Self-select groups increase stability but reduce collaboration in high school physics

Diego Ramírez,[1,2,3,*] Eugenio J. Guzmán-Lavín,[1]
Javier Pulgar,[4,3] and Cristian Candia[1,2,5,6]

[1]*Centro de Investigación en Complejidad Social (CICS), Facultad de Gobierno, Universidad del Desarrollo 7550000, Chile*
[2]*Computational Research in Social Science Laboratory, School of Engineering and School of Government, Universidad del Desarrollo, Las Condes 7550000, Chile*
[3]*Grupo de Investigación en Didáctica de las Ciencias y Matemática (GIDICMA), Universidad del BÍO BÍO, Concepción 4051381, Chile*
[4]*Departamento de Física, Universidad del Bío Bío, Concepción 4051381, Chile*
[5]*Instituto de Data Science, Facultad de Ingeniera, Universidad del Desarrollo, Las Condes 7610658, Chile*
[6]*Northwestern Institute on Complex Systems (NICO), Northwestern University, Evanston, Illinois 60208, USA*

(Received 8 August 2024; accepted 30 April 2025; published 9 June 2025)

Group configurations shape social skills and learning outcomes by fostering student interaction and collaboration. Yet, the factors influencing group formation, social cohesion, and group stability have remained understudied in physics education. Here, we examine these factors in a cohort of 90 students from ninth and tenth grades at a private school in Chile. For a semester-long physics project, students were assigned to groups either through random allocation (control group) or given the autonomy to form their own groups (experimental group). Our findings indicate that the groups formed by choice developed stronger friendship bonds throughout the semester, while the groups formed randomly demonstrated greater levels of cooperation and more frequent changes in membership, suggesting weaker social cohesion. Additionally, tenth-grade students were more likely to stay in their original groups, possibly indicating increased social maturity. These findings emphasize the ongoing challenge in education of balancing student choice in group formation with the need for diverse group compositions that enhance learning. This study provides practical recommendations for physics teachers and researchers.

DOI: 10.1103/PhysRevPhysEducRes.21.010157

I. INTRODUCTION

The 21st-century skills framework highlights the growing importance of collaboration, problem solving, and emotional intelligence alongside traditional academic content [1,2]. These competencies do not emerge in isolation; rather, they are acquired through sustained interactive and collaborative learning opportunities. In this context, group work represents a fundamental means by which students can develop and internalize such skills [3–6]. Although numerous studies underscore the value of group activities in education, comparatively less is known about how different group-formation strategies can be applied to optimize collaborative learning in physics specifically.

*Contact author: dieramirez@udd.cl

To address this gap, it is helpful to consider how groups are commonly formed in everyday classroom settings. Educators may create student groups randomly [7], allow students to self-select [8], assign groups based on teacher judgments [9], or combine these approaches. One key indicator of an effective group is its stability over time [10,11], which fosters members' commitment, trust, and shared accountability [12]. Stable group membership has been linked to heightened motivation and a stronger sense of belonging, thereby reducing student anxiety and encouraging active participation [13]. It also provides opportunities for students to practice crucial interpersonal skills such as conflict resolution and effective communication—competencies that extend beyond the classroom [12]. In such cohesive groups, students are more inclined to share ideas freely and take academic risks, resulting in deeper learning and the development of critical-thinking skills [14–17]. Nonetheless, not all groups spontaneously achieve high levels of stability, underscoring the importance of ongoing monitoring and timely instructional support [12].

In light of the central role that group stability plays in student engagement and skill development, active learning methodologies—particularly project-based learning (PBL) and collaborative problem solving—have gained prominence for their potential to foster both group cohesion and productive teamwork [18–21]. By integrating complex problem-solving tasks with authentic, real-world applications, PBL has been found to boost engagement and conceptual understanding [22–25]. Similarly, collaborative problem-solving methods encourage sustained student interaction and accountability, thereby enhancing academic performance and promoting more positive attitudes toward learning [26–29]. In an educational context that places increasingly values on social cooperation and collaboration, understanding how group formation influences these outcomes becomes especially salient [6,15,30–32].

Despite the growing emphasis on group-based pedagogical approaches, the effects of self-select versus randomly assigned groups on friendship, cooperation, and academic prestige—particularly students' recognition of their peers' physics expertise—remain unclear. Notably, few studies have examined whether these contrasting strategies influence students' willingness to continue working together over time, underscoring the need for further empirical research to advance this field of study. Addressing this gap in physics education research, the present study investigates how ninth- and tenth-grade students form, maintain, or change groups across a semester and how such decisions impact their collaborative, friendship, and prestige networks within the physics classroom. Specifically, we compare self-select groups with randomly assigned ones in terms of students' social connections and academic outcomes, considering these social connections at both the class and group levels. To capture these relational dimensions, we employ a social network survey that measures friendship, collaboration, and perceived physics expertise. This dual approach accounts for the directed nature of social ties, group-level variability, and the structure of our outcome measures, offering a comprehensive understanding of how group-formation strategies shape social and academic dynamics.

II. LITERATURE REVIEW

Previous findings have shown that positive group dynamics, such as small group work, improve collaboration, increase engagement, and sharpen problem-solving skills [33,34]. In physics education, where conceptual complexity demands deeper cognitive processing, collaborative structures become especially beneficial [35,36]. Indeed, effective group work in physics classrooms not only increases conceptual understanding but also cultivates critical thinking and advanced problem-solving abilities [37]. Equally important, these cooperative environments foster key social and emotional skills—including communication, empathy, and conflict resolution—that are central to success in team-based learning [11,38,39].

The cognitive and social benefits of group participation are highly dependent on team cohesion [40]. Cohesion promotes unity among group members, reducing interpersonal barriers and strengthening group identity [41]. By facilitating open communication, reinforcing group norms, and enabling collective responses to challenges, cohesive teams are generally better equipped to maintain productive working relationships [26]. This unity also encourages friendships, which can enhance trust, communication, and mutual accountability [42,43]. The result is a shared sense of purpose that drives students toward collective goals [44,45].

Further research highlights that stable, cohesive groups promote a deeper sense of belonging and community, which are essential for student motivation and long-term achievement [46,47]. Structured teaching methods, such as PBL and collaborative problem solving, not only enhance academic performance but also build the social competencies essential for scientific collaboration [6,48]. Notably, group stability helps address many common challenges in collaborative learning by providing a consistent peer network that strengthens both individual and group accountability [49].

Modern instructional approaches increasingly center on group-based tasks, attributing them to improvements in critical thinking, heightened creativity, and stronger social connections among learners [34,48]. Understanding how groups form, and why students choose to remain in or leave these groups, can substantially enrich the educational experience [30,50]. Yet, each method of group formation—whether random, teacher-guided, or self-select—entails its own benefits and trade-offs. Randomly assigned groups can stimulate new social connections but may produce imbalances in skill levels and decrease overall team efficacy [7]. Conversely, self-select groups often promote cohesive interactions, yet they may limit diversity of perspectives and risk alienating those not chosen by their peers [51–53]. Teacher-guided groups aim to mitigate these issues by strategically pairing students based on strengths and weaknesses, although this practice demands substantial pedagogical insight and a nuanced grasp of classroom dynamics [54].

Given these varied approaches, research offers mixed findings on which factors most strongly shape group dynamics. Some studies suggest that instructor-organized or hybrid methods may yield more favorable learning experiences [55]. Nonetheless, there is broad consensus that the ultimate objective is to create an optimal group environment that fosters robust learning and intellectual diversity [56,57]. Achieving this balance necessitates that educators weigh student autonomy against the need for diversity, all while recognizing the importance of acknowledging each member's expertise as a cornerstone of effective collaboration [58]. This social awareness ensures that teams benefit from specialized skills while also making students feel valued for their unique contributions [53].

In this context, the present study focuses on three social constructs considered key to understanding team dynamics in the physics classroom: friendship, cooperation, and academic prestige. From a social network perspective [3], these concepts are operationalized through self-nominated ties (in-degree) collected via student surveys. Friendship is conceived as bonds of closeness and trust that facilitate interaction and idea sharing [42,43]; cooperation refers to the willingness and practice of working together and providing mutual support when solving problems [12,30]; and academic prestige reflects the recognition of each student's expertise in physics, influencing how their contributions are valued [11,39]. With these definitions in place, a social network approach allows for a detailed understanding of how different group-formation strategies—either self-select or randomly assigned—affect group cohesion and the decision to keep working with the same peers throughout the semester. The next section describes the specific methods used to measure and analyze these constructs in the context of physics instruction.

Building on the diverse factors that influence effective group performance—particularly within the context of PER—this study investigates how two specific group-formation strategies (self-select and random assignment) influence students' friendships, cooperation, and recognition of their peers' physics expertise (prestige). Specifically, this work is framed around the following research questions:

1. Do differences exist in team friendship, cooperation, and recognition of academic prestige between students in randomly assigned groups and those in self-select groups after a semester?
2. Are there variations in the number of team members with whom a student chooses to continue working, between those in randomly assigned groups and those in self-select groups, after a semester?

III. DATA AND METHODS

A. Research context and participants

The study was conducted during the 2021 academic year at a private school in Chile. Participants (see Table I) were ninth and tenth-grade students (ages 14–16) enrolled in four physics courses. The focus was on physics lessons taught by the same teacher. For clarity, the ninth-grade classes are referred to as C1 (Control group 1) and E1 (Experimental group 1), and the tenth-grade classes as C2 and E2. Active learning methodologies were implemented in each of the four classes, designed for group work, and aligned with the physics curriculum content for ninth- (e.g., oscillations and waves) and tenthth-grade students (e.g., particle motion and dynamics). These group activities (see Table II) were specifically designed to foster active student participation in class content, including activities such as discussing and analyzing conceptual errors in others' responses, creating posters, and designing practical experiments to illustrate physical phenomena.

TABLE I. School courses information.

Course	Condition	Grade level	Female	Male	Total
C1	Control	Ninth	14	10	24
C2	Control	Tenth	11	9	20
E1	Experimental	Ninth	14	10	24
E2	Experimental	Tenth	9	13	22
Total			48	42	90

In 2021, as COVID-19 cases decreased and vaccination rates increased, schools began offering in-person learning again. While some students returned to the classroom, others continued with online instruction. To manage health protocols, the school implemented a weekly rotation system based on alphabetical order, dividing each class into two groups. One half attended in-person sessions during a given week, while the other participated online, alternating weekly. The teacher observed that group activities were conducted smoothly in both formats, with all students having the necessary resources to participate online, ensuring consistent peer interaction. However, specific differences in group participation between online and in-person settings were not formally assessed in this study, as the focus was on understanding group dynamics and stability across experimental and control conditions.

B. Data collection and intervention

The study received approval from the school's institutional committee, and participation was voluntary. Data were collected during the first and second semesters of

TABLE II. Description of activities conducted during the first semester. Each activity lasted five sessions, after which students presented their final products (poster, flyer, book, video, and guide).

C1 and E1 (ninth grade)
Activity 1: Based on the content of waves and sound (e.g., classification, characteristics, properties) covered in classes, they developed conceptual questions and provided answers. They then used these questions to survey people in their immediate surroundings, analyze the respondents' answers, and select the one that presented the most conceptual errors. Considering these errors, they created a poster to appropriately address the selected question.
Activity 2: They created an informative and problem-solving poster on the effects of pandemics on society, rating history, physics, and biology.
C2 and E2 (tenth grade)
Activity 1: They developed a practical work (book, video, magazine, poster, script, etc.) to explain some topics related to astronomy (e.g., Kepler's laws, telescopes, and universal gravitation), describing it through the use of themes indicated by their teachers (chemistry and physics) and those selected by their group.
Activity 2: Based on a scientific text related to the kinematics unit, they answered a list of questions. They then used data tables to develop calculations and graphs to respond to a series of questions.

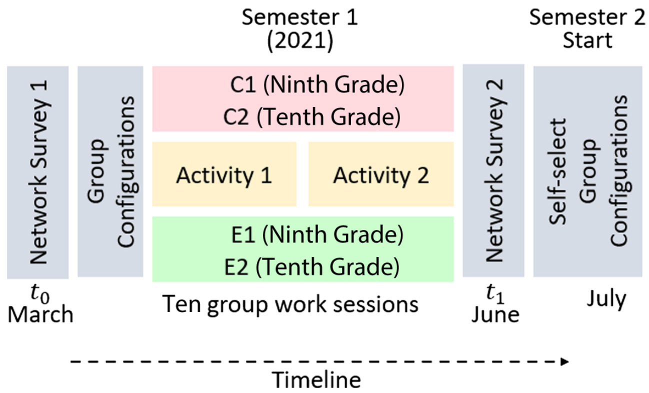

FIG 1. Diagram of the process adopted in the research. Before forming the groups (t_0), the first network survey was conducted, addressing exclusively friendship relationships among students. After the teamwork sessions (activities 1 and 2) at time t_1, the instrument was reapplied, incorporating two items related to perceptions of collaboration and recognition of prestige in the subject.

2021, from March to July, in line with Chile's academic calendar (see Fig. 1). Table I summarizes the total number of participants for each course. Data on students' group membership, grades in group activities, and gender were collected. At the beginning of the first semester (see Fig. 1, time t_0), a social network survey was administered to measure friendship levels among students [11]. The survey was repeated at the end of the semester (see Fig. 1, time t_1) to include questions on physics prestige and collaboration (see Sec. III C). The network surveys were administered during regular class hours to ensure full participation. All students in the study ($N = 90$) completed the surveys at both time points (t_0 and t_1), resulting in a 100% response rate. This approach minimized the likelihood of missing data and ensured that the collected information accurately represented the social dynamics of the entire sample.

For the intervention, students in the control courses (C1 and C2) were randomly assigned to groups by the teacher at the beginning of the first semester, using a standard web-based tool for group creation. Conversely, students in the experimental courses (E1 and E2) formed their own groups autonomously, without teacher intervention. At the start of the second semester, all students had the option to change groups and independently define their new groups. Six work groups were formed in each course, and the group composition remained stable throughout the first semester, with adjustments allowed at the beginning of the second semester (see Fig. 1).

The teams participated in group activities aligned with the national curriculum. Each activity spanned 5 to 7 sessions of 45 min, totaling at least 10 collaborative sessions per semester (see Table II).

Students were informed that their grades would be based on group activities, with specific evaluation instructions and detailed rubrics provided. The activities were designed to be completed during class time, in compliance with school policies discouraging group work outside the classroom.

Each week, students had two 45-min sessions exclusively dedicated to these activities. The teacher, who taught all four courses, implemented the group activities in parallel to ensure consistency. The activities were aligned with institutional capacities, available class time, and the national curriculum, focusing on teamwork.

C. Instruments

We used an online network survey to capture student relations [11,20]. This survey was designed to collect information about three key social dimensions: friendships, physics-related prestige, and collaborative ties. The survey questions were as follows:

1. From the students in the classroom registration, indicate those who are your friends.
2. From the students in the classroom roster, indicate those you consider high-achieving physics students.
3. From the students in the classroom roster, indicate those you collaborated with during the classroom activities.

The decision to use a directed network survey is grounded in its ability to capture asymmetric relationships, which are essential for understanding the dynamics of educational group settings. Directed ties, such as those of friendship or collaboration, provide valuable insights into students' social interactions and their perception of others' academic prestige. Prior studies in social network analysis have demonstrated that these directed relationships play a crucial role in fostering group cohesion, trust, and collaboration, all of which are fundamental to effective group functioning in educational environments [59].

Friendship networks were particularly emphasized in this study because of their influence on students' sense of belonging, group stability, and academic motivation. Research suggests that a strong sense of community, characterized by students' acceptance and inclusion, is crucial to their well-being and performance [46]. In addition, supportive interpersonal relationships among students and between students and teachers foster intrinsic motivation and self-regulation, which can contribute to greater group cohesion. Moreover, friendship networks have been found to mitigate the challenges of group dynamics by fostering trust and a shared sense of purpose [42,43].

The inclusion of physics-related prestige and collaborative ties in the survey aligns with prior literature highlighting the importance of recognizing individual contributions to group performance and understanding how collaborative behaviors evolve [60]. Such an approach allows for a more comprehensive analysis of the interplay between cognitive, social, and emotional factors in group dynamics.

Finally, the use of social network surveys to study group formation is supported by extensive literature in network science, which underscores their utility in capturing the complex interactions that underpin group-based learning. This methodology enables educators and researchers to identify patterns of interaction, measure the strength of relational ties, and evaluate the impact of group configurations on learning outcomes [61].

D. Data analysis

Network surveys were conducted to analyze student group configurations, focusing on friendship, prestige, and collaboration at both the group and overall course levels. These surveys captured indirect networks, including nonmutual relationships. Additionally, a membership network was created to identify each student's group affiliation. By combining these networks, we examined perceptions of friendship, prestige, and collaboration within and between groups. For instance, the variable "team friendship" represents relationships among members within the same group, while the variable "friendship" captures relationships among all students within the same course, as illustrated in Fig. 2. Gephi 0.9 software was used to measure key network characteristics, such as student centrality, including degree, in-degree (how others perceive the individual), and out-degree (how the individual perceives others). Additional details on these metrics are provided in Table VI in the Appendix.

To analyze the relationships captured by the network surveys, we employed ordinal mixed-effects regression models (clmm) using the ordinal package in R [62]. These models were particularly suitable for our data, as

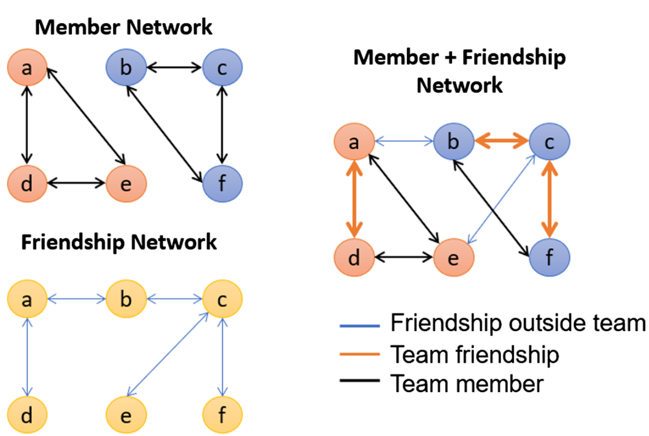

FIG 2. This diagram illustrates how we identify key factors within and between student work groups. Black lines show who belongs to each group, with two example sets of members highlighted in different colors. The friendship network is also shown, where lines connect students who consider each other friends. When we combine these two networks, we get a comprehensive view of friendships both inside (in orange) and outside (in sky blue) of each group. The same approach is used to analyze prestige and cooperation among students. Note that the black lines serve only to show group membership and do not add extra information.

the dependent variables—friendship, cooperation, and prestige—were measured on ordinal scales. The inclusion of random effects at the group level allowed us to account for unobserved heterogeneity between groups. In contrast to linear models, the coefficients in ordinal regression models represent changes in the log-odds of being in a higher category of the dependent variable, rather than linear changes in the outcome itself. As our analysis focused on the predictors' effects on social dynamics, we do not interpret the threshold parameters. These thresholds, which define the boundaries between ordinal categories, are included in the Appendix (see Tables XVI and XVII) for reference but are not central to the study's objectives.

To address the first research question, we modeled within-group friendship, cooperation, and prestige as dependent variables, while controlling for factors such as group size, gender, grade level, and average grades. The predictors included measures of initial friendship within groups (team friendship) and across the entire course (friendship), as well as the experimental condition (self-select groups vs randomly assigned). These models allowed us to evaluate how preexisting ties and group characteristics influenced the final social dynamics within the teams.

For the second research question, focused on group stability, we characterized the groups formed in each course by identifying the total number of groups, the number of members per group, and the percentage of women in each semester. Additionally, a flow diagram was created to understand and visualize the conditions under which students from groups formed at the beginning of semester 1 continued working with more or fewer peers when given the option to switch groups at the start of semester 2. These analyses were complemented by ordinal mixed-effects regression models (models 9 to 11) to examine the number of original team members (0–3) with whom a student continued working. The same covariates and predictors used for the first research question were included, allowing for a comparison of the factors influencing both initial social dynamics and longer-term group cohesion.

For a set of models (models 10 and 11), we included friendship at t_0 as a predictor instead of team friendship at t_0. This decision was made to avoid potential collinearity, as the dependent variable directly reflects how many of the original group members were retained. Including team friendship would have introduced conceptual and statistical redundancy. In contrast, friendship at t_0 captures students' broader social preferences across the class and serves as a more independent predictor of group retention outcomes.

Model comparisons were conducted within each outcome set using the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), which evaluate both model fit and complexity. These criteria were used to assess whether additional predictors or interaction terms improved explanatory power while maintaining parsimony. AIC and BIC values are reported for each model group (friendship, cooperation, prestige, and group retention), and guided—but did not solely determine—the interpretation of results in the subsequent sections.

To ensure the robustness of our results, we conducted complementary analyses. First, we used the check collinearity function from the performance package to confirm the absence of multicollinearity issues in our ordinal models. Second, we fitted linear regression (lm) models as proxies, treating the ordinal dependent variables as continuous, to compare collinearity diagnostics and ensure consistent patterns of significance. Third, we tested alternative link functions (logit, probit, and cloglog) and fitted clm models with nominal terms to relax the proportional odds assumption, assessing their fit using AIC and BIC criteria. Fourth, residual analysis was performed to evaluate the models' goodness-of-fit and verify assumptions of independence and normality in random effects. These robustness checks, detailed in the Appendix, demonstrated that our findings were stable across different model specifications and assumptions.

This multifaceted analytical approach provided a comprehensive examination of the social dynamics captured by the network surveys, ensuring that our conclusions are both statistically and conceptually robust.

IV. RESULTS

A. Social networks

We used ordinal mixed-effects regression models to analyze team-level network characteristics at the end of the first semester (Table III). These characteristics include in-degree of team friendship (models 1–4), in-degree of team cooperation (models 5 and 6), and in-degree of team prestige (models 7 and 8). These models incorporated random effects at the group level to account for unobserved heterogeneity between teams and controlled for variables such as group size, gender, grade level, and average grades. Additionally, initial team friendship and overall friendship before the intervention were included as predictors, along with the experimental condition (self-select vs randomly assigned groups).

Model 1 indicates that initial team friendship [team friendship (t_0)] is a strong positive predictor ($p < 0.001$) of team friendship at the end of the semester, highlighting the influence of preexisting ties. Model 2 includes an interaction term between the experimental condition and initial team friendship, which significantly improves the model's fit. The interaction term is negative ($p < 0.01$), suggesting that the positive effect of initial team friendship on final

TABLE III. Regression models: predicting within-group friendship, prestige, and cooperation links at time t_1. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

	Dependent variable							
	Team friendship (t_1) (in-degree)				Team cooperation (t_1) (in-degree)		Team prestige (t_1) (in-degree)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exp. (self-select groups)	1.51	4.25**	4.16***	5.09***	−3.76**	−4.96**	−0.69	−0.68
	(1.13)	(1.57)	(1.20)	(1.06)	(1.19)	(1.56)	(0.75)	(1.10)
Group size (3)	−2.02	−2.18†	−2.23	−2.26	−6.46***	−6.49***	−0.56	−0.57
	(1.39)	(1.32)	(1.54)	(1.55)	(1.64)	(1.71)	(0.80)	(0.81)
Gender (male)	0.49	0.52	−0.34	−0.34***	−0.61	−0.50	−1.03*	−1.03*
	(0.64)	(0.66)	(0.60)	(0.59)	(0.66)	(0.69)	(0.49)	(0.49)
Grade level (tenth)	0.68	0.67	1.96	1.97†	4.64***	4.77**	−0.44	−0.44
	(1.16)	(1.09)	(1.23)	(1.08)	(1.38)	(1.45)	(0.68)	(0.68)
Average grades	0.04	0.05	−0.02	−0.02***	−0.16**	−0.17*	0.14**	0.14**
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.07)	(0.05)	(0.05)
Team friendship (t_0)	2.87***	4.77***			0.72†	−0.26	0.86**	0.86†
(in-degree)	(0.55)	(1.02)			(0.38)	(0.80)	(0.33)	(0.50)
Exp.* team friendship (t_0)		−2.61**				1.29		−0.01
(in-degree)		(1.01)				(0.94)		(0.64)
Friendship (t_0)			0.32**	0.43***				
(in-degree)			(0.11)	(0.00)				
Exp.* friendship (t_0)				−0.17**				
(in-degree)				(0.00)				
Group effect	Random	Random	Random	Random	Random	Random	Random	Random
Number of observations	90	90	90	90	90	90	90	90
AIC	155.1	149.3	186.4	187.8	161.3	161.3	209.6	211.6
BIC	180.1	176.8	211.4	215.3	186.3	188.8	234.6	239.1
Log likelihood	−67.529	−63.658	−83.194	−82.900	−70.632	−69.632	−94.822	−94.822
edf	10.00	11.00	10.00	11.00	10.00	11.00	10.00	11.00

team friendship is weaker in self-select groups compared to randomly assigned groups. Models 3 and 4 replace team friendship (t_0) with a broader measure of friendship (t_0), showing consistent patterns: initial friendships positively predict final team friendship ($p < 0.01$), but the interaction term with the experimental condition is negative ($p < 0.01$).

In model 4, gender (male) emerges as a significant predictor ($p < 0.001$), indicating that male students are less likely to receive incoming friendship ties within their teams compared to female students. This result underscores the potential influence of social dynamics and gender-based interactions within groups, suggesting that female students may foster stronger interpersonal connections, regardless of the grouping condition. These results highlight that while self-select groups start with stronger initial ties, these do not translate proportionally into higher levels of final team friendship compared to randomly assigned groups. Furthermore, the role of gender in shaping social ties within teams reflects the nuanced interplay between demographic factors and team dynamics.

Models 5 and 6 reveal a significant negative effect of the experimental condition ($p < 0.01$), indicating that self-select groups displayed fewer cooperative ties by the end of the semester. Group size ($p < 0.01$) negatively predicts cooperation, indicating that smaller groups (of three members) tend to report lower levels of cooperation compared to groups of four. This result may reflect a structural limitation, as 3-member teams have fewer possible cooperation ties (maximum in-degree of 2) than 4-member teams (maximum in-degree of 3). In contrast, grade level ($p < 0.001$) positively predicts cooperation, highlighting the role of maturity in fostering collaborative efforts. Initial team friendship [team friendship (t_0)] has a marginally positive effect in model 5 ($p < 0.1$), but this becomes nonsignificant when interactions are included in model 6. Average grades show a small but significant negative effect ($p < 0.05$) across both models, suggesting that higher academic performance may not necessarily align with cooperative behavior in group settings.

In models 7 and 8, the experimental condition does not have a significant impact on team prestige. Gender (male) is negatively associated with prestige ($p < 0.05$), indicating that female students tend to receive more recognition for academic abilities. Average grades positively predict prestige ($p < 0.01$), suggesting that students with higher academic performance are more likely to be recognized for their abilities. Initial team friendship [team friendship (t_0)] is positively associated with prestige in model 7 ($p < 0.01$), but its significance diminishes slightly in model 8. No significant interaction terms are observed, indicating that the effects of academic performance and initial ties on prestige are consistent across both group configurations.

Overall, these models highlight the critical role of initial friendships and group characteristics in shaping final team dynamics. Self-select groups show stronger initial friendships but fail to translate these into cooperative or prestigious relationships within teams, as seen in randomly assigned groups. The findings suggest that grouping strategies can significantly influence social relationships and collaboration within educational settings, with different configurations fostering distinct dynamics of team friendship, cooperation, and prestige.

B. Group changes

Six work groups were formed in each of the ninth- and tenth-grade classrooms in the first semester. In the ninth grade, each group had four members, while in the tenth grade, the group sizes ranged from three to four members (Table IV). In the second semester, ninth-grade groups remained at four members each, but tenth-grade groups varied in size from two to four members. Regarding gender composition, experimental groups in both grades showed a tendency for same-gender grouping: four all-female and five all-male groups in the first semester. Only three groups had a mix of genders. This pattern continued into the

TABLE IV. Description of work team members, including the number of members and the percentage of women in each group, for both first and second semesters.

	Semester 1		Semester 2	
Group	Members	% Women	Members	% Women
Group C1 (ninth grade)				
1	4	50	4	75
2	4	25	4	50
3	4	40	4	100
4	4	50	4	25
5	4	75	4	0
6	4	100	4	100
Group C2 (tenth grade)				
1	3	33.3	2	100
2	3	33.3	3	0
3	3	66.7	3	33.3
4	3	66.7	4	0
5	4	50	4	100
6	4	75	4	75
Group E1 (ninth grade)				
1	4	75	4	100
2	4	0	4	0
3	4	100	4	75
4	4	0	4	50
5	4	75	4	100
6	4	100	4	25
Group E2 (tenth grade)				
1	3	0	4	100
2	4	50	3	100
3	4	100	3	66.7
4	3	100	4	0
5	4	0	4	0
6	4	0	4	0

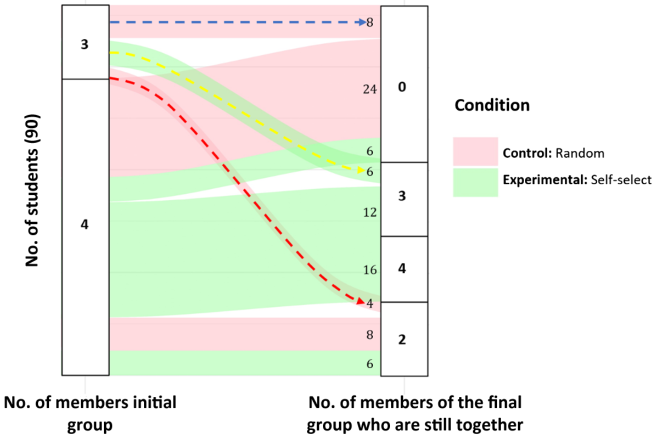

FIG 3. Student group transitions in control and experimental settings. In the control condition, students are randomly assigned to groups (shown in pink), while in the experimental condition, they form groups based on self-select (indicated in green). The arrows, sized proportionally, show the number of students who chose to either stay with their original group members or switch to new peers when given the option.

second semester. In contrast, control groups had only one all-female group in the first semester. However, when students had the chance to change groups at the beginning of the second semester, the number of single-gender groups increased to four all-female and three all-male groups across both grades.

To provide a detailed view of group changes, Fig. 3 shows the student flow in both experimental and control classrooms across the first and second semesters. Out of 90 students, 38 ended up in new groups without any of their original team members; 32 of these (84%) were from control (randomly assigned groups) classrooms C1 and C2. On the other hand, 52 students remained with at least one original group member in the second semester, with 40 of these (76.9%) being from the experimental condition (self-select groups).

The flow diagram in Fig. 3 focuses exclusively on the analysis of groups that were initially formed with three students, both in the experimental and control conditions. This choice allows for a clearer comparison of group stability across the two conditions. In the experimental condition, six students who started in groups of three maintained the same group composition in the second semester, as indicated by the yellow dashed arrow. In the control condition, groups also began with three members in the first semester. Of these, eight students (marked by the blue dashed arrow) transitioned to completely new groups in the second semester, while four students (highlighted by the red dashed arrow) remained with two members from their original group. It is important to note that Fig. 3 does not indicate whether the new groups formed in the second semester were the same size as the original groups from the first semester.

TABLE V. Regression models: predicting the number of students with whom an individual continues to work when given the option to change groups. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

	Dependent variable		
	n students with whom a person continues to work		
	Model (9)	Model (10)	Model (11)
Exp. (self-select groups)	7.35^{***}	7.42^{***}	5.19^{***}
	(1.51)	(0.00)	(0.00)
Group size	−2.41	−2.36	−2.71
	(1.92)	(1.89)	(2.09)
Gender (male)	−0.51	-0.53^{***}	-0.74^{***}
	(0.71)	(0.00)	(0.00)
Grade level (tenth)	3.33^{*}	3.44^{*}	3.99^{*}
	(1.55)	(1.44)	(1.58)
Average grades	0.08	0.08^{***}	0.09^{***}
	(0.06)	(0.00)	(0.00)
Friendship (t_0) (in-degree)		0.11^{***}	-0.24^{***}
		(0.00)	(0.00)
Exp.* friendship (t_0) (in-degree)			0.57^{***}
			(0.00)
Group effect	Random	Random	Random
Number of observations	90	90	90
AIC	157.0	158.1	155.8
BIC	179.5	183.1	183.3
Log likelihood	−69.49	−69.07	−66.88
edf	9.00	10.00	11.00

After examining how students moved between groups, we used ordinal mixed-effects regression models (clmm) to examine how the treatment influenced the number of team members a student chose to continue collaborating with when offered the opportunity to form new groups. The models in Table V reveal that the treatment (self-select groups) was a strong positive predictor ($p < 0.001$) across all three models, indicating that students in self-select groups were significantly more likely to retain their original team members compared to those in randomly assigned groups.

Model 9 establishes the baseline, showing that the experimental condition had the strongest overall effect. Grade level (tenth grade) also exhibited a significant positive impact ($p < 0.05$), suggesting that older students were more inclined to maintain their initial group composition. However, gender (male) and group size did not show significant effects in this model, indicating that neither the number of group members nor the gender of the participants were key predictors of retention at this stage. Average grades, while positively associated, did not reach significance ($p = 0.08$), suggesting a weaker relationship with group retention when considered independently.

Model 10 incorporates friendship (t_0) as a predictor, revealing a significant positive association ($p < 0.001$) between initial friendship ties and group retention.

This inclusion also strengthens the role of average grades, which become a significant predictor ($p < 0.001$). This suggests that students with stronger preexisting friendships and higher academic performance were more likely to continue working with their original peers. Gender becomes significant in this model ($p < 0.001$), indicating that male students were slightly less likely to retain their original group members compared to female students. As in model 9, group size remains nonsignificant.

Model 11 introduces an interaction term between the experimental condition and friendship (t_0), which is significant and positive ($p < 0.001$). This finding indicates that the positive effect of initial friendships on group retention was amplified in self-select groups compared to randomly assigned groups. Gender and average grades remain significant in this model ($p < 0.001$), while group size continues to show no significant effect. Interestingly, the direct effect of friendship (t_0) changes direction in this model ($p < 0.001$), suggesting that the interaction term captures nuances in how initial social ties function under different grouping conditions.

Overall, these results underscore the critical role of the experimental condition and initial friendship ties in shaping students' group retention decisions. The interaction between self-select grouping and friendship ties highlights the importance of social dynamics in fostering group stability. While academic performance and gender also play roles, their effects are secondary to relational factors, demonstrating that decisions to retain peers are influenced more by preexisting friendships and group autonomy than by individual characteristics such as grades or group size.

V. DISCUSSION

The present study aims to investigate two primary research questions. First, we examined whether students in self-select groups differed from students in randomly assigned groups in terms of friendship networks, collaboration, and recognition of physics-related prestige. Second, we explored whether these group-formation strategies affected students' decisions to remain with the same teammates or switch teammates after a semester. We first focus on the contrast between experimental (self-select) and control (random) conditions, then consider grade-level variations.

A. Random and self-select groups

Our findings reveal significant differences in team dynamics depending on whether groups were formed randomly or based on self-select. These last groups showed higher levels of final friendship ties (t_1) ($p < 0.01$, models 2, 3, and 4), consistent with research suggesting that preexisting friendships often lead to greater initial cohesion and trust [11,42,43]. However, the negative interaction between initial friendship (t_0) and the experimental condition ($p < 0.01$, models 2 and 4) indicates that while self-select groups start with higher initial cohesion, this advantage may limit the formation of new ties. This aligns with studies suggesting that highly cohesive groups may become less open to new strategies or connections [41,63]. In contrast, randomly assigned groups exhibited a stronger relationship between initial and final friendship ties, which may reflect an environment more conducive to building new social connections.

In terms of cooperation, randomly assigned groups achieved higher levels of collaborative ties by the end of the semester ($p < 0.01$, models 5 and 6). This result suggests that the need to develop collaborative strategies from scratch in these groups may have fostered a more open and adaptive dynamic. Previous research has shown that the absence of preexisting ties can lead groups to work more actively toward developing cooperative dynamics [64]. In contrast, self-select groups, despite starting with higher levels of social cohesion, did not translate this advantage into cohesive collaboration, possibly due to complacency in highly cohesive groups [41], which could have encouraged off-topics conversations rather than work-related interactions. Additionally, group size had a significant negative impact on cooperation, suggesting that smaller groups may face more challenges in maintaining coordination and cohesion, possibly due to structural constraints on the number of interaction ties that can be formed in teams of only three members.

Regarding academic prestige, results show that female students received greater recognition for their academic abilities than male students ($p < 0.05$, models 7 and 8). This finding highlights the complexity of gender biases, which do not always favor men in STEM contexts [65,66]. While this study does not identify the specific conditions that contributed to this shift, prior research has suggested that the inclusion of more equitable pedagogical norms may help mitigate these biases [67]. Interestingly, this finding contrasts with recent research in university-level physics education reporting that women are less likely to receive peer recognition in terms of academic prestige [68,69]. This divergence may reflect important contextual differences between secondary and tertiary education. Age-related social norms, classroom dynamics, and teacher interventions in high school settings might promote more equitable peer perceptions. These contrasting patterns highlight the importance of examining how gender recognition processes unfold across educational stages and institutional contexts. Additionally, average grades were a significant predictor of academic prestige ($p < 0.01$), indicating that academic performance remains a key criterion for recognition, regardless of gender or experimental condition.

These findings have important implications for the design of educational environments in physics and other STEM disciplines. Group configurations should balance initial cohesion with opportunities to form new connections and foster collaborative behaviors. Self-select groups offer advantages in terms of initial cohesion, but randomly assigned groups seem more effective at promoting dynamic collaborative networks. Furthermore, strategies such as rotating leadership roles and providing constructive peer feedback can help ensure that all students, regardless of gender or background, contribute meaningfully [33,70,71].

B. Group stability and member loyalty

Our analyses show that the experimental condition (self-select groups) significantly predicted group stability. Students were much more likely to continue working with the same teammates when given the option to switch ($p < 0.001$, models 9–11). This finding highlights how strong initial social ties, combined with the autonomy afforded by self-select group formation, strengthen group loyalty. This dynamic aligns with Osterman's [46] perspective that a sense of belonging is a key motivator for sustained collaboration, underscoring the role of relational bonds in maintaining group cohesion.

However, while self-select groups demonstrated higher stability, randomly assigned groups reported greater levels of cooperation ($p < 0.01$, models 5 and 6). This suggests a balance between relational cohesion and task-oriented collaboration. Consistent with research on group functioning, these findings illustrate the dual nature of team cohesion. While close relationships foster trust and comfort, they do not always lead to adaptive strategies or task-focused behaviors [12,40]. These results emphasize the importance of considering both relational and functional aspects of cohesion when designing collaborative learning environments.

Our findings also align with broader studies exploring how group dynamics influence participation and retention, particularly among female students in STEM disciplines [65]. While our analysis did not explicitly address inclusive practices, the findings regarding academic recognition of female students in self-select groups highlight the need to further investigate how group dynamics shape equity in recognition and participation. Previous research suggests that group norms and teacher interventions play a critical role in fostering engagement and perceptions of fairness in these contexts [15,48].

Although students with higher academic performance were slightly more likely to remain in their initial group ($p = 0.08$, model 10), this effect was secondary to the influence of friendship ties and group dynamics. In other words, relational factors outweighed academic considerations in determining group loyalty, consistent with prior research on the interplay between social connections and educational outcomes [4,12]. These patterns highlight the importance of fostering stable social networks alongside academic competencies in educational settings. Effective group configurations and collaborative norms can play a key role in achieving these dual goals, providing a foundation for both collaborative success and equitable participation [30].

Taken together, these results highlight the nuanced relationship between social and academic factors in group dynamics. While self-select groups strengthen social bonds, randomly assigned groups seem to foster more task-oriented collaboration. This suggests that each strategy offers specific advantages that may be more effective depending on the educational context and pedagogical goals. Future research should explore how specific classroom practices, such as rotating roles or structuring collaborative tasks, can balance these dynamics to optimize both group stability and productivity.

C. Comparison of ninth and tenth grade classrooms

When examining grade-level differences, we found that tenth-grade students demonstrated higher levels of teamwork compared to ninth graders ($p < 0.05$, models 5 and 6). This difference may be related to their more developed social-emotional skills, as older students often exhibit greater capacity for managing collaborative tasks. Similarly, tenth graders are less likely to change teammates after one semester, relative to their younger counterparts ($p < 0.05$, models 9 and 10). This suggests that their cooperation may have been primarily task oriented and with an important value on the stability and the effectiveness achieved by their original teams. Conversely, ninth-grade students are more drawn to reconfigure their teams based on their initial friendship relationships when given the opportunity in the second semester, and regardless of the potential work-related effectiveness achieved in the previous academic term.

This finding highlights the potential benefits of strategic scaffolding during early developmental stages. Research suggests that younger learners benefit significantly from explicit group building activities that foster trust and interpersonal growth [72]. Furthermore, tailoring team configuration to stimulate social exploration and adaptability might offer students stronger opportunities toward team monitoring, communication, and coordination, along with the possibility to diversify their support networks. Previous studies emphasize the value of creating structured and age-appropriate opportunities for collaboration, which can enhance teamwork and social development simultaneously [15]. With well-planned interventions, educators can support both the collaborative and interpersonal dimensions of learning, ensuring that students at all grade levels receive the scaffolding needed for effective teamwork.

D. Implications and limitations

These findings underscore the need for careful planning when configuring group formations in physics classrooms. For instance, instructors may consider a balanced approach that harnesses the stability benefits of self-select groups while also promoting the diversity and cooperative behaviors observed in randomly assigned groups. Incorporating scaffolded team-building exercises, setting explicit collaborative goals, and rotating certain roles within each group can help ensure that strong preexisting social bonds do not overshadow task-oriented collaboration. To address these challenges, educators could implement structured activities that build trust in randomly assigned groups and employ explicit role assignments or collaborative frameworks to enhance cooperation within self-select groups [34,48]. Furthermore, the study highlights the importance of tailoring group-formation strategies to the developmental and social needs of students. For instance, younger students may benefit more from guided scaffolding that fosters trust and interpersonal connections, while older students might excel with activities that leverage their advanced cooperative and problem-solving skills. By aligning group configurations with these developmental considerations, educators can better support both relational and task-oriented dynamics in classrooms.

Additionally, this study was conducted in a hybrid learning environment, where some students participated in person while others joined online sessions. This unique context may have influenced group dynamics in ways not explicitly measured. While the teacher observed consistent peer interactions across formats, physical presence might facilitate stronger relational bonds, whereas online participation could pose challenges to communication and coordination [48]. Interestingly, self-select groups demonstrated higher group stability despite the hybrid format, suggesting that strong initial social ties may provide resilience in mixed learning environments. However, the higher levels of cooperation observed in randomly assigned groups raise questions about how hybrid contexts might influence task-oriented collaboration. Future research should investigate how differences between online and in-person participation impact group dynamics, particularly in hybrid or transitional educational settings.

Despite these promising results, several limitations warrant consideration. First, the relatively small sample size and single-institution setting may constrain the generalizability of our conclusions. Second, self-report measures could introduce bias, underscoring the importance of triangulating with observational data or longitudinal interviews. Third, the unique sociocultural and curricular context of the school may not reflect conditions in other educational systems. Future work should replicate this study across diverse school types and subject areas to capture broader nuances in group dynamics. In addition, incorporating hybrid grouping techniques (e.g., partial self-selection coupled with teacher guidance) could yield new insights into optimizing both social cohesion and task-oriented collaboration. Understanding the nuanced interplay between friendship, cooperation, and group stability across different educational contexts will enable the development of more effective teaching strategies. Such strategies can foster meaningful collaboration, enhance learning outcomes, and prepare students with the essential 21st-century skills of teamwork and social-emotional competence.

VI. CONCLUSIONS

This study advances our understanding of how different group-formation strategies influence collaboration and social dynamics in physics classrooms—a critical context for cultivating 21st-century skills. Overall, our findings showed that self-select groups were likelier to maintain their original membership over time, reflecting stronger relational bonds, whereas randomly assigned groups exhibited higher levels of task-focused cooperation. Additionally, tenth-grade students appeared more stable in their grouping choices compared to ninth graders, suggesting that maturity and social experience can shape students' collaborative behaviors.

These insights underscore the importance of designing learning environments that strike a balance between leveraging existing friendships and exposing students to diverse perspectives. Educators could, for instance, intersperse random and self-chosen teams or integrate structured activities that promote both cooperative skills and social cohesion. By aligning group-formation strategies with students' developmental needs, teachers can support not only academic success but also key interpersonal competencies such as communication and conflict resolution.

Even so, this study has limitations. It was confined to a single institution and relied on student self-reports, potentially narrowing the generalizability of the results. Moreover, we did not account for motivational factors or individual preferences that may also influence group work dynamics. Future research should thus explore how these additional variables, as well as different cultural or curricular settings, interact with group-formation strategies. Longitudinal work that tracks students across multiple semesters or subjects may further illuminate how social ties, collaborative behaviors, and academic achievement coevolve.

In addressing these gaps, future investigations can guide the development of more nuanced, evidence-based teaching approaches that foster meaningful collaboration and equip students with the social and cognitive capacities essential for success—both in STEM fields and in an increasingly interconnected global environment.

ACKNOWLEDGMENTS

This study was supported by project FONDECYT Iniciación 11220385, sponsored by Agencia Nacional de Investigación y Desarrollo (ANID), Chile. C. C. was supported by FONDECYT Iniciación 11200986, sponsored by Agencia Nacional de Investigación y Desarrollo (ANID), Chile.

APPENDIX: SUPPLEMENTARY MATERIAL

This section presents descriptive statistics and additional analyses that complement the findings reported in the main

TABLE VI. Descriptive statistics of directed networks at the end of semester 1: friendship ties, physics prestige nomination, and physics collaboration. Note that * represents average number of incoming and outgoing links per students in the class.

		Friendship (t_0)			Friendship (t_1)		
Course	Nodes	Links	Degree*	Density	Links	Degree*	Density
E1	24	175	7.29	0.32	151	6.29	0.27
E2	22	125	5.68	0.27	118	5.36	0.26
C1	24	117	4.88	0.21	112	4.67	0.20
C2	20	117	5.85	0.31	95	4.75	0.25

		Prestige in physics			Collaboration in physics		
Course	Nodes	Links	Degree*	Density	Links	Degree*	Density
E1	24	127	5.29	0.23	113	4.71	0.21
E2	22	108	4.91	0.23	71	3.22	0.15
C1	24	147	6.13	0.27	134	5.58	0.24
C2	20	101	5.05	0.27	57	2.85	0.15

article. These supplementary materials provide deeper insights into the characteristics of the social networks studied, as well as the validity and robustness of the statistical models used in the analysis.

1. Descriptive statistics of social networks

Detailed descriptive statistics of the social networks analyzed at the end of the first semester are included in Tables VI–VIII. These statistics highlight structural differences between experimental and control groups, offering additional evidence.

The data presented in Table VI show that the experimental groups (E1 and E2) maintained levels of friendship tie density that are comparable to or slightly higher than those of the control groups (C1 and C2). This trend may suggest that self-assigned groups help sustain preexisting social ties. Regarding academic collaboration, the densities vary across groups: one control group (C1) exhibits the highest collaboration density, while the others show similar values. These descriptive patterns offer additional context for interpreting the main findings and illustrate the nuanced nature of group interactions in classroom settings.

Table VII presents group-level averages for key variables, including friendship, cooperation, and prestige, at different time points. The results indicate that self-assigned groups report stronger internal cohesion (e.g., *higher friendship levels*) and greater stability in group composition (e.g., *reconfiguration metric*). Randomly assigned groups show a moderate level of cooperation, which aligns with their less cohesive structure.

TABLE VII. Group-level averages for key variables.

Group	Treatment	Team friendship (t_1)	Team cooperation (t_1)	Team prestige (t_1)	Team friendship (t_0)	Friendship (t_0)	Dependent variable in Table V (t_1)
1	Random	0.25	1.75	0	0.5	6	0
2	Random	1.25	3	1.75	0.75	6	0.5
3	Random	0	2.25	0.75	0	3.75	0
4	Random	0.75	3	1.25	0.75	4.5	0.5
5	Random	0.25	3	0.75	0	4.75	0
6	Random	0.75	2.25	0	1	3	0.5
7	Random	0.333	2	0.333	0.333	3	0.667
8	Random	0	3	0.5	0	4.5	0
9	Random	0	1.33	0.333	0.333	3.67	0
10	Random	1.67	2	1	1.33	6.33	0
11	Random	0.667	1.33	1	0.667	4.67	0.667
12	Random	2.5	3	2	1.5	6	0.5
13	Self-selected	2	2.5	0.75	2.25	9.75	1
14	Self-selected	1.5	1	0.25	1.5	7.25	0.5
15	Self-selected	0.5	2.25	0.75	0.5	1.25	1.5
16	Self-selected	3	1	0.75	2.25	5.5	1.5
17	Self-selected	2.5	1.75	1.25	2.5	7.25	1.5
18	Self-selected	2.5	2.25	2.25	2	6.75	3
19	Self-selected	1.67	1.67	0.333	1.67	5.67	2
20	Self-selected	1.5	2.75	0.5	2	2.75	1.5
21	Self-selected	3	3	2	2.75	4.75	3
22	Self-selected	1.33	2	0.333	2	2.67	2
23	Self-selected	2.5	3	1.75	3	8.75	3
24	Self-selected	2.5	2.5	0.25	2	7	3

TABLE VIII. Treatment-level averages for key variables.

Treatment	Team friendship (t_1)	Team cooperation (t_1)	Team prestige (t_1)	Team friendship (t_0)	Friendship (t_0)	Dependent variable in Table V (t_1)
Random	0.705	2.39	0.818	0.591	4.70	0.273
Self-selected	2.07	2.15	0.957	2.04	5.85	1.96

Finally, Table VIII summarizes the treatment-level averages for the same variables. This aggregation highlights significant differences between treatments: self-assigned groups consistently show higher friendship and stability, while randomly assigned groups tend to exhibit slightly higher levels of cooperation. These aggregated results reinforce the trends observed at the group level and provide a broader perspective on the dynamics fostered by different grouping strategies.

2. Relationship between gender, grades, and prestige nominations

Figure 4 provides additional details complementing the findings reported in Table III (models 7 and 8) and discussed in the Discussion section. It is observed that there is no significant difference between men and women regarding the relationship between gender and grades [see Fig. 4(a)]. However, when characterizing the differences associated with physics prestige nominations by gender, significant differences are evident, with women receiving more nominations both at the general level [see Fig. 4(b)] and the group level [see Fig. 4(c)].

3. Evaluation of the robustness of mixed models

a. Random effects metrics

Table IX presents descriptive metrics of the random effects used in the mixed models. These metrics are essential to ensure that the models meet the necessary

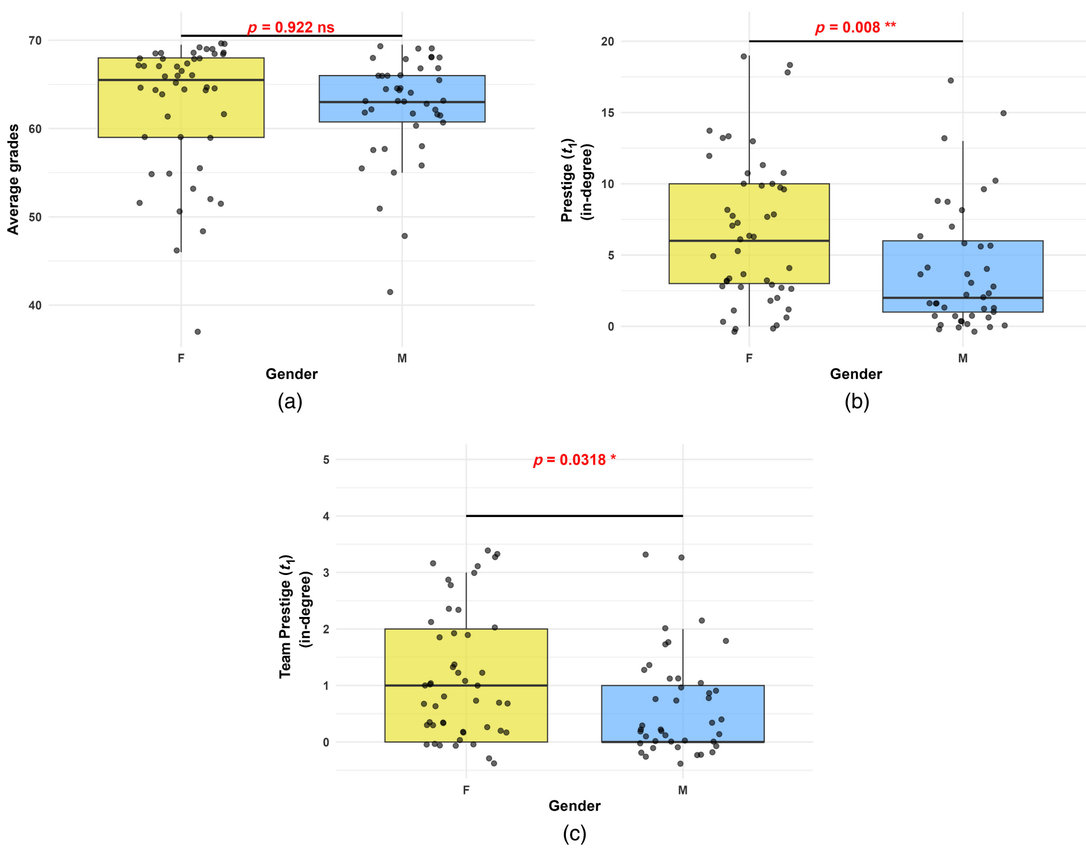

FIG. 4. Comparison of average grades and in-degree for general and group-level prestige by gender. In (a), the distribution of average grades by gender is shown, where no significant differences are observed ($p = 0.922$). In (b), the in-degree for general prestige is presented, showing significant differences between genders ($p = 0.008^{**}$). Finally, in (c), the in-degree for group-level prestige is shown, with significant differences favoring women ($p = 0.0318^{*}$).

TABLE IX. Results of mixed models: descriptive metrics of random effects.

Model	Shapiro p	Asimetra	Curtosis	Media	Std. dev.	Min	Max	Varianza min	Varianza max
1	0.4051	0.443	−0.537	0.029	1.373	−2.220	2.983	0.088	0.682
2	0.6115	0.164	−0.817	0.004	1.171	−2.096	2.520	0.049	0.314
3	0.8002	0.271	−0.598	0.033	1.859	−3.246	3.729	0.056	0.962
4	0.8353	0.169	−0.640	0.029	1.879	−3.257	3.707	0.104	0.907
5	0.0578	−0.087	−1.418	−0.069	1.127	−1.811	1.548	0.037	0.228
6	0.0590	−0.093	−1.409	−0.080	1.240	−2.034	1.726	0.053	0.349
7	0.8111	−0.423	−0.063	0.009	0.398	−0.907	0.743	0.002	0.069
8	0.8083	−0.424	−0.063	0.009	0.398	−0.907	0.742	0.002	0.069
9	0.2240	−0.029	−0.850	0.080	2.081	−3.660	3.916	0.006	1.327
10	0.1878	−0.043	−0.836	0.081	2.161	−3.549	3.989	0.008	1.428
11	0.5682	−0.157	−0.624	0.098	2.418	−4.554	4.418	0.044	2.204

assumptions and that coefficient estimates are reliable. Most models show Shapiro p values greater than 0.05, indicating that the normality hypothesis for the random effects cannot be rejected. Additionally, skewness and kurtosis parameters are mostly close to 0, suggesting a symmetric and normal distribution of the random effects. The mean and standard deviation of the random effects are also consistent with theoretical expectations, indicating no significant biases in the estimates.

b. Collinearity analysis

Table X evaluates the collinearity of the models presented in the main text among the predictor variables using the variance inflation factor (VIF) and tolerance. The results indicate an absence of severe collinearity, with all VIF values below 10 and tolerances above 0.1. This ensures that the coefficient estimates are independent and that the predictor variables are not highly correlated with each

TABLE X. Collinearity in the models. The VIF values indicate low, moderate, or high collinearity among the predictor variables.

Model	Term	VIF	VIF 95% CI	Increased SE	Tolerance	Tolerance 95% CI
1	Exp. (self-select groups)	1.95	[1.54, 2.67]	1.40	0.51	[0.37, 0.65]
	Group size (3)	1.53	[1.26, 2.09]	1.24	0.65	[0.48, 0.80]
	Gender (male)	3.38	[2.53, 4.68]	1.84	0.30	[0.21, 0.39]
	Grade level (tenth)	1.71	[1.38, 2.34]	1.31	0.58	[0.43, 0.72]
	Average grades	1.19	[1.05, 1.73]	1.09	0.84	[0.58, 0.95]
	Team friendship (t_0)	1.95	[1.54, 2.67]	1.40	0.51	[0.37, 0.65]
2	Exp. (self-select groups)	3.98	[2.94, 5.57]	1.99	0.25	[0.18, 0.34]
	Group size (3)	1.54	[1.27, 2.09]	1.24	0.65	[0.48, 0.79]
	Gender (male)	1.02	[1.01, 1.30]	1.01	0.98	[0.75, 0.99]
	Grade level (tenth)	1.72	[1.38, 2.33]	1.31	0.58	[0.43, 0.72]
	Average grades	1.19	[1.05, 1.72]	1.09	0.84	[0.58, 0.95]
	Team friendship (t_0)	5.76	[4.18, 8.13]	2.40	0.17	[0.12, 0.24]
	Exp.* team friendship (t_0)	10.69	[7.59, 15.23]	3.27	0.09	[0.07, 0.13]
3	Exp. (self-select groups)	1.16	[1.03, 1.74]	1.08	0.86	[0.58, 0.97]
	Group size (3)	1.51	[1.24, 2.06]	1.23	0.66	[0.48, 0.80]
	Gender (male)	1.04	[1.00, 5.38]	1.02	0.96	[0.19, 1.00]
	Grade level (tenth)	1.66	[1.34, 2.26]	1.29	0.60	[0.44, 0.75]
	Average grades	1.25	[1.08, 1.76]	1.12	0.80	[0.57, 0.93]
	Friendship (t_0)	1.15	[1.03, 1.75]	1.07	0.87	[0.57, 0.97]
4	Exp. (self-select groups)	5.05	[3.69, 7.11]	2.25	0.20	[0.14, 0.27]
	Exp.* friendship (t_0)	8.46	[6.05, 12.03]	2.91	0.12	[0.08, 0.17]
	Group size (3)	1.51	[1.25, 2.05]	1.23	0.66	[0.49, 0.80]
	Gender (male)	1.05	[1.00, 3.91]	1.02	0.95	[0.26, 1.00]
	Grade level (tenth)	1.67	[1.35, 2.26]	1.29	0.60	[0.44, 0.74]
	Average grades	1.25	[1.08, 1.75]	1.12	0.80	[0.57, 0.92]

(Table continued)

TABLE X. *(Continued)*

Model	Term	VIF	VIF 95% CI	Increased SE	Tolerance	Tolerance 95% CI
5	Exp. (self-select groups)	1.95	[1.54, 2.67]	1.40	0.51	[0.37, 0.65]
	Group size (3)	1.53	[1.26, 2.09]	1.24	0.65	[0.48, 0.80]
	Gender (male)	1.72	[1.38, 2.10]	1.31	0.58	[0.48, 0.72]
	Grade level (tenth)	1.71	[1.38, 2.34]	1.31	0.58	[0.43, 0.72]
	Average grades	1.19	[1.05, 1.73]	1.09	0.84	[0.58, 0.95]
	Team friendship (t_0)	1.95	[1.54, 2.67]	1.40	0.51	[0.37, 0.65]
6	Exp. (self-select groups)	3.98	[2.94, 5.57]	1.99	0.25	[0.18, 0.34]
	Group size (3)	1.54	[1.27, 2.09]	1.24	0.65	[0.48, 0.79]
	Gender (male)	1.43	[1.15, 1.80]	1.20	0.70	[0.56, 0.87]
	Grade level (tenth)	1.72	[1.38, 2.33]	1.31	0.58	[0.43, 0.72]
	Average grades	1.19	[1.05, 1.72]	1.09	0.84	[0.58, 0.95]
	Team friendship (t_0)	5.76	[4.18, 8.13]	2.40	0.17	[0.12, 0.24]
	Exp.* team friendship (t_0)	10.69	[7.59, 15.23]	3.27	0.09	[0.07, 0.13]
7	Exp. (self-select groups)	1.95	[1.54, 2.67]	1.40	0.51	[0.37, 0.65]
	Group size (3)	1.53	[1.26, 2.09]	1.24	0.65	[0.48, 0.80]
	Gender (male)	1.52	[1.20, 2.10]	1.23	0.66	[0.48, 0.83]
	Grade level (tenth)	1.71	[1.38, 2.34]	1.31	0.58	[0.43, 0.72]
	Average grades	1.19	[1.05, 1.73]	1.09	0.84	[0.58, 0.95]
	Team friendship (t_0)	1.95	[1.54, 2.67]	1.40	0.51	[0.37, 0.65]
8	Exp. (self-select groups)	3.98	[2.94, 5.57]	1.99	0.25	[0.18, 0.34]
	Group size (3)	1.54	[1.27, 2.09]	1.24	0.65	[0.48, 0.79]
	Gender (male)	2.10	[1.50, 3.00]	1.45	0.48	[0.33, 0.67]
	Grade level (tenth)	1.72	[1.38, 2.33]	1.31	0.58	[0.43, 0.72]
	Average grades	1.19	[1.05, 1.72]	1.09	0.84	[0.58, 0.95]
	Team friendship (t_0)	5.76	[4.18, 8.13]	2.40	0.17	[0.12, 0.24]
	Exp.* team friendship (t_0)	10.69	[7.59, 15.23]	3.27	0.09	[0.07, 0.13]
9	Exp. (self-select groups)	1.10	[1.01, 1.91]	1.05	0.90	[0.52, 0.99]
	Group size (3)	1.50	[1.24, 2.06]	1.22	0.67	[0.49, 0.81]
	Gender (male)	1.02	[1.01, 1.50]	1.01	0.98	[0.67, 0.99]
	Grade level (tenth)	1.64	[1.33, 2.25]	1.28	0.61	[0.44, 0.75]
	Average grades	1.18	[1.04, 1.74]	1.08	0.85	[0.57, 0.96]
10	Exp. (self-select groups)	1.16	[1.03, 1.74]	1.08	0.86	[0.58, 0.97]
	Group size (3)	1.51	[1.24, 2.06]	1.23	0.66	[0.48, 0.80]
	Gender (male)	1.04	[1.00, 5.38]	1.02	0.96	[0.19, 1.00]
	Grade level (tenth)	1.66	[1.34, 2.26]	1.29	0.60	[0.44, 0.75]
	Average grades	1.25	[1.08, 1.76]	1.12	0.80	[0.57, 0.93]
	Friendship (t_0)	1.15	[1.03, 1.75]	1.07	0.87	[0.57, 0.97]
11	Group size (3)	1.51	[1.25, 2.05]	1.23	0.66	[0.49, 0.80]
	Gender (male)	1.05	[1.00, 3.91]	1.02	0.95	[0.26, 1.00]
	Grade level (tenth)	1.67	[1.35, 2.26]	1.29	0.60	[0.44, 0.74]
	Average grades	1.25	[1.08, 1.75]	1.12	0.80	[0.57, 0.92]
	Friendship (t_0)	3.48	[2.60, 4.85]	1.87	0.29	[0.21, 0.38]
	Exp. (self-select groups)	5.05	[3.69, 7.11]	2.25	0.20	[0.14, 0.27]
	Exp.* friendship (t_0)	8.46	[6.05, 12.03]	2.91	0.12	[0.08, 0.17]

other. The low collinearity observed strengthens the interpretation of the individual effects of each predictor variable in the models. This means that the estimated coefficients reflect the influence of each predictor on the dependent variable without significant interference from other correlated variables.

c. Comparison of AIC between models with and without random effects

Tables XI and XII present a comparison of Akaike information criterion (AIC) values between the random-effects models (clmm) presented in the main text (see Tables III and V) and basic models without random

TABLE XI. Comparison of AIC between models with random effects (clmm) and models without random effects (clm). Models with Nominal: "Variable" indicate that a nominal variable is being used in the model, meaning the categorical variable is treated as a nominal predictor instead of an ordinal or continuous predictor. This allows the modeling of relationships between the categories of the nominal variable without assuming a specific order among them.

Tipo modelo	Modelo	AIC	Modelo	AIC	Modelo	AIC	Modelo	AIC
clmm	1	155.06	3	186.39	5	161.26	7	209.64
clmm	2	149.32	4	187.80	6	161.26	8	211.64
clm	Nominal: Exp.	162.54	Nominal: Exp.	208.54	Nominal: Exp.	168.42	Nominal: Exp.	211.75
clm	Nominal: Group size	162.25	Nominal: Group size	203.84	Nominal: Group size	165.34	Nominal: Group size	212.42
clm	Nominal: Gender	161.99	Nominal: Gender	207.97	Nominal: Gender	167.97	Nominal: Gender	212.54
clm	Nominal: Grade level	164.49	Nominal: Grade level	208.94	Nominal: Grade level	168.75	Nominal: Grade level	213.36
clm	Nominal: Average grades	163.15	Nominal: Average grades	206.95	Nominal: Average grades	170.90	Nominal: Average grades	213.53
clm	Nominal: Team friendship	156.49	Nominal: friendship	199.55	Nominal: Team friendship	169.65	Nominal: Team friendship	213.22

effects (clm), which differ in how each variable is treated. For example, in Table XI, the first column "Model" shows the results for models 1 and 2 from Table XII in the main text. Subsequently, a series of clm models with the same structure as model 1 are presented, but they vary in how the variable specified as "Nominal: Variable" is defined. Another example is column 6, which shows models 5 and 6. In the rows below, the models correspond to model 5 in clm format, varying the specification of the indicated variable.

We note that the label "Nominal: Variable" does not imply that the variable is not inherently nominal (e.g., gender or experimental condition), but rather that it was explicitly treated as a nominal factor in that model specification. This labeling was applied systematically to indicate the modeling approach used in each variant, regardless of whether the original variable was ordinal or nominal. These clm variants were included as part of a robustness check to examine whether model fit was affected by the treatment of categorical predictors.

In most cases, clmm models show significantly lower AIC values compared to clm models, indicating a better fit to the data. This superiority arises because clmm models incorporate random effects, allowing them to capture group-specific variability, which is not possible with clm models. The inclusion of random effects in clmm models enables the modeling of intergroup heterogeneity, resulting in a more accurate representation of the dynamics of friendship, prestige, and academic collaboration. Therefore, the improved fit of the clmm models supports the methodological choice of using mixed models.

This approach is crucial for the robustness of the analysis for several reasons. First, treating a variable as nominal avoids assuming an order that may not exist or be relevant in the data, ensuring that the estimated effects for each category are independent and not influenced by an assumed hierarchy. Second, by fitting multiple clm models where different variables are treated as nominal, it is verified whether the results observed in clmm models remain consistent even when certain specifications of categorical variables are relaxed. If clmm models continue to show better fit (lower AIC values) compared to nominal clm models, this reinforces the choice of using mixed models, demonstrating that their superiority does not depend on the specific coding of categorical variables.

d. Consistency of results across different links

Tables XIII–XV present the main models from the manuscript along with variations associated with the

TABLE XII. Comparison of AIC between models with random effects (clmm) and models without random effects (clm). Models with Nominal: "Variable" indicate that a nominal variable is being used in the model, meaning the categorical variable is treated as a nominal predictor instead of an ordinal or continuous predictor. This allows the modeling of relationships between the categories of the nominal variable without assuming a specific order among them.

Model type	Model	AIC	Model	AIC
clmm	9	156.98	10	158.14
clmm			11	155.76
clm	Nominal: Average grades	161.09	Nominal: Average grades	163.09
clm	Nominal: Group size	172.99	Nominal: Group size	174.90
clm	Nominal: Grade level	179.08	Nominal: Grade level	180.85
clm	Nominal: Gender	179.77	Nominal: Gender	181.68
clm	Nominal: Exp.	180.06	Nominal: Exp.	181.94

TABLE XIII. Coefficients of the models with logit, probit, and cloglog link functions, including threshold coefficients, number of observations (nobs), logLik, AIC, and BIC. Standard errors are shown in parentheses below the coefficients, and significance is indicated to the right of each coefficient: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

	Model 1			Model 2			Model 3			Model 4		
Variable	Logit	Probit	Cloglog	Logit	Probit	Cloglog	Logit	Probit	Cloglog	Logit	Probit	Cloglog
Exp. (self-select groups)	1.51	1.02	1.18	4.25**	2.48**	3.46**	4.16***	2.41***	2.78**	5.09***	2.88**	3.39**
	(1.13)	(0.66)	(0.86)	(1.57)	(0.85)	(1.15)	(1.20)	(0.72)	(1.06)	(1.06)	(0.96)	(1.18)
Group size (3)	−2.02	−1.04	−1.27	−2.18†	−1.12	−1.44	−2.23	−1.29	−1.45	−2.26	−1.28	−1.41
	(1.39)	(0.80)	(1.04)	(1.32)	(0.75)	(1.04)	(1.53)	(0.79)	(1.07)	(1.54)	(0.90)	(1.02)
Gender (male)	0.49	0.31	0.51	0.52	0.30	0.57	−0.34	−0.19	−0.20	−0.34***	−0.19**	−0.18**
	(0.64)	(0.37)	(0.47)	(0.66)	(0.37)	(0.48)	(0.60)	(0.36)	(0.45)	(0.59)	(0.36)	(0.42)
Grade level (tenth)	0.68	0.31	0.31	0.67	0.33	0.36	1.96	1.15	1.40	1.97†	1.13	1.36
	(1.16)	(0.68)	(0.88)	(1.09)	(0.63)	(0.88)	(1.23)	(0.71)	(0.91)	(1.08)	(0.71)	(0.87)
Average grades	0.04	0.02	0.02	0.05	0.03	0.03	−0.02	−0.01	−0.02	−0.02***	−0.01**	−0.02*
	(0.05)	(0.03)	(0.03)	(0.05)	(0.03)	(0.03)	(0.05)	(0.03)	(0.04)	(0.05)	(0.03)	(0.03)
Team friendship (t_0)	2.87***	1.52***	1.86***	4.77***	2.61***	3.65***						
(in-degree)	(0.55)	(0.28)	(0.39)	(1.02)	(0.52)	(0.79)						
Exp.* team friendship (t_0)				−2.61**	−1.46**	−2.21**						
(in-degree)				(1.01)	(0.54)	(0.76)						
Friendship (t_0)							0.32**	0.18*	0.25**	0.43***	0.18**	0.33*
(in-degree)							(0.11)	(0.06)	(0.08)	(0.001)	(0.06)	(0.14)
Exp.* friendship (t_0)										−0.17**	−0.09*	−0.12*
(in-degree)										(0.00)	(0.13)	(0.17)
Threshold 0|1	4.19	2.38	3.08	5.95	3.33	4.70	0.64	0.51	0.85	0.96	0.74	1.16
Threshold 1|2	7.05	4.01	5.12	9.05	5.22	6.88	2.96	1.85	2.39	3.33	2.10	2.72
Threshold 2|3	11.95	6.70	8.43	13.65	7.76	10.08	6.88	4.12	5.12	7.19	4.12	5.37
nobs	90	90	90	90	90	90	90	90	90	90	90	90
logLik	−67.53	−67.91	−68.94	−63.66	−63.95	−63.80	−83.19	−82.38	−83.00	−82.90	−82.12	−82.73
AIC	155.06	155.82	157.88	149.32	149.90	149.60	186.39	184.75	185.99	187.80	186.24	187.46
BIC	180.10	180.82	182.87	176.81	177.39	177.11	211.39	209.75	210.99	215.29	213.74	214.95

TABLE XIV. Coefficients of the models with logit, probit, and cloglog link functions, including threshold coefficients, number of observations (nobs), logLik, AIC, and BIC. Standard errors are shown in parentheses below the coefficients, and significance is indicated to the right of each coefficient: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

	Model 5			Model 6			Model 7			Model 8		
Variable	Logit	Probit	Cloglog	Logit	Probit	Cloglog	Logit	Probit	Cloglog	Logit	Probit	Cloglog
Exp. (self-select groups)	−3.76**	−1.91**	−1.75**	−4.96**	−2.83**	−3.03**	−0.69	−0.37	−0.70	−0.68	−0.30	−0.59
	(1.19)	(0.62)	(0.64)	(1.56)	(0.86)	(1.01)	(0.75)	(0.42)	(0.55)	(1.10)	(0.59)	(0.80)
Group size (3)	−6.46***	−3.62***	−3.66***	−6.49***	−3.75***	−4.05***	−0.56	−0.35	−0.35	−0.57	−0.45	−0.36
	(1.64)	(0.88)	(0.95)	(1.71)	(0.96)	(1.20)	(0.80)	(0.45)	(0.60)	(0.81)	(0.45)	(0.60)
Gender (female)	−0.61	−0.36	−0.38	−0.50	−0.30	−0.31	−1.03*	−0.62*	−0.77*	−1.03*	−0.62*	−0.77*
	(0.66)	(0.36)	(0.38)	(0.69)	(0.37)	(0.41)	(0.49)	(0.28)	(0.37)	(0.49)	(0.28)	(0.37)
Grade level (tenth)	4.64***	2.65***	2.83***	4.77**	2.78**	3.10**	−0.44	−0.27	−0.46	−0.44	−0.26	−0.45
	(1.38)	(0.76)	(0.81)	(1.45)	(0.76)	(1.01)	(0.68)	(0.38)	(0.48)	(0.68)	(0.38)	(0.48)
Average grades	−0.16**	−0.08*	−0.08*	−0.17*	−0.09**	−0.10*	0.14**	0.09***	0.10**	0.14**	0.09***	0.10**
	(0.06)	(0.03)	(0.03)	(0.07)	(0.03)	(0.03)	(0.05)	(0.03)	(0.04)	(0.05)	(0.03)	(0.04)
Team friendship (t_0)	0.72†	0.35	0.34	−0.26	−0.27	−0.37	0.86**	0.48**	0.67**	0.86†	0.52**	0.72**
(in-degree)	(0.38)	(0.20)	(0.21)	(0.80)	(0.40)	(0.41)	(0.33)	(0.18)	(0.23)	(0.50)	(0.27)	(0.35)
Exp.* team friendship (t_0)				1.29	0.87†	1.07*				−0.01	−0.06	−0.19
(in-degree)				(0.94)	(0.49)	(0.52)				(0.64)	(0.37)	(0.45)
Threshold 0|1	−16.46	−8.61	−7.76	−17.41	−8.61	−7.76	8.73	5.42	6.83	8.74	5.42	6.83
Threshold 1|2	−13.97	−7.23	−6.38	−14.87	−7.23	−6.38	10.50	6.45	8.12	10.50	6.45	8.12

(Table continued)

TABLE XIV. *(Continued)*

	Model 5			Model 6			Model 7			Model 8		
Variable	Logit	Probit	Cloglog	Logit	Probit	Cloglog	Logit	Probit	Cloglog	Logit	Probit	Cloglog
Threshold 2|3	−10.75	−5.46	−4.52	−11.54	−5.46	−4.52	11.73	7.16	9.02	11.73	7.16	9.02
nobs	90	90	90	90	90	90	90	90	90	90	90	90
logLik	−70.63	−71.23	−74.41	−69.63	−69.54	−71.91	−94.82	−94.13	−94.96	−94.82	−94.12	−94.94
AIC	161.26	162.46	168.82	161.26	161.08	165.83	209.64	208.27	209.91	211.64	210.24	211.87
BIC	186.26	187.46	193.82	188.76	188.58	193.33	234.64	233.27	234.91	239.14	237.74	239.37

TABLE XV. Coefficients of the models with logit, probit, and cloglog link functions, including threshold coefficients, number of observations (nobs), logLik, AIC, and BIC. Standard errors are shown in parentheses below the coefficients, and significance is indicated to the right of each coefficient: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

	Model 9			Model 10			Model 11		
Variable	Logit	Probit	Cloglog	Logit	Probit	Cloglog	Logit	Probit	Cloglog
Exp. (self-select groups) 1	7.35***	4.05***	4.49***	7.42***	4.06***	4.54***	5.19***	2.75*	2.84*
	(1.51)	(0.95)	(1.04)	(0.00)	(0.98)	(1.09)	(0.00)	(1.22)	(1.34)
Group size (3) 1	−2.41	−1.27	−1.20	−2.36	−1.25	−1.17	−2.71	−1.44	−1.44
	(1.92)	(1.07)	(1.17)	(1.89)	(1.10)	(1.22)	(2.09)	(1.23)	(1.36)
Gender (male) M	−0.51	−0.41	−0.76	−0.53***	−0.37**	−0.70*	−0.74***	−0.49*	−0.73*
	(0.71)	(0.40)	(0.48)	(0.00)	(0.41)	(0.48)	(0.00)	(0.42)	(0.49)
Grade level (tenth) 1	3.33*	1.82*	2.02*	3.44*	1.89*	2.13*	3.99*	2.22*	2.56*
	(1.55)	(0.96)	(1.04)	(1.44)	(1.00)	(1.12)	(1.58)	(1.14)	(1.23)
Average grades	0.08	0.05	0.06	0.08***	0.05**	0.07	0.09***	0.05	0.06**
	(0.06)	(0.03)	(0.03)	(0.00)	(0.03)	(0.07)	(0.00)	(0.03)	(0.04)
Friendship (t_0)				0.11***	0.07*	0.12*	−0.24***	−0.13*	−0.17
				(0.00)	(0.07)	(0.08)	(0.00)	(0.12)	(0.15)
Exp. (self-select groups) 1: friendship (t_0)							0.57***	0.33*	0.42*
							(0.00)	(0.16)	(0.19)
Threshold 0|1	8.19	4.64	6.07	8.71	4.88	6.39	8.14	4.37	5.55
Threshold 1|2	10.55	5.92	7.38	11.08	6.18	7.01	10.66	6.18	7.01
Threshold 2|3	15.10	8.38	10.07	15.70	8.38	10.07	15.62	8.70	10.11
nobs	90	90	90	90	90	90	90	90	90
logLik	−69.49	−69.84	−74.79	−69.07	−69.32	−73.74	−66.88	−67.00	−71.02
AIC	156.98	157.67	167.57	158.14	158.64	167.49	155.76	156.00	164.04
BIC	179.47	180.17	190.07	183.13	183.63	192.49	183.25	183.50	191.54

type of link function specified in the model (logit, probit, and cloglog). It is observed that, regardless of the link function used, the main predictor variables exhibit similar patterns in terms of direction and significance. For instance, membership in self-assigned groups [Exp. (self-select groups)] consistently shows a positive and significant effect across all link functions, reinforcing the robustness of this finding. Similarly, group size [group size (3)] displays negative and significant effects across all links, indicating that larger groups tend to have lower cohesion.

This consistency suggests that the results are robust to different model specifications, increasing confidence in the conclusions drawn. Furthermore, the stability of effects across different link functions confirms that the observed relationships between predictor variables and group dynamics are not artifacts of a particular choice of link function but rather represent true underlying associations in the data.

4. Models with thresholds

The Tables XVI and XVII report the threshold parameters estimated in each of the cumulative link mixed models (CLMMs) presented in the main manuscript. While these threshold values are not interpreted directly in the results or discussion sections, they are essential components of ordinal models and are included here for completeness and transparency.

TABLE XVI. Regression models: predicting within-group friendship, prestige, and cooperation links at time t_1. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

	Dependent variable							
	Team friendship (t_1) (in-degree)				Team cooperation (t_1) (in-degree)		Team prestige (t_1) (in-degree)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exp. (self-select groups)	1.51	4.25^{**}	4.16^{***}	5.09^{***}	-3.76^{**}	-4.96^{**}	−0.69	−0.68
	(1.13)	(1.57)	(1.20)	(1.06)	(1.19)	(1.56)	(0.75)	(1.10)
Group size (3)	−2.02	$-2.18^{\dagger}$	−2.23	−2.26	-6.46^{***}	-6.49^{***}	−0.56	−0.57
	(1.39)	(1.32)	(1.54)	(1.55)	(1.64)	(1.71)	(0.80)	(0.81)
Gender (male)	0.49	0.52	−0.34	-0.34^{***}	−0.61	−0.50	-1.03^{*}	-1.03^{*}
	(0.64)	(0.66)	(0.60)	(0.59)	(0.66)	(0.69)	(0.49)	(0.49)
Grade level (tenth)	0.68	0.67	1.96	$1.97^{\dagger}$	4.64^{***}	4.77^{**}	−0.44	−0.44
	(1.16)	(1.09)	(1.23)	(1.08)	(1.38)	(1.45)	(0.68)	(0.68)
Average grades	0.04	0.05	−0.02	-0.02^{***}	-0.16^{**}	-0.17^{*}	0.14^{**}	0.14^{**}
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.07)	(0.05)	(0.05)
Team friendship (t_0)	2.87^{***}	4.77^{***}			$0.72^{\dagger}$	−0.26	0.86^{**}	$0.86^{\dagger}$
(in-degree)	(0.55)	(1.02)			(0.38)	(0.80)	(0.33)	(0.50)
Exp.* team friendship (t_0)		-2.61^{**}				1.29		−0.01
(in-degree)		(1.01)				(0.94)		(0.64)
Friendship (t_0)			0.32^{**}	0.43^{***}				
(in-degree)			(0.11)	(0.00)				
Exp.* friendship (t_0)				-0.17^{**}				
(in-degree)				(0.00)				
Threshold 0|1	4.19	$5.95^{\dagger}$	0.64	0.96	-16.46^{***}	-17.41^{***}	8.73^{**}	8.74^{**}
	(3.02)	(3.14)	(2.96)	(0.66)	(4.60)	(4.89)	(3.00)	(3.02)
Threshold 1|2	7.05^{*}	9.05^{**}	2.96	3.33^{***}	-13.97^{**}	-14.86^{**}	10.50^{***}	10.50^{***}
	(3.10)	(3.24)	(2.96)	(0.59)	(4.45)	(4.72)	(3.09)	(3.11)
Threshold 2|3	11.95^{***}	13.65^{***}	6.88^{*}	7.19^{***}	-10.75^{*}	-11.54^{**}	11.73^{***}	11.73^{***}
	(3.53)	(3.59)	(3.06)	(0.00)	(4.22)	(4.47)	(3.15)	(3.17)
Group effect	Random	Random	Random	Random	Random	Random	Random	Random
Number of observations	90	90	90	90	90	90	90	90
AIC	155.1	149.3	186.4	187.8	161.3	161.3	209.6	211.6
BIC	180.1	176.8	211.4	215.3	186.3	188.8	234.6	239.1
Log likelihood	−67.529	−63.658	−83.194	−82.900	−70.632	−69.632	−94.822	−94.822
edf	10.00	11.00	10.00	11.00	10.00	11.00	10.00	11.00

TABLE XVII. Regression models: predicting the number of students with whom an individual continues to work when given the option to change groups. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

	Dependent variable		
	n students with whom a person continues to work		
	Model (9)	Model (10)	Model (11)
Exp. (self-select groups)	7.35^{***}	7.42^{***}	5.19^{***}
	(1.51)	(0.00)	(0.00)
Group size	−2.41	−2.36	−2.71
	(1.92)	(1.89)	(2.09)
Gender (male)	−0.51	-0.53^{***}	-0.74^{***}
	(0.71)	(0.00)	(0.00)
Grade level (tenth)	3.33^{*}	3.44^{*}	3.99^{*}
	(1.55)	(1.44)	(1.58)
Average grades	0.08	0.08^{***}	0.09^{***}
	(0.06)	(0.00)	(0.00)

(Table continued)

TABLE XVII. *(Continued)*

	Dependent variable		
	n students with whom a person continues to work		
	Model (9)	Model (10)	Model (11)
Friendship (t_0) (in-degree)		0.11^{***} (0.00)	-0.24^{***} (0.00)
Exp.* friendship (t_0) (in-degree)			0.57^{***} (0.00)
Threshold 0|1	8.19^{*} (4.05)	8.71^{***} (0.00)	8.14^{***} (0.00)
Threshold 1|2	10.55^{*} (4.13)	11.08^{***} (0.00)	10.66^{***} (0.00)
Threshold 2|3	15.10^{***} (4.49)	15.70^{***} (0.95)	15.62^{***} (1.04)
Group effect	Random	Random	Random
Number of observations	90	90	90
AIC	157.0	158.1	155.8
BIC	179.5	183.1	183.3
Log likelihood	−69.49	−69.07	−66.88
edf	9.00	10.00	11.00

[1] Esther Care, Helyn Kim, Alvin Vista, and Kate Anderson, Education system alignment for 21st century skills: Focus on assessment (2018), https://eric.ed.gov/?id=ED592779.

[2] Chris Joynes, Serena Rossignoli, and Esi Fenyiwa Amonoo-Kuofi, 21st century skills: Evidence of issues in definition, demand and delivery for development contexts (2019), https://opendocs.ids.ac.uk/articles/report/21st_Century_Skills_Evidence_of_Issues_in_Definition_Demand_and_Delivery_for_Development_Contexts/26427490.

[3] Stephen P. Borgatti, Ajay Mehra, Daniel J. Brass, and Giuseppe Labianca, Network analysis in the social sciences, Science **323**, 892 (2009).

[4] J. Ashwin Rambaran, Andrea Hopmeyer, David Schwartz, Christian Steglich, Daryaneh Badaly, and René Veenstra, Academic functioning and peer influences: A short-term longitudinal study of network-behavior dynamics in middle adolescence, Child Dev. **88**, 523 (2017).

[5] Lev Semenovich Vygotsky and Michael Cole, *Mind in Society: Development of Higher Psychological Processes* (Harvard University Press, Cambridge, MA, 1978).

[6] Olli-Pekka Heinimäki, Simone Volet, Cheryl Jones, Eero Laakkonen, and Marja Vauras, Student participatory role profiles in collaborative science learning: Relation of within-group configurations of role profiles and achievement, Learn. Cult. Soc. Inter. **30**, 100539 (2021).

[7] Kenneth J. Chapman, Matthew Meuter, Dan Toy, and Lauren Wright, Can't we pick our own groups? The influence of group selection method on group dynamics and outcomes, J. Manag. Educ. **30**, 557 (2006).

[8] Mary L. Connerley and Fred A. Mael, The importance and invasiveness of student team selection criteria, J. Manag. Educ. **25**, 471 (2001).

[9] Paul Blowers, Using student skill self-assessments to get balanced groups for group projects, Coll. Teach. **51**, 106 (2003).

[10] J. R. Hackman and N. Katz, Group behavior and performance, in *Handbook of Social Psychology*, edited by S. T. Fiske, D. T. Gilbert, and G. Lindzey (John Wiley and Sons, Inc., New York, 2010), pp. 1208–1251.

[11] Javier Pulgar, Diego Ramírez, Abigail Umanzor, Cristian Candia, and Iván Sánchez, Long-term collaboration with strong friendship ties improves academic performance in remote and hybrid teaching modalities in high school physics, Phys. Rev. Phys. Educ. Res. **18**, 010146 (2022).

[12] David W. Johnson and Roger T. Johnson, Cooperative learning in 21st century [Aprendizaje cooperativo en el siglo XXI], Ann. Phys. (Berlin) **30**, 841 (2014).

[13] Terrell L. Strayhorn, *College Students' Sense of Belonging*, 2nd ed. (Routledge, New York, NY, 2018).

[14] Brigid Barron, When smart groups fail, J. Learn. Sci. **12**, 307 (2003).

[15] Robyn M. Gillies, The effects of cooperative learning on junior high school students during small group learning, Learn. Instr. **14**, 197 (2004).

[16] Rupert Wegerif, *Dialogic Education and Technology: Expanding the Space of Learning* (Springer Science & Business Media, Heidelberg, 2007), Vol. 7.
[17] Neil Mercer, Talk and the development of reasoning and understanding, Hum. Dev. **51**, 90 (2008).
[18] Mohammed Abdullatif Almulla, The effectiveness of the project-based learning (PBL) approach as a way to engage students in learning, SAGE Open **10** (2020).
[19] Dimitra Kokotsaki, Victoria Menzies, and Andy Wiggins, Project-based learning: A review of the literature, in *Improving Schools* (SAGE Publications Sage UK, London, England, 2016), Vol. 19, pp. 267–277.
[20] Javier Pulgar, Cristian Candia, and Paul M. Leonardi, Social networks and academic performance in physics: Undergraduate cooperation enhances ill-structured problem elaboration and inhibits well-structured problem solving, Phys. Rev. Phys. Educ. Res. **16**, 010137 (2020).
[21] Javier Pulgar, Diego Ramírez, and Cristian Candia, Enhanced social connectivity in hybrid classrooms versus academic centrality in online settings, Phys. Rev. Phys. Educ. Res. **19**, 020155 (2023).
[22] Abdelkrim Hasni, Fatima Bousadra, Vincent Belletête, Ahmed Benabdallah, Marie-Claude Nicole, and Nancy Dumais, Trends in research on project-based science and technology teaching and learning at K–12 levels: A systematic review, Stud. Sci. Educ. **52**, 199 (2016).
[23] I. Wayan Santyasa, Ni Ketut Rapi, and I. Sara, Project based learning and academic procrastination of students in learning physics, Int. J. Instr. **13**, 489 (2020).
[24] Jon-Chao Hong, Mei-Yung Chen, Ashley Wong, Tsui-Fang Hsu, and Chih-Chi Peng, Developing physics concepts through hands-on problem solving: A perspective on a technological project design, Int. J. Technol. Des. Educ. **22**, 473 (2012).
[25] Sofi Hanif, Agus Fany Chandra Wijaya, and Nanang Winarno, Enhancing students' creativity through STEM project-based learning, J. Sci. Learn. **2**, 50 (2019).
[26] Joseph S. Krajcik and Charlene M. Czerniak, *Teaching Science in Elementary and Middle School: A Project-Based Learning Approach*, 5th ed. (Routledge, New York, 2018); 4th ed. (Routledge, 2013); 1st ed. (McGraw-Hill, 1999); 1st ed. (2018).
[27] Elif Ince, An overview of problem solving studies in physics education, J. Educ. Learn. **7**, 191 (2018).
[28] Norazlilah Md Nordin and Kamisah Osman, Students' generated animation: An innovative approach to inculcate collaborative problem solving (CPS) skills in learning physics, J. Educ. Sci. Environ. Health **4**, 206 (2018).
[29] Arthur C. Graesser, Stephen M. Fiore, Samuel Greiff, Jessica Andrews-Todd, Peter W. Foltz, and Friedrich W. Hesse, Advancing the science of collaborative problem solving, Psychol. Sci. Public Interest **19**, 59 (2018).
[30] Robyn M. Gillies, The behaviors, interactions, and perceptions of junior high school students during small-group learning, J. Educ. Psychol. **95**, 137 (2003).
[31] Cristian Candia, Javier Pulgar, and Flavio Pinheiro, Interconnectedness in education systems, in *Handbook of Computational Social Science*, edited by Taha Yasseri (Edward-Elgar Publishing, Cheltenham, UK).
[32] Cristian Candia, Sara Encarnação, and Flávio L. Pinheiro, The higher education space: Connecting degree programs from individuals' choices, EPJ Data Sci. **8**, 39 (2019).
[33] Yiping Lou, Philip C. Abrami, and Sylvia d'Apollonia. Small group individual learning with technology: A meta-analysis, Rev. Educ. Res. **71**, 449 (2001).
[34] Leonard Springer, Mary Elizabeth Stanne, and Samuel S. Donovan, Effects of small-group learning on undergraduates in science mathematics, engineering, and technology: A meta-analysis, Rev. Educ. Res. **69**, 21 (1999).
[35] Eric Brewe, Laird Kramer, and George O'Brien, Modeling instruction: Positive attitudinal shifts in introductory physics measured with CLASS, Phys. Rev. ST Phys. Educ. Res. **5**, 013102 (2009).
[36] Louis Deslauriers and Carl Wieman, Learning and retention of quantum concepts with different teaching methods, Phys. Rev. ST Phys. Educ. Res. **7**, 010101 (2011).
[37] Louis Deslauriers, Ellen Schelew, and Carl Wieman, Improved learning in a large-enrollment physics class, Science **332**, 862 (2011).
[38] Patricia Heller, Ronald Keith, and Scott Anderson, Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving, Am. J. Phys. **60**, 627 (1992).
[39] David T. Brookes, Yuehai Yang, and Binod Nainabasti, Social positioning in small group interactions in an investigative science learning environment physics class, Phys. Rev. Phys. Educ. Res. **17**, 010103 (2021).
[40] Daniel J. Beal, Robin R. Cohen, Michael J. Burke, and Christy L. McLendon, Cohesion and performance in groups: A meta-analytic clarification of construct relations, J. Appl. Psych. **88**, 989 (2003).
[41] Won-Woo Park, Mee Sook Kim, and Stanley M. Gully, Effect of cohesion on the curvilinear relationship between team efficacy and performance, Small Group Res. **48**, 455 (2017).
[42] JinYoung Han and Anat Hovav, To bridge or to bond? Diverse social connections in an IS project team, Int. J. Project Manag. **31**, 378 (2013).
[43] Jordan D. Herbison, Alex J. Benson, and Luc J. Martin, Intricacies of the friendship-cohesion relationship in children's sport, Sport Exerc. Psychol. Rev. **13**, 10 (2017).
[44] Hsiao-Yun Liang, Hsi-An Shih, and Yun-Haw Chiang, Team diversity and team helping behavior: The mediating roles of team cooperation and team cohesion, Eur. Manag. J. **33**, 48 (2015).
[45] Anil Kumar Goswami and Rakesh Kumar Agrawal, Explicating the influence of shared goals and hope on knowledge sharing and knowledge creation in an emerging economic context, J. Knowl. Manag. **24**, 172 (2020).
[46] Karen F. Osterman, Students' need for belonging in the school community, Rev. Educ. Res. **70**, 323 (2000).
[47] Nicole L. Ennen, Emily Stark, and Andrea Lassiter, The importance of trust for satisfaction, motivation, and academic performance in student learning groups, Soc. Psychol. Educ. **18**, 615 (2015).
[48] R. T. Johnson and D. W. Johnson, Active learning: Cooperation in the classroom, Annu. Rep. Educ. Psychol. Jpn. **47**, 29 (2008).

[49] Noreen M. Webb, Kariane Mari Nemer, and Stephen Zuniga, Short circuits or superconductors? Effects of group composition on high-achieving students' science assessment performance, Am. Educ. Res. J. **39**, 943 (2002).

[50] Elizabeth G. Cohen, Restructuring the classroom: Conditions for productive small groups, Rev. Educ. Res. **64**, 1 (1994).

[51] IvanSrba and Maria Bielikova. Dynamic group formation as an approach to collaborative learning support, IEEE Trans. Learn. Technol. **8**, 173 (2015).

[52] Katherine Deibel, Team formation methods for increasing interaction during in-class group work, in *Proceedings of the 10th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education, Caparica Portugal* (ACM, 2005), pp. 291–295.

[53] Barbara Oakley, Richard M. Felder, Rebecca Brent, and Imad Elhajj, Turning student groups into effective teams, J. Student Centered Learn. **2**, 9 (2004).

[54] Larry K. Michaelsen, Arletta Bauman Knight, and L. Dee Fink, *Team-Based Learning: A Transformative Use of Small Groups* (Greenwood Publishing Group, Westport, CT, 2002).

[55] Susan Brown Fiechtner and Elaine Actis Davis, Republication of "Why some groups fail: A survey of students' experiences with learning groups", J. Manag. Educ. **40**, 12 (2016).

[56] Dragan Gašević, Amal Zouaq, and Robert Janzen, Choose your classmates, your GPA is at stake!": The association of cross-class social ties and academic performance, Am. Behav. Sci. **57**, 1460 (2013).

[57] Hichang Cho, Geri Gay, Barry Davidson, and Anthony Ingraffea, Social networks, communication styles, and learning performance in a CSCL community, Comput. Educ. **49**, 309 (2007).

[58] J. Richard Hackman and Nancy Katz, Group behavior and performance, in *Handbook of Social Psychology*, 1st ed., edited by Susan T. Fiske, Daniel T. Gilbert, and Gardner Lindzey (Wiley, New York, 2010)

[59] C. Reffay and T. Chanier, How social network analysis can help to measure cohesion in collaborative distance-learning, in *Designing for Change in Networked Learning Environments*, edited by Barbara Wasson, Sten Ludvigsen, and Ulrich Hoppe (Springer Netherlands, Dordrecht, 2003), pp. 343–352.

[60] Ling Cen, Dymitr Ruta, Leigh Powell, Benjamin Hirsch, and Jason Ng, Quantitative approach to collaborative learning: Performance prediction, individual assessment, and group composition, Int. J. Comput. Support. Collab. Learn. **11**, 187 (2016).

[61] Mohammed Saqr, Uno Fors, and Jalal Nouri, Using social network analysis to understand online problem-based learning and predict performance, PLOS One **13**, e0203590 (2018).

[62] R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2020).

[63] David W. Johnson, Roger T. Johnson, and Karl Smith, The state of cooperative learning in postsecondary and professional settings, Educ. Psychol. Rev. **19**, 15 (2007).

[64] David W. Johnson and Roger T. Johnson, An educational psychology success story: Social interdependence theory and cooperative learning, Educ. Res. **38**, 365 (2009).

[65] Sarah-Jane Leslie, Andrei Cimpian, Meredith Meyer, and Edward Freeland, Expectations of brilliance underlie gender distributions across academic disciplines, Science **347**, 262 (2015).

[66] Adrienne L. Traxler, Ximena C. Cid, Jennifer Blue, and Ramón Barthelemy, Enriching gender in physics education research: A binary past and a complex future, Phys. Rev. Phys. Educ. Res. **12**, 020114 (2016).

[67] Danny Doucette and Chandralekha Singh, Gender equity in physics labs, Phys. Rev. Phys. Educ. Res. **20**, 010102 (2024).

[68] Meagan Sundstrom and N. G. Holmes, Bias in physics peer recognition does not explain gaps in perceived peer recognition, Nat. Phys. **21**, 524 (2025).

[69] L. N. Simpfendoerfer Meagan Sundstrom, Annie Tan, Ashley B. Heim, and N. G. Holmes, Who and what gets recognized in peer recognition, Phys. Rev. Phys. Educ. Res. **20**, 010127 (2024).

[70] John Hattie and Helen Timperley, The power of feedback, Rev. Educ. Res. **77**, 81 (2007).

[71] Richard M. Felder and Rebecca Brent, Cooperative learning, in *Active Learning: Models from the Analytical Sciences* (American Chemical Society, Washington, DC, 2007), Vol. 970, pp. 34–53.

[72] Maria-Elena Osiceanu and MioaraDanac. The importance of personal development and group cohesion within young learners classroom, in *Education Facing Contemporary World Issues* (2019), pp. 1080–1089, 10.15405/epsbs.2019.08.03.131.