designers, and behavioral scientists running the Tandem platform to make the results of research studies impact the platform in concrete ways: for example, we are creating a new lesson in Tandem based on a review we conducted about best practices for teams with neurodivergent members, and we are embedding the seven facets of team equity identified in project 1 above into the design of Tandem's interventions. While Tandem was initially built for engineering design courses, it is currently used in team-based courses by about 3,500 unique students per year, a number that is growing steadily, especially as Tandem is now being used at other universities.

The project falls within our long-term goal of fostering equitable learning processes and outcomes for all students and specifically within our research program aiming to improve equity and inclusion in student engineering teams. Team-based pedagogies are common across higher education, but they are a space where patterns of marginalization common in our disciplines and the larger society are sometimes reproduced. Faculty need to be actively working to interrupt these patterns, but faculty have many competing demands and priorities and are not always adequately equipped to foreground inclusion and equity. We are interested in understanding patterns of student experiences, and how students' social identities shape these experiences. Our research strives to acknowledge and incorporate the complexity of intersectionality in student identities, though sample sizes make this challenging, especially in quantitative research. We aim to identify patterns of privilege and power, with the goal of developing and testing interventions that support faculty in providing equitable and inclusive team experiences to their students. Our overall goal is that student teamwork can be a high-impact pedagogical practice for all students.

References

[1] M. Fathi, M. Ghobakhloo, and A. Syberfeldt, "An interpretive structural modeling of teamwork training in higher education", *Education Sciences*, vol. 9, no. 2, article 16, Jan. 2019. https://doi.org/10.3390/educsci9010016

[2] L. Riebe, A. Girardi, and C. Whitsed, "A systematic literature review of teamwork pedagogy in higher education", *Small Group Research*, vol. 47, no. 6, pp. 619-664, Sept. 2016. https://doi.org/10.1177/1046496416665221

[3] R. R. Fowler, and M. P. Su, "Gendered risks of team-based learning: A model of inequitable task allocation in project-based learning", *IEEE Transactions on Education*, vol. 61, no. 4, pp. 312–318, Mar. 2018. https://doi.org/10.1109/TE.2018.2816010

[4] T. S. Henderson, "Understanding access to learning opportunities in collaborative projects: Gendered social hierarchies in student teams", *Studies in Engineering Education*, vol. 4, no. 1, pp. 90–114, July 2023. https://doi.org/10.21061/see.101

[5] R. Fowler, L. K. Alford, S. Sheffield, C. Hayward, T.S. Henderson, & R. L. Matz, "Supporting equitable team experiences using tandem, an online assessment and learning tool", *2021 ASEE Virtual Annual Conference Content Access,* July 2021. https://doi.org/10.18260/1-2--37787

[6] A. C. Tricco et al., “PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation,” *Ann. Intern. Med.*, vol. 169, no. 7, pp. 467–473, Sept. 2018. https://doi.org/10.7326/M18-0850
[7] A. D. Moffat, R. R. Fowler, R. L. Matz, & M. P. Jeffery, “Is an effective team an equitable team? Protocol for a scoping review” *Institute of Electrical and Electronics Engineers (IEEE) Frontiers in Education (FIE) Conference, October 8-11, 2022, Uppsala, Sweden.* Nov. 2022.
[8] A. D. Moffat, R. R. Fowler, R. L. Matz, & M. P. Jeffery, “Facets of team equity: A scoping review”, *Small Group Research* (forthcoming).
[9] B. M. Ferdman, and B. R. Deane, *Diversity at work: The practice of inclusion.* San Francisco, CA: Jossey-Bass, 2014.
[10] M. Slattery, and P. Hutchison, “Assessing equity in collaborative learning situations: A comparison of methods”, *Proceedings of International Conference of the Learning Sciences, ICLS, London, UK, June 23-30, 2018,* vol. 2, pp. 1149–1152, July 2018. https://repository.isls.org//handle/1/582
[11] L. J. Hirshfield, “Equal but not equitable: Self-reported data obscures gendered differences in project teams”, *IEEE Transactions on Education*, vol. 61, no. 4, pp. 305–311, Nov. 2018. https://doi.org/10.1109/TE.2018.2820646
[12] D. T. Brookes and Y. Yang, “Social positioning in small group interactions in an investigative science learning environment physics class”, *Physical Review Physics Education Research*, vol. 17, no. 1, p. 10103-1 - 10103-13, Jan. 2021. https://doi.org/10.1103/PhysRevPhysEducRes.17.010103
[13] R. L. Boyd, A. Ashokkumar, S. Seraj and J. W. Pennebaker, *The development and psychometric properties of LIWC-22*, Austin, TX: University of Texas at Austin, 2022. https://www.liwc.app/
[14] N. M. M. Dowell, T. M. Nixon, and A. C. Graesser, “Group communication analysis: A computational linguistics approach for detecting sociocognitive roles in multiparty interactions”, *Behavior Research Methods*, vol. 51, no. 3, pp. 1007–1041, June 2019. http://doi.org/10.3758/s13428-018-1102-z
[15] A. D. Moffat, R. R. Fowler, R. L. Matz, & S. JaQuay, “Approaches to evidencing intra-team equity in student collaborative design decision-making interactions”, *American Society for Engineering Education (ASEE) Annual Conference & Exposition, Baltimore, MD,* June 25-28, 2023. https://doi.org/10.18260/1-2--42285