IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 7 November 2025, accepted 3 January 2026, date of publication 12 January 2026, date of current version 16 January 2026.
Digital Object Identifier 10.1109/ACCESS.2026.3653114

APPLIED RESEARCH

# The Formation of Collaborative Learning Teams in Schools Through Mathematical Optimization for Improving the Classroom Climate

**ALFREDO CANDIA-VÉJAR[1], ÁLVARO FAÚNDEZ[2], CAMILA ROJAS[2], ISIDORA JERIA[2], NATACHA BENÍTEZ[2], MARÍA IGNACIA TUPPER[2], ROMINA INOSTROZA[2], ETIENNE BELLENGER[2], AND FRANCISCO PÉREZ-GALARCE[3]**
[1]School of Civil Engineering, Universidad Finis Terrae, Santiago 7501015, Chile
[2]BRAVE-UP, Santiago 7510186, Chile
[3]Núcleo de Investigación en Data Science, Facultad de Ingeniería y Negocios, Universidad de Las Américas, Santiago 8320000, Chile
Corresponding author: Francisco Pérez-Galarce (fperezg@udla.cl)

**ABSTRACT** School climate plays a fundamental role in teaching and learning processes, directly influencing the achievement of educational objectives. Moreover, climate-related issues often manifest as bullying, which remains a persistent challenge for education systems worldwide. In this context, teamwork emerges as a key strategy to foster student interaction, strengthen friendship networks, and reduce bullying and aggression. However, traditional team assignment methods—such as student self-selection, teacher allocation, or random distribution—fail to account for the complexity of student interactions and individual characteristics. This study introduces a web-based decision support system designed to assist in team formation within educational settings by utilizing optimization models and algorithms, thereby promoting a more interconnected and inclusive student network. The models incorporate factors such as team diversity, victim protection, and student preferences. Its primary input is a social network along with derived metrics (e.g., betweenness centrality and triadic relationships), which feed into three non-linear integer optimization models that aim to consolidate, create, or enhance student interpersonal relationships. Additionally, the models were evaluated through a large quasi-experiment using both quantitative and qualitative analyses across four dimensions particularly relevant to educational activities: team dynamics, attitude toward the team, team cohesion, and team performance. Empirical results demonstrate that it is possible to improve the classroom climate without compromising the four team-related dimensions. This collaborative learning technology serves as a valuable resource for enhancing collaborative educational strategies and fostering a more inclusive and supportive classroom environment.

**INDEX TERMS** Collaborative learning, team formation problem, mathematical optimization, classroom climate.

## I. INTRODUCTION

School climate is a fundamental pillar of the education system, containing all relationships and interactions among educational community members within and beyond the school. Promoting a healthy classroom climate is challenging due to the many interacting factors that influence student behavior and relationships. Various concepts are closely associated with school climate, including the UNESCO goal of 'learning to live together', which emphasizes the critical role of education in promoting social camaraderie from the early stages of schooling [45]. Moreover, school climate plays a crucial role in mitigating bullying, a persistent and concerning issue in educational environments. Recent studies indicate that one in five students reports being a victim of bullying, and 41.3% of these students believe they will be targeted again [9]. The consequences of bullying can range from social exclusion to extreme acts of violence, including

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad J. Abdel-Rahman.

school shootings, and it can manifest in diverse forms, such as rumors, physical aggression, and psychological intimidation. An intuitive teaching practice to promote a better climate is to incorporate collaborative or cooperative activities. In designing such activities, team formation constitutes a key decision; however, in practice, simplistic or greedy processes are often employed, such as student self-selection, teacher allocation, or random distribution, without considering the rich structure of interpersonal networks or previous data from each classroom.

From a mathematical modeling perspective, the Team Formation Problem (TFP) emerges as an alternative to provide optimal solutions. This problem has attracted interest across various research fields [10], [21], [22], [24], primarily due to its relevance in diverse application areas, ranging from human resource management to the design of working groups in educational settings. In academic contexts, studies have addressed the problem of team formation [22], [49]; however, to our knowledge, the school climate has not been considered a central focus. Academic performance remains the primary dimension explored. However, school climate should gain greater prominence in teaching and learning processes, given its recognized importance in contributing to educational success as extensively evaluated in the literatures [42] and [46]. In particular, the discipline domain, which encompasses bullying and aggression, and the relationships domain, which involves students' interactions with peers, are considered two of the main domains of school climate [20].

We propose rethinking the TFP in schools to enhance student interaction. The objective is to create teams that promote positive relationships among students. To achieve this, we model a classroom using a directed graph $G(N, A)$, where $N$ represents the set of students and $A$ denotes the set of weighted links (positive, negative, or neutral) between students. To improve the course climate, we use network cohesion as a measure; that is, the set of students $N$ will exhibit a better climate if we can induce a denser network with positive arcs. The collaborative activity can change the type of arc between students who work together, generating new links, improving negative links, and strengthening positive links. Consequently, the network $G$ will evolve into a new network $G'$, as the formation of subgraphs or teams will modify the types of links. Figure 1 shows an example of the modification of a social network after the formation of the teams. In this study, we refer to $G$ as the course, level, and classroom, and $T$ as the team. Some research questions arise from these definitions: *To what extent must we compromise team cohesion or team climate in order to improve the broader room climate? Can we improve the classroom climate without compromising the team's performance? How can we design teams to increase classroom climate without reducing team performance or team climate?*

A key assumption in this research is the availability of detailed information about social networks. This assumption is made possible through the use of sociometric instruments, particularly those developed and automated by BRAVE-UP!,[1] a start-up specializing in web-based educational solutions for school climate and a crucial partner in this research. Among these, the enriched sociogram provides a comprehensive view of classroom dynamics, offering global and individual-level metrics such as mediation, isolation, rejection, and popularity [27], [36], [40]. These sociometric survey also capture valuable information on victimization, gender identity, and well-being, factors known to shape school climate.

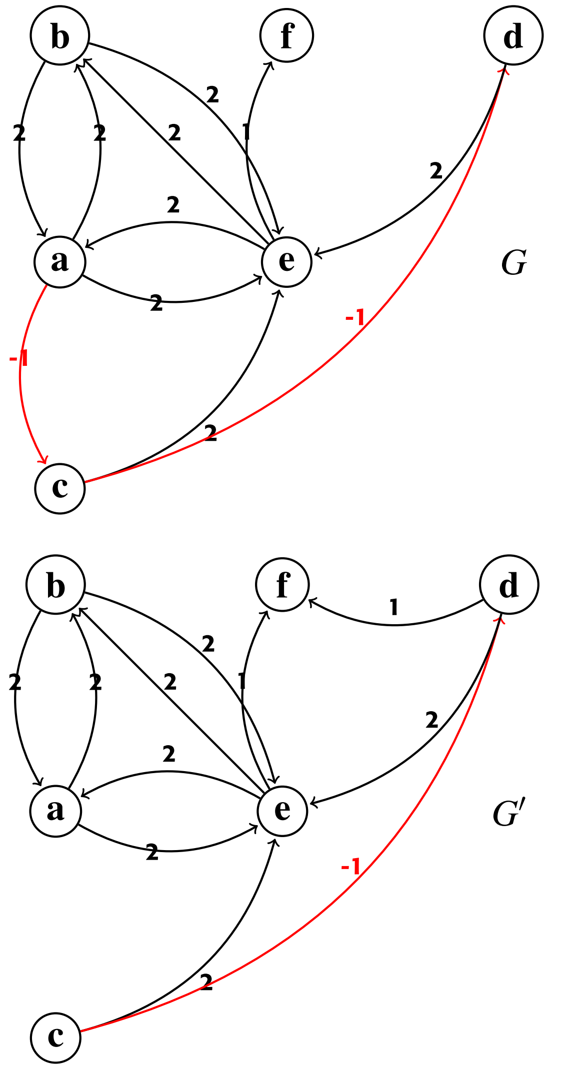



**FIGURE 1.** Top: Social network $G$ before the team activity. Bottom: Social network $G'$ after the activity. Team $T_1$ was composed of a, b, and c, while team $T_2$ included d, e, and f. The negative link between a and c was repaired, and a new link between d and f was induced. Red links indicate a negative relationship between students. The weight represents the strength of the relationship; higher values indicate better relations, while a weight of zero represents a neutral relationship with no edge.

Using these sociometric data as the primary source of information, we introduce a novel technology, which is a web-based decision support system (DSS), designed to

[1] https://app.braveup.co/

enhance cohesion in educational structures. This technology is built on three optimization models that incorporate individual and collective interpersonal relationships derived from sociometric surveys to create teams that enhance the room climate. Empirical evaluations conducted in real-world educational settings have yielded promising results, underscoring the model's potential to inform and enhance team formation practices. Specifically, the contributions of this paper are as follows:

- We mathematically modeled a set of innovative constraints and objective functions aimed at fostering a positive climate within school environments. These elements are grounded in interaction data from sociometric surveys and form the foundation of novel non-linear integer optimization models that incorporate specific information from the social graph (e.g., betweenness centrality and triadic relationships). The models include mechanisms to promote inclusive team dynamics, such as safeguards against bullying, strategies to prevent segregation, constraints to avoid strong mutual rejection, and limitations on the formation of exclusive friendship cliques. In contrast to previous studies, our work explicitly integrates social network structures into the team formation problem by incorporating sets of positive and negative links with varying intensities, as well as constraints such as cliques and victim protection. Moreover, we propose three novel objective functions designed to enhance classroom cohesion by fortifying, creating, and repairing social links, an approach not previously explored in the literature. To address conflicts directly, we further incorporate the betweenness centrality index. Overall, our formulation extends existing models by aligning optimization with multidisciplinary insights on classroom climate, emphasizing compatibility, segregation, and protection throughout the team assignment process.
- We developed strategies from an exact algorithmic perspective to enhance the performance of an open-source solver. These strategies included the design of symmetry-breaking constraints, a linearization analysis of two formulations, and the integration of valid inequalities. In addition, we proposed a heuristic based on the random key genetic algorithm (RKGA) to solve the models, which is deployed in a web application. This algorithm is integrated into an automated data pipeline spanning from survey collection to team proposal generation.
- We validated our technology through a large-scale quasi-experiment across 26 courses involving 880 students. The results yielded promising quantitative and qualitative insights, statistically demonstrating the feasibility of enhancing classroom cohesion without compromising team-related performance dimensions.

This article continues as follows. Section II presents a comprehensive literature review divided into two subsections: methodological contributions to solving the TFP and a summary of educational applications. Section III-A details the optimization problem, emphasizing its educational relevance. Section III-B provides the mathematical formulation, including objective functions and constraints. Section III-C presents an illustrative example and computational experiments with exact algorithms. Section IV outlines the heuristic based on RKGA. Section V offers insights into the quasi-experiment from both qualitative and quantitative perspectives. Finally, conclusions and future research directions are presented in Section VI.

## II. RELATED WORK

The TFP is an optimization problem that focuses on creating teams from a given set of individuals. This problem arises in various contexts, including education, sports, industry, and others. Recent applications can be found in education, [22]; online games, [2]; robotic competitions, [16]; and soccer, [50]. In its various applications, the TFP includes a set of basic constraints that require each individual to participate in exactly one team and that the teams have predefined sizes. Additional constraints are frequently added based on the specific field of application. Due to the NP-hardness of the TFP [26], various strategies have been used to try to obtain optimal solutions. We first review methodological issues, discussing formulations and algorithms for several variants of the problem. Secondly, some existing research gaps in the literature close to the TFP in educational settings are presented.

### A. ADVANCES IN TFP VARIANTS AND SOLUTION STRATEGIES

The literature on mathematical programming and heuristic approaches for modeling and solving the TFP is well-established and has experienced a notable rebirth in recent years. Fitzpatrick and Askin [17] formulated a model for grouping workers to form multiple teams in a manufacturing setting. In this case, workers were evaluated according to specific skills, and each was categorized based on the role they could fulfill. The objective was to select workers for each role to maximize synergy and minimize inertia in team performance. A heuristic was developed to address large problem instances. Ani et al. [3] proposed a genetic algorithm-based method to form student teams for software projects according to their programming skills. Similarly, [1] developed heuristics to group students in a way that maximizes aggregate student gain. They specifically examined the effect of combining the weakest and strongest students to increase group diversity. Sadeghi and Kardan [41] addressed team formation in collaborative learning environments by proposing an integer programming model that incorporates the concept of justice.

In [14], the role of social network structure in team formation is explored. They formulated a mixed-integer linear programming (MILP) model that quantitatively accounts for social dependencies and used a Lin-Kernighan-based heuristic to solve the problem. Gutiérrez et al. [21] introduced a MILP model for the problem of multiple team formation, where the objective function is derived from a sociometric matrix that captures interpersonal relationships. They proposed a constraint programming approach and a variable neighborhood search metaheuristic to solve the model. Note that [21] proposed a team-formation problem without climate-related constraints, and the application is focused on project management. They considered only a sociometric coefficient in the objective function; in other words, they did not exploit the structure of a sociogram. We explicitly cover this gap by incorporating constraints that leverage social network structures, such as the clique constraints.

Fathian et al. [15] formulated a model to maximize the reliability of the team by considering the probability that unreliable experts leave the team. Their approach also incorporated the experts' skill profiles within a collaboration network. Zheng et al. [51] studied team formation in collaborative learning environments, proposing a MILP formulation and a genetic algorithm to address the problem. The model aimed to form balanced groups by minimizing intra-group diversity. Magnanti and Natarajan [31] analyzed a variant of the TFP in which the objective was to allocate final projects to senior-year undergraduate students. Their results indicated that the proposed method provided practical support for academic staff managing the group formation process.

Moreno et al. [35] proposed several grouping criteria, allowing for arbitrary parameters related to team size and grouping attributes. To solve the problem, they developed a hybrid solution algorithm based on a randomized heuristic method. Rehman et al [39] addressed the TFP within the context of social networks, introducing an algorithmic approach that leverages communication costs and graph reduction techniques. Vallés-Catala and Palau [48] proposed an algorithm grounded in network theory to more effectively form heterogeneous groups. Their method demonstrated strong performance when applied to real-world data from a Master's program in teacher training.

Hill and Peuker [22] first constructed a student social network within an educational environment and then formulated an integer linear model to form groups to maximize the potential for new social ties. The proposed model outperformed classical grouping approaches such as random assignment and self-selection. It is important to underscore that this contribution also leverages the network structure and proposes several constraints related to classroom climate. However, it only models positive links, overlooking negative relationships that can be highly relevant for effective team formation. Considering both sides of social interactions is crucial, not only to prevent teams with strong rejections but also to address relationships characterized by mild tension. Our work aims to bridge this gap by explicitly incorporating these attributes into the team formation model. Finally, note that [22] addresses a different formulation that does not consider peer-wise parameters between students, which limits the types of objective functions that can be optimized.

Muniz and Flamand [37] studied a new variant of the problem in which pre-existing teams are reorganized through member exchanges. The goal is to achieve balanced teams in terms of efficiency, ensure equitable distribution of success, and minimize training costs by optimally reallocating the current workforce. Lv et al. [30] introduced a reinforcement learning approach to tackle the project team formation problem in real-world organizational networks. The objective was to identify a subset of individuals that satisfy the skill requirements while minimizing disruption to the existing network structure.

Although the reviewed papers differ in their objective functions and constraints within the TFP framework, they share a common characteristic: a high level of computational complexity that makes finding the optimal solution challenging. Many optimization models involve a quadratic objective function, which further increases the difficulty of solving them efficiently. To address this complexity, various strategies have been proposed. Some works implement reformulations to design tailored branch-and-bound algorithms. These are further strengthened by incorporating valid inequalities to reduce the solution space. Campelo and Figueiredo [10] followed this approach, and [6] structured the feasible region as a polyhedron defined by set-covering constraints associated with skills, along with conflict constraints that prevent incompatible pairs from being assigned to the same team [6]. Depending on the number of participants, the proposed algorithms required only a few seconds to several minutes to reach a solution, [6].

Most works begin by defining a specific TFP variant and then develop an algorithm tailored to that formulation. For example, [13] proposed a particle swarm optimization algorithm incorporating a novel swap operator, which demonstrates improved performance over standard benchmarks in terms of solution quality and computational time across datasets with varying numbers of experts and skills. In learning environments, [25] considered social, cognitive, and academic dimensions within a social network-based learning setting. After categorizing student attributes, they applied a genetic algorithm (GA) with a modified two-point crossover to promote the generation of better offspring. Their method was tested in a Software Technology course with 48 students at a Greek university. Despite 71% of students expressing satisfaction, 44% preferred self-selection, indicating a preference for working with friends. Similarly, [34] addressed team formation by optimizing heterogeneous intragroup and homogeneous intergroup structures using a multi-objective GA. Each objective sought to minimize the difference between group and overall population means for student characteristics. The method was applied in an engineering course with 135 students at the National University of Colombia, where groups of five students were formed.

Lin et al. [29] proposed a two-objective optimization model based on benefit and cost criteria related to student attributes. A web-based system was developed for practical use by teachers, allowing input of student characteristics and generating group solutions. Although the method improved learning outcomes compared to a control group, it did not consider social cohesiveness, only academic factors. Extending this idea, [33] included empathy as a third objective alongside heterogeneous intragroup and homogeneous intergroup diversity. Using NSGA-II, they evaluated their model in two educational settings: a face-to-face computer programming course and an online e-commerce course. The proposed grouping method outperformed both random and self-assigned methods, particularly in terms of group empathy. Finally, [18] also used NSGA-II to solve the optimal learning group formation problem. Their formulation incorporated five distinct student characteristics for pre-qualification but, like others, did not account for social cohesion.

### *B. RESEARCH GAPS FOR THE TFP IN EDUCATION*

A common criterion across many of the previously presented studies is the formation of internally heterogeneous and externally homogeneous teams. Furthermore, some works introduce climate-related attributes into their models. However, to our knowledge, none of the reviewed studies explicitly focus on student climate as the central objective. In particular, we found no models that simultaneously integrate interpersonal relationships (positive and negatives at different levels), network attributes (e.g., betweenness centrality and triadic relationships), personal attributes (e.g., nationality and gender identity), and victimization constraints within each team. It is also crucial to highlight that the literature does not consider bullying as a key factor in team formation, despite its relevance to school climate and student well-being. In contrast, our work aims to fill this gap by proposing a team formation model that prioritizes room climate. Additionally, we introduce three distinct objective functions that reinforce climate-related criteria as a central component of the grouping process.

## III. A TEAM FORMATION PROBLEM TO IMPROVE THE SOCIAL NETWORK COHESION

### *A. PROBLEM DEFINITION*

The team conditions were defined in collaboration with experts in educational climate through a series of working meetings to understand the problem, define variables and assumptions, and validate the proposed models. Below, we outline the key conditions that define a feasible team.

#### 1) TEAM SIZE

There is a minimum and maximum number of participants in each team.

#### 2) TEAM MEMBERSHIP

Each student must belong to exactly one team.

#### 3) EXTREME INCOMPATIBILITIES

Each team cannot consist of students with a very high mutual rejection level (minimizing the high probability of conflict). This condition would allow students with low-level rejection to participate in a team without mutual rejection.

#### 4) EXTREME COMPATIBILITIES

Each team cannot be made up solely of students with a high level of mutual acceptance (minimizing the high probability of team friendship). A parameter can be defined that determines the number of pairs of students with mutual acceptance allowed. This condition helps generate new bonds.

#### 5) INCLUSION

A set of characteristics is defined (e.g., gender or ethnicity), so that, for each characteristic, no team has all its students sharing the same characteristic.

#### 6) VICTIM PROTECTION

This constraint is about protecting students known to have been victims of aggression. For this purpose, the following are defined. Each victim must belong to a team with at least one of their friends. Additionally, care must be taken that none of their friends reject the victim.

The objective functions defined during this phase are presented below:

#### 7) MAXIMIZE STUDENT PREFERENCES

This goal aims to improve network cohesion by strengthening existing links.

#### 8) MAXIMIZE THE PROBABILITY OF CREATING NEW LINKS

This goal aims to enhance network cohesion through the emergence of new links.

#### 9) MAXIMIZE THE MEDIATION INDICATOR

This goal considers that a mediator can impact the strengthening of connections within a team. In particular, in cases of negative links, a mediator can help to solve this tension.

### *B. A MATHEMATICAL PROGRAMMING FORMULATION FOR THE OPTIMIZATION PROBLEM*

We define the problem of team formation on a network. The network is a complete directed weighted graph $G = (N, A, c)$, where $N$ is the set of students, $A$ is the set of arcs in the network, which define relationships between students, and $c$ is a weight function assigned to the arcs of the network and will be defined later. It is important to note that the arcs of the network allow for the specification of unidirectional relationships from student $i$ to student $j$, denoted as $(i, j)$, in the case where one also wants to define the reciprocal relationship from student $j$ to student $i$, a new arc $(j, i)$ must be defined. The acceptance/rejection levels of the network and the rest of the parameters and sets presented below are estimated through the analysis of a sociometric instrument that considers a sociogram and questions linked to aggressions in the school context.

### 1) SETS

The set of students in the course is denoted by $N$, and the set of teams to be created is represented by $K$. For each student $i \in N$, we define several preference-based sets describing the interpersonal relations identified in the sociometric analysis:

- $\mathrm{NT}(i)$: set of students perceived as neutral by student $i$.
- $\mathrm{R}(i)$: set of students weakly rejected by $i$.
- $\mathrm{RF}(i)$: set of students strongly rejected by $i$.

The set of student pairs with mutual strong rejection is defined as:

$$\mathrm{RFM} = \{(i,j) : i \in \mathrm{RF}(j),\ j \in \mathrm{RF}(i)\}.$$

Conversely, the sets of students positively evaluated by student $i$ are:

- $\mathrm{AW}(i)$: set of students weakly accepted by $i$.
- $\mathrm{AF}(i)$: set of students strongly accepted by $i$.

The set of mutually strong acceptance pairs is:

$$\mathrm{AFM} = \{(i,j) : i \in \mathrm{AF}(j),\ j \in \mathrm{AF}(i)\}.$$

Let $\Omega$ denote the set of student attributes (e.g., nationality, gender identity, academic achievement). For each attribute $l \in \Omega$, let $\Gamma_l$ represent the set of possible categories $t$, and let $\Phi_{tl}$ denote the set of students belonging to category $t \in \Gamma_l$ in attribute $l$.

The subset $\mathrm{VB} \subset N$ includes students reported as victims of aggression. For each victim $i \in \mathrm{VB}$, we define $\mathrm{AR}(i)$ as the set of classmates $j$ accepted as friends (either weakly or strongly) who did not reject $i$ in return:

$$\mathrm{AR}(i) = \{j \in N : j \in \mathrm{AW}(i) \cup \mathrm{AF}(i),\ i \notin R(j) \cup \mathrm{RF}(j)\}.$$

If $\mathrm{AR}(i)$ is empty, a fallback mechanism selects students with high betweenness centrality. The betweenness of student $i$, denoted $b_i$, is defined as the sum of the fractions of all pairs of shortest paths that pass through $i$ [8].

Finally, let $M$ denote the set of mediation triplets $(i, j, h)$ representing potential interventions between students $i$ and $j$—connected by an arc in $R(j)$—through a third student $h \in N$. These mediation triplets are identified by enumerating triads in the social network [4], retaining those that include the arc $(i,j) \in R(j)$.

### 2) PARAMETERS

$\gamma$: A positive integer that defines the maximum size of the teams.
$S_k$: Size of team $k$, $k \in K$. $\gamma - 1 \le S_k \le \gamma$, $\gamma$ is a positive integer. It is computed based on $\gamma$ and the course size.
$b_i$: Betweenness centrality index of student $i$, where $i \in N$, provides information about the number of friendship paths in which student $i$ is involved [7], thereby indicating their potential role as a mediator.
$v_i$: Number of connections student $i$ has in the network, whether they are incoming or outgoing arcs.
$v_{ij}$: Value of joint linkage. For each $i, j$, calculate $v_{ij} = v_i + v_j$.
$w_{ij}$: Elements of the weighting matrix for the objective of new links. The idea is that pairs $(i, j)$ with high $v_{ij}$ indicate that both students have a high connection. We prioritize that the model places pairs with low linkage on a team, and then at least one weakly linked student will improve their level. From the joint linkage values $v_{ij}$, we define $w_{ij}$ which is their multiplicative inverses $1/(v_{ij} + 1)$.
$c_{ij}$: Definition of the weighting function associated with the network connections. The value $c_{ij}$ then defines the level of acceptance or rejection that student $i$ has towards student $j$. Five levels are defined in the network: $c^{--}$, $c^{-}$, $c^{o}$, $c^{+}$, and $c^{++}$, where at the extremes, $c^{--}$ represents a strong rejection from $i$ to $j$, and $c^{++}$ represents a strong acceptance from $i$ to $j$. Specifically, $c_{ij} = c^{--}$ if $j \in \mathrm{RF}(i)$; $c_{ij} = c^{-}$ if $j \in R(i)$; $c_{ij} = c^{0}$ if $j \in NT(i)$; $c_{ij} = c^{+}$ if $j \in AW(i)$; and $c_{ij} = c^{++}$ if $j \in \mathrm{AF}(i)$.
$L$: This is the maximum size of a clique allowed in a team. A clique of size $q$ in a network is a subset of nodes such that all pairs of nodes are connected, and it contains $q$ nodes.
$CQ_L$: A clique of size less than or equal to $L$. The cliques are defined in this work, in an undirected network $G^U$ generated from the network $G$ as follows. $G^U = \{\{i,j\} : (i,j) \in A, c_{i,j} = c_{ij}^{++}\}$. The notation $\{i, j\}$ is used here to emphasize that the connections in $G^U$ are edges and not directed arcs.
$\alpha_{tl}, \beta_{tl}$: These define the upper and lower bounds, respectively, for the number of students within each team with the category $t \in \Gamma_l$ in the attribute $l \in \Omega$. To balance the teams in terms of the categories of each attribute, $\beta_{tl}$ and $\alpha_{tl}$ can be estimated as $\left\lfloor \frac{|\Phi_{tl}|S_k}{|N|} \right\rfloor$ and $\left\lceil \frac{|\Phi_{tl}|S_k}{|N|} \right\rceil$, respectively. When $\alpha$ and $\beta$ are equal, i.e., $\frac{|\Phi_{tl}|S_k}{|N|}$ is an integer, they are redefined as $\alpha = \alpha + 1$ and $\beta = \beta - 1$.

### 3) DECISION VARIABLE

To formulate the optimization models, it is only necessary to define a set of decision variables $\mathbf{x} \in \{0, 1\}^{|\mathrm{N} \times \mathrm{K}|}$, where each element $x_{ik}$ takes the value 1 if student $i$ is assigned to team $k$, and 0 otherwise.

### 4) THE OPTIMIZATION MODELS

In this section, we define three distinct models, each sharing the same feasible solution space but differing in their objective functions, denoted as $\mathrm{M}_1$, $\mathrm{M}_2$, and $\mathrm{M}_3$. The three objective functions considered, along with the corresponding constraints, are presented below.

$$\max Z_1 : \quad Z_1 = \sum_{(i,j)\in A} c_{ij} \sum_{k\in K} x_{ik} x_{jk} \tag{M$_1$}$$

$$\max Z_2 : \quad Z_2 = \sum_{(i,j)\in A} w_{ij} \sum_{k\in K} x_{ik} x_{jk} \tag{M$_2$}$$

$$\max Z_3: \quad Z_3 = \sum_{(i,j,h)\in M} b_h \sum_{k\in K} x_{ik}x_{jk}x_{hk} \tag{M$_3$}$$

subject to:

$$\sum_{i\in N} x_{ik} = S_k, \quad \forall k \in K \tag{1}$$

$$\sum_{k\in K} x_{ik} = 1, \quad \forall i \in N \tag{2}$$

$$x_{ik} + x_{jk} \leq 1, \quad \forall (i,j) \in \mathrm{RF}(i), \forall k \in K \tag{3}$$

$$\sum_{i\in CQ_L} x_{ik} \leq L, \quad \forall k \in K \tag{4}$$

$$\sum_{i\in\Phi_{cl}} x_{ik} \leq \alpha_{cl}, \quad \forall k \in K, \forall c \in \Gamma_l, \forall l \in \Omega \tag{5}$$

$$\sum_{i\in\Phi_{cl}} x_{ik} \geq \beta_{cl}, \quad \forall k \in K, \forall c \in \Gamma_l, \forall l \in \Omega \tag{6}$$

$$x_{ik} - \sum_{p\in AR(i)} x_{pk} \leq 0, \quad \forall i \in \mathrm{VB}, \forall k \in K \tag{7}$$

$$x_{ik} \in \{0,1\} \forall i \in N, k \in K, \tag{8}$$

Model $M_1$ maximizes student preferences, aiming to create work teams that enhance network cohesion by reinforcing existing social links. Accordingly, the objective function $Z_1$ sums the weights of connections between student pairs within the formed teams, where the strength of each link is quantified and maximized. Model $M_2$ seeks to maximize the number of new social links. This function encourages the formation of teams that foster new interactions within the classroom network. Analogous to $Z_1$, the objective function $Z_2$ prioritizes student pairings with higher potential for new link formation, as defined by the linking matrix and the associated weights $w_{ij}$. Finally, model $M_3$ aims to increase the presence of effective mediators in teams that include students with strained but potentially repairable relationships. The objective function $Z_3$ rewards configurations in which a student with a high mediation score (betweenness centrality) is assigned to the same team as a student involved in such a strained relationship, thereby leveraging the mediator's positive influence.

Constraint (1) enforces the predefined team size $S_k$, as determined by the decision maker for each team $k$. Constraint (2) ensures that every student is assigned to precisely one team, thereby avoiding both unassigned students and duplicate assignments. To handle interpersonal conflicts, Constraint (3) prevents students who have firmly rejected each other from being assigned to the same team. This constraint includes both unidirectional and mutual rejections. Conversely, Constraint (4) limits the number of students with strong mutual acceptance who are placed on the same team. This condition is intended to discourage the formation of strong cliques beyond a permissible size, defined by the parameter $CQ_L$, which depends on the total team size. Constraints (5) and (6) prevent the formation of homogeneous teams based on sensitive attributes such as gender identity, academic performance, or nationality, ensuring diversity within and across teams. Lastly, Constraint (7) provides safeguards for students identified as victims of aggression, ensuring that each student is assigned to a team with at least one member from their trusted set $AR(i)$, promoting a safe team environment.

Note that, in order to study the impact of the constraints, we introduce $M_0$, which corresponds to model $M_1$ without climate-related constraints, allowing for comparative analysis. The exact solution approaches used to solve these models—namely, linearization strategies, symmetry-breaking constraints, and the development of novel valid inequalities—are detailed in the Appendix and supported by computational experiments.

**TABLE 1. Students. ∗ identifies students reported or self-reported as victims of aggression. Students with positive betweenness centrality, namely, 4, 5, and 8, could potentially play a protective role for victim 6.**

| Student | Betweenness | Nationality | Gender Identity | Grade |
|---|---|---|---|---|
| 1 | 0.00 | local | male | high |
| 2 | 0.00 | local | male | low |
| 3 | 0.00 | local | female | low |
| 4 | 0.11 | local | male | low |
| 5 | 0.16 | local | non-binary | low |
| 6* | 0.15 | local | female | medium |
| 7 | 0.00 | immigrant | female | medium |
| 8 | 0.09 | local | female | medium |
| 9 | 0.00 | immigrant | non-binary | high |
| 10 | 0.00 | local | female | medium |
| 11 | 0.00 | immigrant | male | high |
| 12 | 0.00 | local | male | high |
| 13 | 0.00 | local | male | high |

### C. ILLUSTRATIVE EXAMPLE

The following example considers a set of 13 students, denoted by the letters $\{1, 2, \ldots, 12, 13\}$. The relationships among the students are represented by a directed graph shown in Figure 2, where only non-neutral relationships are visualized; that is, edges with $c_{ij} = 0$ are omitted. Within this classroom scenario, the objective is to form four work teams of sizes 3, 3, 3, and 4, respectively ($\gamma = 4$). Table 1 presents the individual characteristics associated with each student. Based on this information, the sets $V$, $\Omega$, $\Gamma$, and $\phi$ are constructed, along with the parameter $b$ corresponding to betweenness centrality.

Table 2 presents the parameters $\alpha$ and $\beta$ that define expected ranges for each category of student attributes. In practice, $\beta = 2$ and $\alpha = 3$ mean that for that category, the model requires at least two students and at most three students from that category in each team. Tables 3 and 4 illustrate the sociometric preference matrix and the joint disengagement matrix.

Table 5 shows the solutions obtained by the three models. The teams obtained and the respective values for each objective are detailed. An analysis of the results is then presented, which helps to understand the impact of the three models on team formation. It is also worth noting that the analysis includes the results provided by the base model $M_0$, which considers the objective function $Z_1$ and includes

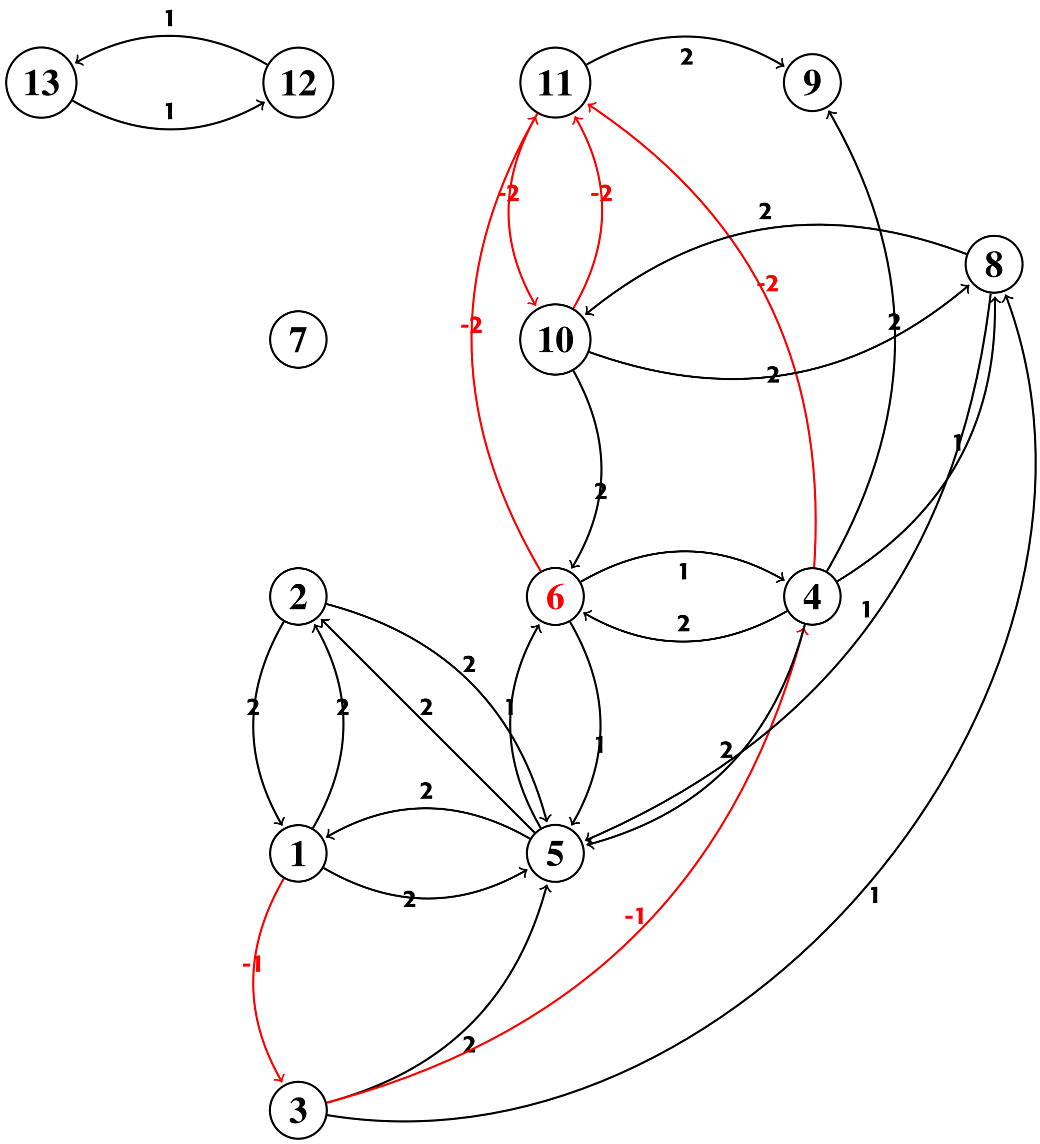



**FIGURE 2.** The number associated with each arc represents the type of link, where $c^{--} = -2$, $c^{-} = -1$, $c^{+} = 1$, and $c^{++} = 2$. The network consists of 28 arcs and thirteen nodes, including four strongly negative arcs ($c^{--}$), two weakly negative arcs ($c^{-}$), eight weak acceptance arcs ($c^{+}$), and fourteen strong acceptance arcs ($c^{++}$). Four mutual strong acceptances can also be found ({2, 1}, {2, 5}, {1, 5}, and {10, 8}) and one mutual strong rejection ({10, 11}). According to the graph, the weakly negative arc (1, 3) can be mediated by student 5, and the weakly negative arc (3, 4) can be also mediated by student 5.

**TABLE 2.** Expected ranges for the number of students per team with each attribute and category.

| Attribute | Category | $\beta$ | $\alpha$ |
|---|---|---|---|
| Nationality | Local | 2 | 3 |
| | Foreign | 0 | 1 |
| Gender Identity | Female | 1 | 2 |
| | Male | 1 | 2 |
| | Non-binary | 0 | 1 |
| Grade Range | High | 1 | 2 |
| | Medium | 0 | 2 |
| | Low | 0 | 2 |

basic constraints but not the fundamental constraints of the classroom climate. Therefore, this analysis allows us to appreciate the impact of the constraints and the different objective functions considered. These results appear in the first row of Table 5.

### 1) STUDY OF CONSTRAINTS

1. Strongly negative relations: In the groups related to $M_1$, $M_2$, and $M_3$, we note that in each of them there is no strong negative arc ($c^{--}$) between two nodes. This result is consistent with avoiding this type of linkage during team generation. Consequently, we observe from the results that in no case do the two nodes 10 and 11, which have strong mutual rejection in the social network, belong to the same team.

2. Victim students: In the social network, there is a victim student, node 6. It can be observed that student 6 belongs

**TABLE 3.** Sociometric matrix of preferences, $c_{ij}$.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 2 | -1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | -1 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 1 | 2 | 0 | -2 | 0 | 0 |
| 5 | 2 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | -2 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | -2 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | -2 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |

to teams where either student 4 or 5 is present. This effect aligns with the intended behavior of the victim-protection constraints. Both 4 and 5 were indicated as students accepted by 6.

3. Segregation: We study segregation in three dimensions, gender identity, nationality, and grade level. Regarding gender identity, all teams exhibit diversity in this attribute; however, in $M_0$, one group ({ 1, 2, 5 }) does not include female students. This issue is resolved in models $M_1$, $M_2$, and $M_3$. In terms of nationality, team $\{3, 9, 11\}$ in $M_0$ shows a high concentration of immigrant students, with two out of three members belonging to that category. This situation improves in $M_1$, $M_2$, and $M_3$. Finally, no teams are composed entirely of students from the same grade level, although team ({ 4, 6, 8, 10 }) includes three members with medium grades and none with high grades.

4. Compatibility between students: These constraints limit the size of the largest clique in a team. The social network contains a clique of three nodes with nodes $\{1, 2, 5\}$. The results show that this clique does not appear in any team with climate-room constraints. However, this clique appears in the teams created in $M_0$, as this model does not include climate constraints.

### 2) STUDY OF OBJECTIVE FUNCTIONS

From an optimization perspective, it is important to compare the objective function values $Z_1$, $Z_2$, and $Z_3$. For the analysis of Table 5, we adopt the following convention: each row corresponds to an optimal solution $T$ obtained using the model associated with that row. For example, the second row shows that $M_1$ yields the optimal solution $T_1$, with the following team configuration: $T_1 = \{\{4, 6, 9\}, \{8, 10, 12\}, \{1, 2, 7\}, \{3, 5, 11, 13\}\}$. When this solution is evaluated under the objective function $Z_2$, the result is 4.38, whereas under $Z_3$ it is 0.0. In contrast, the third row presents the optimal solution $T_2$, for which the value of $Z_1$ is 10.0—lower than the 15.0 achieved by $T_1$. These comparisons highlight the effect of employing different objective functions while maintaining the same set of constraints. They demonstrate that varying optimization criteria can lead to substantially different team configurations and performance outcomes, depending on the prioritization of school climate.

Base model without climate constraints ($M_0$): When working with the preference-based objective function but without including climate constraints, the teams $\{3, 9, 11\}$, $\{7, 12, 13\}$, $\{1, 2, 5\}$, and $\{4, 8, 10, 6\}$ emerge, as shown in Table 5. In this solution, the previously defined climate constraints are not enforced, which explains the high objective function value of $Z_1 = 26$ reported in Table 5.

Preferences ($M_1$): This model, unlike the previous, includes climate constraints. On the one hand, we obtained a very different optimal solution for teams from the last case since none of them was repeated. Additionally, we can see that the climate constraints, by restricting the model $M_0$, worsen the value of the objective function by decreasing from a value of 26 to 15, which makes sense.

New links ($M_2$): we observe that, unlike the analysis of $M_1$, there is a group, $\{2, 8, 11\}$, in which there are no connections between them on the social network. Therefore, new connections can be interpreted as being generated in this team. This behavior is also observed in the team $\{3, 7, 12, 13\}$, where connections exist only in a pair of nodes, so new links are also generated in this team. We can note that these links are due to the definition of $M_2$, which weighs the joint disconnection weights $w_{ij}$. In the team $\{3, 7, 12, 13\}$, we note that in the weights $w_{ij}$, the pairs (3, 7), (7, 12), (7, 13) have the highest value of 0.33 in the matrix. Consequently, $M_2$ was able to prioritize the formation of this team based on the $w_{ij}$ value of its pairs of elements in the team. In this way, by using this objective function, grouping highly linked students is avoided.

Mediators ($M_3$): we note that in team $\{1, 3, 5, 7\}$, a weakly negative arc is obtained from student 1 to 3. As both students are in the same group, student 5 acts as a mediator. Furthermore, that mediation was only achieved using $M_3$, and only affected one team. It is very likely that a higher level of mediation can be found in an example where there is a high level of negative relationships. Subsequently, a computational study will be carried out, and we will work with many instances where we can appreciate the performance of the various models. All code and datasets for reproducing these results are available at https://github.com/frperezgalarce/team-formation-problem.

## IV. RANDOM KEY GENETIC ALGORITHM

The RKGA is a well-established method for solving complex combinatorial optimization problems [5]. We implement this algorithm to address our specific team formation problem as follows. Given an ordered sequence of $N$ nodes, denoted by $A_N = (a_n)_{n \leq N} = (a_1, \ldots, a_N)$, a number $K$ of teams to be formed, and an ordered sequence of team sizes $S_K = (s_k)_{k \leq K}$, where $s_k > 0$ and $\sum_{k=1}^{K} s_k = N$, a feasible team assignment solution under basic constraints is represented by:

$$T(A_N, S_K) = \{t_k\}_{k \leq K}, \quad \text{where} \quad t_k = \{a_l, \ldots, a_u\},$$

with $l = \sum_{i<k} s_i + 1$ and $u = \sum_{i \leq k} s_i$. This representation ensures that the solution satisfies two fundamental constraints: each team has a predefined size, and each

**TABLE 4.** Joint unlinking matrix, $w_{ij}$.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.11 | 0.11 | 0.14 | 0.10 | 0.07 | 0.10 | 0.20 | 0.10 | 0.14 | 0.12 | 0.17 | 0.14 | 0.14 |
| 2 | 0.11 | 0.11 | 0.14 | 0.10 | 0.07 | 0.10 | 0.20 | 0.10 | 0.14 | 0.12 | 0.17 | 0.14 | 0.14 |
| 3 | 0.14 | 0.14 | 0.20 | 0.12 | 0.08 | 0.12 | 0.33 | 0.12 | 0.20 | 0.17 | 0.25 | 0.20 | 0.20 |
| 4 | 0.10 | 0.10 | 0.12 | 0.09 | 0.07 | 0.09 | 0.17 | 0.09 | 0.12 | 0.11 | 0.14 | 0.12 | 0.12 |
| 5 | 0.07 | 0.07 | 0.08 | 0.07 | 0.05 | 0.07 | 0.10 | 0.07 | 0.08 | 0.08 | 0.09 | 0.08 | 0.08 |
| 6 | 0.10 | 0.10 | 0.12 | 0.09 | 0.07 | 0.09 | 0.17 | 0.09 | 0.12 | 0.11 | 0.14 | 0.12 | 0.12 |
| 7 | 0.20 | 0.20 | 0.33 | 0.17 | 0.10 | 0.17 | 1.00 | 0.17 | 0.33 | 0.25 | 0.50 | 0.33 | 0.33 |
| 8 | 0.10 | 0.10 | 0.12 | 0.09 | 0.07 | 0.09 | 0.17 | 0.09 | 0.12 | 0.11 | 0.14 | 0.12 | 0.12 |
| 9 | 0.14 | 0.14 | 0.20 | 0.12 | 0.08 | 0.12 | 0.33 | 0.12 | 0.20 | 0.17 | 0.25 | 0.20 | 0.20 |
| 10 | 0.12 | 0.12 | 0.17 | 0.11 | 0.08 | 0.11 | 0.25 | 0.11 | 0.17 | 0.14 | 0.20 | 0.17 | 0.17 |
| 11 | 0.17 | 0.17 | 0.25 | 0.14 | 0.09 | 0.14 | 0.50 | 0.14 | 0.25 | 0.20 | 0.33 | 0.25 | 0.25 |
| 12 | 0.14 | 0.14 | 0.20 | 0.12 | 0.08 | 0.12 | 0.33 | 0.12 | 0.20 | 0.17 | 0.25 | 0.20 | 0.20 |
| 13 | 0.14 | 0.14 | 0.20 | 0.12 | 0.08 | 0.12 | 0.33 | 0.12 | 0.20 | 0.17 | 0.25 | 0.20 | 0.20 |

**TABLE 5.** $M_0$ represents the baseline model, which does not consider climate-related constraints but does include the optimization of preferences $M_1$.

| | $Z_1$ | $Z_2$ | $Z_3$ | team 1 | team 2 | team 3 | team 4 |
|---|---|---|---|---|---|---|---|
| $M_0$ | 26.00 | 4.82 | 0.00 | 3, 9, 11 | 7, 12, 13 | 1, 2, 5 | 4, 8, 10, 6 |
| $M_1$ | 15.00 | 4.38 | 0.00 | 4, 6, 9 | 8, 10, 12 | 1, 2, 7 | 3, 5, 11, 13 |
| $M_2$ | 10.00 | 5.28 | 0.00 | 2, 8, 11 | 1, 5, 6 | 4, 9, 10 | 3, 7, 12, 13 |
| $M_3$ | 8.00 | 4.40 | 0.16 | 9, 10, 12 | 2, 8, 11 | 4, 6, 13 | 1, 3, 5, 7 |

student is assigned to exactly one team. We refer to these as the basic restrictions. Climate-related constraints defined in the optimization model are not yet enforced at this stage. Instead, during the operation of the genetic algorithm, these climate restrictions are incorporated through a penalization mechanism applied to the fitness function. This approach guides the search toward feasible, high-quality solutions that progressively satisfy basic and climate constraints.

A reordered sequence $A_N^R$ can be easily obtained using random keys, a random key $R$, and a vector of random numbers within 0 and 1. If a random key vector with identical length to $A_N$ is generated, the ordered sequence $A_N$ can be reordered according to the keys in increasing order to obtain a new valid sequence. For example, given $A_N = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10)$ and $S_K = (4, 3, 3)$, the team assignment would be $T(A_N, S_K) = \{\{1, 2, 3, 4\}, \{5, 6, 7\}, \{8, 9, 10\}\}$. Considering the same sequence $A_N = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10)$ and a vector of random numbers of the same length, $R = (0.10, 0.50, 0.20, 0.30, 0.90, 0.25, 0.35, 0.55, 0.95, 0.15)$, these elements could be reordered by these values, obtaining the new sequence: $A_N^R = (1, 10, 3, 6, 4, 7, 2, 8, 5, 9)$, then, $A_N$ "induces" the solution $S = \{\{1, 2, 3, 4\}, \{5, 6, 7\}, \{8, 9, 10\}\}$, and after applying the transformation according to the random number vector $R$ we obtain the new sequence $A_N^R$, which induces the new solution $S' = \{\{1, 10, 3, 6\}, \{4, 7, 2\}, \{8, 5, 9\}\}$. Therefore, if $A_N$ is fixed, then the application of different vectors of random numbers to $A_N$ will generate different solutions to the problem. In this way, for computational purposes in the genetic algorithm that we will study in the next section, a solution $S$ is represented only by the vector of random numbers $R$ applied to $A_N$. Using the notation already defined, $S = T(I(A_N, R), S_K)$.

Given the solution $S$ defined by the vector of random numbers $R$, the chromosome associated with $S$ is represented as: $C^S = (R^s)$. Given $A_N$ and $S_K$, the initialization of the population of size $P$ is given by the generation of $P$ vectors of random numbers of length $N$. Three well-known crossover methods were implemented: mid-point, one-point, and uniform [28], [44], [47].

Two alternatives are proposed to carry out mutations. First, replacing a gene with another random number, i.e., a gene is randomly selected (or $k$ genes in general), and a new random number is created for that position. Second, a simple exchange of two random numbers is performed; two genes are randomly selected, and their values are swapped. This strategy can be adapted so that the selected genes belong to different teams, using the information in $S_k$. Let $F$ represent the fitness function for the problem under investigation, comprising two essential components: the objective function and a penalization component. The selection of individuals considers the direct passage of the $B$ individuals with the best $F$ to the next population, and the elimination of the $W$ worst individuals evaluated in $F$. The rest of the individuals that move on to the next population are obtained from the crossover between individuals (excluding the $W$ worst individuals) through linear rank selection [19], [23], which is focused on avoiding premature convergence.

Using penalty functions is one of the most commonly used strategies for managing constraints in applying GA to constrained optimization problems, [12]. In our case, we partition the constraint set into two parts: basic constraints and climate constraints. The GA algorithm we propose can require that all the solutions it generates comply with the so-called basic constraints. These are linearization, team size, and team membership constraints. For the other group of climate constraints, we use the method of static penalties [43].

**TABLE 6. Descriptive statistics for standardized factor scores.**

| Dimension | Minimum | Maximum | Q1 | Q2 | Q3 |
|---|---|---|---|---|---|
| Team dynamics | -3.3033 | 1.4411 | -0.4992 | 0.2038 | 0.7652 |
| Attitude toward the Team | -4.2846 | 1.5019 | -0.7026 | 0.0553 | 0.9164 |
| Team cohesion | -1.9153 | 1.9416 | -0.9511 | 0.0131 | 0.8256 |
| Team performance | -4.2496 | 1.4845 | -0.7729 | 0.0510 | 0.8749 |

**TABLE 7. Descriptive statistics for standardized scores by treatment type and dimension.**

| Dimension | Treatment | Mean | Std. Dev. | Q1 | Q2 | Q3 |
|---|---|---|---|---|---|---|
| Team dynamics | Link strengthening | 0.1096 | 0.9990 | -0.3074 | 0.2323 | 0.7970 |
| | New links | -0.0890 | 1.0686 | -0.6504 | 0.0212 | 0.6701 |
| | Conflict mediation | -0.1226 | 1.1023 | -0.6788 | 0.0212 | 0.6643 |
| | Control | 0.0794 | 1.0503 | -0.2969 | 0.2550 | 0.7938 |
| Attitude toward the team | Link strengthening | 0.0270 | 1.0622 | -0.7028 | 0.0553 | 1.0539 |
| | New links | -0.0705 | 1.1633 | -0.8058 | 0.0553 | 0.7961 |
| | Conflict mediation | -0.1082 | 1.2881 | -0.8058 | 0.0553 | 1.0539 |
| | Control | 0.0551 | 1.1063 | -0.5302 | 0.0728 | 0.7359 |
| Team cohesion | Link Strengthening | 0.0845 | 1.0682 | -0.5449 | 0.0131 | 0.8635 |
| | New Links | 0.0197 | 1.0982 | -0.9511 | 0.0131 | 0.8501 |
| | Conflict Mediation | -0.1338 | 1.0284 | -0.9511 | 0.0131 | 0.4193 |
| | Control | 0.1095 | 1.1422 | -0.8373 | 0.0131 | 0.9774 |
| Team performance | Link Strengthening | 0.1112 | 1.0932 | -0.5586 | 0.0510 | 0.8749 |
| | New Links | -0.2571 | 1.3779 | -1.3825 | 0.0510 | 0.8749 |
| | Conflict Mediation | -0.0853 | 1.1977 | -0.7729 | 0.0510 | 0.8749 |
| | Control | 0.1123 | 1.1651 | -0.5586 | 0.0510 | 0.8749 |

Computational experiments evaluating this approach are presented in Appendix IX. The DSS module, utilizing this algorithm, was embedded into BRAVE-UP! Platform, prioritizing usability and scalability.

## V. IMPACT AND DISCUSSION

Beyond computational experiments comparing algorithmic performance, we addressed our research questions through a large-scale quasi-experiment designed to assess the impact of our team formation approach on four dimensions: team dynamics, attitude toward the team, team cohesion, and team performance. Note that it can be quite challenging to compete with the control group (self-selection or teacher selection), particularly in aspects such as team dynamics and cohesion. Therefore, we consider it a good result when there are no significant differences, as it indicates that we can improve classroom cohesion without compromising team cohesion.

The intervention consisted of three structured sessions in which students produced a report on the ''Digital school climate''. Guidelines for developing the activity were provided to the teacher responsible for its implementation. A mixed-methods approach was employed, combining surveys with non-participant observation rubrics. The study was conducted in six schools and involved a census-type sample of 880 students in 7th and 8th grades.

### A. QUASI-EXPERIMENT

The sample was divided into two groups of courses: a control group, consisting of teams formed by self-selection or teacher selection, and a treatment group, consisting of teams formed through the DSS. This approach enabled the exploration of effects on:

Team dynamics: Refers to the interactions and processes that take place within the work groups.
Attitude toward the team: refers to each student's attitude regarding the formation of the work groups and the work conducted within them.
Team cohesion: focuses on understanding whether the formation of work groups and the activities carried out promote new ties or relationships among students who do not typically interact.
Team performance: reflects the students' perceptions of both individual and group performance during the news report creation activity.

To ensure comparability across the four groups (control and treatments with $M_1, M_2$ and $M_3$), team compositions were balanced based on three indicators, whose data were obtained from sociograms collected in the classrooms. These indicators are: (1) conflict rate, (2) cohesion rate, and (3) percentage of aggression. Specifically, care was taken to ensure that the four groups had similar means across these three indicators. It is worth noting that, to facilitate the group formation process at the beginning of the activity, the classroom was used as the unit of analysis for treatment assignment. Appendix VIII provides a detailed explanation of this procedure.

#### 1) QUANTITATIVE PHASE

In the quantitative phase, surveys were administered to 693 (out of 880) students in grades 7 and 8, resulting in a representative sample with a 95% confidence level. The analysis included using ANOVA and ANCOVA to assess differences between control and treatment teams regarding team dynamics, attitudes toward the team, team cohesion, and

**TABLE 8. Performance metrics for exact strategies at different instance size when $M_1$ is solved. The table displays statistics for the 'gap' and time (seconds) achieved by each strategies, including mean, minimum, maximum, and standard deviation.**

| n | strategy | gap | | | | time (seconds) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | mean | min | max | std | mean | min | max | std |
| 10 | base-m1-f | 0.00 | 0.00 | 0.00 | 0.00 | 1.91 | 0.41 | 3.85 | 1.26 |
| | base-m1-g | 0.00 | 0.00 | 0.00 | 0.00 | **1.12** | 0.84 | 1.82 | 0.27 |
| | symmetry-m1-f | 0.00 | 0.00 | 0.00 | 0.00 | 1.94 | **0.36** | 3.08 | 1.07 |
| | symmetry-m1-g | 0.00 | 0.00 | 0.00 | 0.00 | 1.22 | 1.00 | **1.45** | **0.16** |
| | validineq-m1-f | 0.00 | 0.00 | 0.00 | 0.00 | 2.11 | 0.48 | 3.27 | 1.09 |
| | validineq-m1-g | 0.00 | 0.00 | 0.00 | 0.00 | 1.39 | 1.05 | 1.79 | 0.20 |
| 15 | base-m1-f | 0.00 | 0.00 | 0.00 | 0.00 | 17.70 | 9.75 | 24.28 | 4.21 |
| | base-m1-g | 0.00 | 0.00 | 0.00 | 0.00 | 11.95 | **7.78** | 16.46 | 2.85 |
| | symmetry-m1-f | 0.00 | 0.00 | 0.00 | 0.00 | 16.22 | 13.00 | 21.37 | 2.78 |
| | symmetry-m1-g | 0.00 | 0.00 | 0.00 | 0.00 | **11.85** | 9.00 | **16.39** | **2.35** |
| | validineq-m1-f | 0.00 | 0.00 | 0.00 | 0.00 | 19.15 | 9.88 | 25.67 | 4.84 |
| | validineq-m1-g | 0.00 | 0.00 | 0.00 | 0.00 | 13.22 | 9.55 | 17.25 | 2.49 |
| 20 | base-m1-f | 6.81 | 0.00 | 51.94 | 16.65 | 446.05 | 0.85 | 900.97 | 355.12 |
| | base-m1-g | 3.67 | 0.00 | 25.48 | 8.43 | 380.65 | **0.71** | 900.88 | 347.14 |
| | symmetry-m1-f | **0.00** | 0.00 | **0.00** | 0.00 | 368.30 | 0.97 | 792.94 | **279.14** |
| | symmetry-m1-g | 2.16 | 0.00 | 21.60 | 6.83 | **357.59** | 0.76 | 900.87 | 303.89 |
| | validineq-m1-f | 5.36 | 0.00 | 53.60 | 16.95 | 382.91 | 1.04 | 901.98 | 300.77 |
| | validineq-m1-g | 12.31 | 0.00 | 75.59 | 24.88 | 408.71 | 0.99 | 901.74 | 357.89 |
| 25 | base-m1 | 64.62 | 47.10 | 89.17 | 14.78 | 902.63 | 902.43 | 903.03 | 0.20 |
| | base-m1-g | **34.10** | **6.24** | 63.71 | 18.26 | 902.33 | 901.60 | 904.51 | 0.79 |
| | symmetry-m1-f | 84.05 | 53.42 | 110.87 | 20.04 | 908.77 | 902.55 | 961.67 | 18.60 |
| | symmetry-m1-g | 40.84 | 10.64 | 88.55 | 25.12 | 902.64 | 902.08 | 903.54 | 0.47 |
| | validineq-m1-f | 68.03 | 48.98 | 82.22 | 11.62 | 908.60 | 902.22 | 961.54 | 18.61 |
| | validineq-m1-g | 39.59 | 27.66 | **54.34** | 8.06 | 902.72 | 902.32 | 903.11 | 0.26 |
| 30 | base-m1 | **95.52** | **67.24** | **124.48** | 16.97 | 905.95 | 904.37 | 909.81 | 1.59 |
| | base-m1-g | 132.18 | 95.17 | 193.19 | 34.16 | 904.92 | 904.04 | 908.29 | 1.32 |
| | symmetry-m1-f | 152.93 | 112.37 | 188.09 | 22.31 | 905.51 | 903.95 | 906.72 | 0.93 |
| | symmetry-m1-g | 215.07 | 171.51 | 252.54 | 23.89 | 905.00 | 903.40 | 907.34 | 1.22 |
| | validineq-m1-f | 96.36 | 69.49 | 131.06 | 21.53 | 904.57 | 903.93 | 906.47 | 0.85 |
| | validineq-m1-g | 106.59 | 69.09 | 163.89 | 31.22 | 904.08 | 903.52 | 904.85 | 0.42 |

team performance. Each of these dimensions was estimated using factor analysis based on an adapted version of the questionnaire presented by [38]. The factorial structure identified during the exploratory phase was validated through confirmatory factor analysis. The maximum likelihood method was used for estimation, along with the fit indices: root mean square error of approximation, comparative fit index and Tucker–Lewis index. When examining the aggregated results in Table 6 across the different team formation methods by dimension, it is observed that team performance shows the greatest variability in scores. In contrast, team cohesion is the dimension with the most concentrated scores and the highest minimum value.

With respect to team dynamics (Table 7 ), it was observed that the control groups and those formed by prioritizing the conflict mediation criterion showed slightly lower minimum values compared to the other treatments. Conversely, the groups formed by prioritizing link strengthening and the control groups exhibited, on average, better dynamics than teams formed by the other methods. This observation is further supported by the fact that these groups also reached the highest values in the first, second, and third quartiles. To verify whether the differences observed between the groups according to their formation method were substantive, an ANOVA test was conducted. The results indicated that, at a 99% confidence level, no statistically significant differences were found. In this regard, previous research suggested that self-selected groups tend to achieve higher scores in social dimensions (dynamics, cohesion, and attitude) than those formed through external assignment [11], [32]. Therefore, the absence of significant differences between the control and treatment groups suggests that the assignment methods based on the tool succeeded in forming teams that enhanced the quality of team dynamics, achieving results comparable to those of the control teams, which typically showed better perceptions in this dimension. Additionally, an analysis of covariance (ANCOVA) was performed to control and adjust the variability associated with conflict rates, cohesion, and the percentage of aggression in the results obtained for each dimension—thus improving the precision and sensitivity of the analysis. The ANCOVA results indicated that team dynamics did not differ significantly between the groups formed by the different methods, supporting the ANOVA findings.

Regarding the participant's attitude toward the team, it was observed that the control groups and those formed by prioritizing the conflict mediation criterion showed slightly lower minimum values than the other treatments. At the same time, the groups formed by prioritizing link strengthening and the control groups recorded, on average, higher scores than teams formed using the other methods. However, the quartile values did not reveal a clear predominance of any particular method in terms of the obtained scores. ANOVA test results indicated that, at a 99% confidence level,

**TABLE 9.** As Table 8 but solving $M_2$.

| n | strategy | gap mean | gap min | gap max | gap std | time (seconds) mean | time (seconds) min | time (seconds) max | time (seconds) std |
|---|---|---|---|---|---|---|---|---|---|
| 10 | base-m2-f | 0.00 | 0.00 | 0.00 | 0.00 | 4.93 | 2.47 | 6.32 | 1.22 |
| | base-m2-g | 0.00 | 0.00 | 0.00 | 0.00 | 3.22 | 1.78 | 4.55 | 0.89 |
| | symmetry-m2-f | 0.00 | 0.00 | 0.00 | 0.00 | 5.14 | 3.40 | 8.11 | 1.30 |
| | symmetry-m2-g | 0.00 | 0.00 | 0.00 | 0.00 | 2.79 | 1.50 | 3.94 | 0.74 |
| | validineq-m2-f | 0.00 | 0.00 | 0.00 | 0.00 | 2.81 | **0.43** | 4.12 | 1.27 |
| | validineq-m2-g | 0.00 | 0.00 | 0.00 | 0.00 | **1.95** | 1.22 | **2.69** | **0.42** |
| 15 | base-m2-f | 9.67 | 0.00 | 27.27 | 10.85 | 674.25 | 324.12 | 901.07 | 259.71 |
| | base-m2-g | 7.27 | 0.00 | 26.30 | 10.04 | 554.44 | 161.70 | 900.69 | 337.41 |
| | symmetry-m2-f | 6.24 | 0.00 | 23.49 | 8.65 | 615.05 | 180.11 | 900.79 | 302.84 |
| | symmetry-m2-g | 0.00 | 0.00 | 0.00 | 0.00 | 342.16 | 118.13 | 813.73 | 230.83 |
| | validineq-m2-f | 0.00 | 0.00 | 0.00 | 0.00 | 74.69 | 27.88 | 124.26 | 29.11 |
| | validineq-m2-g | 0.00 | 0.00 | 0.00 | 0.00 | **33.67** | **21.35** | **58.90** | **10.58** |
| 20 | base-m2-f | 143.24 | 0.00 | 218.61 | 59.42 | 813.95 | 28.89 | 901.48 | 275.84 |
| | base-m2-g | 124.57 | 0.00 | 182.11 | 55.28 | 811.21 | 0.95 | 901.51 | 284.70 |
| | symmetry-m2-f | 152.41 | 0.00 | 242.61 | 67.56 | 811.15 | **0.80** | 901.67 | 284.73 |
| | symmetry-m2-g | 111.03 | 0.00 | 173.27 | 57.44 | 811.26 | 1.05 | 901.82 | 284.68 |
| | validineq-m2-f | **8.76** | 0.00 | **22.90** | **7.97** | **771.87** | 1.10 | 902.19 | 292.40 |
| | validineq-m2-g | 26.85 | 0.00 | 37.70 | 14.71 | 803.35 | 0.93 | 901.04 | **282.94** |
| 25 | base-m2-f | 253.82 | 211.39 | 294.52 | 24.21 | 902.72 | 901.79 | 903.73 | 0.52 |
| | base-m2-g | 393.42 | 302.02 | 460.46 | 62.93 | 904.81 | 903.50 | 906.64 | 1.11 |
| | symmetry-m2-f | 256.90 | 208.63 | 294.75 | 25.21 | 902.68 | 901.75 | 903.33 | 0.45 |
| | symmetry-m2-g | 382.49 | 313.85 | 508.89 | 80.75 | 905.34 | 903.81 | 908.06 | 1.47 |
| | validineq-m2-f | 31.19 | 22.56 | 44.59 | 7.42 | 902.53 | 901.96 | 903.15 | 0.39 |
| | validineq-m2-g | **25.53** | **20.15** | **37.46** | **4.69** | 903.03 | 902.51 | 903.89 | 0.42 |
| 30 | base-m2-f | 324.57 | 279.33 | 367.73 | 25.47 | 906.61 | 905.34 | 908.83 | 1.38 |
| | base-m2-g | 636.85 | 511.44 | 706.38 | 62.25 | 907.68 | 905.51 | 912.09 | 2.45 |
| | symmetry-m2-f | 322.80 | 290.25 | 356.71 | 21.72 | 905.02 | 903.79 | 906.18 | 0.99 |
| | symmetry-m2-g | 688.57 | 660.42 | 713.71 | 16.95 | 905.92 | 904.13 | 906.80 | 0.84 |
| | validineq-m2-f | **37.83** | **24.97** | 51.29 | 8.08 | 904.10 | 903.47 | 905.21 | 0.63 |
| | validineq-m2-g | 43.40 | 30.69 | **50.55** | **7.81** | 904.24 | 903.61 | 906.05 | 0.75 |

no statistically significant differences were found. These findings suggest that the team assignment methods based on the tool successfully formed groups in which students demonstrated attitudes as positive as those who selected their own teammates. Moreover, ANCOVA results for this dimension also revealed no significant differences among the groups, and the covariates showed no significant effects on the dependent variable.

In team cohesion dimension, the control groups obtained, on average, higher scores than teams formed by other methods. Likewise, groups formed with our DSS prioritizing link strengthening also exhibited high mean and quartile values compared with the remaining methods. However, ANOVA results indicated that, at a 99% confidence level, there were no statistically significant differences. Thus, the findings suggest that the DSS-based methods succeeded in forming groups where students displayed levels of team cohesion comparable to those of self-selected teams. The ANCOVA test also showed no significant differences or effects of the covariates.

Similar to other dimensions, in team performance dimension both the control groups and those formed using the DSS with link strengthening prioritization recorded, on average, higher scores than the other treatments. In this regard, ANOVA results at a 99% confidence level suggested the existence of statistically significant differences in the perceived quality of team performance among the groups formed by the different methods. However, the ANCOVA test indicated no significant differences in performance between groups, and the covariates had no significant effects on the dependent variable. Specifically, the control groups and those formed by prioritizing link strengthening achieved significantly better performance than those formed by prioritizing new link creation. Although prior evidence [38] suggests that academic performance tends to be higher in groups formed by instructors than in self-selected ones, it is worth noting that in this study, performance was measured based on students' perceptions rather than grades, making it difficult to establish a standardized view of team outcomes.

The main insights obtained from this quantitative analysis are presented below:

Team dynamics: The analysis revealed no statistically significant differences between the control and treatment team in terms of team dynamics, with a 99% confidence level. However, teams formed through conflict mediation showed slightly lower scores than those formed through link strengthening and self-selection. The latter teams showed notable improvements in team dynamics, achieving higher values in performance quartiles.

Attitude towards the team: Regarding attitude towards the team, no significant differences were found between the teams. However, control groups and those formed through link strengthening showed more positive attitudes, especially in terms of collaboration and enthusiasm during sessions. Teams focused on new links exhibited higher levels of

**TABLE 10.** As Table 8 but solving $M_3$.

| n | strategy | gap mean | gap min | gap max | gap std | time (seconds) mean | time min | time max | time std |
|---|---|---|---|---|---|---|---|---|---|
| 10 | base-m3-f | 0.00 | 0.00 | 0.00 | 0.00 | 0.85 | 0.22 | 2.95 | 0.89 |
| | base-m3-g | 0.00 | 0.00 | 0.00 | 0.00 | **0.64** | 0.13 | **1.36** | **0.47** |
| | symmetry-m3-f | 0.00 | 0.00 | 0.00 | 0.00 | 1.47 | 0.39 | 2.50 | 0.79 |
| | symmetry-m3-g | 0.00 | 0.00 | 0.00 | 0.00 | **0.64** | 0.11 | 1.42 | 0.58 |
| | validineq-m3-f | 0.00 | 0.00 | 0.00 | 0.00 | 1.92 | 0.20 | 3.66 | 1.30 |
| | validineq-m3-g | 0.00 | 0.00 | 0.00 | 0.00 | 0.77 | **0.10** | 1.95 | 0.73 |
| 15 | base-m3-f | 0.00 | 0.00 | 0.00 | 0.00 | 12.74 | **1.44** | 21.10 | 5.43 |
| | base-m3-g | 0.00 | 0.00 | 0.00 | 0.00 | **6.36** | 3.37 | **8.77** | **1.71** |
| | symmetry-m3-f | 0.00 | 0.00 | 0.00 | 0.00 | 14.91 | 6.29 | 22.95 | 5.69 |
| | symmetry-m3-g | 0.00 | 0.00 | 0.00 | 0.00 | 7.22 | 3.19 | 12.46 | 2.89 |
| | validineq-m3-f | 0.00 | 0.00 | 0.00 | 0.00 | 11.97 | 6.46 | 17.47 | 4.07 |
| | validineq-m3-g | 0.00 | 0.00 | 0.00 | 0.00 | 7.04 | 1.88 | 12.58 | 2.72 |
| 20 | base-m3-f | 0.00 | 0.00 | 0.00 | 0.00 | **262.58** | 0.99 | 809.71 | **250.71** |
| | base-m3-g | 7.03 | 0.00 | 70.32 | 22.24 | 321.66 | **0.72** | 900.88 | 270.13 |
| | symmetry-m3-f | 0.00 | 0.00 | 0.00 | 0.00 | 293.41 | 0.91 | 819.49 | 285.78 |
| | symmetry-m3-g | 0.00 | 0.00 | 0.00 | 0.00 | 291.76 | 0.79 | **710.32** | 252.41 |
| | validineq-m3-f | 19.58 | 0.00 | 195.79 | 61.92 | 277.94 | 0.94 | 903.01 | 254.33 |
| | validineq-m3-g | 7.96 | 0.00 | 79.56 | 25.16 | 325.67 | 0.89 | 901.63 | 281.58 |
| 25 | base-m3-f | 128.91 | 53.55 | 195.08 | **42.82** | 903.37 | 902.64 | 904.27 | 0.63 |
| | base-m3-g | 109.48 | 17.77 | 371.63 | 110.23 | 903.21 | 901.85 | 904.79 | 1.04 |
| | symmetry-m3-f | 123.25 | **0.00** | 200.67 | 58.21 | 897.01 | 833.26 | 905.42 | 22.41 |
| | symmetry-m3-g | 127.56 | 13.55 | 479.92 | 134.56 | 904.64 | 903.00 | 910.52 | 2.26 |
| | validineq-m3-f | 132.22 | **0.00** | 198.89 | 65.32 | 896.41 | 830.42 | 904.96 | 23.20 |
| | validineq-m3-g | **102.43** | 55.81 | **173.96** | 44.47 | 904.23 | 903.22 | 905.74 | 0.78 |
| 30 | base-m3-f | 250.01 | 153.29 | 739.76 | 185.82 | 907.88 | 906.15 | 910.84 | 1.44 |
| | base-m3-g | 540.96 | 350.96 | 871.60 | 154.56 | 906.55 | 904.80 | 908.71 | 1.32 |
| | symmetry-m3-f | 263.74 | 219.31 | 431.90 | 64.16 | 907.43 | 904.44 | 911.70 | 2.19 |
| | symmetry-m3-g | 813.43 | 424.75 | 1550.98 | 436.26 | 907.90 | 906.06 | 909.70 | 1.31 |
| | validineq-m3-f | **247.95** | **152.54** | **326.12** | **47.30** | 906.90 | 905.43 | 912.31 | 2.11 |
| | validineq-m3-g | 720.26 | 435.87 | 1186.37 | 258.85 | 908.15 | 905.25 | 912.45 | 2.44 |

collaboration, while those formed under conflict mediation had lower enthusiasm scores.

Team cohesion: In terms of team cohesion, the results also indicated no significant differences between teams, although the control groups and those formed via the DSS with a focus on link strengthening had better scores. An interesting finding was that control groups and those formed through conflict mediation showed higher percentages of agreement with the statement, "I made new friends in the team assigned for the activity". This suggests that these methods favored the creation of new social ties.

#### a: TEAM PERFORMANCE

In analyzing team performance, control groups and those in link-strengthening conditions showed higher average scores. In terms of performance, significant differences were observed between DSS teams that prioritized link strengthening and those that prioritized new links. However, ANCOVA did not detect significant differences in performance between the control and treatment teams, suggesting that the students' perceptions of performance were similar, regardless of the method used.

To sum up, the results of this large-scale quasi-experiment provide encouraging evidence that algorithmic team allocation methods can achieve levels of classroom cohesion, teamwork attitudes, and performance comparable to those of self- or teacher-assigned groups. This finding is particularly valuable given the inherent complexity of school environments, where social relationships, classroom dynamics, and teacher practices interact in nontrivial ways. The absence of significant differences between control and treatment groups suggests that the proposed approach can promote inclusive and balanced teams without compromising the natural dynamics of classroom collaboration. Nevertheless, obtaining stronger statistical evidence requires addressing several methodological challenges. Future studies should incorporate larger and more diverse samples across schools to increase statistical power, apply longitudinal designs to capture temporal effects, and employ multilevel models capable of distinguishing within- and between-classroom variance. In addition, improving measurement reliability through refined psychometric instruments and reducing potential biases from non-random assignment or missing data will further enhance the validity of the results. Overall, these findings support the feasibility and potential of algorithmic DSS in education, while underscoring the importance of rigorous statistical design to consolidate robust, generalizable evidence in real-world school settings.

### 2) QUALITATIVE PHASE

The qualitative phase focused on non-participant observation, using specific rubrics to evaluate team dynamics, attitude towards the team, and results in the first and third sessions. The groups were observed under the same categories used in the quantitative analysis, allowing for the identification of emerging patterns in student interactions over time.

**TABLE 11.** Summary statistics related to the relative optimality gap and execution time using the RKGA are presented below.

| Model | Size | Gap | | | Time (seconds) | | |
|---|---|---|---|---|---|---|---|
| | | Min | Mean | Max | Min | Mean | Max |
| | 10 | 0.00 | 0.00 | 0.00 | 7.00 | 7.13 | 7.38 |
| | 15 | 0.00 | 3.50 | 15.78 | 7.96 | 8.21 | 10.08 |
| $M_1$ | 20 | 0.00 | 5.75 | 20.39 | 11.14 | 12.19 | 14.79 |
| | 25 | 0.00 | 7.63 | 23.40 | 22.95 | 25.36 | 27.19 |
| | 30 | -2.09 | 8.68 | 24.45 | 27.30 | 30.39 | 31.82 |
| | 10 | 0.00 | 0.00 | 0.00 | 6.99 | 7.13 | 7.41 |
| | 15 | 0.00 | 0.14 | 0.95 | 7.85 | 8.10 | 10.06 |
| $M_2$ | 20 | -0.09 | 0.15 | 1.51 | 11.41 | 12.24 | 14.96 |
| | 25 | -0.92 | -0.75 | 0.02 | 22.77 | 25.05 | 26.70 |
| | 30 | -2.28 | -1.97 | -0.68 | 27.37 | 30.42 | 31.69 |
| | 10 | 0.00 | -0.00 | 0.00 | 7.03 | 7.16 | 7.40 |
| | 15 | 0.00 | 2.89 | 21.57 | 7.84 | 8.02 | 9.85 |
| $M_3$ | 20 | 0.00 | 8.58 | 32.29 | 11.71 | 12.90 | 15.59 |
| | 25 | 0.00 | 13.12 | 41.15 | 22.29 | 24.24 | 25.43 |
| | 30 | -2.42 | 15.77 | 45.41 | 27.56 | 30.62 | 31.94 |

**TABLE 12.** Formation method algorithm.

| | Group | Conflict | Cohesion | Aggression |
|---|---|---|---|---|
| **Control** | Self-selection and/or teacher selection | 0.016 | 0.030 | 11.24 |
| **Treatment** | DSS: $Z_1$ | 0.013 | 0.029 | 10.45 |
| | DSS: $Z_2$ | 0.015 | 0.029 | 11.24 |
| | DSS: $Z_3$ | 0.015 | 0.030 | 11.39 |

Team dynamics: Regarding team dynamics, control groups showed more disorganized communication in the first session, with some teams successfully resolving disagreements constructively. In the third session, all control groups successfully resolved disagreements, although active listening regressed in some cases due to distractions and the teacher's continuous intervention to maintain order. Model-based teams focused on link strengthening showed improvements in assertive communication and disagreement resolution, with most courses achieving better fluidity in interactions.

Attitude towards the Team: this dimension was positive in most treatment and control groups, with some teams showing resistance in the first session due to the assignment of less familiar peers. However, as the activity progressed, most students maintained a positive attitude, particularly in the link-strengthening teams. In contrast, teams focused on conflict mediation showed somewhat more negative attitudes than those using the other methods.

### B. DISCUSSION

In terms of results, control groups, such as those formed through conflict mediation, were more successful in creating new connections, although not necessarily at a deep level. In DSS teams, especially those focusing on link strengthening, students improved the quality of their interactions and got to know peers with whom they usually did not collaborate.

The quantitative analysis did not reveal significant differences between the control and treatment groups in team dynamics, cohesion, or attitude. However, teams focusing on link strengthening showed slight improvements in performance and cohesion. The qualitative results, on the other hand, indicated a positive evolution in team dynamics and attitudes towards assigned teams, particularly in model-based teams, which prioritized conflict mediation and link strengthening.

No significant differences in performance were observed in the quantitative analysis, and the qualitative findings suggest that the model-based intervention, particularly those focused on link strengthening and conflict mediation, positively impacted school cohesion and the quality of interactions within the teams. These results underscore the importance of considering factors such as cohesion, team dynamics, and attitudes towards collaborative work when designing team formation methods in educational settings. Hence, classroom cohesion can be improved without compromising desirable team characteristics; even more so, in some cases, we can improve both by forming an appropriate team.

It is important to emphasize that the success of the intervention depends on two key components: the quality of the team formation and the effective planning of the collaborative activity. Both elements must be carefully aligned to foster meaningful student participation and achieve the desired climate outcomes. Furthermore, using a recent sociogram when designing the teams is essential, as student relationships can be highly volatile, particularly at certain educational levels. Sociometric data collected before the intervention may no longer accurately reflect the actual structure of the social network, leading to suboptimal or even counterproductive team configurations.

The integration of algorithmic systems into educational contexts entails ethical and pedagogical challenges. First, student privacy must be guaranteed through rigorous data protection protocols and anonymization, given the sensitive nature of sociometric information on friendship, rejection, and victimization. Second, interpretability is essential for pedagogical transparency. Teachers must understand the rationale behind algorithmic recommendations, which should be presented in an interpretable, traceable way, allowing educators to combine these insights with their professional judgment. Finally, teacher acceptance depends on trust and perceived usefulness. Initial skepticism often shifts to acceptance once teachers observe how the system promotes balanced, inclusive teams. Thus, algorithmic team formation should support, not replace, pedagogical expertise.

## VI. CONCLUSION

This research focuses on an aspect closely related to effective learning: the school climate. The literature highlights the importance of student relationships in shaping the school climate, as these relationships contribute significantly to educational success. Therefore, learning technologies that foster improved student relationships are essential.

We introduced a new technology that incorporates sociometric data and optimization models to support team formation in educational environments. Unlike previous approaches to the TFP, the proposed system explicitly integrates variables related to school climate, such as interpersonal acceptance and rejection, diversity, and the protection of vulnerable students. It has been tested in real educational settings, with 880 students in 26 classrooms. The results demonstrate that classroom cohesion can be improved without compromising team-related dimensions—and that, in some cases, both aspects can be enhanced simultaneously. These findings suggest that the proposed approach fosters a favorable classroom climate and promotes inclusive, collaborative learning, paving the way for improved outcomes when new learning technologies are applied.

The technology has been integrated into the BRAVE-UP! Platform, offering schools a practical and data-driven tool for forming student teams based on formal criteria. It provides an alternative to subjective or arbitrary grouping methods by incorporating climate-oriented constraints and objective functions.

Future research is expected to explore dynamic extensions of the optimization problem, enabling team configurations to adjust as classroom conditions evolve. A promising direction involves applying multi-objective optimization, where suitable solutions from the Pareto frontier can be selected through data-driven or expert-knowledge-based mechanisms aligned with classroom profiles and institutional objectives. Additional efforts may explore adaptive mechanisms that incorporate longitudinal data and user feedback. Finally, an expanded evaluation framework is planned to assess the long-term impact of the DSS on educational outcomes and student well-being.

## APPENDIX

### A. LINEARIZATION STRATEGIES

It is necessary to consider the linearization of function objectives to solve the problem with commercial solvers for mixed-integer linear programming. Two linearization approaches are implemented,

*Fortet:* Constraints 9 and 10 help linearize $x_{ik}x_{jk}$ through the creation of a new variable $y_{ijk}$. Then, constraints 11 and 12 help to linearize $x_{ik}x_{jk}x_{hk}$, which is equal to $y_{ihk}y_{jhk}$, through $z_{ijhk}$.

$$(x_{ik} + x_{jk} - 2y_{ijk}) > 0, \quad i, j \in N : i < j,\ k \in K \tag{9}$$
$$(x_{ik} + x_{jk} - y_{ijk}) < 1, \quad i, j \in N : i < j,\ k \in K \tag{10}$$
$$(y_{ihk} + y_{jhk} - 2z_{ijhk}) > 0, \quad (i, j, h) \in M,\ k \in K \tag{11}$$
$$(y_{ihk} + y_{jhk} - z_{ijhk}) < 1, \quad (i, j, h) \in M,\ k \in K \tag{12}$$

where $y_{ijk}$ is a binary variable that takes the value 1 if student $i$ and student $j$ are assigned to team $k$, and 0 otherwise, $i < j$; and $z_{ijhk}$ is also binary variable that takes the value 1 if student $i$, student $j$, and student $h$ are assigned to team $k$, and 0 otherwise, $(i, j, h) \in M$.

*Groove:* In the linearization proposed by Grove, the variables $y$ and $z$ are relaxed to continuous and positive, as they are bounded by the binary variables $x$. Sets 13-17 model quadratic non-linearity, and constraints 18-20 model cubic non-linearity.

$$y_{ijk} - x_{ik} \leq 0, \quad i, j \in N : i < j,\ k \in K \tag{13}$$
$$y_{ijk} - x_{jk} \leq 0, \quad i, j \in N : i < j,\ k \in K \tag{14}$$
$$y_{ijk} - x_{ik} - x_{jk} \geq -1, \quad i, j \in N : i < j,\ k \in K \tag{15}$$
$$y_{ijk} - y_{jik} = 0, \quad i, j \in N : i > j,\ k \in K, \tag{16}$$
$$y_{ijk} = 0, \quad i, j \in N : i = j,\ k \in K \tag{17}$$
$$z_{khij} - y_{kjh} \leq 0, \quad (i, j, h) \in M,\ k \in K \tag{18}$$
$$z_{khij} - y_{kih} \leq 0, \quad (i, j, h) \in M,\ k \in K \tag{19}$$
$$z_{khij} - y_{kjh} - y_{kih} + 1 \geq 0, \quad (i, j, h) \in M,\ k \in K \tag{20}$$

### B. SYMMETRY-BREAKING CONSTRAINTS

A clear symmetry is observed in the group index within the context of the group formation problem. Two different solutions that are identical in terms of objective function values can only be obtained by changing the group index $(k)$ between two groups $k$, $k'$ with same size, it is to say, $S_k == S_{k'}$. To address this, we propose a symmetry-breaking constraint set aimed at introducing differentiation in the order of this index, as outlined below:

$$\sum_{i \in N} i\, x_{ik} \leq \sum_{i \in N} i\, x_{ik'}, \ \forall k, k' \in K : k < k',\ S_k = S_{k'} \tag{21}$$

This constraint ensures that, among groups of identical size, groups with lower indices are assigned students with lower aggregate index values, thereby eliminating symmetric solutions without excluding any optimal solution.

### C. VALID INEQUALITIES

In order to provide a better formulation we propose the following valid inequalities:

$$\sum_{k \in K} y_{ijk} \leq 1, \quad \forall i, j \in N, \tag{22}$$
$$\sum_{k \in K} y_{ijk} \leq S_k(S_k - 1), \quad \forall k \in K, i, j \in N, \tag{23}$$

Equation 22 restricts the number of interactions between students $i$ and $j$, ensuring that they are assigned to at most one group. Although this constraint is not obligatory, it remains valid given the requirement of one-group assignment for each student (see constraint 2). Similarly, Constraint 23 limits the total number of interactions within each group, represented by the summation of $y_{ijk}$ over all students $i$ and $j$ in group $k$.

### D. EXPERIMENTS WITH EXACT METHODS

Tables 8, 9, and 10 present computational results for the three single-objective models, namely $M_1$, $M_2$, and $M_3$, respectively. Each table presents results for different sets of instances, and each row represents a set of ten instances. The parameter $n$ indicates the number of students considered in each family. The tables compare six approaches to solving the models. The prefix ''base'' denotes a model that includes only linearizations. The suffix ''f'' indicates the linearization provided by Fortet, and the suffix ''g'' denotes the linearization proposed by Groove. The prefixes "symmetry" and "validineq" refer to the inclusion of symmetry-breaking constraints and valid inequalities proposed in Section III-B.

From a general perspective, we observe in Tables 8, 9, and 10 that all exact approaches are not practical when instances consider 30 students. The smallest average gap is 37%, but some strategies have gaps of up to 600%. From our perspective, the most notable contributions are observed in Table 9, where both breaking-symmetry constraints and valid inequalities significantly reduce running times and gaps, sometimes reducing the gap by up to one order of magnitude for instances with 20, 25, and 30 students. Some optimal solutions were found for instances that were not solved by the base approaches.

Despite these improvements, in Tables 8 and 9, a dominant strategy is not evident, with a performance difference that is not very significant. However, in Table 10, we observe slightly better performance in approaches that include valid inequalities. Finally, it is important to highlight that, according to our experiments, the superiority of Groove linearization over Fortet linearization is not clear in this model.

## VII. EXPERIMENTS WITH RKGA

Table 11 presents summary statistics for the relative optimality gap and execution time obtained using RKGA. Each row corresponds to ten problem instances, with 100 experiments conducted for each instance. The performance of the RKGA varies depending on the objective function. For Objective Function 1, an average gap of 8% was observed. Objective Function 2, which is the easiest to solve, consistently achieved a relative optimality gap below 2% across all instances. In the case of Objective Function 3, the average gap reached 15.77%. In general, when the RKGA was limited to a run time of less than 30 seconds, the average gap remained below 15.77%, which is considered acceptable for this application.

## VIII. CONTROL/TREATMENT FORMATION

Each participating class was assigned a group: either the control condition or selection via DSS. For the latter, three prioritization criteria were proposed for assigning students to groups: (1) the formation of new ties between students, (2) conflict mediation, and (3) the strengthening of pre-existing positive ties. Accordingly, each class was assigned one of the four available group formation alternatives (Table 9). To ensure comparability among the four groups, the assignment of group formation mechanisms for the activity was balanced based on three indicators, which were derived from sociograms administered in the classes. These indicators were: (1) conflict rate, (2) cohesion rate, and (3) percentage of aggression. Specifically, care was taken to ensure that the four groups had similar averages across the three indicators. It is important to note that, to facilitate the implementation of team formation at the beginning of the activity, the class was used as the unit of analysis for assigning treatments. As a result of the final assignment, seven groups were formed according to the control criterion, six using the DSS tool prioritizing the strengthening of ties, six using DSS prioritizing the formation of new ties, and six using DSS prioritizing conflict mediation. Table 9 presents the average values for each group according to the assigned group formation method. An ANOVA test was conducted to detect differences between the groups for each indicator. The results showed no statistically significant differences at the 99% confidence level: the values obtained were 1.092 (F-statistic) and 0.362 (p-value) for conflict rate, 0.306 (F-statistic) and 0.999 (p-value) for cohesion rate, and 0.235 (F-statistic) and 1.000 (p-value) for percentage of aggression.

## ACKNOWLEDGMENT

The authors sincerely thank the anonymous reviewers for their insightful comments and constructive suggestions, which helped improve the clarity and quality of this work. They also gratefully acknowledge the valuable contributions of José Massaro, Oscar Acevedo, Bárbara Cáceres, Álvaro Carrasco, and Felipe Zanoni to this project. In addition, they acknowledge CORFO for its support and commitment to promoting innovation and development in Chile.

## ETHICAL APPROVAL

All procedures were carried out in accordance with the Declaration of Helsinki, and informed consent was obtained from all participants. Formal approval was also obtained from the participating schools. All analyses were conducted using anonymized data, and the mathematical models were developed to support teachers in making informed decisions.

## REFERENCES

[1] R. Agrawal, B. Golshan, and E. Terzi, "Grouping students in educational settings," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. discovery data mining*, Aug. 2014, pp. 1017–1026.

[2] E. Alhazmi, S. Horawalavithana, J. Skvoretz, J. Blackburn, and A. Iamnitchi, "An empirical study on team formation in online games," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Jul. 2017, pp. 431–438.

[3] Z. C. Ani, A. Yasin, M. Z. Husin, and Z. A. Hamid, "A method for group formation using genetic algorithm," *Int. J. Comput. Sci. Eng.*, vol. 2, pp. 3060–3064, Aug. 2010.

[4] V. Batagelj and A. Mrvar, "A subquadratic triad census algorithm for large sparse networks with small maximum degree," *Social Netw.*, vol. 23, no. 3, pp. 237–243, Jul. 2001.

[5] J. C. Bean, "Genetic algorithms and random keys for sequencing and optimization," *ORSA J. Comput.*, vol. 6, no. 2, pp. 154–160, May 1994.

[6] N. Berktaş and H. Yaman, "A branch-and-bound algorithm for team formation on social networks," *INFORMS J. Comput.*, vol. 33, no. 3, pp. 1162–1176, Jul. 2021.

[7] U. Brandes, "A faster algorithm for betweenness centrality," *J. Mathematical Sociol.*, vol. 25, no. 2, pp. 163–177, 2001.

[8] U. Brandes, "On variants of shortest-path betweenness centrality and their generic computation," *Social Netw.*, vol. 30, no. 2, pp. 136–145, May 2008.

[9] E. Burns, R. Mann, and C. Yanez, "Student reports of bullying: Results from the 2019 school crime supplement to the national crime victimization survey," National Center for Education Statistics, Washington, DC, USA, Tech. Rep. NCES 2022-031, 2022. [Online]. Available: https://eric.ed.gov/?id=ED622371

[10] M. Campelo and T. F. Figueiredo, "Integer programming approaches to the multiple team formation problem," *Comput. Operations Res.*, vol. 133, Sep. 2021, Art. no. 105354.

[11] K. J. Chapman, M. Meuter, D. Toy, and L. Wright, "Can't we Pick our own groups? The influence of group selection method on group dynamics and outcomes," *J. Manage. Educ.*, vol. 30, no. 4, pp. 557–569, Aug. 2006.

[12] C. A. C. Coello, "Constraint-handling techniques used with evolutionary algorithms," in *Proc. Genetic Evol. Comput. Conf. Companion*, Jul. 2022, pp. 1310–1333.

[13] W. H. El-Ashmawi, A. F. Ali, and M. A. Tawhid, "An improved particle swarm optimization with a new swap operator for team formation problem," *J. Ind. Eng. Int.*, vol. 15, no. 1, pp. 53–71, Mar. 2019.

[14] A. Farasat and A. G. Nikolaev, "Social structure optimization in team formation," *Comput. Oper. Res.*, vol. 74, pp. 127–142, Oct. 2016.

[15] M. Fathian, M. Saei-Shahi, and A. Makui, "A new optimization model for reliable team formation problem considering Experts' collaboration network," *IEEE Trans. Eng. Manag.*, vol. 64, no. 4, pp. 586–593, Nov. 2017.

[16] L. Feola and V. Trianni, "Adaptive strategies for team formation in minimalist robot swarms," *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 4079–4085, Apr. 2022.

[17] E. L. Fitzpatrick and R. G. Askin, "Forming effective worker teams with multi-functional skill requirements," *Comput. Ind. Eng.*, vol. 48, no. 3, pp. 593–608, May 2005.

[18] S. Garshasbi, Y. Mohammadi, S. Graf, S. Garshasbi, and J. Shen, "Optimal learning group formation: A multi-objective heuristic search strategy for enhancing inter-group homogeneity and intra-group heterogeneity," *Expert Syst. Appl.*, vol. 118, pp. 506–521, Mar. 2019.

[19] D. E. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," in *Foundations of Genetic Algorithms*, 1991, pp. 69–93.

[20] C. Gonzálvez, V. Bacon, and C. A. Kearney, "Systematic and evaluative review of school climate instruments for students, teachers, and parents," *Psychol. Schools*, vol. 60, no. 6, pp. 1781–1836, Jun. 2023.

[21] J. H. Gutiérrez, C. A. Astudillo, P. Ballesteros-Pérez, D. Mora-Melià, and A. Candia-Véjar, "The multiple team formation problem using sociometry," *Comput. Oper. Res.*, vol. 75, pp. 150–162, Nov. 2016.

[22] A. Hill and S. Peuker, "Expanding students' social networks via optimized team assignments," *Ann. Operations Res.*, vol. 332, nos. 1–3, pp. 1107–1131, Jan. 2024.

[23] K. Jebari and M. Madiafi, "Selection methods for genetic algorithms," *Int. J. Emerg. Sci.*, vol. 3, pp. 333–344, Sep. 2013.

[24] J. Juárez, C. Santos, and C. A. Brizuela, "A comprehensive review and a taxonomy proposal of team formation problems," *ACM Comput. Surv.*, vol. 54, no. 7, pp. 1–33, Sep. 2022.

[25] A. Krouska and M. Virvou, "An enhanced genetic algorithm for heterogeneous group formation based on multi-characteristics in social-networking-based learning," *IEEE Trans. Learn. Technol.*, vol. 13, no. 3, pp. 465–476, Jul. 2020.

[26] T. Lappas, K. Liu, and E. Terzi, "Finding a team of experts in social networks," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. discovery data mining*, Jun. 2009, pp. 467–476.

[27] B. P. Leung and J. Silberling, "Using sociograms to identify social status in the classroom," *California School Psychologist*, vol. 11, no. 1, pp. 57–61, Jan. 2006.

[28] S. M. Lim, A. B. M. Sultan, M. N. Sulaiman, A. Mustapha, and K. Y. Leong, "Crossover and mutation operators of genetic algorithms," *Int. J. Mach. Learn. Comput.*, vol. 7, no. 1, pp. 9–12, Feb. 2017.

[29] Y.-S. Lin, Y.-C. Chang, and C.-P. Chu, "Novel approach to facilitating tradeoff multi-objective grouping optimization," *IEEE Trans. Learn. Technol.*, vol. 9, no. 2, pp. 107–119, Apr. 2016.

[30] B. Lv, J. Jiang, L. Wu, and H. Zhao, "Team formation in large organizations: A deep reinforcement learning approach," *Decis. Support Syst.*, vol. 187, Dec. 2024, Art. no. 114343.

[31] T. L. Magnanti and K. Natarajan, "Allocating students to multidisciplinary capstone projects using discrete optimization," *Interfaces*, vol. 48, no. 3, pp. 204–216, Jun. 2018.

[32] S. Mahenthiran and P. J. Rouse, "The impact of group selection on student performance and satisfaction," *Int. J. Educ. Manage.*, vol. 14, no. 6, pp. 255–265, Nov. 2000.

[33] P. B. C. Miranda, R. F. Mello, and A. C. A. Nascimento, "A multi-objective optimization approach for the group formation problem," *Expert Syst. Appl.*, vol. 162, Dec. 2020, Art. no. 113828.

[34] J. Moreno, D. A. Ovalle, and R. M. Vicari, "A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics," *Comput. Educ.*, vol. 58, no. 1, pp. 560–569, Jan. 2012.

[35] J. Moreno, J. D. Sánchez, and A. F. Pineda, "A hybrid approach for composing groups in collaborative learning contexts," *Heliyon*, vol. 7, no. 6, Jun. 2021, Art. no. e07249.

[36] J. L. Moreno, "Sociometry and the cultural order," *Sociometry*, vol. 6, no. 3, p. 299, Aug. 1943.

[37] M. Muniz and T. Flamand, "A column generation approach for the team formation problem," *Comput. Oper. Res.*, vol. 161, Jan. 2024, Art. no. 106406.

[38] M. L. Post, A. Barrett, M. Williams, and L. Scharff, "Impact of team formation method on student performance, attitudes, and behaviors," *J. Scholarship Teach. Learn.*, vol. 20, no. 1, pp. 1–21, Feb. 2020.

[39] M. Z. Rehman, K. Z. Zamli, M. Almutairi, H. Chiroma, M. Aamir, M. A. Kader, and N. M. Nawi, "A novel state space reduction algorithm for team formation in social networks," *PLoS ONE*, vol. 16, no. 12, Dec. 2021, Art. no. e0259786.

[40] M. Rytioja, K. Lappalainen, and H. Savolainen, "Behavioural and emotional strengths of sociometrically popular, rejected, controversial, neglected, and average children," *Eur. J. Special Needs Educ.*, vol. 34, no. 5, pp. 557–571, Oct. 2019.

[41] H. Sadeghi and A. A. Kardan, "A novel justice-based linear model for optimal learner group formation in computer-supported collaborative learning environments," *Comput. Hum. Behav.*, vol. 48, pp. 436–447, Jul. 2015.

[42] J. Shindler, A. Jones, A. D. Williams, C. Taylor, and H. Cardenas, "The school climate–student achievement connection: If we want achievement gains, we need to begin by improving the climate," *J. School Admin. Res. Develop.*, vol. 1, no. 1, pp. 9–16, May 2016.

[43] A. E. Smith, D. W. Coit, T. Baeck, D. Fogel, and Z. Michalewicz, "Penalty functions," in *Handbook of Evolutionary Computation*, vol. 97. Oxford Univ. Press, 1997.

[44] W. Spears and K. D. Jong, "On the virtues of parameterised uniform crossover," in *Proc. ICGA*, 1991, pp. 230–236.

[45] S. Tawil and R. Locatelli. (2015). *Rethinking Education: Towards a Global Common Good.* [Online]. Available: https://www.norrag.org/rethinkingeducation-towardsa-global-common-good

[46] W. Tomaszewski, N. Xiang, and Y. Huang, "School climate, student engagement and academic achievement across school sectors in Australia," *Austral. Educ. Researcher*, vol. 51, no. 2, pp. 667–695, Apr. 2024.

[47] A. J. Umbarkar and P. D. Sheth, "Crossover operators in genetic algorithms: A review," *ICTACT J. Soft Computing*, vol. 6, no. 1, pp. 1–15, Oct. 2015.

[48] T. Vallès-Catala and R. Palau, "Minimum entropy collaborative groupings: A tool for an automatic heterogeneous learning group formation," *PLoS ONE*, vol. 18, no. 3, Mar. 2023, Art. no. e0280604.

[49] X. Wang, G. Song, and R. Ghannam, "Enhancing teamwork and collaboration: A systematic review of algorithm-supported pedagogical methods," *Educ. Sci.*, vol. 14, no. 6, p. 675, Jun. 2024.

[50] Y. Wu, X. Xie, J. Wang, D. Deng, H. Liang, H. Zhang, S. Cheng, and W. Chen, "ForVizor: Visualizing spatio-temporal team formations in soccer," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 1, pp. 65–75, Jan. 2019.

[51] Y. Zheng, C. Li, S. Liu, and W. Lu, "An improved genetic approach for composing optimal collaborative learning groups," *Knowl.-Based Syst.*, vol. 139, pp. 214–225, Jan. 2018.


**ALFREDO CANDIA-VÉJAR** received the bachelor's degree in mathematics from the University of Concepción, Chile, and the M.Sc. and Ph.D. degrees in computer science and systems engineering from the Federal University of Rio de Janeiro, Brazil, in 1988 and 1991, respectively. From 1998 to 2021, he was an Associate Professor with the Department of Industrial Engineering, University of Talca, Chile. From 2001 to 2003, he was the President of the Chilean Institute of Operations Research (ICHIO). From November 2016 to June 2020, he was the Director of the Logistics Center (CETLOG), Chile. He joined Universidad Finis Terrae, in 2022. He has published several articles in top operations research journals, such as EJOR, C&OR, *Information Sciences*, and IJPE. His expertise in mathematical problem modeling and his experience in solving complex models are valuable assets to the team.

**ÁLVARO FAÚNDEZ** received the degree in computer science from the University of Chile and the master's degree in data science from the City University of New York. He is a Computer Scientist. He has 12 years of experience in agile software development, focusing on high-availability services and big data processing.

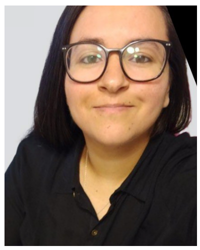
**CAMILA ROJAS** is currently pursuing the M.Sc. degree in computer science with the Pontificia Universidad Católica de Chile (PUC Chile). She is a Sociologist. She has experience in Analytics with BRAVE UP and the Millennium Institute for Foundational Research on Data (IMFD). Her work focuses on applied data analytics for education and social systems.

**ISIDORA JERIA** is a Sociologist from the Pontificia Universidad Católica de Chile and currently works as a Sociologist at BRAVE UP. Her research interests focus on gender inclusion, educational assessment, and mixed-methods research. She has experience in academic research and has authored publications in prestigious peer-reviewed journals, contributing to evidence-based approaches in education and social analysis.

**NATACHA BENÍTEZ** received the degree in sociology from Universidad Diego Portales (UDP), Chile. She currently leads operational coordination with schools at BRAVE-UP.

**MARÍA IGNACIA TUPPER** is the Director of Education at BRAVE UP. She is a Psychologist with the Universidad del Desarrollo with a specialization in school coexistence from the University of Los Andes. She has eight years of experience in education, coordinating the implementation of programs in educational institutions, specifically addressing school attendance and absenteeism, student socio-emotional education, and school coexistence. During her time at BRAVE UP, she has specialized in developing content and tools for both measuring and supporting the work of coexistence teams to improve coexistence in their educational institutions. Her expertise in school coexistence is vital to the successful execution of the project.

**ROMINA INOSTROZA** is the Head of the Education Department at BRAVE UP. She is a Sociologist from Universidad Diego Portales. She has nine years of experience in education, particularly in outreach to educational institutions and in content creation and development, working with both quantitative and qualitative methodologies. Her work at BRAVE UP! has focused on school coexistence. Her experience and expertise will help us build strong relationships with schools and establish effective feedback mechanisms.



**ETIENNE BELLENGER** received the degree in computer science from Pontificia Universidad Católica de Valparaíso (PUCV), Chile. She is currently a Data Scientist.

**FRANCISCO PÉREZ-GALARCE** received the master's degree in operations management from the University of Talca and the Ph.D. degree in engineering sciences from the Pontifical Catholic University of Chile. He is currently an Assistant Professor with the Faculty of Engineering and Business, University of the Americas. With more than a decade of experience leading advanced analytics projects in both research and industry, he has developed and implemented innovative analytical methods to address complex problems across diverse domains. He has more than 15 articles in important journals, covering applications of *Operations Research* and *Machine Learning*.