Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 25-1313

Can Peer Group Design Improve Engagement in Online STEM Courses? The Role of Motivation to Lead

XunFei Li
University of California, Irvine

Qiujie Li
Nanyang Technological University

Di Xu
University of California, Irvine

Xuehan Zhou
Peking University

Amanda Holton
University of California, Irvine

Peer interaction is important for student engagement and success in higher education and becomes even more critical in online STEM education, where limited interaction can undermine motivation and belonging—key factors for success in rigorous STEM coursework. Despite the widespread use of peer group activities to foster collaboration, there is limited understanding of how to effectively form peer groups to optimize interaction and engagement. Prior research on group dynamics suggests that leadership disposition—a non-cognitive trait reflecting a student's tendency to lead in teamwork—can lower coordination costs and improve group engagement. Thus, we hypothesize that diversifying peer groups based on leadership dispositions enhances group interaction and learning outcomes. To test this hypothesis, a randomized controlled trial was conducted in a large online STEM course, where students studied together in groups. Students were assigned either to a control condition with randomly assigned peers or to a treatment condition with peers diversified by leadership dispositions. The results indicate that students in the treatment condition demonstrated significantly higher levels of participation, interactions, and satisfaction. Analysis of conversational data from group interactions further shows that treatment students displayed higher levels of leadership and a more positive tone in discussions. Finally, underrepresented minority (URM) students in the treatment condition showed particularly pronounced improvements in group participation, satisfaction, and sense of belonging. These findings suggest that structuring peer groups to include diverse leadership dispositions can enhance engagement. More broadly, this low-cost, scalable intervention offers a promising approach to improving collaboration in various contexts.

VERSION: October 2025

Suggested citation: Li, XunFei, Qiujie Li, Di Xu, Xuehan Zhou, and Amanda Holton. (2025). Can Peer Group Design Improve Engagement in Online STEM Courses? The Role of Motivation to Lead. (EdWorkingPaper: 25-1313). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/tmmg-0128

Can Peer Group Design Improve Engagement in Online STEM Courses? The Role of Motivation to Lead

XunFei Li[a,1], Qiujie Li[b], Di Xu[a], Xuehan Zhou[c], and Amanda Holton[a]

[a]University of California, Irvine

Address: University of California, Irvine, Irvine, CA, USA 92697

[b]National Institute of Education, Nanyang Technological University

Address: 1 Nanyang Walk, Singapore 6376168

[c]Institute of Economics of Education, Graduate School of Education, Peking University

Address: Yi Heyuan Road No.5, Haidian District, Beijing, China 100871

Author Note

XunFei Li Email: xunfeil@uci.edu Phone: (510) 926-5305

Qiujie Li Email: qiujie.li@nie.edu.sg

Di Xu Email: dix3@uci.edu

Xuehan Zhou Email: xuehanz@pku.edu.cn

Amanda Holton Email: abrindle@uci.edu

Correspondence concerning this article should be addressed to XunFei Li.

Acknowledgments

We thank the students who participated in the intervention and the UCI Postsecondary Education Research & Implementation Institute (PERI) for providing administrative data. We are grateful to Dr. Rachel Baker, Dr. Greg Duncan, Dr. Andrew Penner, and the 2023 APPAM participants for their valuable feedback, as well as to Yuhan Xiao and Waverly Tseng for their technical support.

Abstract

Peer interaction is important for student engagement and success in higher education and becomes even more critical in online STEM education, where limited interaction can undermine motivation and belonging—key factors for success in rigorous STEM coursework. Despite the widespread use of peer group activities to foster collaboration, there is limited understanding of how to effectively form peer groups to optimize interaction and engagement. Prior research on group dynamics suggests that leadership disposition—a non-cognitive trait reflecting a student's tendency to lead in teamwork—can lower coordination costs and improve group engagement. Thus, we hypothesize that diversifying peer groups based on leadership dispositions enhances group interaction and learning outcomes. To test this hypothesis, a randomized controlled trial was conducted in a large online STEM course, where students studied together in groups. Students were assigned either to a control condition with randomly assigned peers or to a treatment condition with peers diversified by leadership dispositions. The results indicate that students in the treatment condition demonstrated significantly higher levels of participation, interactions, and satisfaction. Analysis of conversational data from group interactions further shows that treatment students displayed higher levels of leadership and a more positive tone in discussions. Finally, underrepresented minority (URM) students in the treatment condition showed particularly pronounced improvements in group participation, satisfaction, and sense of belonging. These findings suggest that structuring peer groups to include diverse leadership dispositions can enhance engagement. More broadly, this low-cost, scalable intervention offers a promising

approach to improving collaboration in various contexts.

Keywords: Peer Interaction; Online Learning; Group Learning; Group Composition; Classroom Intervention; Motivation to lead

Can Peer Group Design Improve Engagement in Online STEM Courses?

The Role of Motivation to Lead

1. Introduction

Peer interaction is arguably one of the most important components of both social and academic integration. Seminal theories of student integration, such as Tinto's (2012) framework on college persistence, emphasize the importance of peer interaction for student retention and academic success. According to these frameworks, students who are more socially and academically integrated through meaningful peer interactions are more likely to persist in college and achieve higher academic outcomes. These theoretical arguments are further supported by empirical evidence, which converge to show that effective peer interactions are associated with higher levels of content comprehension, critical thinking, and learning motivation (Eagan et al., 2017; Lavy & Schlosser, 2011; Tenenbaum et al., 2020; Tsai et al., 2023).

The need for effective peer interaction becomes more pronounced in the context of online STEM education. In response to the growing demand for a skilled and diverse STEM workforce, institutions have increasingly adopted online formats to expand access to high-demand introductory STEM courses. However, the flexibility of the online format also introduces its unique challenges. The limited interaction in online formats not only diminishes student engagement, which is crucial for navigating demanding STEM courses, but also weakens students' sense of belonging—a key factor affecting students' persistence in STEM fields (Baldwin 2009; Bambara et al. 2009; Peacock et al., 2020; Peacock & Cowan, 2019). These challenges are particularly pronounced among students from underserved communities, who experience lower levels of sense of belonging and

satisfaction as well as poorer academic performance in online STEM courses (Ke & Kwak, 2013; Rainey et al., 2018; Xu & Jaggars 2013).

Given the consensus on the critical role of peer interactions, instructors at higher education institutions have used various strategies to foster them. In particular, there has been increasing use of peer group activities during the teaching and learning process to enhance interaction, such as through peer-led discussions, study groups, and collaborative projects (Barkley et al., 2014; Herrera-Pavo, 2021; Springer et al., 1999; Zhou et al., 2023). Yet, despite the prevalence of peer group activities, a significant gap remains regarding how to effectively structure these groups to optimize productivity and engagement. Successful group activities typically require shared regulation, where members collaboratively plan, monitor, and coordinate their efforts (Grau & Whitebread, 2012; Hadwin et al., 2017). This process hinges on effective communication and coordination among group members. While research suggests that group member characteristics, such as gender and racial composition, may influence the dynamics and effectiveness of peer interactions (Hoxby, 2000; Zimmerman, 2003), there is limited empirical evidence to guide educators on how to strategically form peer groups to optimize engagement and learning outcomes (Carrell et al., 2013; Oakley et al., 2004).

This study explores one specific strategy for improving the effectiveness of peer group interaction by focusing on group members' leadership tendencies—their motivation to lead group activities. Research across multiple contexts consistently shows that noncognitive skills, such as leadership disposition, play a crucial role in shaping group interaction dynamics and performance (e.g., Barron et al., 2000; Deming et al., 2017; Woolley et al., 2010). Existing evidence indicates that individuals who are inclined

to take on leadership roles within a group often take the initiatives to coordinate group activities and facilitate discussions, which, in turn, foster more positive group relationships, engagement, and performance (Deming et al., 2017; Riedl et al., 2021; Woolley et al., 2010).

In a similar vein, empirical studies in higher education settings suggest that peer leaders play a key role in driving group success by organizing tasks, ensuring participation, and keeping the group focused. These individuals often take on responsibilities such as setting objectives, delegating tasks, and fostering engagement, which help maintain momentum and cohesion within the group (Arendale, 2004; Barry & Stewart, 1997; Johnson & Johnson, 2002). However, much of the existing literature focuses on structured models, where peer leaders are often intentionally selected, trained, and assigned to support other students. One challenge with these approaches is that they often require substantial resources for recruitment, training, supervision, and sometimes financial compensation (Colvin & Ashman, 2010), which makes them difficult to scale in large, lower-division STEM courses that enroll hundreds of students across multiple sections.

The goal of this study is to identify a scalable strategy to effectively form study groups to optimize interaction and engagement. Drawing on prior literature, we propose a low-cost approach—forming groups based on students' leadership dispositions from the outset. Instead of assigning or developing formal peer leaders, we strategically compose groups with individuals with varying levels of motivation to lead teamwork. By distributing leadership tendencies across groups, we aim to enhance group dynamics organically without formally assigning peer leaders to groups. We hypothesize that

diversifying peer groups based on leadership dispositions enhances group interaction and learning outcomes. To test this hypothesis, we conducted a randomized controlled trial in a large online introductory STEM course at a selective research university. Students were assigned to study groups of three, with treatment groups deliberately composed to include one student with a high, one with a medium, and one with a low motivation to lead, based on pre-survey measures. Control groups were formed through random assignments. All group participants engaged in weekly study sessions and completed surveys to assess their group experiences and interactions. Our research question is whether students in groups with peers of diverse leadership dispositions have higher levels of group peer interactions, satisfaction and learning outcomes, as compared with students in groups with randomly assigned peers.

2. Relevant literature

2.1. Leadership and its relationship with online study groups and group engagement

Leadership is defined as a reciprocal process of influence between a leader and the group members. Its key purpose is to coordinate and guide the group toward achieving shared goals related to a particular task (Hollander, 1985). Empirical evidence from group learning contexts has consistently demonstrated a strong correlation between group-level leadership and group engagement and outcomes (Kahai et al., 2013; Lim & Liu, 2006). For instance, Kahai et al. (2013) conducted experiments with students at a public university in the northeastern United States and found that group-level leadership positively influenced both the quality of decisions and students' satisfaction with discussions. Leadership may play a particularly crucial role in online collaborative

learning environments, where the absence of constant, close instructor supervision places greater emphasis on learners' autonomy in organizing and managing group activities to achieve shared goals. (Huang et al., 2010; Kim et al., 2020).

Prior research has identified two main approaches to facilitating effective leadership in these settings. One commonly used strategy involves assigning peer leaders within each group and equipping them with appropriate training or support to guide discussions, encourage participation, and help members stay focused on the task at hand (Northouse, 2025). However, implementing this approach often requires substantial resources for peer leader selection, training, and supervision, posing challenges for scalability in large-enrollment courses. In contrast, **emergent leadership**—where a group member informally takes on a leadership role without formal authority—has also been shown to play a crucial role in successful peer collaboration (Ehrman & Dornyei, 1998; Leeming, 2019; Northouse, 2025). Research in higher education suggests that emergent leaders can enhance group performance. For example, Taggar et al. (1999) conducted a 13-week study involving 480 undergraduates in small groups and found that groups with higher levels of emergent leadership performed significantly better.

Emergent leadership presents a valuable opportunity when exploring effective, scalable, and sustainable strategies to enhance engagement in online study groups. First, unlike assigned leadership, it does not rely on formal selection or instructor oversight, which is particularly challenging in online learning environments, where formal structures and real-time instructor supervision are often limited. Second, rather than through institutional authority, emergent leaders gain influence organically by earning the trust and respect of their peers through demonstrated competence, initiative, and

engagement, thereby fostering sustained participation within the group (Bandura et al., 2021).

2.2. Motivation to Lead (MTL) and its relevance for Online Student Study Group

Given the promising role of emergent leadership in enhancing engagement in online study groups, it is instrumental to identify an effective way to foster it. Extensive prior research has established that **motivation to lead** is a key factor influencing emergent leadership (Badura et al., 2022; Barry & Stewart, 1997; Hong, et al., 2011; Luria & Berson, 2013; Taggar et al., 1999; Pepper, 2009; Van Iddekinge et al., 2009). Developed to help explain why some individuals seek out leadership roles, MTL refers to an individual's internal desire to take on leadership roles and responsibilities (Chan & Drasgow 2001).[1] A meta-analysis of MTL has indicated that individuals who have high MTL are more likely to emerge as leaders and display higher levels of transformational leadership behaviors such as inspiring team members to perform beyond their perceived capabilities and align individual and the team goals (Badura et al., 2020).

Regulation literature provides conceptual insights into how students with high MTL may contribute to group engagement within the context of online learning (Hayes et al., 2015; Higgins, 1997, 1998). Online learning requires higher levels of regulation skills to stay engaged with the learning process and connected with peers (Banson, 2022). Specifically, effective online groups depend on students collaboratively taking

[1] Conceptualized as a multidimensional construct, MTL includes three distinct but related components: Affective-Identity MTL that captures an individual's intrinsic enjoyment of leading and identification as a leader, Social-Normative MTL that reflects a sense of responsibility to lead, and non-calculative MTL that captures an individual's altruistic willingness to lead without expecting personal benefit. This study leverages the Affective-Identity MTL as an important instrument in building study groups based on conceptual and empirical foundation. Conceptually, it captures an individual's intrinsic enjoyment of leading and identification as a leader. Empirically, it is a strong predictor for emergent leadership in leaderless groups (Badura et al., 2020; Hong, et al., 2010; Pepper, 2009; Van Iddekinge et al., 2009).

responsibility for planning, monitoring, and reflecting on their activities—a process commonly referred to as shared regulation (Hayes et al., 2015). Shared regulation is often achieved through co-regulation, where one member of the group with higher motivation and leadership skills scaffold support for other members (Grau & Whitebread, 2012; Hadwin et al., 2017). Students with higher MTL are more likely to engage actively in these co-regulation processes, through which they support the group in achieving shared regulation (Dobbs et al., 2019). Kark and Van Dijk (2007) provided insights into such processes by stating that high MTL individuals tend to emphasize growth, goal attainment, and positive group outcomes. Through group interactions, these priorities of high MTL students can be transmitted to their peers, who in turn may become more motivated, engaged, and focused on group tasks. This kind of social influence may be particularly pronounced for peers who are less confident or less inclined to take initiatives.

Direct empirical studies examining how high MTL students enhance online study group engagement remain limited. However, existing empirical evidence in higher education shows that high MTL students exhibit behaviors similar to co-regulation processes, indicating their potential to improve online study group engagement (Badura et al., 2020; Hong et al., 2011; Volet et al. 2009). Students with high MTL often show behaviors that support other group members; they tend to take initiative, organize tasks, and maintain group momentum. Hong et al. (2011) demonstrates that college students with high MTL are more likely to become leaders in discussions, persuading team members to resolve differences of opinions and reach agreements. These behaviors not only coordinate group communication more effectively, but can also influence how other

group members behave and engage by modeling productive behaviors and establishing group norms (Volet et al., 2009).

Given that motivation to lead is a key driver of leadership emergence and has the potential to enhance group engagement through the co-regulation mechanisms, this highlights the potential for leveraging naturally occurring motivational traits in shaping group dynamics. However, relatively few studies have translated these insights into practical strategies for forming and supporting effective online study groups. Building on the existing literature on MTL, we examine whether group dynamics and performance can be improved by structuring study groups to include individuals with varying levels of MTL. This approach leverages naturally occurring variation in students' leadership dispositions as a low-cost, scalable alternative to formal leadership training. While prior research highlights the benefits of strong leaders within groups, little is known about how the distribution of leadership motivation across group members shapes collaboration and engagement. By intentionally composing groups with a mix of high-, medium-, and low-MTL students, our study explores whether leadership can emerge more organically and whether this kind of composition can foster more inclusive and effective group interactions. This has important implications for educators seeking practical ways to support student collaboration, especially in large courses where assigning or training leaders is often not feasible.

3. Study design and methods

3.1. Study context

We conducted the intervention in a large online introductory chemistry course in winter 2023 at a selective, public four-year research university, which is recognized as a Hispanic-Serving and Asian American and Native American Pacific Islander-Serving Institution. The course is a critical prerequisite for several STEM majors, including Biological Sciences, Engineering, and Nursing Science.

All the course content was delivered virtually; students could learn all the material by watching the asynchronous lecture videos independently. However, synchronous interactions are also available: the instructor offered two optional online sessions to review important topics and to answer students' questions; teaching assistants held both online and in-person discussion sessions, with five of each. The course included three midterms (40% of the course grade) and a final exam (30% of the course grade). Required assignments included online synchronous homework offered through an adaptive learning platform (20% of the course grade), video accountability assignments (5% of the course grade), and quizzes given on weeks that did not have a midterm (six in total). The average of the five highest quizzes could be used to replace one midterm exam grade. Additionally, a study group component—where our intervention took place—was integrated into the course and accounted for 5% of the overall grade.

Starting in the second week, students were assigned to a fixed study group consisting of three members and were required to have a synchronous meeting with their assigned peers either virtually or in person for at least one uninterrupted hour each week to discuss any course-related topic. Students are also provided with a link to the online communication platform, Discord, where they could communicate with their assigned group members. For example, they could coordinate meeting times, discuss course

content, or talk about anything else. The research team collected students' conversational data from Discord, which will be further discussed in the following section. To earn the 5% course grade, students were required to complete a short survey about their group study sessions each week, including questions such as how long they met, how satisfied they were with the meeting, and upload proof of their meeting (see Appendix A Sample of the weekly survey). In addition to the weekly surveys, all students were invited to complete two comprehensive surveys: one at the beginning of the course, which included questions about students' demographic characteristics and leadership dispositions, and another at the end of the course that gathered information about students' experiences and perceptions regarding study groups and the course. Each of these surveys offered an additional 2.5% credit toward the final exam grade (refer to Appendix B Pre-course survey and Appendix C Post-course survey for more details). In the following section of "Key outcome measures", we detail how different sources of data are used in the analysis.

3.2. Experimental design

We conducted a two-arm randomized controlled trial to investigate how modifying the composition of group members' motivation to lead affects students' engagement and learning outcomes, as shown in Fig. 1. Specifically, at the beginning of week 2, study participants were assigned to either the treatment or control condition via stratified randomization to ensure balance in baseline demographic characteristics.[2] The

[2] Non-participants, who either did not participate in the pre-course survey or enrolled in the class after the survey period, were assigned to separate study groups to prevent potential interference with the main participant sample. Since non-participants did not complete the pre-course survey, we are unable to investigate whether student characteristics differ significantly between non-participants and those in our analytic sample.

key difference between the treatment and control conditions was the specific composition of their study group peers:

Treatment condition (Diversified Leadership Dispositions Peers): Students were categorized into high, medium, or low motivation to lead based on their responses to the Motivation to Lead items in the pre-course survey (Chan & Drasgow, 2001) (see Appendix B Pre-course survey Q18). Specifically, four items were used to assess students' leadership preferences, interest in leading, perceived contributions as a follower versus a leader, and tendencies to take charge in group settings. Students responded to these items based on a five-point Likert scale ranging from "Very untrue of me" to "Very true of me". Responses were coded from 1 to 5, with higher values indicating higher levels of MTL. Items were reverse-coded where necessary. We then added the scores for these four items and calculated an MTL score for each student. Among students in the treatment group, we ranked students based on their MTL scores from highest to lowest. Students in the top third were categorized as high MTL, those in the middle third as medium MTL, and those in the bottom third as low MTL. One student from each category was then randomly selected to form a study group of three, ensuring varied levels of leadership dispositions within each group. This grouping procedure relies on the relative levels of students' MTL and can be adapted to different student pools by leveraging those with relatively higher levels of MTL.

Control condition (Randomly Assigned Peers): Students were randomly assigned to study groups of three, resulting in groups that could consist of any combination of motivation to lead levels, such as low-low-high or medium-high-high.

3.3. Analytic sample

A total of 342 students who completed the pre-course survey by the end of Week 1 were included in the random assignment process, where both the treatment and control conditions had 171 students distributed across 57 study groups, for a total of 114 study groups. An additional 45 students, who either did not participate in the pre-course survey or enrolled in the class after the survey period, were assigned to similar but separate study groups to prevent potential interference with the main participant sample. These students were excluded from our analytical sample. The final analytic sample included 288 students, excluding those who dropped out of the course or did not complete the post-course survey. Attrition tests and student characteristics balance tests are presented in Appendix Table D1 and Table D2. No significant differences were found between students who remained in the analytic sample from pre-course to post-course surveys.

Descriptive statistics for both the treatment and control conditions, along with the differences between them, are detailed in Table 1. The sample consisted of 69% female students and 44% underrepresented minority students. No statistically significant differences were observed between the treatment and control conditions regarding any of the observable student characteristics. These characteristics include demographic and academic features—such as gender, first-generation college student status, and high school GPA —as well as student perceptions and peer connections, including course motivation, peer interaction preference, and number of known classmates and friends. The comparable distribution of these various student characteristics between treatment and control conditions provides evidence of successful randomization implementation.

3.4. Key outcome measures

The study utilized four domains of outcome measures to assess the impact of

diversifying study group composition based on students' leadership dispositions. These measures covered both behavioral and perception outcomes related to both the study group activities as well as the overall classroom. For outcomes that are calculated from multiple related items as composite scores, we report Cronbach's alpha for each composite score to show their reliability and adopt a standard of 0.7 or higher as the acceptable range (Tavakol & Dennickm 2011).

3.4.1. Behavioral outcomes: participation and engagement in study group activities

To continuously monitor student engagement within study groups, we administered weekly brief surveys (refer to Appendix A Sample of the weekly survey). These surveys assessed various aspects of participation, including the total hours students spent with study group peers and the frequency of discussions about course-related and non-course-related content. For students who did not submit their surveys, we code them as "0" for these items. We also tried coding it as missing, and the results are qualitatively similar. Additionally, we measured study group satisfaction on a weekly basis for most of the weeks. These outcomes are measured by a five-point Likert scale.

3.4.2. Student interaction within study groups: Conversational data through the communication platform

At the beginning of the study group intervention, we provided students with a link to the online communication platform, Discord, to facilitate discussions among group members. Following the intervention, we collected data from each study group's conversations on Discord and employed text analysis techniques to examine the extent to which the conversations across different treatment conditions exhibited specific sentiments and themes. Each message was treated as a separate observation. An example

of the Discord data collected is a message like, "Hi, should we find a time to meet?"

To gain a deeper understanding of the students' interactions within the study groups, we examined the extent to which the conversations reveal different sentiments and themes. We utilized the Linguistic Inquiry and Word Count (LIWC)—a software with a built-in dictionary—to focus on two content-relevant measures and one structural feature among the summary variables provided by LIWC: Leadership, Emotional Tone, and Words Per Sentence (Boyd et al., 2022; Cohn et al., 2004; Kacewicz et al., 2014). LIWC outputs the percentage indicating to what extent the text message shows the content of the intent measure. Take leadership, for instance, 50% show a higher level of leadership-relevant content in the text message than 20%. In the official model analysis, each of the three outcomes is standardized. The theme of leadership is directly relevant to our experimental design of modified group composition. We hypothesize that the treatment condition—study groups of peers with high, medium, and low leadership dispositions—will exhibit a higher level of leadership in group conversations compared with groups composed of randomly assigned peers. In addition, the high leadership disposition peer in each group is expected to initiate more engaging group dynamics, contributing to group satisfaction and positive interactions.

3.4.3. Study-group perceptions

In the post-course survey, we collected items regarding students' experience and perception regarding study groups (see Appendix C Post-course survey). These items are learned from existing survey items in five aspects: involvement in group activities, group process, relationships among group members, perceived academic and social supports (Arbaugh et al., 2007; Järvenoja et al., 2013; Jung & Sosik, 2002). Involvement (alpha =

0.86) included questions aimed at assessing the individual contributions of participants within study group settings. Specifically, the items sought to evaluate the extent to which members coordinated activities, facilitated engaging and productive discussions, and maintained focus on assigned tasks. For example, we asked students to evaluate the frequency of the following behavior "I helped keep the study group members on task." Group process (alpha = 0.89) assessed aspects of group dynamics within study groups, focusing on their understanding of group purpose, effectiveness in scheduling activities, productivity in meetings, and flexibility to adapt to changing needs. These items aimed to measure the functional attributes that contribute to the success of group collaboration. For example, we asked students to what extent they agree with the following statement about their group: "My group has a clear understanding of the study group purpose." Group relationships (alpha = 0.89) evaluated the interpersonal dynamics within study groups, examining aspects such as connectivity, listening skills, conflict resolution, cooperation, and mutual respect among members. These items were designed to measure the quality of relationships and collaborative efficacy within the groups. Perceived academic (alpha = 0.88) and social support (alpha = 0.73) collected information about the support students get from the study group experience. For example, we asked to what extent they agreed with "The study group helped keep me on task in a way that helped me to learn."

3.4.4. Course social and learning experience, and academic outcomes

In addition to study group outcomes, we also gathered data on students' social and learning experiences related to the overall course. Social experiences (alpha = 0.89) were quantified through variables such as the number of friends students reported, the number of peers from whom they could seek help, and their perceived sense of belonging

(Garrison et al., 2008). For instance, we asked students how true the following statement was to them: “I felt connected to others in this course. ” For learning experiences, we assessed students’ course motivation (Vo et al., 2017), time management skills (Thomas & Brown, 2011), and overall satisfaction with the course. Regarding academic performance, we collected students’ average quiz grades, average midterm grades, raw final exam scores, and final course grades.

3.5. Analytic method

To assess the impact of eligibility for the treatment condition on various student outcomes, we employed a linear regression model as described in equation [1]. In this model, Y_i represents the various outcomes for student i, as mentioned above. Given that students were randomly assigned to the treatment condition, the coefficients β represent our primary estimates of the impact of treatment conditions—diversified group members’ leadership dispositions—on student outcomes. We controlled for covariates collected before the course in X_i, including gender, race/ethnicity, high school GPA, and composite scores for course motivation, peer interaction preference, motivation to lead, general stress level, friends known in the course before, and students that they can seek help within the course. ε_i denotes the error term.

$$Y_i = \alpha + \beta Treat + X_i + \varepsilon_i \quad [1]$$

4. Results

Overall, we observed significant behavioral differences between students assigned to the treatment and control conditions. Students in groups with Diversified Leadership Dispositions Peers demonstrated higher levels of participation in group activities and

more frequent contact with their group members compared with students in groups with Randomly Assigned Peers. These differences were further corroborated by interaction dynamics observed through conversational data.

4.1. Diversifying leadership dispositions in study groups promotes student participation and satisfaction

Students in the treatment condition, characterized by peers with varied levels of leadership dispositions, showed significantly higher levels of behavioral engagement compared with students in study groups with randomly assigned peers. Specifically, as shown in Fig. 2, students in the treatment condition dedicated more time to studying with their group members ($b = 0.277$, $SE = 0.111$, $P = 0.013$). They also showed a higher frequency of contact for both course-related ($b = 0.316$, $SE = 0.110$, $P = 0.005$) and non-course-related topics ($b = 0.212$, $SE = 0.113$, $P = 0.060$). Additionally, students in the treatment condition reported greater satisfaction with their study groups ($b = 0.327$, $SE = 0.119$, $P = 0.006$). More detailed results are available in Table 2. The non-aggregated weekly outcomes, as detailed in Appendix Table D3, consistently indicated a positive overall impact on students' weekly participation.

These findings demonstrate the significant effects of modified group composition on students' behavioral outcomes. The substantial effect size indicates that diversifying study groups based on students' leadership dispositions can serve as an effective strategy in educational settings to enhance satisfaction among students participating in group activities.

4.2. Enhanced leadership and a more positive tone in treatment group conversations

To further investigate the impact of diversifying leadership dispositions in study

groups on students' interaction dynamics, we analyzed the conversational data from each study group, focusing on two categories of measures. First, we used the frequency of messages and total word count to capture the overall quantity of communication. In addition, we used leadership tone, emotional tone, and words per sentence to achieve a more nuanced understanding of the quality and nature of the conversations. Our results indicated that students in the treatment and control conditions sent a similar number of messages that had comparable total word counts, suggesting that the overall volume of communication was similar between the treatment and control conditions (detailed information can be found in Appendix Table D4).

However, a deeper analysis of tone and sentence complexity revealed significant differences in the quality and nature of the conversations (detailed information can be found in Table 3). Specifically, the treatment groups exhibited higher levels of leadership in their conversations, as demonstrated by a significantly higher leadership tone (b = 0.046, SE = 0.018, P = 0.010). This suggests that students in the treatment groups, who were organized based on their leadership dispositions, were more likely to take on leadership roles to facilitate more organized and focused discussions. This aligns with our hypothesis that diversifying leadership dispositions in study groups may foster a more structured and directed group interaction, helping the whole group to stay on track and achieve objectives more efficiently.

Furthermore, the treatment groups demonstrated a more positive emotional tone in their conversations than the control groups (b = 0.045, SE = 0.018, P = 0.010). A positive tone in group discussions is often associated with better cooperation and a more supportive learning environment, which can contribute to improved group performance

and individual satisfaction. This positivity in communication suggests that the intervention fostered a more supportive and encouraging environment.

Finally, messages from groups in the treatment condition had a higher average word count per sentence ($b = 0.046$, $SE = 0.018$, $P = 0.010$). This finding indicates that students in the treatment condition were engaging in more complex and detailed discussions. Despite having a similar total word count, the treatment groups' higher average words per sentence suggest that their conversations were more elaborate and detailed. This means that students in the treatment condition were likely engaging in deeper, more complex discussions despite not sending significantly more messages overall.

4.3. Modest positive effects on course social and learning experience and course grades.

To fully understand the impact of diversifying leadership dispositions in study groups, it is crucial to assess not only the immediate group dynamics but also broader outcomes, such as students' overall social and course learning experiences and their academic performance. These outcomes provide a comprehensive picture of how changes in group composition affect students beyond their interactions within study groups. Therefore, we examined students' overall perceptions of their study groups, social experiences in the course, and course learning experiences using post-course surveys (for detailed measures, refer to Appendix C Post-course survey). We also analyzed their course grades.

Fig. 3 presents the average scores of items related to students' perceptions of their study groups, social experiences, and learning experiences. Students in the treatment

condition reported marginally significant positive overall study group perceptions (b = 0.164, SE = 0.096, P = 0.091), which include items such as study group satisfaction, member relations, and perceived social support. With regard to both social and learning experiences in the course, students in the treatment condition exhibited higher, though not statistically significant, coefficients for the average scores. In terms of course grades, students in the treatment condition demonstrated a marginally significant increase in average quiz grades (b = 0.190, SE = 0.107, P = 0.075). A detailed table that includes a more comprehensive set of outcome measures can be found in Table 4.

Overall, these findings suggest that diversifying leadership dispositions in study groups had a modest positive impact on students' overall study group perception. While the effects on broader social experiences, learning outcomes, and academic performance were more modest and not consistently significant, the positive trends observed are encouraging.

4.4. Underrepresented minorities[3] (URM) benefited equally from the intervention, if not more.

Finally, we investigated whether URM—including Hispanic, Black, American Indian, Alaska Native, Native Hawaiian, or Other Pacific Islander students—and non-URM students benefited differently from this intervention. Sample description indicated that White students were more likely to have higher levels of leadership disposition than Black, Hispanic, and Asian students (as shown in Appendix Table D5). We are concerned

[3] Underrepresented minorities: Races or ethnicities whose representation in STEM employment and S&E education is smaller than their representation in the U.S. population. This includes Blacks or African Americans, Hispanics or Latinos, and American Indians or Alaska Natives. Source: https://ncses.nsf.gov/pubs/nsf23315/report/glossary

that non-URM students may dominate group interactions, potentially overshadowing the contributions and engagement of URM students. Thus, understanding whether the intervention benefits both groups equally is essential to ensure that URM students are not marginalized. This approach will help ensure that their unique perspectives and needs are adequately addressed during group interactions, fostering a more inclusive and supportive learning environment for all students.

Fig. 4 presents the estimated impact of the intervention on weekly group interaction and satisfaction for URM and non-URM students, respectively, showing comparable and sometimes larger coefficients for URM students compared with the non-URM students. For example, the intervention improved the amount of time URM students spent with their study groups by 0.450 standard deviations (SE = 0.018, P = 0.009), compared with an increase of 0.132 standard deviations for non-URM students (SE = 0.152, P = 0.385). In a similar vein, the intervention improved study group satisfaction by 0.480 standard deviations for the URM students (SE = 0.179, P = 0.008), compared with 0.210 for non-URM students with the study groups (SE = 0.158, P = 0.183). While the interaction terms are not significant, these results suggest that the intervention holds the potential for scaling up to provide broader support for non-URM students in STEM courses.

Fig. 5 shows the impact of diversifying leadership dispositions on overall study group perception, course social and learning experiences and academic performance for URM and non-URM students respectively, and the results revealed similar patterns: the treatment significantly improved URM students' Study Group Perception (b = 0.365, SE = 0.145, p = 0.012) and Social Experience in the Course (b = 0.387, SE = 0.125, p =

0.002), while also showing a positive but non-significant impact on URM students' Learning Experience in the Course (b = 0.098, SE = 0.122, p = 0.421) and their Average Quiz Score (b = 0.120, SE = 0.163, p = 0.463). The promising results indicate that the current intervention holds strong potential to improve URM students' experiences in these dimensions of college STEM coursework. Detailed model results are provided in Table 5.

5. Discussion

Effective peer interaction is essential for students' academic and social integration, especially in STEM fields that have been persistently associated with lower retention rates. Yet, the current literature provides limited guidance on effective, low-cost, and scalable strategies for forming peer groups to enhance group engagement and productivity in high-stakes STEM courses. Our study addresses this gap by providing strong causal evidence for the benefits of modifying group composition based on group members' motivation to lead, which positively impacted student behavioral and psychological outcomes that are crucial for STEM identity and persistence (Baldwin, 2009; Bambara et al., 2009). Conducted within a challenging online STEM gateway course, our randomized controlled trial revealed that students with Diversified Leadership Dispositions Peers exhibited significantly higher levels of group participation, engagement, and satisfaction than those in Randomly Assigned Peer groups. These effects were particularly pronounced among URM students and is validated by the more significant impact on their Study Group Perception (e.g., group member relationships, perceived support), and extends to their course-level experience (e.g., increased number

of friends, help-seeking peers within the course, enhanced sense of belonging).

This study contributes to the growing literature on developing supportive learning communities in challenging online STEM courses by demonstrating the effectiveness of modifying study group composition based on students' motivation to lead. Our work also speaks to the broader learner leadership and peer grouping literature. First, to our knowledge, this is the first study to leverage individuals' motivation to lead as a grouping strategy to achieve better group outcomes. We provide further empirical support for the critical role of social skills, particularly learner leadership, in the dynamics of individual interaction within group processes (Barry & Stewart, 1997; Deming et al., 2017; Hara et al., 2000; Leeming, 2019; Riedl et al., 2021). We note that there is a rich body of literature on leveraging other individual features to improve group outcomes such as prior knowledge and personalities (e.g., Barry & Stewart; Zambrano et al., 2019). We justify our focus on motivation to lead through its suitability for low-stakes study group tasks and its potential to be scaled up and adapted to other contexts. Future research could explore the effectiveness of other individual features based on the nature of group tasks and the available information. Second, building on the well-established literature on peer effects (e.g., Carrell et al., 2013; Hoxby, 2000; Mathieu et al., 2014; Zimmerman, 2003), we gain insights about the potential of modifying peer composition for better student outcomes and the benefits of diverse peer composition. We note that, although this study demonstrates the effectiveness of grouping students with high-, medium-, low-motivation to lead relative to randomly assigned groups, we do not fully examine the effectiveness of all possible group compositions based on individuals' motivation to lead. We elaborate on this limitation and provide suggestions in the following section.

This study presents educators with a practical and scalable approach that addresses the critical challenges of fostering a sense of belonging and engagement in challenging gateway STEM courses. The scalability of this intervention across various course contexts and student populations is feasible for several reasons. First, the minimal group size requirement of three students makes it highly adaptable, from small seminars to large lectures. Second, the approach is simple and intuitive, requiring minimal additional resources. This not only makes it cost-effective, but also allows educators to adopt it without needing extensive training or changes to the existing curriculum. Tools such as discussion boards, video conferencing, and collaborative documents are already familiar to most educators and students. Instructors can easily categorize students by leadership dispositions using straightforward surveys or self-assessment tools, integrating this step seamlessly into the course setup process. Third, because the group tasks are designed to be general rather than subject-specific, this approach can be easily applied to a wide range of disciplines, from humanities to STEM fields.

6. Limitation and future studies

While our study provides critical insights into the benefits of diversifying leadership dispositions in study groups, several limitations should be considered. First, our study primarily focused on short-term outcomes, such as satisfaction, group participation, and sense of belonging during the course. It is unclear whether these immediate benefits may translate into sustained academic success and long-term retention in STEM disciplines, which require longitudinal studies that follow students over several semesters or years. Students' study group participation was self-reported and

supplemented with supporting evidence of group meetings. However, this approach may not fully capture their interactions in an objective manner. Future studies could consider alternative methods, such as video recordings, to collect more detailed data, but should also acknowledge that such methods may influence students' behavior.

Second, as groups were randomly assigned in the control condition, they might inadvertently include a mix of high, medium, and low leadership dispositions that resemble the group composition in the treatment condition. As a result, the control groups' composition may not represent a true baseline, potentially underestimating the treatment effect. Additionally, other group compositions that were naturally formed in the randomly assigned condition, such as high-low-low or high-high-low leadership dispositions, might yield comparable or even better results than the high-medium-low group composition in the current treatment condition. Exploring the relative effects of various group compositions would require much larger sample sizes and more complex experimental designs, which was beyond the scope of the current study. Future studies may wish to compare specific group compositions to provide a more comprehensive and accurate understanding of the optimal group composition regarding leadership dispositions.

7. Conclusion

This study demonstrates that forming study groups by diversifying member's motivation to lead (high, medium, and low) is an effective, low-cost, and scalable strategy to enhance student engagement and satisfaction. In a randomized controlled trial conducted in an online STEM gateway course, we observed significant behavioral and

perception differences between the treatment and control conditions. Students in the treatment condition showed significantly higher levels of participation in group activities, dedicated more time to group study, had higher frequencies of course-related and non-course-related peer interactions, reported greater satisfaction with their study groups, and perceived greater social support from peers. Analysis of conversational data further revealed higher levels of leadership and a more positive tone in conversations within the treatment condition. Moreover, heterogeneity analysis indicated that underrepresented minority (URM) students benefited more from the intervention, where the treatment effects were more pronounced in terms of group participation, satisfaction, and sense of belonging compared with their non-URM counterparts.

These results make several important contributions to the literature. First, they provide clear experimental evidence that strategically structuring peer groups based on students' motivation to lead can improve engagement and satisfaction in collaborative learning. This approach offers a practical and scalable way for educators to foster more effective and interactive learning environments. Second, the greater benefits observed among URM students highlight the intervention's potential to address equity gaps. Lastly, this study also contributes to the broader discussions of the important role of noncognitive skills in improving collaborative learning outcomes. These insights advocate for a more nuanced approach to educational interventions, one that considers the interplay of cognitive and noncognitive factors in optimizing student course participation, satisfaction, and learning.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT in order to improve language and readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Arbaugh, J. B., Cleveland-Innes, M., Díaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., Shea, P., & Swan, K. P. (2007). Community of inquiry survey instrument (draft v14). *Community of Inquiry Survey*.

Arendale, D. R. (2004). Pathway to Persistence: A Review of Postsecondary Peer Cooperative Learning Programs. *Best Practices for Access and Retention in Higher Education*, 27-40.

Badura, K. L., Galvin, B. M., & Lee, M. Y. (2022). Leadership emergence: An integrative review. *Journal of Applied Psychology, 107*(11), 2069–2100. https://doi.org/10.1037/apl0000997

Badura, K. L., Grijalva, E., Galvin, B. M., Owens, B. P., & Joseph, D. L. (2020). Motivation to lead: A meta-analysis and distal-proximal model of motivation and leadership. *Journal of Applied Psychology*, *105*(4), 331. https://doi.org/10.1037/apl0000439

Baldwin, R. G. (2009). The climate for undergraduate teaching and learning in STEM fields. *New Directions for Teaching and Learning*, *117*. https://doi.org/10.1002/tl.340

Bambara, C. S., Harbour, C. P., Davies, T. G., & Athey, S. (2009). Delicate engagement: The lived experience of community College students enrolled in high-risk online courses. *Community College Review*, *36*(3). https://doi.org/10.1177/0091552108327187

Banson, J. (2022). Co-regulated learning and online learning: A systematic review. In *Social Sciences and Humanities Open* (Vol. 6, Issue 1). https://doi.org/10.1016/j.ssaho.2022.100376

Barkley, E. F., Major, C. H., & Cross, K. P. (2014). *Collaborative learning techniques: A handbook for college faculty*. John Wiley & Sons.

Barron, J. M., Ewing, B. T., & Waddell, G. R. (2000). The effects of high school athletic participation on education and labor market outcomes. *Review of Economics and Statistics*, *82*(3). https://doi.org/10.1162/003465300558902

Barry, B., & Stewart, G. L. (1997). Composition, process, and performance in self-managed groups: The role of personality. *Journal of Applied Psychology*, *82*(1). https://doi.org/10.1037/0021-9010.82.1.62

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*, *10*, 1–47.

Carrell, S. E., Sacerdote, B. I., & West, J. E. (2013). From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica*, *81*(3),

855–882.

Chan, K. Y., & Drasgow, F. (2001). Toward a theory of individual differences and leadership: Understanding the motivation to lead. *Journal of Applied Psychology*, *86*(3). https://doi.org/10.1037/0021-9010.86.3.481

Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science*, *15*(10). https://doi.org/10.1111/j.0956-7976.2004.00741.x

Colvin, J. W., & Ashman, M. (2010). Roles, risks, and benefits of peer mentoring relationships in higher education. *Mentoring and Tutoring: Partnership in Learning*, *18*(2). https://doi.org/10.1080/13611261003678879

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

Dobbs, J. M., Lindsay, D. R., & Jackson, R. J. (2019). Antecedents of the motivation to lead: What type of leader wants to step up. *American Research Journal of Humanities & Social Sciences*, *2*(6), 54-63.

Eagan, M. K., Stolzenberg, E. B., Zimmerman, H. B., Aragon, M. C., Whang Sayson, H., & Rios-Aguilar, C. (2017). The American freshman: National norms fall 2016 Higher Education Research Institute. *UCLA, Los Angeles*.

Ehrman, M. E., & Dörnyei, Z. (1998). Interpersonal dynamics in second language education: The visible and invisible classroom. In *Interpersonal dynamics in second language education: The visible and invisible classroom*.

Garrison, D. R., & Vaughan, N. D. (2008). *Blended learning in higher education: Framework, principles, and guidelines*. John Wiley & Sons. https://doi.org/10.1002/9781118269558

Grau, V., & Whitebread, D. (2012). Self and social regulation of learning during collaborative activities in the classroom: The interplay of individual and group cognition. *Learning and Instruction*, *22*(6). https://doi.org/10.1016/j.learninstruc.2012.03.003

Hadwin, A., Järvelä, S., & Miller, M. (2017). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In *Handbook of Self-Regulation of Learning and Performance, Second Edition* (2nd ed., pp. 83–106). https://doi.org/10.4324/9781315697048-6

Hara, N., Bonk, C. J., & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. *Instructional Science*, *28*(2). https://doi.org/10.1023/A:1003764722829

Hayes, S., Smith, S. U., & Shea, P. (2015). Expanding learning presence to account for the direction of regulative intent: Self-, co- and shared regulation in online learning.

Journal of Asynchronous Learning Network, *19*(3), 15-31. https://doi.org/10.24059/olj.v19i3.666

Herrera-Pavo, M. Á. (2021). Collaborative learning for virtual higher education. *Learning, Culture and Social Interaction*, *28*. https://doi.org/10.1016/j.lcsi.2020.100437

Higgins, E. T. (1997). Beyond pleasure and pain. *American psychologist*, *52*(12), 1280. https://doi.org/10.1037/0003-066x.52.12.1280

Higgins, E. T. (1998). Promotion and Prevention: Regulatory Focus as A Motivational Principle. *Advances in Experimental Social Psychology*, *30*(C). https://doi.org/10.1016/S0065-2601(08)60381-0

Hollander, E. P. (1985). Leadership and power. In G. Lindzey & E. Aronson (Eds.), The handbook of social psychology (Vol. II, pp.485– 537). New York, NY: Random House.

Hong, Y., Catano, V. M., & Liao, H. (2011). Leader emergence: The role of emotional intelligence and motivation to lead. *Leadership & Organization Development Journal*, *32*(4). https://doi.org/10.1108/01437731111134625

Hoxby, C. M. (2000). Peer Effects in the Classroom: Learning from Gender and Race Variation. *NBER Working Paper*, *7867*.

Huang, R., Kahai, S., & Jestice, R. (2010). The contingent effects of leadership on team collaboration in virtual teams. *Computers in Human Behavior*, *26*(5), 1098-1110.

Järvenoja, H., Volet, S., & Järvelä, S. (2013). Regulation of emotions in socially challenging learning situations: An instrument to measure the adaptive and social nature of the regulation process. *Educational Psychology*, *33*(1). https://doi.org/10.1080/01443410.2012.742334

Johnson, D. W., & Johnson, R. T. (2002). Learning Together and Alone: Overview and Meta-analysis. *Asia Pacific Journal of Education*, *22*(1). https://doi.org/10.1080/0218879020220110

Jung, D. I., & Sosik, J. J. (2002). Transformational leadership in work groups: The role of empowerment, cohesiveness, and collective-efficacy on perceived group performance. *Small Group Research*, *33*(3). https://doi.org/10.1177/10496402033003002

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun Use Reflects Standings in Social Hierarchies. *Journal of Language and Social Psychology*, *33*(2). https://doi.org/10.1177/0261927X13502654

Kahai, S., Jestire, R., & Huang, R. (2013). Effects of transformational and transactional leadership on cognitive effort and outcomes during collaborative learning within a virtual world. British Journal of Educational Technology, 44(6), 969–985.

https://doi.org/10.1111/bjet.12105

Kark, R., & Van Dijk, D. (2007). Motivation to lead, motivation to follow: The role of the self-regulatory focus in leadership processes. *Academy of Management Review*, *32*(2). https://doi.org/10.5465/AMR.2007.24351846

Ke, F., & Kwak, D. (2013). Online learning across ethnicity and age: A study on learning interaction participation, perception, and learning satisfaction. *Computers and Education*, *61*(1). https://doi.org/10.1016/j.compedu.2012.09.003

Kim, M. K., Lee, I. H., & Wang, Y. (2020). How students emerge as learning leaders in small group online discussions. *Journal of Computer Assisted Learning*, *36*(5), 610-624. https://doi.org/10.1111/jcal.12431

Lavy, V., & Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, *3*(2). https://doi.org/10.1257/app.3.2.1

Leeming, P. (2019). Emergent Leadership and Group Interaction in the Task-Based Language Classroom. *TESOL Quarterly*, *53*(3). https://doi.org/10.1002/tesq.506

Lim, J., & Liu, Y. (2006). The role of cultural diversity and leadership in computer-supported collaborative learning: A content analysis. Information and Software Technology, 48(3), 142–153. https://doi.org/10.1016/j.infsof.2005.03.006

Luria, G., & Berson, Y. (2013). How do leadership motives affect informal and formal leadership emergence? *Journal of Organizational Behavior*, *34*(7), 995-1015. https://doi.org/10.1002/job.1836

Mathieu, J. E., Tannenbaum, S. I., Donsbach, J. S., & Alliger, G. M. (2014). A Review and Integration of Team Composition Models: Moving Toward a Dynamic and Temporal Framework. *Journal of Management*, *40*(1), 130-160. https://doi.org/10.1177/0149206313503014

Northouse, P. G. (2025). *Leadership: Theory and practice*. Sage publications.

Oakley, B., Felder, R. M., & Brent, R. (2004). Turning Student Groups into Effective Teams. *Journal of Student Centered Learning*, *2*(1), 9–34.

Peacock, S., Cowan, J., Irvine, L., & Williams, J. (2020). An exploration into the importance of a sense of belonging for online learners. *International Review of Research in Open and Distance Learning*, *21*(2). https://doi.org/10.19173/irrodl.v20i5.4539

Peacock, S., & Cowan, J. (2019). Promoting sense of belonging in online learning communities of inquiry in accredited courses. *Online Learning Journal*, *23*(2). https://doi.org/10.24059/olj.v23i2.1488

Pepper, R. C. (2009). *The impact of motivation to lead on college students' cocurricular*

involvement. Regent University.

Rainey, K., Dancy, M., Mickelson, R., Stearns, E., & Moller, S. (2018). Race and gender differences in how sense of belonging influences decisions to major in STEM. *International Journal of STEM Education*, *5*(1). https://doi.org/10.1186/s40594-018-0115-6

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 of the United States of America*, *118*(21). https://doi.org/10.1073/pnas.2005737118

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, *69*(1). https://doi.org/10.3102/00346543069001021

Taggar, S., Hackett, R., & Saha, S. (1999). Leadership emergence in autonomous work teams: Antecedents and outcomes. *Personnel Psychology*, *52*(4). https://doi.org/10.1111/j.1744-6570.1999.tb00184.x

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. In *International journal of medical education* (Vol. 2), 53. https://doi.org/10.5116/ijme.4dfb.8dfd

Tenenbaum, H. R., Winstone, N. E., Leman, P. J., & Avery, R. E. (2020). How effective is peer interaction in facilitating learning? A meta-analysis. *Journal of Educational Psychology*, *112*(7). https://doi.org/10.1037/edu0000436

Thomas, D., & Brown, J. S. (2011). *A new culture of learning: Cultivating the imagination for a world of constant change* (Vol. 219). Lexington, KY: CreateSpace.

Tinto, V. (2012). *Leaving college: Rethinking the causes and cures of student attrition*. University of Chicago press.

Tsai, C. W., Lin, M. Y. C., Cheng, Y. P., Lee, L. Y., Chyr, W. L., Lin, C. H., Lin, J. W., & Tsai, M. C. (2023). The effects of online peer-facilitated learning and distributed pair programming on students' learning. *Computers and Education*, *203*. https://doi.org/10.1016/j.compedu.2023.104849

Van Iddekinge, C. H., Ferris, G. R., & Heffner, T. S. (2009). Test of a multistage model of distal and proximal antecedents of leader performance. *Personnel Psychology*, *62*(3), 463-495. https://doi.org/10.1111/j.1744-6570.2009.01145.x

Vo, H. M., Zhu, C., & Diep, N. A. (2017). The effect of blended learning on student performance at course-level in higher education: A meta-analysis. *Studies in Educational Evaluation*, *53*. https://doi.org/10.1016/j.stueduc.2017.01.002

Volet, S., Summers, M., & Thurman, J. (2009). High-level co-regulation in collaborative

learning: How does it emerge and how is it sustained? *Learning and Instruction*, *19*(2), 128-143. https://doi.org/10.1016/j.learninstruc.2008.03.001

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, *330*(6004). https://doi.org/10.1126/science.1193147

Xu, D., & Jaggars, S. S. (2013). The impact of online learning on students' course outcomes: Evidence from a large community and technical college system. *Economics of Education Review*, *37*. https://doi.org/10.1016/j.econedurev.2013.08.001

Zhou, X., Li, Q., Xu, D., Holton, A., & Sato, B. K. (2023). The promise of using study-together groups to promote engagement and performance in online courses: Experimental evidence on academic and non-cognitive outcomes. *Internet and Higher Education*, *59*. https://doi.org/10.1016/j.iheduc.2023.100922

Zambrano R., J., Kirschner, F., Sweller, J., & Kirschner, P. A. (2019). Effects of prior knowledge on collaborative and individual learning. *Learning and Instruction*, *63*. https://doi.org/10.1016/j.learninstruc.2019.05.011

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. In *Review of Economics and Statistics* (Vol. 85, Issue 1). https://doi.org/10.1162/003465303762687677

Table 1. Student characteristics by experimental condition.

	Treatment group	Control group	p-value
Gender			
Male	0.286	0.318	0.558
	(0.453)	(0.467)	
Female	0.693	0.676	0.755
	(0.463)	(0.47)	
Non-binary/third gender	0.014	0.000	0.146
	(0.119)	(0.000)	
Prefer not to say	0.007	0.007	0.969
	(0.085)	(0.082)	
Race/ethnicity			
American Indian, Alaska Native, Native Hawaiian, or Other Pacific Islander	0.043	0.041	0.922
	(0.203)	(0.198)	
Black or African American	0.043	0.047	0.857
	(0.203)	(0.213)	
Hispanic, Latino, or Spanish origin	0.357	0.324	0.558
	(0.481)	(0.47)	
Asian	0.443	0.466	0.692
	(0.499)	(0.501)	
White	0.114	0.115	0.988
	(0.319)	(0.32)	
missing	0.000	0.007	0.332
	(0.000)	(0.082)	
Low-income status	0.429	0.453	0.681
	(0.497)	(0.499)	
First-generation college student status	0.500	0.500	1.000
	(0.502)	(0.502)	
High school GPA	3.840	3.820	0.531
	(0.225)	(0.307)	
Composite score of course motivation	3.524	3.590	0.367
	(0.636)	(0.607)	
Composite score of peer interaction preference	3.331	3.173	0.094
	(0.802)	(0.791)	
Composite score of motivation to lead	3.109	3.155	0.600
	(0.743)	(0.756)	
Composite score of stress level	2.822	2.763	0.356
	(0.551)	(0.533)	
Number of classmates known in the class	1.671	1.642	0.778
	(0.933)	(0.841)	
Number of friends known in the class	1.386	1.311	0.221
	(0.544)	(0.493)	
Number of students that can ask for help in the class	1.407	1.324	0.191
	(0.574)	(0.498)	
Entry term			
Fall 2016	0.007	0.000	0.305
	(0.085)	(0.000)	
Fall 2017	0.007	0.000	0.305
	(0.085)	(0.000)	
Fall 2019	0.014	0.007	0.531

	(0.119)	(0.082)	
Fall 2020	0.043	0.061	0.495
	(0.203)	(0.240)	
Fall 2021	0.221	0.243	0.663
	(0.417)	(0.430)	
Fall 2022	0.707	0.689	0.741
	(0.457)	(0.464)	
Observations	140	148	Total: 288

Note: Standard deviation in parentheses.

Table 2. Treatment effect on aggregated weekly outcomes (post-survey sample).

	Study group interaction frequency (course-related)	Study group interaction frequency (non-course-related)	Total study group study times	Study group satisfaction (weekly)
Treatment condition (no covariates)	0.318**	0.228*	0.247*	0.300*
	(0.109)	(0.112)	(0.112)	(0.117)
Treatment condition (with covariates)	0.316**	0.212+	0.277*	0.327**
	(0.110)	(0.113)	(0.111)	(0.119)
Observations	288	288	288	283
R-squared	0.126	0.136	0.152	0.127

Note: Sample size varies as missing satisfaction data weren't replaced with zeros. Given a successful RCT, the estimates from the model with no covariates and with covariates should yield similar outcomes. By controlling covariates, we aim to increase the precision of the model. Covariates include gender, race/ethnicity, high school GPA, and composite scores for course motivation, peer interaction preference, motivation to lead, general stress level, friends known in the course beforehand, and the number of peers from whom students can seek help within the course. All outcomes have been standardized. Standard errors in parentheses. Significant level: + $p<0.10$ * $p<0.05$ ** $p<0.01$ *** $p<0.001$.

Table 3. Impact of treatment condition on conversational outcomes.

	Standardized Conversational Outcomes		
	Language of leadership, status	Degree of positive tone	Average words per sentence
Treatment	0.046**	0.045*	0.046**
	(0.018)	(0.018)	(0.018)
Constant	-0.023+	-0.022+	-0.023+
	(0.012)	(0.012)	(0.012)
Mean	45.481	39.639	6.151
SD	38.683	34.544	5.068
Observations	12,854	12,854	12,854

Note: In this analysis, each data point represents a message. We use LIWC to generate the outcome for each message and then assess whether the treatment condition predicts the sentiment of the message. All outcomes have been standardized. Standard errors in parentheses. Significant level: + $p<0.10$ * $p<0.05$ ** $p<0.01$ *** $p<0.001$.

Table 4. Impact of treatment condition on study group perception, social and learning experience in the course, and course grade.

	No Covariates	With Covariates
Study group perception		
Study group satisfaction	0.224+	0.240*
	(0.117)	(0.120)
Composite score on involvement	0.110	0.113
	(0.118)	(0.115)
Composite score on group process	0.116	0.118
	(0.118)	(0.120)
Composite score on member relations	0.139	0.151
	(0.118)	(0.119)
Composite score on academic support	0.122	0.150
	(0.118)	(0.118)
Composite score on social support	0.223+	0.211+
	(0.117)	(0.121)
Average score for study group perception	0.156	0.164+
	(0.096)	(0.096)
Social experience in the course		
Number of Friends in Course	0.181	0.073
	(0.118)	(0.109)
Number of Advisers in Course	0.098	0.058
	(0.118)	(0.118)
Composite Score of Sense of Belonging	0.131	0.117
	(0.118)	(0.118)
Average scores for social experience in the course	0.136	0.083
	(0.087)	(0.084)
Learning experience in the course		
Composite Score of Motivation	0.043	0.101
	(0.105)	(0.090)
Composite Score of Self-Regulated Learning	0.073	0.063
	(0.118)	(0.114)
Course Satisfaction	0.136	0.209+
	(0.118)	(0.119)
Average scores for learning experience in the course	0.084	0.124
	(0.086)	(0.081)
Course grade		
Average quiz grade	0.157	0.190+
	(0.109)	(0.107)
Average midterm grade	0.037	0.032
	(0.101)	(0.095)
Raw final exam	0.043	0.044
	(0.107)	(0.100)
Final grade	0.049	0.043
	(0.096)	(0.091)
Observations	288	288

Note: The left column lists the different outcomes for each model. Given a successful RCT, the estimates from the model with no covariates and with covariates should yield similar outcomes. By controlling covariates, we aim to increase the precision of the model. Covariates include gender, race/ethnicity, high school GPA, and composite scores for course motivation, peer interaction preference, motivation to lead, general stress level, friends known in the course beforehand, and the number of peers from whom students can seek help within the course. All outcomes have been standardized. Standard errors in parentheses. Significant level: + $p<0.10$ * $p<0.05$ ** $p<0.01$ *** $p<0.001$.

Table 5. Heterogeneous analysis with covariates (n = 287, URM = 123 non-URM = 167).

	URM	Non-URM	Diff.
Panel A: Behavioral outcomes/ Aggregated weekly outcomes			
Study group interaction frequency (course-related)	0.397*	0.247	0.149
	(0.169)	(0.150)	(0.227)
Study group interaction frequency (non-course-related)	0.222	0.197	0.024
	(0.173)	(0.153)	(0.232)
Total study group study times	0.450**	0.132	0.319
	(0.171)	(0.152)	(0.230)
Study group satisfaction (weekly)	0.479**	0.211	0.268
	(0.179)	(0.158)	(0.240)
Panel B: Study group perception			
Study group satisfaction	0.441*	0.083	0.358
	(0.181)	(0.160)	(0.243)
Composite score on involvement	0.283	-0.029	0.312
	(0.174)	(0.154)	(0.234)
Composite score on group process	0.318+	-0.042	0.359
	(0.180)	(0.159)	(0.242)
Composite score on member relations	0.424*	-0.065	0.489*
	(0.178)	(0.158)	(0.239)
Composite score on academic support	0.353*	-0.011	0.364
	(0.178)	(0.157)	(0.239)
Composite score on social support	0.372*	0.085	0.286
	(0.183)	(0.161)	(0.245)
Average score for study group perception	0.365*	0.004	0.361+
	(0.145)	(0.128)	(0.195)
Panel C: Course social experience			
Number of Friends in Course	0.360*	-0.152	0.511*
	(0.164)	(0.145)	(0.220)
Number of help-seeking peers	0.376*	-0.188	0.564*
	(0.178)	(0.157)	(0.239)
Composite score of sense of belonging	0.427*	-0.135	0.562*
	(0.177)	(0.157)	(0.238)
Average scores for social experience in the course	0.387**	-0.158	0.546**
	(0.125)	(0.110)	(0.167)
Panel D: Course learning experience			
Composite score of motivation	0.035	0.145	-0.110
	(0.137)	(0.121)	(0.184)
Composite score of self-regulated learning	0.128	0.011	0.117
	(0.173)	(0.153)	(0.233)
Course satisfaction	0.133	0.271+	-0.139
	(0.180)	(0.159)	(0.242)
Average scores for learning experience in the course	0.098	0.142	-0.044
	(0.122)	(0.108)	(0.164)
Panel E: Course grade			
Average quiz grade	0.120	0.235	-0.116
	(0.163)	(0.144)	(0.218)
Average midterm grade	-0.065	0.103	-0.169

	(0.144)	(0.128)	(0.194)
Raw final grade	-0.007	0.074	-0.081
	(0.153)	(0.136)	(0.206)
Final grade	-0.024	0.089	-0.113
	(0.139)	(0.123)	(0.187)
Covariates	Yes	Yes	Yes

Note: Given a successful RCT, the estimates from the model with no covariates and with covariates should yield similar outcomes. By controlling covariates, we aim to increase the precision of the model. Covariates include gender, race/ethnicity, high school GPA, and composite scores for course motivation, peer interaction preference, motivation to lead, general stress level, friends known in the course beforehand, and the number of peers from whom students can seek help within the course. All outcomes have been standardized. Standard errors in parentheses. Significant level: + $p<0.10$ * $p<0.05$ ** $p<0.01$ *** $p<0.001$.

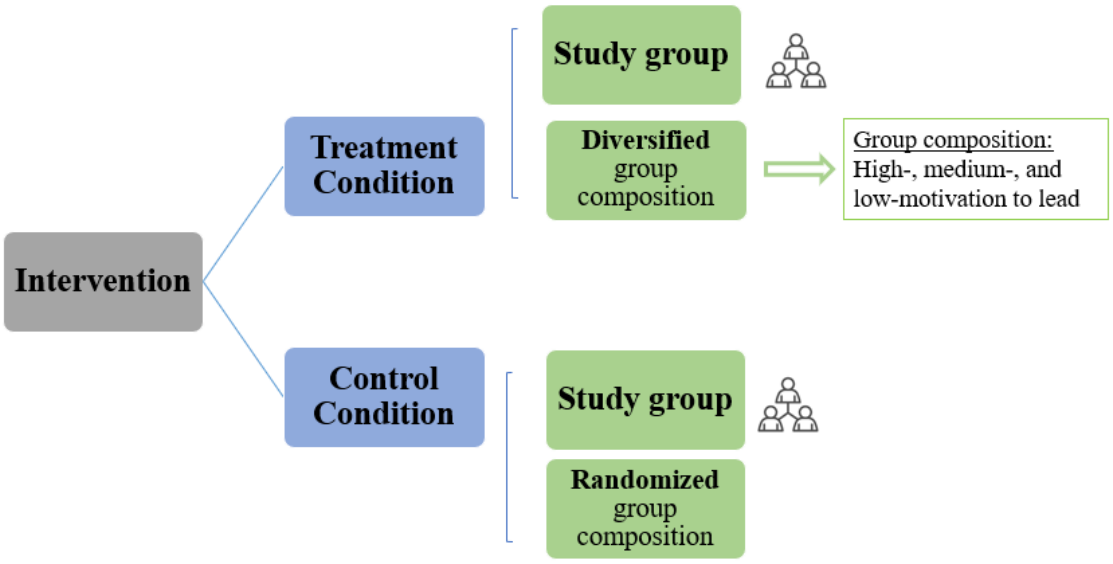

Fig. 1. Experimental design and assignment process.

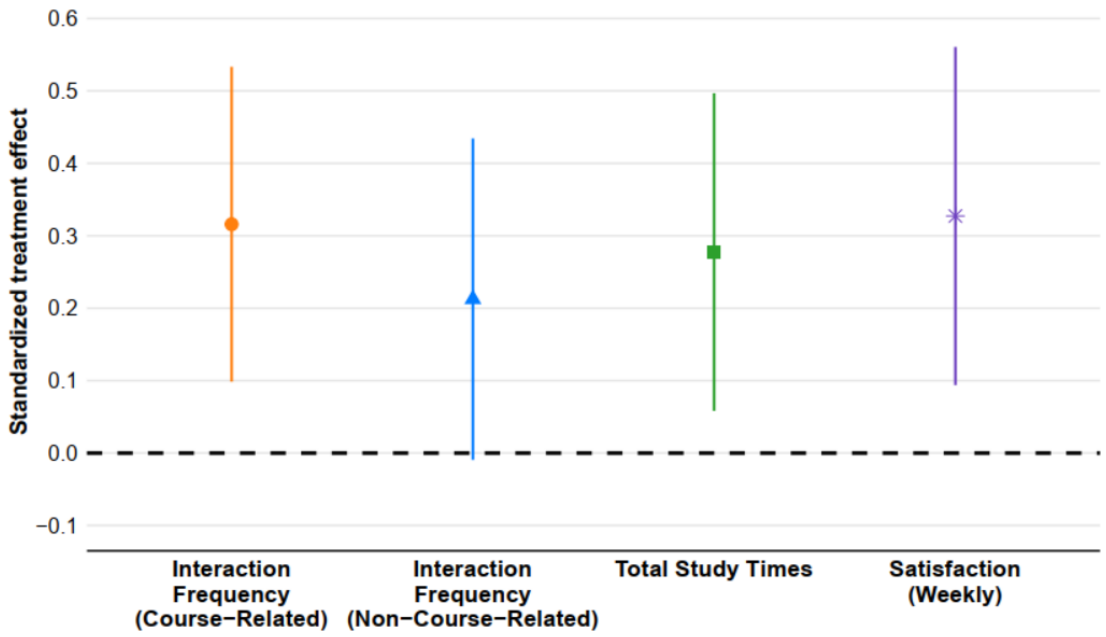

Fig. 2. Treatment effects on study group participation and engagement outcomes (at the 95% confidence interval).

Note: The outcome of Interaction Frequency (Non-Course-Related) is marginally significant at the 90% confidence interval.

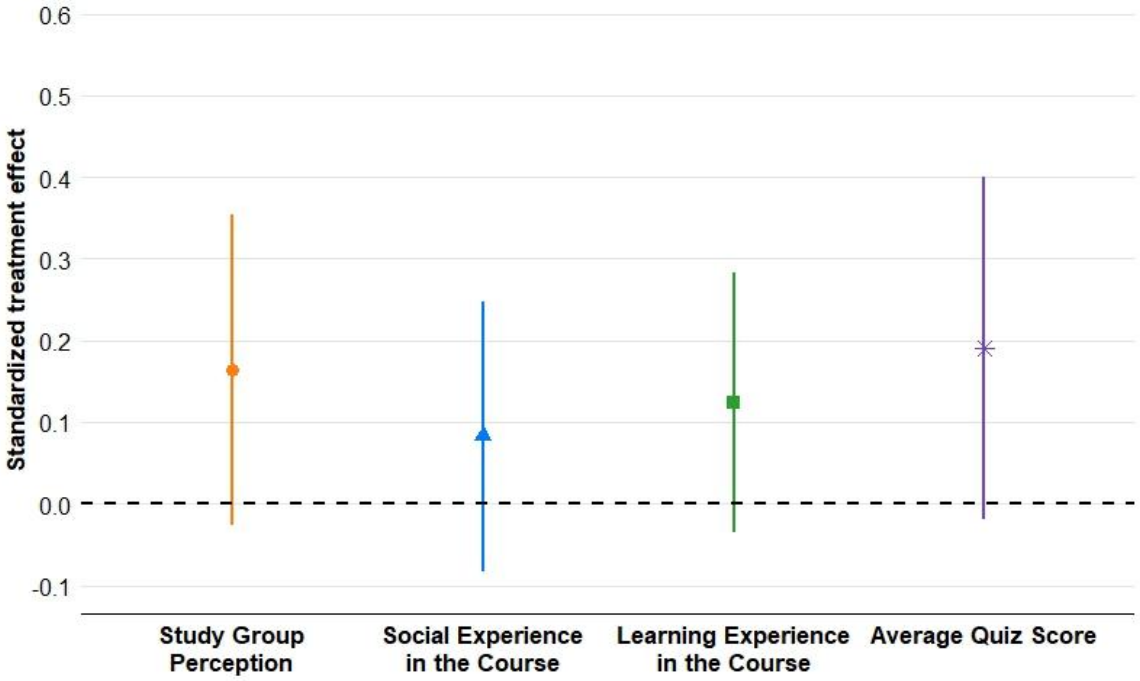

Fig. 3. Treatment effects on the average scores of Study Group Perception, Social Experience in the Course, Learning Experience in the Course, and Average Quiz Score (at the 95% confidence interval).

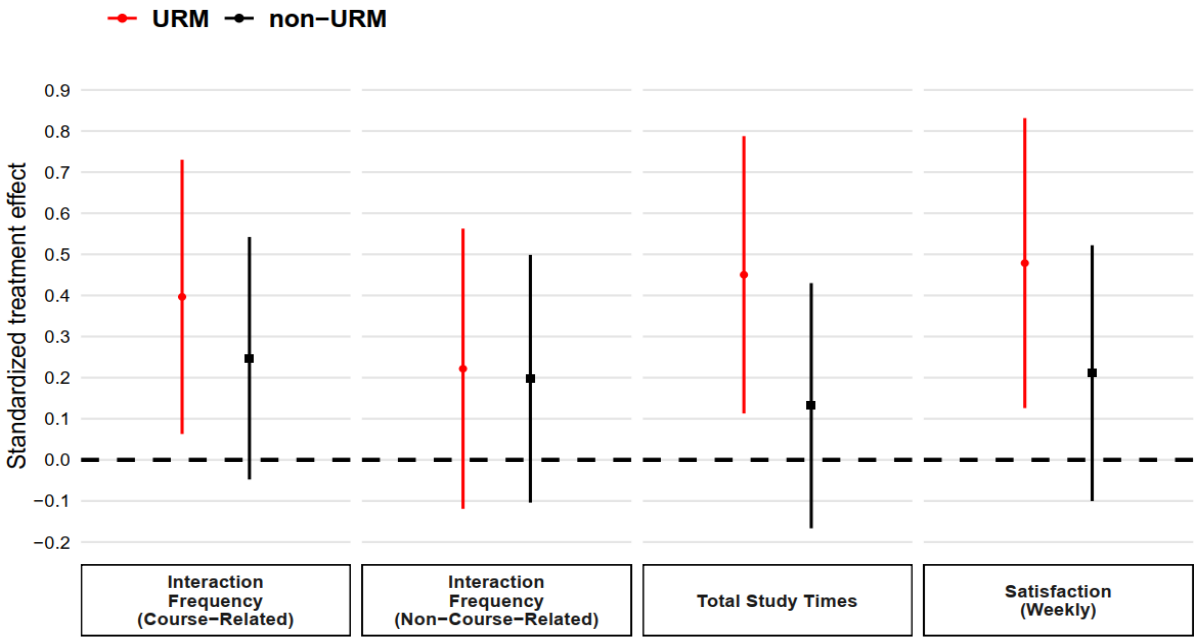

Fig. 4. Treatment effects on study group participation and engagement outcomes for URM and non-URM students (at the 95% confidence interval).

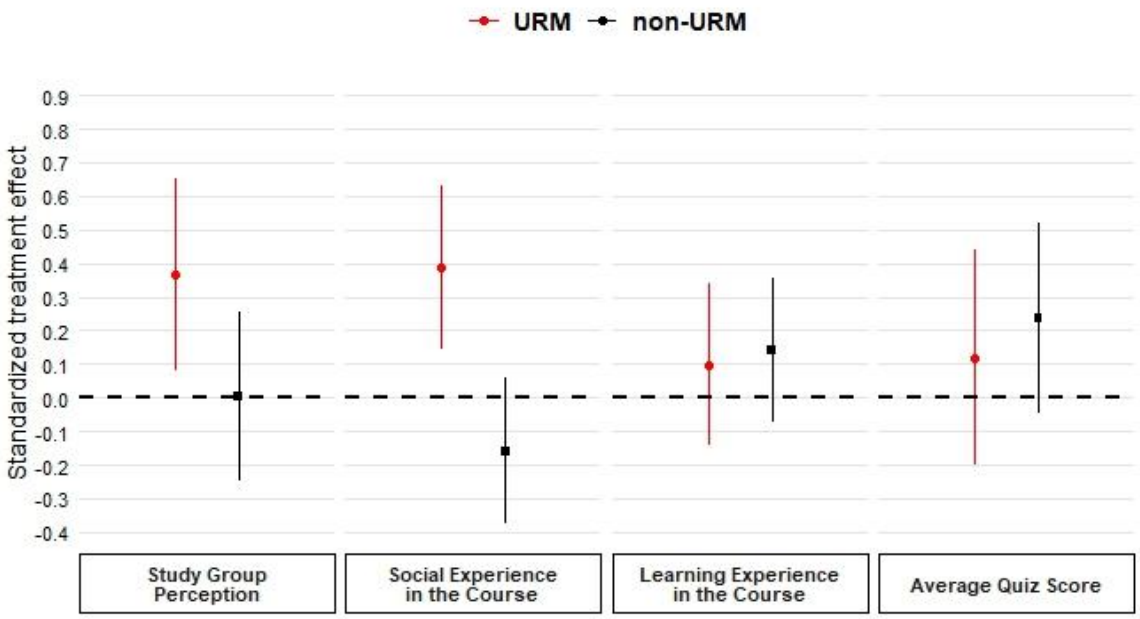

Fig. 5. Treatment effects on the average scores of Study Group Perception, Social Experience in the Course, Learning Experience in the Course, and Average Quiz Score for URM and non-URM students (at the 95% confidence interval).

Appendices

Appendix A

Sample of the weekly survey

Study Group Homework 4
Q0 Study Group Homework 4
What do you need to do?
To get the full credit, you need to:
(1) Answer questions about your study group experience you completed for *Study Group Homework 3*. (2) Make a plan to study with your group member(s) during Week 5. (By "Week 5", we mean anytime from Wednesday, Feb 8 to Tuesday, Feb 14.)
How to find my group members?
(1) Here are your group members in Group X:
Group member 1, Group member 2, Group member 3, Group member 4
(2) Here is the link to your Discord group chat (if it doesn't work after you click it, try to copy and paste it to your browser):
(3) Here is the Zoom link for your group to meet virtually:
Due Date & Grades
Please complete this homework by Tuesday, Feb 7 at 11:59 PM.
For questions about this assignment, contact peer assistant: (email listed)

Q1 Please answer the questions below about your study group experience you completed for Study Group Homework 3 (6 points)

Q2 Did your group study together during Week 4 (from Wednesday, Feb 1 to Tuesday, Feb 7)? (3 points)
It is ok if it is not the same time period as you planned.

- ○ No, we didn't (0 point)
- ○ Yes, we just met with each other briefly (1 point)
- ○ Yes, we studied for one hour (2 points)

Q2.1 Please upload a screenshot of your virtual meeting as proof. You could provide other proof if your group met in person.
Here are some examples:

Q2.2 Did your group meet virtually or in person?

- ○ meet virtually
- ○ in person

Q2.3 What did your group do during the Week 4 meeting?

You can get full points for this question as long as you answer this question.

- [ ] Watching lecture videos/Reading course materials
- [ ] Doing homework
- [ ] Having discussions
- [ ] Others _____

Q3 In the past seven days, how many days did you ever contact your study group peers from this class to discuss course-related content? (1 point)
You can get full points for this question as long as you answer this question.

- ○ 0 (1 point)
- ○ 1 - 2 days (1 point)
- ○ 3 -4 days (1 point)
- ○ 5 -6 days (1 point)
- ○ Everyday (1 point)

Q4 In the past seven days, how many days did you ever contact your study group peers from this class to discuss non-course-related content (e.g., chit-chat, student organizations, campus life)? (1 point)
You can get full points for this question as long as you answer this question.

- ○ 0 (1 point)
- ○ 1 - 2 days (1 point)
- ○ 3 -4 days (1 point)
- ○ 5 - 6 days (1 point)
- ○ Everyday (1 point)

Q5 In the past seven days, how many hours in total did you spend studying with your study group peers from this class? (1 point)

You can get full points for this question as long as you answer this question.

- ○ 0 - 1 hour (1 point)
- ○ 1 - 3 hours (1 point)
- ○ 3 - 6 hours (1 point)
- ○ 6 - 10 hours (1 point)
- ○ More than 10 hours (1 point)

Q6 Please answer the questions below to make a plan with your group members to study together for one uninterrupted hour during week 5 (from Wednesday, Feb 8 to Tuesday, Feb 14) (4 points)

Q7 Did your group schedule a time to study together during Week 5? (2 points)
It is ok if your plan changes during the week; You can get full points as long as your group scheduled a meeting time.

- ○ No, we haven't scheduled a meeting time and will come back later (0 point)
- ○ Yes, we scheduled a meeting time (2 points)

Q7.1 What time do your group plan to study together during Week 5? (e.g., Monday 8-9 PM)
It is ok if your plan changes during the week.

_____

Q8 What do you plan to do during this time period? (2 points)
Your plan can be the same or different from the other group members' plans.

- ○ We didn't make a plan (0 point)
- ○ Watching lecture videos (2 points)
- ○ Doing homework (2 points)
- ○ Having discussions (2 points)
- ○ Others (2 points) _____

Q10 Reminder for next week's Study Group Homework 4:

Your group needs to study for one uninterrupted hour together in order to earn the credit. In the Study Group homework 4 next week, you will also be asked to upload a screenshot/picture of your virtual/in person meeting as a proof (e.g., Zoom, Discord). Here are some examples：

Appendix B

Pre-course survey

Q0 Welcome to the course X Pre-Course Survey. This survey is part of a study exploring strategies that can increase student success in online courses. It is intended to help us understand your feelings about this course and your learning approaches. You will receive 1 extra credit point (each extra credit point will add 2.5% to your final exam score) by completing this survey. The survey should only take you 5-8 minutes to complete. Your personal information will be kept confidential and will not be shared with the instructor. If you are willing to participate, please log in with your [Institution Name] email address to complete the survey by January 13th at 11:59 PM.

Alternatively, if you choose not to complete the survey, you will have the opportunity to earn the extra credit point by completing an alternate essay assignment (3-4 paragraphs each).

If you have any questions or comments and/or wish to complete the alternative essay, please contact XX (email listed).

Q1 What is your gender?

- ○ Male
- ○ Female
- ○ Non-binary / third gender
- ○ Prefer not to say

Q2 Are you of Hispanic, Latino, or Spanish origin?

- ○ Yes
- ○ No

Q3 What is your race?

- ☐ White
- ☐ Black or African American
- ☐ American Indian or Alaska Native
- ☐ Asian
- ☐ Native Hawaiian or Other Pacific Islander

☐ Some Other Race _____

Q4 Your high school GPA (4.0-scale)?
Note: Please fill in a number.

_____

Q5 How much fun is learning Chemistry to you?

○ None at all

○ A little

○ A moderate amount

○ A lot

○ A great deal

Q6 How important is performing well in this course (COURSE X) to you?

○ Not at all important

○ Slightly important

○ Moderately important

○ Very important

○ Extremely important

Q7 How much does being good at the material taught in this course (COURSE X) mean to you for your life?

○ Not at all

○ A little

○ A moderate amount

○ A lot

○ A great deal

Q8 How many students enrolled in this course (COURSE X) did you know prior to this course?
Note: This link *will show you the roster of COURSE X.*

- ○ 0
- ○ 1-3
- ○ 4-7
- ○ 7-9
- ○ 10 or more

Q9 Friendship Network Questions How many friends do you have in this course (COURSE X)?

- ○ 0
- ○ 1-3
- ○ 4-7
- ○ 7-9
- ○ 10 or more

Q10 Please select students in this course (COURSE X) from the dropdown list that you consider your friends.
Note: You can easily locate your friend by typing their first name.

Friend 1 ▼
Friend 2 ▼
Friend 3 ▼
Friend 4 ▼
Friend 5 ▼
Friend 6 ▼
Friend 7 ▼
Friend 8 ▼
Friend 9 ▼
Friend 10 ▼

Q11 If you have more than 10 friends, or friends in this course (COURSE X) that you cannot find their name, please do your best to list their first and last names here.

_____

Q12 If you did not list any names in the previous question, please tell us why.

- ○ I already listed their name.
- ○ I accidentally skipped the question. (If so, you can go back to answer it)

○ I have friends in this course (COURSE X), but I do not know their names.

○ I have friends in this course (COURSE X), but I do not feel comfortable giving their names.

○ I listed their names manually.

Q13 Help-Seeking Network Questions How many students do you have in this course (COURSE X) you would ask for advice or help?

○ 0

○ 1-3

○ 4-7

○ 7-9

○ 10 or more

Q14 Please select students in this course (COURSE X) from the dropdown list that you would ask for advice or help.
Note: You can easily locate your friend by typing their first name.

Person 1 ▼
Person 2 ▼
Person 3 ▼
Person 4 ▼
Person 5 ▼
Person 6 ▼
Person 7 ▼
Person 8 ▼
Person 9 ▼
Person 10 ▼

Q15 If you have more than 10 students that you would ask for advice or help, or students in this course (COURSE X) that you cannot find their name, please do your best to list their first and last names here.

_____

Q16 If you did not list any names in the previous question, please tell us why.

○ I already listed their name.

○ I accidentally skipped the question. (If so, you can go back to answer it)

○ I have students that I can ask for advice or help in this course (COURSE X), but I do not know their names.

○ I have students that I can ask for advice or help in this course (COURSE X), but I do not feel comfortable giving their names.

○ I listed their names manually.

Q17 The following questions ask about your peer interactions preference in terms of completing coursework.
Please select the response that best describes how you feel.

	Very untrue of me	Somewhat untrue of me	Neutral	Somewhat true of me	Very true of me
I prefer to study with peers in the same course rather than studying individually.	○	○	○	○	○
I think that frequently interacting with my classmates would help me succeed in the course.	○	○	○	○	○
I can learn better when studying individually compare to studying with my peers in the same course.	○	○	○	○	○

Q18 Please select the response that best describes how you feel.

	Very untrue of me	Somewhat untrue of me	Neutral	Somewhat true of me	Very true of me
Most of the time, I prefer being a leader rather than a follower when working in a group.	○	○	○	○	○
I am the type of person who is not interested to lead others.	○	○	○	○	○
I believe I can contribute more to a group if I am a follower rather than a leader.	○	○	○	○	○
I have a tendency to take charge in most groups or teams that I work in.	○	○	○	○	○

Q19 The following questions ask you about your general stress levels with regard to various aspects of university life.
Please select the response that best describes how you feel.

	None of the time	A little of the time	Some of the time	Most of the time	All of the time
Academic / coursework demands	○	○	○	○	○
Study / life balance	○	○	○	○	○
Staying on top of coursework	○	○	○	○	○

University environment	○	○	○	○	○
Daily responsibilities of university life	○	○	○	○	○
Friendships	○	○	○	○	○
Please select "Most of the time" for this particular question	○	○	○	○	○
Family relationships	○	○	○	○	○
Romantic relationships	○	○	○	○	○
Relationship break-down	○	○	○	○	○
Discrimination	○	○	○	○	○
Language / cultural issues	○	○	○	○	○
Sexual orientation issues	○	○	○	○	○

Appendix C

Post-Course Survey

Q0 Welcome to the COURSE X Post-Course Survey. This survey is part of a study exploring strategies to increase student success in online courses. It will help us understand your feelings about this course and learning approaches.

Completing this survey will give you one extra credit point (each extra credit point will add 2.5% to your final exam score). The survey should only take you around 10 minutes to complete. Your personal information will be kept confidential and not shared with the instructor. If you are willing to participate, please log in with your [Institution Name] email address to complete the survey by Sunday, March 19th at 11:59 PM.

Alternatively, if you choose not to complete the survey, you will have the opportunity to earn the extra credit point by completing an alternative essay assignment (3-4 paragraphs each) by Sunday, March 19th at 11:59 PM.

If you have any questions or comments and/or wish to complete the alternative essay, please contact the peer assistant XX (email listed).

Q0 From weeks 2 to 10, you were assigned a study group, studied with your peers, and completed study group assignments. We would like to hear about your personal experiences and feelings regarding your time in the study group:

Q1 The following questions ask about your involvement in the study group activities. Please select the response that best describes your behavior in the study group.

	Never	Rarely	Sometimes	Frequently	Very Frequently
I spent time coordinating the study group activities.	○	○	○	○	○
I helped keep the study group members engaged and participating in productive dialogue.	○	○	○	○	○
I helped keep the study group members on task.	○	○	○	○	○

Q2 The following questions ask about the group process of your study group. Please select the response that best describes your group.

	strongly disagree	disagree	neutral	agree	strongly agree
My group has a clear understanding of the study group purpose.	○	○	○	○	○
My group is effective in scheduling the study group activities.	○	○	○	○	○
Study group meetings are very productive towards achieving our goals.	○	○	○	○	○
My group works with a great deal of flexibility so that we can adapt to changing or different needs.	○	○	○	○	○

Q3 The following questions ask about the relationships among your group members. Please select the response that best describes your group.

	strongly disagree	disagree	neutral	agree	strongly agree
Group members do not connect very well with one another.	○	○	○	○	○
Group members are effective listeners.	○	○	○	○	○
Group members are able to work through differences of opinion without damaging relationships.	○	○	○	○	○
Group members display high levels of cooperation and mutual support.	○	○	○	○	○
Everyone values what each member contributes to the study group.	○	○	○	○	○

Q4 The following questions ask about the academic and social support you get from the study group experience. Please select the response that best describes your feelings.

	strongly disagree	disagree	neutral	agree	strongly agree
The study group helped keep me on task in a way that helped me to learn.	○	○	○	○	○
Study group members provided feedback that helped me understand my strengths and weaknesses relative to the course's goals and objectives.	○	○	○	○	○
Study group discussions increased my interest in course issues.	○	○	○	○	○
I feel connected with study group members.	○	○	○	○	○
When I am with my study group, I feel included.	○	○	○	○	○
I feel as if study group members do not care about me.	○	○	○	○	○

Q5 How satisfied are you with the study group experience this quarter?

○ Very Satisfied

○ Satisfied

○ Neutral

○ Dissatisfied

○ Very Dissatisfied

Q6 How much fun is learning Chemistry to you?

○ None at all

○ A little

○ A moderate amount

○ A lot

○ A great deal

Q7 How important is performing well in this course (COURSE X) to you?

○ Not at all important

○ Slightly important

○ Moderately important

○ Very important

○ Extremely important

Q8 How much does being good at the material taught in this course (COURSE X) mean to you for your life?

○ Not at all

○ A little

○ A moderate amount

○ A lot

○ A great deal

Q9-1 Friendship Network Questions How many friends do you have in this course (COURSE X)?

○ 0

○ 1-3

○ 4-6

○ 7-9

○ 10 or more

Q9-2 Please select students in this course (COURSE X) from the dropdown list that you consider your friends.
Note: You can easily locate your friend by typing their first name.

Friend 1 ▼
Friend 2 ▼
Friend 3 ▼
Friend 4 ▼
Friend 5 ▼
Friend 6 ▼
Friend 7 ▼
Friend 8 ▼
Friend 9 ▼
Friend 10 ▼

Q9-3 If you have more than 10 friends, or friends in this course (COURSE X) that you cannot find their name, please do your best to list their first and last names here.

_____

Q9-4 If you did not list any names in the previous question, please tell us why.

○ I already listed their name.

○ I accidentally skipped the question. (If so, you can go back to answer it)

○ I have friends in this course (COURSE X), but I do not know their names.

○ I have friends in this course (COURSE X), but I do not feel comfortable giving their names.

○ I listed their names manually.

Q10-1 Help-Seeking Network Questions How many students do you have in this course (COURSE X) you would ask for advice or help?

○ 0

○ 1-3

○ 4-6

○ 7-9

○ 10 or more

Q10-2 Please select students in this course (COURSE X) from the dropdown list that you would ask for advice or help.
Note: You can easily locate the student by typing their first name.

Person 1 ▼
Person 2 ▼
Person 3 ▼
Person 4 ▼
Person 5 ▼
Person 6 ▼
Person 7 ▼
Person 8 ▼
Person 9 ▼
Person 10 ▼

Q10-3 If you have more than 10 students that you would ask for advice or help, or students in this course (COURSE X) that you cannot find their name, please do your best to list their first and last names here.

_____

Q10-4 If you did not list any names in the previous question, please tell us why.

○ I already listed their name.

○ I accidentally skipped the question. (If so, you can go back to answer it)

○ I have students that I can ask for advice or help in this course (COURSE X), but I do not know their names.

○ I have students that I can ask for advice or help in this course (COURSE X), but I do not feel comfortable giving their names.

○ I listed their names manually.

Q00 Below, we would like to know your feelings and experience in COURSE X throughout the quarter.
Q11 Please indicate how true each statement is for you.

	Very untrue of me	Somewhat untrue of me	Neutral	Somewhat true of me	Very true of me
I felt that students in this course cared about each other.	○	○	○	○	○
I felt connected to others in this course.	○	○	○	○	○
I felt a spirit of community.	○	○	○	○	○

I felt isolated in this course.	○	○	○	○	○
I felt confident that others would support me.	○	○	○	○	○

Q12 Please indicate how true each statement is for you.

	Very untrue of me	Somewhat untrue of me	Neutral	Somewhat true of me	Very true of me
I made good use of my study time for COURSE X.	○	○	○	○	○
I kept a record of what my COURSE X assignments were and when they were due.	○	○	○	○	○
I made sure I kept up with the weekly readings and assignments for COURSE X.	○	○	○	○	○
I usually plan my work in advance so that I can turn in my COURSE X assignments on time.	○	○	○	○	○
I worked hard to do well in COURSE X even if I didn't like what I was doing.	○	○	○	○	○
I often felt so lazy or bored when I studied that I quit before I finished what I planned to do.	○	○	○	○	○
When I did not fully understand something, I asked other classmates in COURSE X for ideas.	○	○	○	○	○
I shared my problems with my classmates in COURSE X so we knew what we were struggling with and how to solve our problems.	○	○	○	○	○

Q18 How satisfied are you with the COURSE X course experience?

- ○ Very Satisfied
- ○ Satisfied
- ○ Neutral
- ○ Dissatisfied
- ○ Very Dissatisfied

Appendix D

Table D1. OLS regression of treatment impact on post-course survey respondents.

	Post-Course Survey Respondents	
Treatment group	0.037	0.030
	(0.039)	(0.038)
Covariates		Yes
Observations	340	340
R-squared	0.003	0.118

Note: Standard errors in parentheses. Significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001.

Table D2. Balance test of OLS regression results.

	Treatment group *Sample: all students with pre-course survey*	Treatment group *Sample: all students with post-course survey*
Gender (reference group: male)		
Female	0.005	0.020
	(0.062)	(0.068)
Non-binary/third gender	0.178	0.695+
	(0.218)	(0.375)
Prefer not to say	-0.052	0.005
	(0.268)	(0.403)
Race/ethnicity (reference group: White)		
American Indian, Alaska Native, Native Hawaiian, or Other Pacific Islander	0.103	0.011
	(0.157)	(0.193)
Black or African American	0.082	-0.000
	(0.153)	(0.169)
Hispanic, Latino, or Spanish origin	-0.005	0.022
	(0.100)	(0.116)
Asian	-0.003	-0.043
	(0.097)	(0.105)
missing	-0.440	-0.446
	(0.516)	(0.521)
Low-income status		-0.026
		(0.071)
First-generation college student status		0.001
		(0.078)
High school GPA	0.009	0.055
	(0.093)	(0.117)
Composite score of course motivation	-0.055	-0.056
	(0.045)	(0.052)
Composite score of peer interaction preference	0.051	0.064
	(0.034)	(0.040)
Composite score of motivation to lead	-0.015	-0.026
	(0.038)	(0.044)
Composite score of stress level	0.040	0.056
	(0.052)	(0.059)
Number of classmates known in the class	-0.028	-0.013
	(0.036)	(0.040)
Number of friends known in the class	0.002	0.025
	(0.088)	(0.095)
Number of students that can ask for help in the class	0.078	0.052
	(0.077)	(0.086)
Entry term control	+	+
Observations	340	288
R-squared	0.028	0.052

Note: Model in Column 1: F(16, 323) = 0.58, Prob > F = 0.8966. Model in Column 2: F(23, 264) = 0.63, Prob > F = 0.9027. Standard errors in parentheses. Significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001.

Table D3. Treatment impact on weekly outcomes.

Week	Study group interaction frequency (course-related)	Study group interaction frequency (non-course-related)	Total study group study times	Observations	Study group satisfaction	Observations
2	0.326**	0.155	0.169	288		
	(0.109)	(0.111)	(0.116)			
3	0.289*	0.139	0.113	288		
	(0.112)	(0.116)	(0.114)			
4	0.176	0.160	0.131	288		
	(0.110)	(0.114)	(0.115)			
5	0.275*	0.326**	0.266*	288	0.222+	263
	(0.110)	(0.111)	(0.112)		(0.123)	
6	0.212+	0.209+	0.169	288	0.332**	268
	(0.112)	(0.114)	(0.113)		(0.119)	
7	0.213+	0.150	0.311**	288	0.245*	263
	(0.111)	(0.114)	(0.111)		(0.123)	
8	0.092	0.062	0.102	288	0.279*	259
	(0.114)	(0.114)	(0.113)		(0.124)	
9	0.279*	0.183	0.236*	288	0.272*	257
	(0.111)	(0.113)	(0.113)		(0.124)	

Note: Sample size varies as missing satisfaction data weren't replaced with zeros. Standard errors in parentheses. Significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001.

Table D4. Descriptive statistics of conversations for treatment and control groups.

Group Treatment Status		Total messages sent			Total messages length			Total messages word counts		
	Observations	Mean	Std. err.	t statistics	Mean	Std. err.	t statistics	Mean	Std. err.	t statistics
Treatment Group	57	119.983	12.890	0.141	3678.649	387.402		764.421	81.232	0.220
Control Group	57	122.702	14.355		3881.123	454.903	0.339	791.035	89.814	

Note: This table uses data at the study-group level because students use nicknames in their group conversations, allowing us to differentiate treatment status only at the group level.

Table D5. Sorting test of leadership disposition and student characteristics.

	Motivation to Lead Category (1 = Low 2 = Medium 3 = High)	
	(1) Original Sample	(2) Analytic Sample
Gender (reference group: male)		
Female	-0.133	-0.196+
	(0.096)	(0.104)
Non-binary/third gender	-0.540	-0.407
	(0.338)	(0.577)
Prefer not to say	-0.447	-0.230
	(0.420)	(0.621)
Race/ethnicity (reference group: White)		
American Indian, Alaska Native, Native Hawaiian, or Other Pacific Islander	-0.099	-0.308
	(0.244)	(0.287)
Black or African American	-0.336	-0.395
	(0.245)	(0.262)
Hispanic, Latino, or Spanish origin	-0.561***	-0.631***
	(0.150)	(0.160)
Asian	-0.445**	-0.453**
	(0.145)	(0.156)
missing	-0.290	-0.417
	(0.481)	(0.805)
High school GPA	0.227	0.314+
	(0.146)	(0.177)
Composite score of course motivation	0.103	0.146+
	(0.069)	(0.078)
Composite score of peer interaction preference	0.122*	0.120*
	(0.054)	(0.061)
Composite score of stress level	0.065	0.053
	(0.080)	(0.088)
Number of classmates known in the class	-0.071	-0.087
	(0.056)	(0.060)
Number of friends known in the class	0.130	0.139
	(0.136)	(0.147)
Number of students that can ask for help in the class	-0.134	-0.105
	(0.119)	(0.132)
Observations	342	288
R-squared	0.099	0.107

Note: Standard errors in parentheses. Significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001.