IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 30 January 2023, accepted 13 February 2023, date of publication 14 February 2023, date of current version 23 February 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3245325

APPLIED RESEARCH

Relationships Between Social Interactions and Belbin Role Types in Collaborative Agile Teams

DIEGO MONSALVES[1], HÉCTOR CORNIDE-REYES[2], AND FABIÁN RIQUELME[1]
[1]Escuela de Ingeniería Informática, Universidad de Valparaíso, Valparaíso 2362735, Chile
[2]Departamento de Ingeniería Informática y Ciencias de la Computación, Universidad de Atacama, Copiapó 1531772, Chile

Corresponding author: Fabián Riquelme (fabian.riquelme@uv.cl)

This work was supported in part by the Agencia Nacional de Investigacion y Desarrollo de Chile (ANID), Chile, through the Fondecyt de Iniciación, under Grant 11200113.

ABSTRACT Incorporating technologies to support the monitoring, evaluation, and feedback of soft skills attracts great attention in the scientific community. For a team to have a good performance and manage itself, it can be helpful to identify the behavioral characteristics and social interactions established between its members. Here, we explore the relationships between Belbin's behavioral roles, affinity sociograms, and audio interactions collected by the Naira multimodal analysis platform. We present a case study with university students designing with the Lego Serious Play methodology and incorporating agile practices that would stimulate collaborative work. The results are promising, allowing us to define new research hypotheses that will serve as guidelines for future work. The findings indicate that the Naira platform would enable us to recognize behavioral traces for the Social and Action natural roles.

INDEX TERMS Teamwork, human behavior, Belbin, network analysis, multimodal analysis.

I. INTRODUCTION

Nowadays, there is a growing demand in the search for professional skills that allow collaborative work [1], [2], that is, instances that integrate groups of two or more people to work together, intending to solve a problem, complete a task, or create a product. In collaborative work, each participant is responsible not only for their own actions but also for those of their colleagues [3]. It is known that in different contexts, collaborative groups increase the efficiency and flexibility of the work [4]. One example is the agile methodologies, initially created for software development, demonstrating their versatility in different industrial processes [5].

For K. Bruffee, one of the first to treat the theory and basic principles of collaborative learning in the 1980s, collaboration in collaborative environments is related to the social interactions carried out by the members of a collaborative group [6]. These interactions emerge naturally and are a source and reflection of collaboration, allowing synergy and constructive actions among the participants. Thus, the participants of a collaborative team can be represented as actors or nodes of a finite network. Likewise, the different interrelationships between participants can be represented as edges or ties between their respective nodes [7]. The above allows a collaborative activity to be studied as a dynamic social network through social network analysis (SNA) techniques [8].

Researchers agree that behavioral roles are relevant in forming balanced, collaborative groups. An adequate formation of work groups can facilitate collaboration among the participants, thus helping to improve the performance of the tasks for which the group was formed (e.g., academic, work, or creative tasks). The different profiles can arise from the participants' occupations, work experiences, and personal attitudes, among others [9]. In this sense, the profile categorization of potential participants can be relevant to balanced team formation. For example, a group of individuals with leadership profiles might not be as efficient for some tasks requiring action-oriented profiles.

The associate editor coordinating the review of this manuscript and approving it for publication was Anton Kos.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

The Belbin methodology [10] is one of the most used for categorizing behavioral profiles of participants in collaborative groups for the professional field. The methodology is based on self-assessment tests, in which participants are asked how they feel and how they would behave in a collaborative group. The instrument does not measure personality but allows a categorization based on nine roles, grouped into three types of profiles [11]. Thus, the tool allows for identifying the participants' strong and weak points, aiming to maximize the work team's performance [12]. It also provides recommendations for the integration of participants into work teams. Being self-assessment tests, Belbin himself acknowledges that some people could respond biasedly, with a vision of how they would like to be cataloged rather than how they really behave [9]. In this way, the test does not allow an in situ validation (i.e., during the activity) of the person's attitudes.

Nowadays, there is a growing research interest in identifying and monitoring collaboration in educational or professional environments through technologies. The ease of finding economically viable sensors within reach of people allowed us to move from studies of Human-Computer relations to Human-Human relations. Multimodal learning analysis (MMLA) was proposed by Worsley et al. [13] as a triangulation between traditional and non-traditional forms of data to characterize or model student learning in complex learning environments [14]. Outside the educational field, the MMLA has also been used to identify leaders and experts in collaborative activities, allowing visual identification of conversation leaders in collaborative groups [15], [16].

In previous work, we have explored collaborative behaviors of people working together to solve a problem using MMLA techniques [16]. These findings have provided valuable insights into how teams collaborate at different stages of the challenge. However, those experiences do not allow us to infer the behavior of each individual and the roles they play when working together. For this reason, in this research, we have considered using a scientifically validated instrument (Belbin Role Report) to improve the analysis of what we have done.

How are professional roles expressed during the development of collaborative activity with work teams? How do participants with different roles and professional profiles relate at work? Are these relationships detectable and analyzable? In this work, we seek a first approach to the relationship between the social interactions that emerge from a collaborative activity and the Belbin professional roles of the participants. As a case study, we consider 24 undergraduate students from a Chilean university who, through collaborative agile activities, meet in four work groups to solve a common objective with the LEGO®Serious Play (LSP) methodology [17]. Using SNA techniques, two different types of social networks are analyzed. First, a network of preference relationships is considered, in which students are asked whom they would like to work with. After the work teams have already been formed, the activity generates dynamic group networks made up of the speech interactions of the participants. These speech interactions are collected and analyzed in real-time using Naira [16]. The analysis of social relations is contrasted with the individual and group Belbin profiles obtained from the students' self-assessments. In this way, In this way, this work aims to (1) determine if affinity relationships influence the formation of efficient work teams and provide evidence on the analysis of speech interactions as a complement to the Belbin methodology for collaborative team characterization, and (2) determine if it is possible to identify Belbin's roles by analyzing multimodal data generated in real-time from a collaborative activity.

The rest of the article continues as follows. Section II presents a brief literature review related to this work. Section III is devoted to explain some fundaments about Belbin GetSet, Naira and affinity sociograms. Section IV describes the methodology of our work, which is applied on the case study detailed in Section V. Section VI presents and discusses the main results and findings. Finally, Section VII presents the main conclusions and future work.

II. RELATED WORK

There is extensive literature on identifying behavioral roles for improvement in work teams. The Myers–Briggs Type Indicator (MBTI) [18] is an introspective questionnaire created in 1962 to identify personality types. It consists of four dichotomies related to the perception of the world and the ability to make decisions, whose purpose is to associate the individual with one of 16 different personality archetypes. The indicator received updates in 1985 and 1998 [19]. Although currently considered unreliable and of poor validity [20], to date, it enjoys high popularity and trust from its consumers [21]. The role-based collaboration (RBC) methodology provides a set of concepts, models, processes, and algorithms intended to support collaboration and collaborative activities. In [22], the authors propose a generalized three-step process of RBC and describe a formal E-CARGO model that supports the fundamental principles of role-based collaboration and meets the requirements of role-based collaboration. Finally, they propose an architecture for implementing role-based collaboration from E-CARGO. The application of this architecture is quite broad, and is also used as a predictor of team performance [23], [24], [25] Another indicator explicitly designed for identifying work team roles is the Belbin Team Inventory (Belbin, for short) proposed in 1981 [11] and 1993 [26]. It is one of the most used questionnaires by researchers since it offers a positive categorization of nine types of roles from three different personality profiles, in addition to feedback and indications for forming a team compatible with the assigned roles [10].

Identifying personality profiles and work team roles can be an excellent input for studying group activities. These questionnaires and indicators can be applied in interdisciplinary studies together with other tools, such as those provided by multimodal learning analytics (MMLA). Thus, qualitative analysis can be combined with technology to quantify team members' activity through different sensors for audio [27], geospatial location [28], movement [29], microexpressions [30], etc. This interaction data collected by the sensors can be analyzed using traditional statistical techniques, artificial intelligence (AI) [31], and even social network analysis (SNA) [32].

Work teams form dynamic networks in which participants interact in various ways. In [33], academic networks represented as graphs of labor relations and bibliometrics are studied, generating confidence graphs to recommend research teams. The authors developed T-RecS, an application to identify similar profiles and potential candidates.

In [34], the authors use SNA techniques to study the relationships of software developers in an Indian company during the testing phase of a development project. The aim is to study the evolution of project teams and collect information that can be correlated with the health and outcome of the project. For the above, the authors categorize each user with static attributes (assigned role, location, seniority, organizational level) and dynamic attributes (e.g., the role assumed by each person during a specific project). At the group level, a social network connects workers through explicit relationships (email messages, instant messaging) and implicit relationships (working in the same section of the project, test cases, among others). They use the eigenvector centrality measure to identify the most relevant actors in the network.

In [35], different coding collaboration properties were correlated on the Github network to improve developer recommendations and evaluate the formation of development teams. In this case, the network nodes are the developers, who are related if there is a collaboration between them. The edge weights are the collaboration strength among the developers. Some of the centralization measures used were the clustering coefficient, Neighborhood Overlap, and Adamic-Adar coefficient. Also, the authors propose two new collaboration properties based on the number of lines in a commit and the previous social interactions of the developers. Finally, of the 21 properties analyzed, the authors conclude that only five of them are enough to develop computational models to measure collaboration strength.

In [36], the problem of task assignment in a team is addressed, seeking to test whether SNA can be used to improve coordination in the Software industry. A pilot case was carried out in a geographically distributed environment, forming a network with four groups of 7 and 8 participants. The author concludes that SNA techniques can contribute to the SW industry. However, the small amount of data obtained (sample size) and the low participants' experiences in SW development (sample quality) can be relevant limitations.

A larger sample was considered in [37], where data from 55 software development teams was collected to understand the effect of team leaders. It was concluded that leaders are not necessarily the most central network actors. However, when technical leaders are more central (which in this work means that they communicate more information than they receive), a significant improvement in productivity and task quality is observed.

On the other hand, there is extensive literature on the applicability of Belbin in work teams. For instance, the Belbin team roles can be applied in the health area to identify the teamwork capacity of medical students [38]. In software engineering, Belbin is one of the bases for creating an agile methodologies framework called ASEST+ [39].

Despite the above, most research in computing refers to analyzing its effectiveness in work teams. In [40], a method designed for forming balanced work teams based on Belbin's roles has shown good student results. The authors indicate that applying their methodology to students with a high educational commitment reduces their study time outside of class and generates more interest in the course contents. In addition, it helps them improve interpersonal relationships and social skills, as well as enhance the elements of cooperative learning, followed by positive interdependence and individual responsibility. As indicated, applying Belbin's roles to students makes them more aware of the different individual and team skills needed to succeed within a team. The students learn a greater self-understanding of their strengths (and weaknesses) and learn to work in an environment centered on diversity (of roles and skills) more than friendship. In [41], an exploratory study was carried out on the influence of Belbin on integrating software development teams with members who present compatible roles. Although the authors failed to confirm or disprove his hypothesis, they learned lessons to continue with controlled experiments. In [42], Belbin was applied to improve teamwork skills in an educational setting. Through Belbin, students learn about their approach to work and can discuss their strengths and weaknesses in their interactions with others. In this case, the author did find compelling indications that Belbin helped in group work, a conclusion obtained some time after the completion of the activity. In [43], a Belbin-based model for predicting academic performance for engineering students was proposed. The main objective was to provide evidence of the real impact of the questionnaire results on the final students' team performance. A machine learning model based on random forest achieved classifications with an accuracy and F1 score of 80%.

In [44], the authors present a study on the theories of team roles, inquiring about character strengths and positive team roles. They relate character strengths and work-related team roles in the individual, considering job satisfaction, self-esteem, supervisor-rated performance, and team level through teamwork quality data, self-rated team performance, and qualifications by the supervisor. They examine how team composition relates to results, that is, whether balanced teams go hand in hand with desired outcomes and whether there is an overrepresentation of team roles or character strengths. Note that in this case, the authors do not use Belbin's roles, as in our investigation; instead, they use the seven professional roles of the VIA methodology. In addition, the data they manipulate are qualitative aspects obtained through a form. In our case, Naira allows obtaining quantitative data in real-time on how people relate to each other in collaborative activities, intending to establish the feasibility of identifying Belbin roles.

In [45], an approach closer to the present work is presented. The authors use Belbin on a group of Computer Engineering undergraduate students to define the dominant roles and profiles of the participants. They then construct an affinity network or sociogram of the students, who are asked which two students they prefer as co-workers. On this network, specific SNA techniques are applied, such as the degree centrality measure and the identification of triads or triangles of the network. The objective is to facilitate the formation of work teams for the teachers, complementing their experience with Belbin and SNA. According to the authors themselves, the proposal is viable but inconclusive. In addition, unlike our work, the authors only analyze the affinity network of the students, but not the networks of interaction during collaborative work once the teams have been formed. The latter does not allow them to compare the choice of groups with their performance in specific activities.

In light of previous work, there are clear advances in studying group work to improve the performance of collaborative work teams. However, the objectives of analysis, methods, and techniques differ significantly between one work and another. Table 1 shows a synoptic summary that shows some of these differences. The use of technological platforms such as NAIRA, capable of monitoring in real-time the individual and group performance of collaborative work teams, is scarce. In this work, we seek to stimulate collaborative work through agile practices. The novelty lies in the search for a new combined approach to analyze collaborative work teams, namely, the monitoring of collaborative activities in real-time, which are analyzed using social network analysis techniques, but also considering previous profiling of the participants (obtained in this case through the Belbin Role Report), which allows knowing a priori their potential behavior during the activity.

III. PRELIMINARIES

This section presents three techniques used in this research to capture and analyze different kinds of data: i) people's disposition towards collaborative work, ii) their preferences when forming collaborative groups, and iii) their voice interactions during a collaborative activity.

A. BELBIN GetSet

Belbin Team Inventory provides a set of tests, analysis tools, and reports that today have become a powerful and validated instrument to characterize people's behavior in collaborative work [10]. Table 2 describes the three profiles and nine roles proposed by the Belbin methodology [10]. Belbin defines a *team role* as a tendency to behave, contribute and relate in a certain way. He also determines that said behavior is influenced by six factors [26]: (1) Personality, (2) Mental Ability, (3) Values and motivations, (4) Experience, (5) External influences, and (6) Learned Role. The Belbin reports contain personalized guidance and helpful advice for team members who want to improve their performance. All this information provides a broad view of people's strengths and weaknesses and how they will contribute at an individual, relational, and team level.

Belbin GetSet, a tool specially designed for young students, is used in this work. Belbin GetSet reports aim to promote students' self-awareness and self-confidence, help them develop and communicate their skills, improve their employability, empower them to make decisions and resolve conflicts, and teach them how to work in teams successfully [46]. The report contains graphs showing the percentile contribution of each student's team roles, ordered from highest to lowest. These percentiles seek to measure and express the strength of an individual's team role predisposition. Values between 0-30 are considered *rejected roles*, between 31-70 are the *assumable roles*, and between 71-100 are the *natural roles*. The reports are generated by the Interplace 7 expert platform, based on the student's answers to an online behavior-based questionnaire called Belbin Self-Perception Inventory (SPI).

B. AFFINITY SOCIOGRAMS

In social network analysis, a social network is a set of individuals or actors related in some way. A social network is usually represented as a graph or sociogram (V, E), where the set of nodes or vertices V are the network's actors, and the set of ties or edges $E \subseteq V \times V$ are the relationships between actors. Thus, the relationship between two actors a and b is represented by an edge $(a, b) \in E$ [7].

Affinity sociograms are a way to model and analyze groups of people according to their affinity relationships [47]. These relationships can be built from a simple survey of the actors. In that survey, they are asked to indicate with which other actors they would feel comfortable collaborating. Note that this answer does not necessarily require knowing why each person was chosen, e.g., friendship or job affinity. Note also that the affinity relationship is asymmetric, i.e., an actor a could choose b without the affinity being reciprocal. The latter results in a sociogram with directed edges, where $(a, b) \neq (b, a)$. The *in-degree* of an actor is the number of actors pointing to it, and the *out-degree* is the number of actors pointing from it [7].

If, in addition to just indicating whom they would like to work with, we ask the actors to rank their preferences, then we can distinguish the strengths of the affinity relationships by a weight function $w : E \to \mathbb{R}$, thus obtaining an influence graph [48]. Thus, if actor a chooses b as the 1st preference, the edge weight will be $w(a, b) = 1$; if it is the 2nd preference, it will be $w(a, b) = 1/2$; and if it is the k-th preference, the weight will be $w(a, b) = 1/k$. Given an actor $a \in V$, its *weighted in-degree* is the sum of the weights of the actors pointing to it, i.e., $\sum\{w(b, a) \mid (b, a) \in E\}$, and its *weighted*

TABLE 1. Synoptic table.

Ref.	Goal / Scope / Focus / Context	Method	Data Collection / Data Analysis	Focused on agile practices	Real-time visualization
[23]	Team performance predictor model called E-CARGO to generate evidence in the study of role-based collaboration.	RBC, E-CARGO	Team performance predictions according to a given role configuration, based on data from e-learning activities.	No	No
[34]	Study software development teams' evolution and performance in the testing phase to identify relationships between the teams' evolution and the project results.	SNA	Relational data between the members of a development team during the testing phase. Static and dynamic attributes are defined according to the roles defined in the software development team.	No	No
[35]	Collaboration analysis between software development teams to evaluate team formation.	SNA	21 properties are defined to monitor the collaboration of software developers using Github.	No	No
[36]	Analysis of the effects of using social networks to improve coordination and assignment of tasks in software development teams.	SNA	Case study based on 4 teams of master's students co-located geographically between the Netherlands and India.	No	No
[45]	Team building analysis combining SNA with Belbin roles.	Belbin, SNA	Case study of Computer Engineering students to look for interaction patterns between group members. The data obtained from the Belbin test are contrasted with the network analysis results in the groups formed by the teacher.	Yes	No
[44]	Study the relationship between character and team roles proposed by the VIA methodology with individual and group results from work teams.	VIA	Sample of 42 teams, considering individual variables (e.g., job satisfaction, performance self-assessment, supervisor assessment) and group data (e.g., quality of teamwork, team performance, and supervisor evaluation).	No	No
[39]	Evaluate the team cohesion's perception in agile practices as defined by the ASEST+ framework.	Framework ASEST+	Quasi-experiment addressing personality traits, conflict resolution, and the interdependence of critical project tasks.	Yes	No

TABLE 2. Belbin team roles for each profile.

Profile	Belbin team role	ID
Social	Coordinator	CO
	Teamworker	TW
	Resource Investigator	RI
Thinking	Plant	PL
	Monitor Evaluator	ME
	Specialist	SP
Action	Completer Finisher	CF
	Implementer	IMP
	Sharper	SH

out-degree is the sum of the weights of the actors pointing from it, i.e., $\sum\{w(a, b) \mid (a, b) \in E\}$. In particular, the weighted in-degree is a centrality measure that quantifies how required each actor is to form work teams.

Besides the centrality measures that allow understanding the relevance of each actor in the sociogram, there are centralization measures to compare the connectivity and cohesion levels of networks and network components. For instance, the well-known *average clustering coefficient* (ACC) refers to the network cohesion level, while the *average weighted in-degree* (AWI) refers to the affinity relationships homogeneity level between actors. In addition, based on the concepts of clustering and structural balance [7], we define the *balance index* of an affinity sociogram as the average of the balances of each actor. The balance of an actor is its in-degree divided by its degree (i.e., the number of its neighbors). In small sociograms, the balance of an actor is associated with the rate of cycles with positive affinity containing the actor. A sociogram with a high balance index represents a network with a strongly transitive affinity relationship.

C. NAIRA

Naira is an MMLA cloud-based application that allows storage, analysis, and visualization from voice interaction data collected through lavalier microphones in group activities. It was introduced in [16] and is made up of two main components: A cross-platform mobile application that, through the browser client, allows cloud services for storage, processing, and analysis (Naira App) and a dashboard that allows creating activities and monitoring ongoing activities in real-time (Naira Web). Figure 1 illustrates the high-level operation of the various components of Naira while monitoring a collaborative activity.

Naira App captures and processes the data of each speech interaction event made by the student. Among the metrics provided to Naira Web are the speaking time, the number of interactions made by each participant, and their average voice amplitude. The effective speaking time variable that

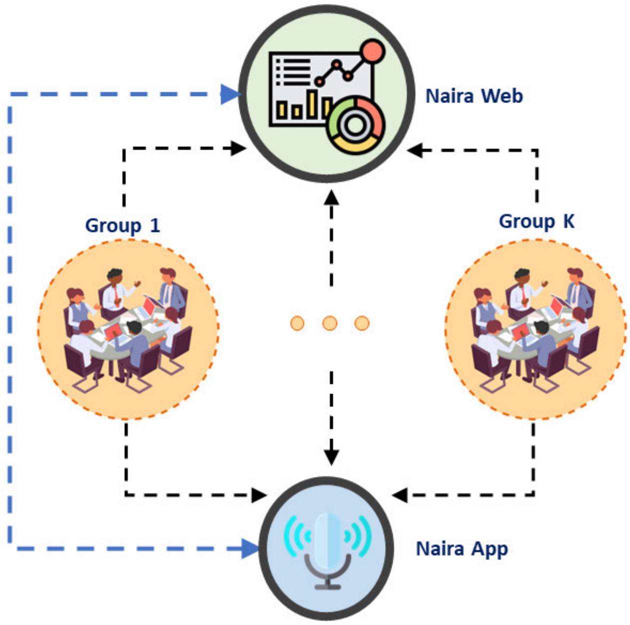

FIGURE 1. Naira components while monitoring a collaborative activity.

Naira has increased its value for each user when, by analyzing the data, a person has at least two consecutive presences at the time of going through the list of interactions, avoiding the sum of speaking time on occasions when two users speak simultaneously. At the group level, statistics and group measures are delivered based on the previous metrics. Additionally, one of the main features of Naira Web is the real-time visualization of group speech interactions through sociograms or graphs. Each graph (V, E) represents the social network for a group, where V are the group's actors, and E are the speech interaction relationships between its members. A (directed) edge $(a, b) \in E$ represents that actor b has spoken immediately after actor a. Analogous to the affinity sociogram of Section III-B, the edge weight $w(a, b)$ (displayed as the thickness of the edge) represents the number of times b has replied to actor a. In this case, a node labeling function $f : V \rightarrow \mathbb{R}$ (displayed as the node size) is also displayed, representing the current talk time $f(i)$ for each actor $i \in V$.

IV. METHODS

This research has an exploratory approach since, to our knowledge, it is the first attempt to compare speech interactions and affinity relationships with Belbin roles in collaborative work teams. Therefore, rather than an experimental approach, a case study is proposed considering the guidelines proposed in [49]. This work integrates SNA, MMLA, and psychometric techniques under a collaborative environment created with the LSP methodology.

A. RESEARCH QUESTIONS

The research questions (RQ) that guide our exploratory process are the following:

RQ1 Is it possible to identify Belbin's behavioral roles through data analysis from a collaborative activity?
RQ2 Does affinity sociograms help to identify work teams with better performance?
RQ3 Is there a relationship between behavioral roles, speaking times, and the number of speech interactions?

B. CASE STUDY DESIGN

The case study should be an activity that stimulates interaction and collaboration among the participants. Thus, we propose using the LEGO® Serious Play (LSP) methodology, incorporating some agile Scrum practices [17] to visualize the artifact construction process iteratively and incrementally. Agile Scrum practices considered are:

- Approach the work iteratively and incrementally, that is, organize the work to be done in sprints of fixed duration.
- Self-organized teams with a maximum of five students.
- Co-located teams.
- Work Planning, run the Sprint Planning ceremony.
- Minimum Viable Product (MVP).

LSP belongs to the Serious Games (SG) category, which allows students to acquire skills through game-based activities, given its playful and interactive nature [50]. The LSP methodology has become an engaging, entertaining, and innovative way to promote social skills development since it allows one to think with hand while tackling the challenge of solving a problem in a group [51], [52]. LSP facilitates the synergy between the participants, managing to generate social interactions that emerge from the activity.

Figure 2 illustrates the steps used to carry out a generic case study (see Section V-B for its application in a real case study).

C. DATA COLLECTION

The following data sources are considered to help answer the research questions:

- Belbin GetSet reports (see Section III-A). Students should not know their results before the practical activity to not condition their natural roles. After the case study is completed, the Belbin reports' results will be sent to each student through the Interplace 7 platform. Subsequently, facilitation will be carried out to guide the interpretation of the individual results. Guidelines will also be provided for the future formation of working groups, given the predominant roles detected in the reports.
- MMLA Naira Platform (see Section III-C). The participants must be familiar with the application before starting the collaborative activity.
- Affinity sociogram (see Section III-B). The survey necessary to build the affinity sociogram must be carried out before the collaborative activity.
- Observer annotations. During the activity, one or more observers should be available to record notes on both individual and group behaviors. The objects of interest to register will depend on the case study.

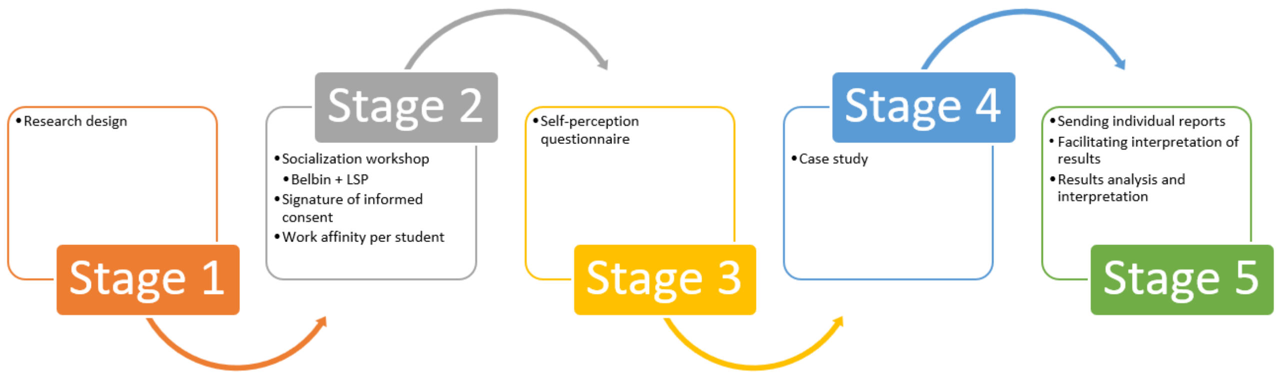

FIGURE 2. Research Model.

- Self-assessment questionnaire together with a satisfaction survey. Both instruments will be applied to each student at the end of the activity to obtain their perceptions of the group work carried out and evaluate their individual performance concerning what was observed in their other groupmates.

D. LIMITATIONS

The most important limitation of this work is related to Naira's data collection process. In previous research experiences, we have found that ambient noise and conversations from members of other groups could harm the data collection. In this sense, Naira App includes functionality to calibrate the microphone of each mobile device, allowing to mitigate possible collection problems. This is the first time the Naira App has been used in a real face-to-face environment for research purposes. Note that in [16], it was used remotely amid the health restrictions derived from the Covid-19 pandemic.

On the other hand, the main limitation identified by Belbin GetSet is the possible abundance of the same role in the same group. The above can lead to internal group competition and leadership conflicts. Belbin advises dividing teams with balanced roles to avoid these types of disputes. Another limitation we can identify is the type of collaborative activity carried out. Indeed, the kind of activity and the duration of its stages could affect the need for specific profiles and roles over others. For this work, a collaborative activity divided into well-differentiated stages of beginning, development, and end was designed, considering the possibility that all roles may be necessary at some point.

Finally, the sample size is relevant from an instrument validation point of view. As we shall see below, this work considers a case study of 24 students. This number is justified at a methodological level as a first exploratory approach. However, due to the sample size, the results obtained cannot be generalized to other contexts, without first being validated through broader experimentation.

V. CASE STUDY

Below we describe our case study based on all the considerations mentioned in Section IV.

A. PARTICIPANTS

As a case study, we considered 24 student volunteers (21 females and 3 males) from the 5th semester of Business Administration and Economics (Commercial Engineering) career at a public university. The activity was taken in the context of the Administrative Information Systems course. This course aims to understand the importance of information systems in organizations to support decision-making.

B. ACTIVITY DESCRIPTION

One week before the LSP activity, an affinity sociogram was constructed, asking the students for a ranking of at most five classmates with whom they would like to form a work group. Thus, all the actors in the resulting affinity sociogram have an out-degree between 1 and 5. Also, the weight of the edges can assume values in the set $\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}\}$. These sociograms were analyzed and visualized with Gephi software.[1]

The chosen LSP activity encourages students to build a model representing their labor specialty's perspective. The resources used were the following:

- Belbin GetSet role questionnaires.
- LEGO materials: LEGO Creative Brick Box, 10696; Gray Baseplate LEGO Classic, 11024; Wild Animal Set, LEGO Education, 45012; Community people Set, LEGO Education, 45010.
- Naira equipment: Naira App, lavalier microphones connected to personal smartphones.
- Group answer sheets to record the work done.
- Individual activity evaluation questionnaires.

Figure 3 describes the activities of the case study with their respective duration times. The schedule of activities was controlled by an online stopwatch[2] that was visible to all participating groups. Each phase is detailed as follows:

1) Organization of the workspace, leaving all the necessary material on each table. The students organize themselves into groups at the tables arranged in the room. The students recognize the delivered material and the activity instructions.

[1] https://gephi.org/
[2] https://www.online-stopwatch.com/spanish/full-screen-stopwatch.php

2) *Phase 1* (10 min): The facilitator (teacher) explains the challenge and defines the objectives. Each student reflects on the challenge, determining elements that could incorporate into the solution. At the end of this activity, data capture begins using the Naira App.
3) *Phase 2* (15 min): Each student makes their proposal without being interrupted by the others. If time is left over, a new round starts so that the student who requires it can continue developing their idea. After listening to all the proposals, the group determines the artifacts to build and distributes the work tasks in sprints (backlog prioritized + workload).
4) *Phase 3* (15 min), *Sprint 1*: The first work cycle is developed, where the group carries out what was planned. In the last 5 minutes, the group reflects and reorganizes the assigned work.
5) *Phase 4* (15 min), *Sprint 2*: The second construction cycle is developed, where the group adjusts the missing work to meet the challenge posed. Once this activity finishes, data collection with Naira App stops.
6) *Phase 5* (20 min): Each group presents their designed model, explaining the metaphors created as a solution. Students complete the self-assessment and satisfaction survey of the activity carried out.

For this case study, two observers were arranged, whose objects of interest were the following:

- Group organization to solve the LSP challenge.
- Group work dynamics.
- Group interactions behavior.
- Group works climate.
- Group behavior for work under pressure (as the end of each sprint approaches).
- Artifact construction speeds with LEGO blocks.
- Use of objects (animals and characters) in the proposed models.
- Any other individual or group behavior that may be of interest.

The questions of the self-assessment questionnaire applied at the end of the activity were the following:

1) How do you evaluate your ability to work as a team in the activity compared to the other group members? (Under, Normal, Influential, Higher)
2) How do you evaluate your ability to collaborate in the activity compared to the other group members? (Under, Normal, Influential, Higher)
3) How do you evaluate your communication skills in the activity compared to the other group members? (Under, Normal, Influential, Higher)
4) How do you evaluate the result obtained by the group?
5) Were you able to recognize leadership characteristics in any group member? (Yes, No. How many?)
6) What was your first positive feeling when working with your colleagues? (Trust, Security, Relaxed, Entertained. Choose only one)
7) What was your second positive feeling when working with your colleagues? (Trust, Security, Relaxed, Entertained. Choose only one)
8) Were you able to work comfortably in the group? (Yes, No. Why?)
9) Can you clearly visualize your contribution to the final result obtained by the group? (Yes, No. Why?)
10) Did the use of monitoring technologies cause any changes in your natural way of working? (Yes, No. Which?)

An additional open question was added to the questionnaire to assess comments and suggestions about the activity:

11) Finally, we ask you for comments and suggestions to help us improve the performance of these activities.

C. ETHICAL CONSIDERATIONS

All students participated in the activity voluntarily. Before requesting their participation, all the activity details were explained, together with the objectives and theoretical research foundations. Those who decided to participate completed and signed an informed consent establishing the confidentiality of the data obtained.

VI. RESULTS AND FINDINGS

This section shows the results and findings from the Belbin GetSet report, the affinity network, and the interaction networks obtained with Naira. The other two data sources, i.e., the observer annotations and the self-assessment questionnaire (see Section IV-C), is used to support discussions of the findings.

A. BELBIN GetSet REPORTS

Table 3 shows the percentiles of the roles obtained by each student according to the Belbin GetSet reports delivered by the Interplace 7 platform. Natural roles are highlighted in bold (see Section III-A).

Note that G1 is a well-balanced group since it is the only one in which all the roles are natural. In G2, only two members could barely assume the IMP (Implementer) role, while all rejected the PL (Plant) role. Likewise, G3 looks balanced except for the PL role, that only could be barely assumed by two members. In G4, TW (Teamworker), ME (Monitor Evaluator), and SP (Specialist) are not natural roles, although they could be assumed by different members, especially the ME role. Finally, in G5, the CO (Coordinator), RI (Resource Investigator), and IMP roles are not natural but can also be assumed by different members.

When analyzing the report's results, close attention should also be paid to groups with excess natural roles. For example, in G1, there are three students (G11, G13, G15) with high percentiles in the TW role. The latter could cause indecision in the group, excessive trying to please, concealment of conflicts, and poor decision-making. Similarly, the IMP role is natural in G11, G12, and G13, which could cause resistance to change, excessive organization, and disputes in the group

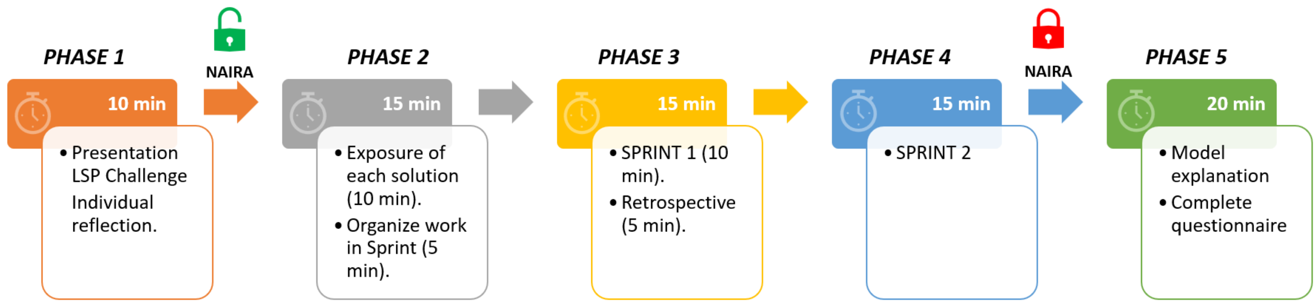

FIGURE 3. Time distribution in case study.

TABLE 3. Role percentile scores (%) obtained per group and for each student.

		Social			Thinking			Action		
Group	Member	CO	TW	RI	PL	ME	SP	CF	IMP	SH
G1	G11	29	**75**	**88**	16	49	21	**74**	**72**	54
	G12	46	39	13	32	6	61	**90**	**78**	**82**
	G13	**90**	**84**	8	**71**	13	61	61	**82**	0
	G14	46	46	43	58	41	14	68	42	**82**
	G15	20	**84**	67	0	**91**	**91**	39	26	37
G2	G21	**96**	69	**74**	16	**93**	5	**84**	34	0
	G22	**82**	62	67	16	33	0	53	26	8
	G23	29	62	0	0	25	**89**	**98**	18	28
	G24	**77**	39	19	16	**76**	37	53	42	**86**
	G25	20	**84**	**88**	16	0	61	24	26	37
G3	G31	8	**75**	67	32	**91**	37	53	18	6
	G32	14	24	67	0	0	**93**	**92**	34	**82**
	G33	**93**	54	**74**	0	8	54	11	18	45
	G34	29	**84**	43	41	25	61	8	34	54
	G35	20	7	0	16	41	**80**	61	**100**	37
G4	G41	46	24	**88**	41	41	37	**87**	58	45
	G42	54	39	67	**71**	49	69	61	58	0
	G43	**77**	12	43	65	70	**85**	53	34	37
	G44	0	46	**74**	0	64	**75**	**87**	**82**	20
	G45	**98**	54	**74**	16	8	54	24	**82**	6
G5	G51	70	39	59	49	**81**	5	30	58	**89**
	G52	37	**79**	59	49	64	**75**	24	42	20
	G53	54	62	43	**71**	57	21	47	42	62
	G54	54	**79**	13	32	70	**91**	**79**	42	8

Natural profiles (> 70) are highlighted in bold

to establish the limits of individual work. The first box plot of Figure 4 displays the results from the variability of each role, the quartiles, averages, median, maximum and minimum values.

Table 4 displays descriptive statistics for all students. The predominant profile is Social, followed by Action, and finally, Thinking. The most natural roles are SP, CF (Completer Finisher), and TW, followed by CO and RI. Excessing these roles in groups can produce excessive tension situations and inadequate task delegation, causing a work overload in some students. Of all the roles, SP and CF are the only natural roles in all the workgroups. On the contrary, the PL and SG (Sharper) are the most rejected roles, so there is a lack of students in the course who contribute to the distinctive characteristics of these roles. A shortage of the PL role can generate a lack of innovation and originality, as well as problems of work jams in crisis situations or in facing complex problems. The absence of the SH role can result in loss of objectives and blockage of group work in the face of difficulties or unexpected changes.

Figure 4 shows the box plot graphs of each participating group. Each box size represents the variability existing in the group. G1 presents a significant variability in all the roles, unlike G2, G4, and G5, where the variability for some roles is much lower. In G2, the results of the PL and IMP roles are striking. For some roles, the gaps regarding the maximum value are very different in each group. The wider this gap, the greater the group's dependence on the student with the characteristics and qualities expected of that role. In this sense, G4 and G5 present small gaps, reinforcing the idea that in both groups, the dominance of the roles is relatively balanced.

From all this information, it is possible to develop assumptions to help anticipate group behaviors. The behavior assumptions of G1 are:

- As this group has good behavior distribution between social (G11, G13), thinking (G15), and action (G12, G14) roles, it could have a greater chance of good performance.
- Since there is high variability in all the roles, with a predominance of the Social and Action profiles, the group should strive to incorporate the characteristics of the G15 Thinking profile.
- There are two students (G11 and G12) with natural IMP-CF roles. Sometimes, they could try to impose their criteria and behave excessively demandingly.
- The low presence of PL in the group can cause the action to be unorganized and inefficient.
- Excessive social and action roles could cause a lack of creativity and originality in group work.

The behavior assumptions of G2 are:

- This group has an excess of social roles (G21, G22, G24, and G25), which can cause group problems such as indecision, bad decisions, superficiality, and excessive conversation.
- G23 is the only student without natural social roles. That student may feel misunderstood and even unheard. Experience indicates that people with roles different

TABLE 4. General descriptive statistics for the Belbin role percentiles (%) obtained in all groups.

	Social			Thinking			Action		
Statistic	CO	TW	RI	PL	ME	SP	CF	IMP	SH
Average	49.5	55.1	51.6	30.2	45.7	53.2	56.7	47.8	38.5
Standar desviation	29.9	23.7	28.6	24.6	30.4	30.1	26.8	23.8	30.1
Mode	28	39	67	16	41	61	53	42	37
Max	98	84	88	71	93	93	98	100	89
Min	0	7	0	0	0	0	8	18	0
Percentiles > 70	7	8	7	3	5	8	8	6	5
Percentiles < 30	9	4	6	12	8	6	5	6	10

from the majority, although highly necessary in the group, are not directly integrated.

- G24, being the only one with a natural and dominant SH role, will be essential for the group not to fall into complacency and work in the right direction.
- Three students (G21, G22, and G24) have a CO natural role. This excess can cause a battle of egos and internal competition for control and delegation.
- The absence of PL and IMP roles can cause a lack of organization in the work group.

The behavior assumptions of G3 are:

- As this group has good behavior distribution between social (G33, G34), thinking (G31, G32), and action (G35) roles, it could have a greater chance of good performance.
- There is an excess of the SP role. Particular attention should be paid to the group gathering too much information or delving too deeply into work-related topics. The lack of another student with a natural SH role can aggravate that situation.
- G35 can lead the group by setting and maintaining goals, the pace of work, and opening up to new opportunities.
- Like G1 and G2, the group lacks the PL role, so the same assumptions apply here.

The behavior assumptions of G4 are:

- It is a balanced group, with high percentile values in the natural roles for the social and action profiles (G41, G44, G45). For the same reason, this group may enjoy doing things well, with quality, in a structured manner, and following the rules.
- The lack of the SH role can make them lose their work pace and not set themselves ambitious challenges.
- IR and PL roles are present so the group can be open to new opportunities and development.
- The lack of TW can cause a lack of mutual support. Eventually, the students could behave with little empathy and low group interaction.

The behavior assumptions of G5 are:

- The lack of CO and RI as natural roles evidences a low social profile at the group level. Internal communication could be the biggest obstacle.
- The G51 and G53 students have the SH role among their two highest values. This situation could generate conditions of competition between them. Their natural roles ME (G51) and PL (G53) could self-regulate them and thus mitigate said competition during the activity.
- Due to the absence of the IMP role, this group may need to structure its actions to achieve better results.

TABLE 5. The three predominant Belbin team roles of each student and their weighted in-degrees.

		Predominant roles			weighted
Group	Member	1st	2nd	3rd	in-degree
G1	G11	RI	TW	CF	2.20
	G12	CF	SH	IMP	2.16
	G13	CO	TW	IMP	1.95
	G14	SH	CF	PL	2.53
	G15	ME	SP	TW	2.20
G2	G21	CO	ME	CF	3.75
	G22	CO	RI	TW	2.00
	G23	CF	SP	TW	1.83
	G24	SH	CO	ME	2.24
	G25	RI	TW	SP	2.16
G3	G31	ME	TW	RI	1.00
	G32	SP	CF	SH	1.00
	G33	CO	RI	SP	1.66
	G34	TW	SP	SH	1.33
	G35	IMP	SP	CF	3.33
G4	G41	RI	CF	IMP	1.49
	G42	PL	SP	RI	0.25
	G43	SP	CO	ME	2.33
	G44	CF	IMP	SP	2.00
	G45	CO	IMP	RI	3.33
G5	G51	SH	ME	CO	2.66
	G52	TW	SP	ME	1.00
	G53	PL	SH	TW	1.00
	G54	SP	CF	TW	1.91

B. STUDENT AFFINITY SOCIOGRAM

The resulting affinity sociogram has 24 nodes and 86 directed edges. Recall that the edge weights can assume values 1.0, 0.5, 0.33, 0.25, or 0.2. This sociogram is illustrated in Figure 5. The thickness of the edges represents their weight. Each node's size represents the actor's weighted in-degree, which varies between 0.25 and 3.75. The node's colors are the predominant Belbin behavior profiles (red for People, green for Action, and blue for Thinking profile). In addition, each node's label represents its subsequent arrangement in the workgroups. A label Gij represents the actor j of the group Gi. Furthermore, Table 5 present the three predominant Belbin team roles for each student and their weighted in-degrees.

1) SOCIOGRAM CONNECTIVITY

Notice that there are three connected components, which are labeled A, B, and C. No isolated nodes exist, so all students

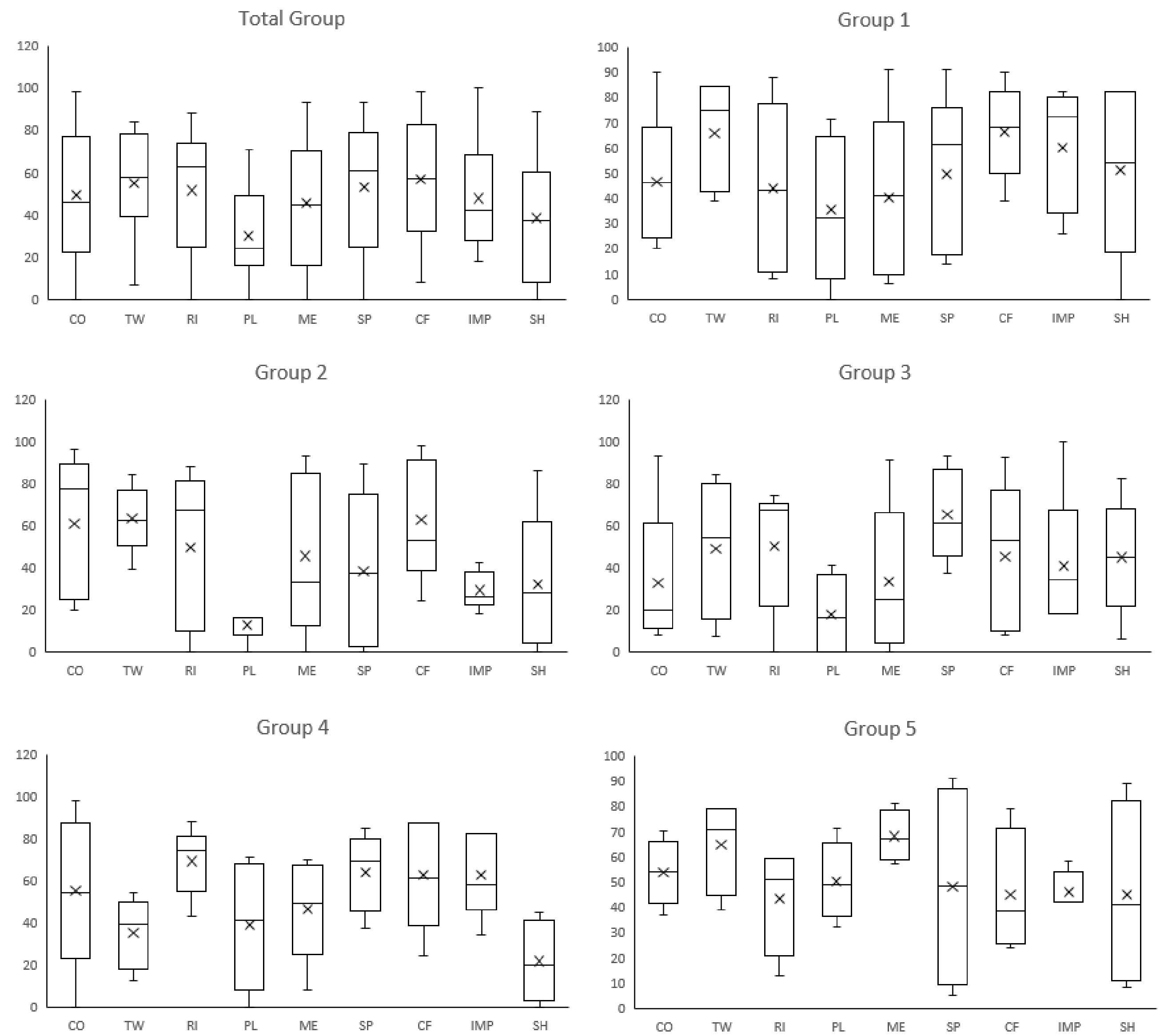

FIGURE 4. Analysis of variability in percentiles of each Belbin's role for each group.

TABLE 6. Affinity sociogram centralization metrics.

centralization measure		full sociogram	components *A*	*B*	*C*
clustering coefficient	avg	0.72	0.93	0.59	0.80
	std	0.30	0.04	0.35	0.11
in-degree	avg	3.58	4.67	3.23	3.20
	std	1.47	0.52	1.64	1.30
weighted in-degree	avg	1.97	2.21	1.89	1.88
	std	0.82	0.19	0.90	1.13

avg: average, std: standard deviation

are related to at least one peer. Table 6 shows three centralization measures for the entire sociogram and their components.

Despite having three components, the global network is quite cohesive. It has a high ACC (0.72), mainly due to the *A* (0.93) and *C* (0.8) components. Component *B* is less cohesive (ACC of 0.59) due to a weak cutting edge from actor G35 to G21, without which component *B* would be divided. Likewise, a strong connection is also observed between the actors G52 and G53, both chosen mutually as the first option but not having other actors who chose them. Actors G52 and G53 form a dyad only linked to the *B* component through the choice of G53 over G22 as the second option.

The in-degree of the students varies between 1 and 5, with an average of 3.58 and a high standard deviation (1.47). Therefore, on average, each student was chosen by between 3 and 4 classmates. These deviations are also high in the *B* and *C* components, where the average in-degree is close to 3. However, in the *A* component, very high connectivity is

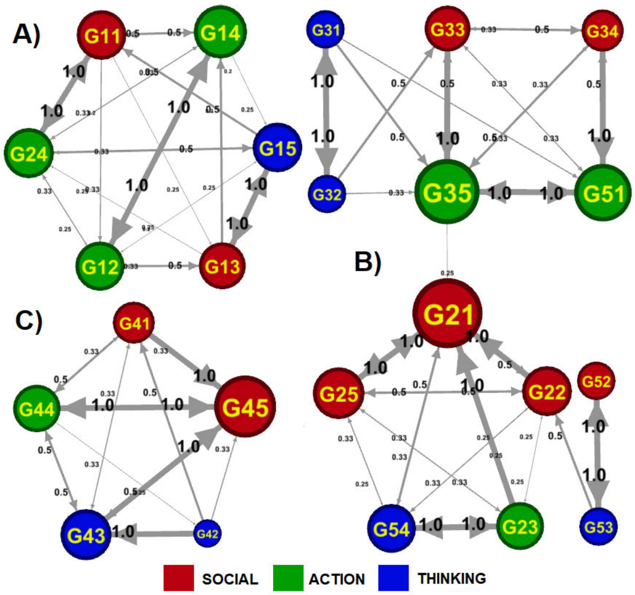

FIGURE 5. Affinity sociogram. The node's size represents the actor's weighted in-degree and the node's colors in the Belbin behavior profile.

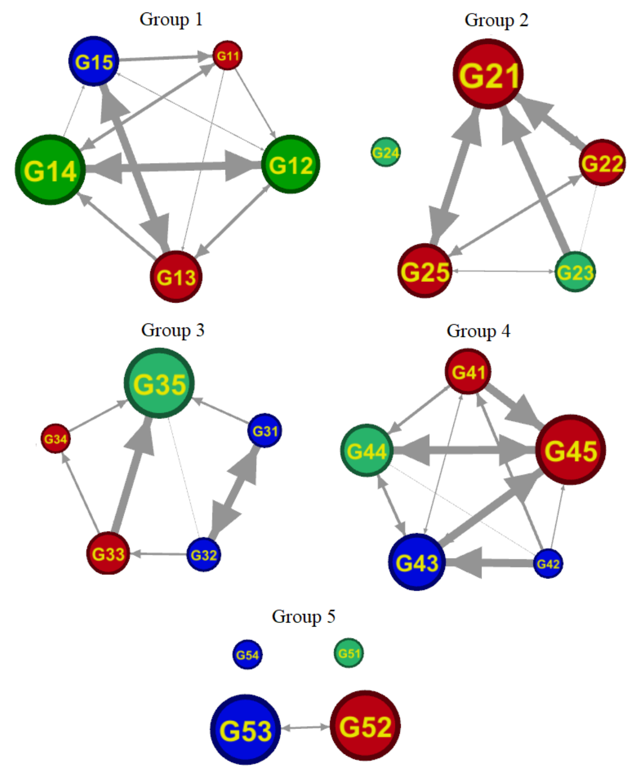

FIGURE 6. Working groups with their affinity relationships.

observed, with an average in-degree of 4.67, with a low standard deviation (0.52). The latter means that the six students of this component coincided in their desire to participate in the same work team.

On the other hand, the global network has an AWI of 1.97, which means that, on average, each student was chosen with a strength equivalent to that of two first-priority choices. The above is valid for all three components. However, of the three, the component *A* (AWI 2.21) is by far the most homogeneous, with a low standard deviation (0.19), which means that the weighted in-degree of its actors is very similar. The component *B* (AWI 1.89) has a relatively high standard deviation (0.90), so despite being a well-connected component, there are different preference orders among the students for team formation. This component includes the favorite student to form work teams, G21 (weighted in-degree of 3.75), and four of the five least favorite students (G31, G32, G52, G53), which form two pairs chosen themselves with first priority (weighted in-degree 1.0). Finally, the component *C* (AWI 1.88) presents the highest standard deviation (1.13), having the actor G45 with the second highest weighted in-degree, together with G35 (3.33), and also the actor with the slightest predilection of the entire sociogram, G42, which, with a weighted in-degree of 0.25, was only chosen by a classmate in fourth place.

2) SOCIOGRAM PREDOMINANT BELBIN PROFILES

Of the five most relevant students in the sociogram, in terms of weighted in-degree, two have a predominant Social profile (G21 and G45) and three a predominant Action profile (G35, G51, and G14). On the other hand, of the five students with the lowest weighted in-degree, four have a Thinking profile (G42, G31, G32, and G53), and only one has a Social profile (G52), with a 79% percentile in the TW role, closely followed by a Thinking profile, with a 75% percentile in the SP role (see Table 5).

Interestingly, all three Belbin profiles are present in all three components. Furthermore, if the weak cutting edge between actors G35 and G21 of component *B* would not exist, the three Belbin profiles would still be present in the two new resulting components.

3) FINAL GROUP COMPOSITIONS

Now we analyze the final formation of the work groups for the activity, maintaining the affinity relationships within each group. Let us remember that the groups were formed by free choice. Some formed quickly according to their affinity ties, while others took longer, having to settle for joining groups with peers who were not among their priorities. Figure 6 shows the five team works formed for the activity. Although the thickness of the edges still represents the weight of the affinity relationship, since some edges disappeared, here we are not concerned with the preference orders but with the groups' structure and cohesiveness.

Note that G4 corresponds to component *C* of the original sociogram. The *A* component, which was also very cohesive, had to leave the G24 student to form the G1 group to meet the requirement of five members maximum, even though G24 was the second with the highest weighted in-degree of the component (see Table 5). Actor G24 joined group G2, made up of one of the three divisions of the *B* component. This component was separated into groups G2, G3, and G5.

TABLE 7. Groups centralization metrics.

	weighted in-degree		clustering coefficient		balance index	
Group	avg	std	avg	std	avg	std
G1	1.75	0.37	0.90	0.07	0.75	0.25
G2	1.38	1.12	0.80	0.45	0.80	0.45
G3	1.30	0.60	0.58	0.24	0.60	0.25
G4	1.88	1.13	0.80	0.11	0.65	0.38
G5	0.50	0.58	0.00	0.00	0.50	0.58

avg: average, std: standard deviation

Groups G2 and G5 are made up of students on each side of the bridge of component B (see edge (G35, G21) in Figure 5). The strong dyad of students G52 and G53 formed the G5 group with two actors in opposite places of the bridge.

Table 7 shows various centralization measures for the different groups, considering their affinity relationships. Groups G1 and G4, derived from components A and C, have the highest AWI (1.75 and 1.88, resp.) and ACC (0.90 and 0.80, resp.). However, the very low weighted in-degree of the G42 student (0.25) explains the high standard deviation of the G4 group (1.13). Despite having an isolated student, the G2 group is also quite cohesive (AWI 1.38, ACC 0.80). Furthermore, it is the most balanced group, with a balance index of 0.80, since all its relationships are symmetric or bidirectional. Hence, it is clear that the G24 student, excluded from the $G1$ group, joined a very close, predominantly social G2 group, which, however, lacks the predominant Thinking profile. Note that the G4 group is not very balanced (0.65) as it has a very dominant actor (G45) and another one very nondominant (G42). Without having isolated actors, of the groups of five students, G3 is the least cohesive and balanced, with several students without affinity relationships and a single unifying student (G35) with an Action predominant profile. Finally, the only group of four students is G5. This group is the least cohesive and the most unbalanced, as it has only two students related to each other and two isolated ones.

C. NAIRA SPEECH INTERACTION NETWORK

Figure 3 shows the different phases developed in the activity. The Naira platform recorded the speech interactions developed in Phases 2, 3, and 4. Figure 7 shows the distribution of (cumulative effective) speaking times for each student at the end of each phase. G43 did not record data on Naira due to a mobile device problem. Since the activity was already in progress, it was not feasible to replace the device.

Furthermore, Figure 8 shows the Naira Web graphs with the accumulated activities of each group at the end of Phase 4 (Sprint 2). With all the data available and Figures 7–8, it is possible to identify the following.

Group G1 has the highest number of interventions and the lowest speaking time, which coincides with a clear predominance of social and action roles. Moreover, the thick edges of the corresponding interaction graph show intense communication with short dialogues between its members. Note that Naira evidences the natural SH role of G14 by maintaining high activity throughout all phases and a strong connection with the rest of the group members. On the other hand, the natural PL role of G13 is observed by its activation in phase 3 (Sprint 1) and then in phase 4 (Sprint 2). G15 has two natural mental roles (Thinking profile) with very high percentiles (91), although it was not possible to identify those behavioral characteristics during the LSP activity.

In G2, a clear leadership of G22 is observed in each activity's phases. The natural role of G22 is CO, a social role whose strengths are showing maturity at work, being very self-confident, helping the group to clarify goals, and delegating effectively [10]. In turn, G21 has the lowest participation in terms of the number of interventions and speaking time, despite having four natural roles in all three categories. G23, G24, and G25 maintain permanent communication throughout the challenge that intensifies near the end of phases.

Unlike the other groups, G3 presents a substantial alternation in the speaking times for each phase. Although students G32 and G35 dominated the conversation in different phases, the thick bidirectional graph interactions show a closer group collaboration. In phase 2, the conversation was dominated by G32; in phase 3, by G35; and in phase 4, both dominated. Both students present low percentile values in the social roles and very high in the SP mental role. Both also dominate action roles.

In G4, the work looks balanced in the three phases of the LSP challenge. G42, having a natural mental role (PL), which is very different from the rest, may have had problems integrating into the group, showing some initial activity, which declines in the two following phases. In contrast, G41 participates relatively little in the first phase but increases its activity in phases 2 and 3, which seems consistent with its RI and CF natural roles.

G5 presents an interaction graph similar to G4, with an even more balanced and stable speech distribution. G51, the only one with the natural mental role ME and the natural action role SH, mostly dominates the activity in all phases. During LSP activity, G51, G52, and G53 maintain close communication with similar values of speech interactions and time. G54 has three natural roles (TW, SP, and CF); however, her results in Naira indicate that this student has the lowest participation and the shortest speaking time in the group.

D. OBSERVER ANNOTATIONS

Observer annotations registered various individual and group behaviors during the LSP activity. Figure 9 shows the final models built by each group. Some characteristics of these models were also registered.

The G1 group was formed very quickly. Although they worked in an orderly manner, at first, it was difficult for them to concentrate and define the business model to be built. In fact, the group asked the most questions before starting the construction phases (Sprints 1 and 2). During Sprint 1, they talked a lot and built very few artifacts. G14 and G12 lead the work and organization of the group. They do not elaborate metaphors in the LSP model.

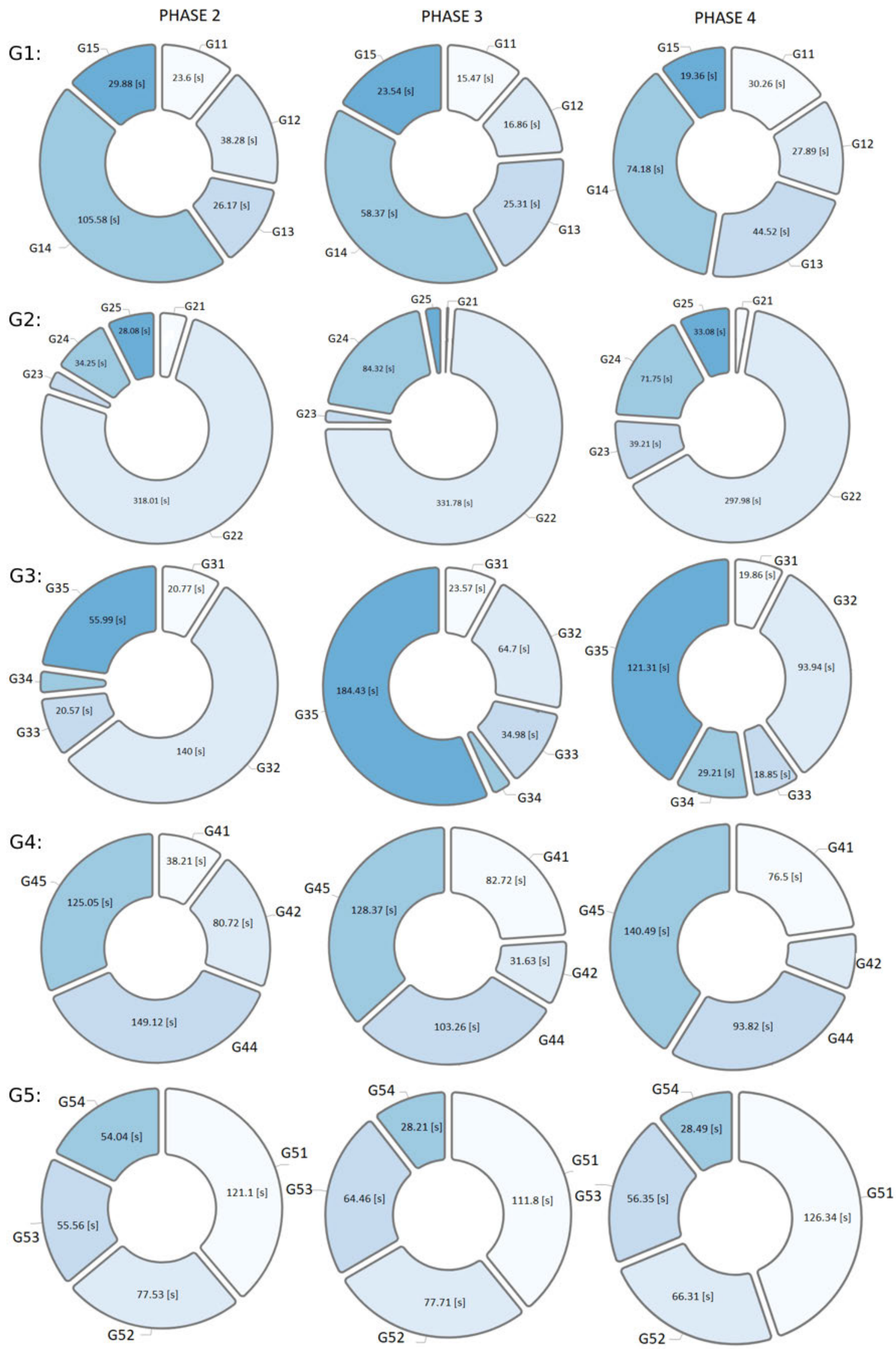

FIGURE 7. Naira platform graphs obtained at the end of phases 2–4.

The G2 group was the most restless and eager to build with Lego blocks. Anxiety causes them to start building artifacts without even having established group agreements. However, they got the most detailed specifications of the models.

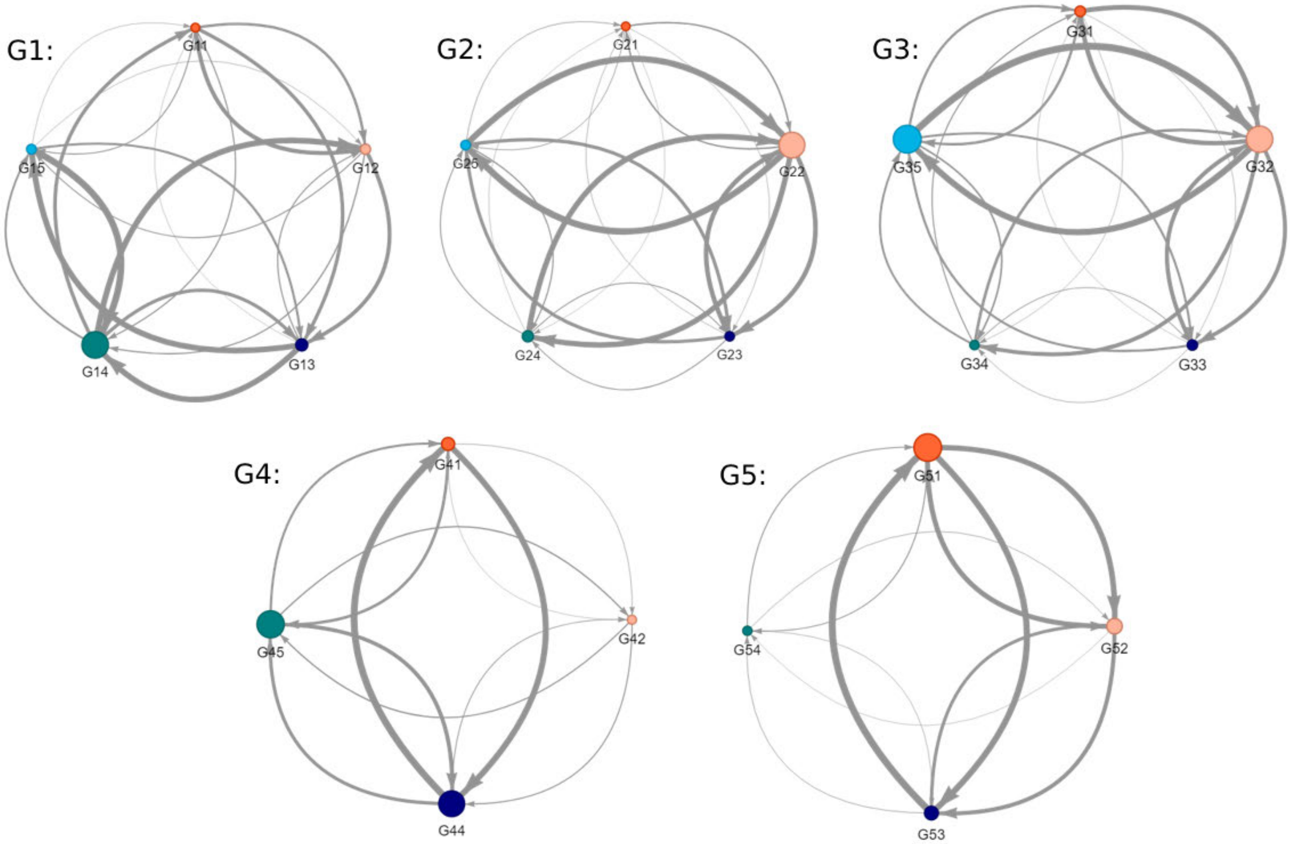

FIGURE 8. Final Naira graphs for each group. The node's size represents speaking time, and the edge's weight is the number of interactions.

FIGURE 9. LSP models built by the working groups.

During both Sprints, they were very enthusiastic and fun. It was the only group where two members worked standing up and not sitting down. Also, they rotated their locations as they built the artifacts. They developed three very clear metaphors:

the giraffe as the company's leader and process control, the polar bear as the company's strength, and modernity as the construction of a heliport.

Group G3 was the most orderly and disciplined group. They started with doubts, but it didn't take them long to define the problem well. They worked well under pressure: they built the most artifacts when they were warned that time was short. In Sprint 2, they copied the behavior of G2, and two of their members (G32 and G35) worked standing up. They also elaborate on three clear metaphors: the boss lying down and accompanied by an elephant, a thief representing finances, and a crocodile as the external variables influencing the company. The final model artifacts are highly integrated.

The G4 group was also very orderly and had no communication problems. Its members also rotate but always work sitting around the table. They developed three metaphors using the animals: the crocodile at the center of the company, representing the work environment; the tiger as the characteristics of the company leader; the penguin as the company's staff, conveying the joy expected in modern organizations.

Finally, the G5 group is the only one where it is clear that the members do not usually work together because they communicate very little with each other. They are the only group with time to spare in Phase 2. Initially, each student works in their own space with little interaction. However, at the end of Sprint 1, the group behavior changes, interacting more with the artifacts placed on the board. In Sprint 2, they seem to have fun, and communication is more fluid. Although their built model is the simplest in terms of the number of artifacts, it is the best explained.

E. SELF-ASSESSMENT QUESTIONNAIRE

The questions posed in the student self-assessment questionnaire seek to capture the students' perceptions regarding individual performance contrasted with the rest of the group members.

Students generally do not perceive differences between individual performance and the rest of the group members according to the skills consulted (teamwork, collaboration, and communication). 54% declares that the ability to work in a team is *Normal* among the group members. In groups G2 and G3, all students indicated that they did not perceive differences in this ability. In G1, G4, and G5, the perception was more varied. Also in general, the perception of collaboration was mostly *Normal* (46%) and *Influential* (46%), while for effective communication, 67% perceived it as *Influential* or *Higher*. We believe that these results reflect the benefits of the LSP development activities because they promote and stimulate the participation of all its members, generating a positive climate among all the participants. For the same reason, individual perception is clearly influenced by the positive climate of the LSP work environment. The latter is reinforced when asked about the first feeling when working with their colleagues, where 71% declare themselves as *Entertained*. Another aspect we were interested in exploring is whether the use of technologies that measure behavior causes any change in how students work. 83% indicate that the use of technology does not alter their natural behavior, a positive result for Naira's non-invasive lavalier microphones.

Finally, consulted about their comments on the activity, the students mostly pointed out very positive and highly motivated aspects. Some comments are the following: *"Entertaining, educational and improves relationships. Distracts from other responsibilities"*, *"good experience that I would repeat"*, *"very fun and easy to relate to the group"*, *"helps communication with more reserved people"*, *"I think it was very good and an entertaining activity that allowed us to gain confidence with the other classmates"*. Regarding the suggestions, the students consider the time the most critical factor, requesting that the times be more relaxed without so many restrictions. The second important suggestion is to have a larger physical space for the activity.

F. DISCUSSION

Next, we will discuss the results from the point of view of the research questions defined in Section IV-A.

1) RQ1. IS IT POSSIBLE TO IDENTIFY BELBIN'S BEHAVIORAL ROLES THROUGH DATA ANALYSIS FROM A COLLABORATIVE ACTIVITY?

To help answer this question, Section VI-A presents group behavioral assumptions based on the individual Belbin Get-Set Reports, which were then contrasted with the other data analysis results.

Most of the students presented more than one natural role. Some showed a predominantly Social (G22, G25, G33, G34), Thinking (G42, G53), and Action (G12, G14) profile. Other students had natural roles in two profiles, namely, Social/Thinking (G15, G31, G43, G52), Social/Action (G11, G41, G45), and Thinking/Action (G23, G32, G35, G51). Finally, other students presented natural roles in the three profiles (G13, G21, G24, G44, G54). If we look at Figure 8, those students who concentrate the greatest number of interactions are G14, G22, G32, G35, G44, G45, and G51. The common factor among all these students (excepting only G22) is that they have at least one natural role in the Action profile. The latter allows us to answer this research question positively since we identified students' behavioral factors in executing the LSP collaborative activity. From this analysis, we have proposed the following hypothesis, which we will seek to validate in future research.

H1: *People who have only Action natural roles in these kinds of activities develop more speech interactions than others.*

Although the individual analysis was very interesting, our interest is focused on the teams' performance and how they interact, communicate and collaborate to solve a problem. Considering the natural roles of the students (Table 3, Figure 4), the groups were formed in a balanced way by having members in the three profiles. G1 was the only group with at least one student in the nine roles with a percentile above 70. Furthermore, G1 was the only group with two students with four natural roles, two with three natural roles, and one with just one natural role. As noted in Section VI-C, G1 was the group with the highest number of interventions but at the same time had the shortest effective speaking time. In Section VI-A, we state the importance of G15 Thinking roles as a behavioral assumption. In this regard, for G15, it was not possible to identify aspects of these roles in the phases of the LSP activity. Due to the above and the results recorded, the collaborative performance of the group may have been affected by the phenomenon called *Team role sacrifice* [10]. This phenomenon occurs when a person sets aside their natural roles to adopt another more absent role in the group. For G2, G3, G4, and G5, the numbers of natural roles are much lower, so it is more feasible to visualize and identify the behavioral characteristics of the group members. From this, we propose the hypothesis:

H2: *Belbin role behaviors are more likely to be identified in groups with a maximum of two natural roles.*

2) RQ2. DOES AFFINITY SOCIOGRAMS HELP TO IDENTIFY WORK TEAMS WITH BETTER PERFORMANCE?

Based on what was developed in Section VI-B and the observations during the activity, it was possible to see that work teams that were well cohesive by affinity were effective and efficient in task completion. Specifically, the groups with higher values in their centralization metrics and lower standard deviations (see Table 7) better fulfilled the main objective of the LSP activity, according to the teachers' observations. Hence, we can also answer this research question positively. In general, self-organized teams seem to be formed not only considering friendship relationships but also indirectly seeking complementarity in their problem-solving abilities. All this leads us to the following hypothesis statement.

H3: *Self-organized work teams are more balanced in behavioral profiles than non-self-organized ones.*

3) RQ3. IS THERE A RELATIONSHIP BETWEEN BEHAVIORAL ROLES, SPEAKING TIMES, AND THE NUMBER OF SPEECH INTERACTIONS?

Table 8 shows the number of interventions, speaking times (in seconds), and the number of natural roles for each group. An intervention is a speech window of at least 0.7 seconds. This small window is handled since this is a collaborative activity, where several students are expected to speak almost simultaneously.

Note that the group with the most action roles (G1) is the one with the highest number of interactions, while two of the groups with the highest social roles (G2 and G4) are the ones with the longer speaking times. Despite the above, no conclusive relationships are observed. In order to collect more evidence, it will be necessary to design controlled experiments in the future, reducing the limitations identified in Section IV-D regarding data capture with Naira. However, we can suggest the following hypotheses:

TABLE 8. Overall results obtained in NAIRA together with the number of Belbin roles present in each group.

Member	Number of interventions	Speaking time [s]	Natural roles number Social	Thinking	Action
G11	1801	65.1	2	0	2
G12	2074	79.2	0	0	3
G13	2512	96.3	2	1	1
G14	2773	234.7	0	0	1
G15	2011	73.8	1	2	0
Total G1	11171	549.2	5	3	7
G21	189	26.4	2	1	1
G22	980	829.5	1	0	0
G23	515	61.6	0	1	1
G24	435	146.1	1	1	1
G25	632	71.5	2	0	0
Total G2	2751	1135.0	6	3	3
G31	1313	71.9	1	1	0
G32	1989	304.4	0	1	2
G33	1249	67.8	2	0	0
G34	1203	51.9	1	0	0
G35	1624	326.9	0	1	1
Total G3	7378	823.0	4	3	3
G41	1350	205.7	1	0	1
G42	598	146.7	0	1	0
G43	0	0.0	1	1	0
G44	1525	391.4	1	1	2
G45	1131	411.5	2	0	1
Total G4	4604	1155.3	5	3	4
G51	1920	378.0	0	1	1
G52	1510	215.7	1	1	0
G53	1674	192.2	0	1	0
G54	614	130.1	1	1	1
Total G5	5728	916.0	2	4	2

H4: *In LSP activities, the groups with predominant Action profile tend to have higher numbers of interactions.*

H5: *In LSP activities, the groups with predominant Social profile tend to have longer speaking times.*

VII. CONCLUSIONS AND FUTURE WORK

This exploratory work was designed to study the relationship between the Belbin roles of participants in collaborative activities and the social interactions that emerge in the context of problem-solving. We used the LSP methodology in a real case study, incorporating agile practices to stimulate collaborative work. As far as we know, this is the first study that relates Belbin's roles with MMLA and SNA techniques. In addition, it allows expanding other works like [45] about the application of affinity sociograms for forming self-organized collaborative work groups.

Participants with several natural roles tend to be more flexible and can choose their dominant role for the benefit or balance of the group. When a participant with many natural roles shares part of these roles with other less diverse participants, the activity of the first diminishes to the benefit of the expression of the others. Thus, the activity levels in Naira allow detection of participants with few natural roles, for both Action and Social profiles, but mainly Action (possibly due to the type of activity). Therefore, according to this work, we can point out that Naira's data analysis tools provided evidence of the student's behavior with natural Action roles.

The survey results confirm that experimental LSP activities provide significant advantages for observing and evaluating soft skills. A methodology based on LSP with agile practices seems useful for studying collaborative work environments. The roles of the Thinking profile were the most difficult to analyze with Naira. We believe designing a storytelling-type activity could leave traces in Naira that help us identify the characteristics of this profile.

Based on the findings, our next works aim to improve the Naira platform to reduce the effects of the stated limitations. As a first exploratory approach, we considered a case study of 24 students. Of course, to validate the results obtained in more general contexts, the experiments should be extended. Therefore, we plan to replicate this activity with other groups of students to enhance the tracking of evidence according to the behavioral roles. Likewise, the controlled experiment design will help evaluate the new hypotheses raised in this research. A possible future work is in the comparative analysis of teams' performance considering the Belbin methodology with the RBC methodology. Finally, we consider the incorporation of variables that allow differentiating behaviors of people with similar natural roles. Among these variables could be those obtained from the DISC test, which provides valuable information on previous experiences and the analysis of leadership styles that appear in work groups.

REFERENCES

[1] J. Voogt and N. P. Roblin, "A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies," *J. Curriculum Stud.*, vol. 44, no. 3, pp. 299–321, Jun. 2012.

[2] P. Thomson, "Nuevas formas de trabajar en la empresa del futuro," in *Reinventar la Empresa en la Era Digital*. New York, NY, USA: Turner, 2015, pp. 247–263.

[3] J. Acosta, *Trabajo en Equipo*. New Delhi, India: Alfa Omega, 2000.

[4] A. López-Alcarria, A. Olivares-Vicente, and F. Poza-Vilches, "A systematic review of the use of agile methodologies in education to foster sustainability competencies," *Sustainability*, vol. 11, no. 10, p. 2915, May 2019.

[5] J. P. P. Lojano, E. A. L. Cedeño, D. B. P. Zurita, and J. V. M. Osejos, "Agile methodologies in the fundamental stages of industrial design," *Universidad Ciencia y Tecnología*, vol. 25, no. 110, pp. 87–96, Aug. 2021.

[6] K. A. Bruffee, "Collaborative learning and the 'conversation of mankind,'" *College English*, vol. 46, no. 7, p. 635, 1984.

[7] S. Wasserman and K. Faust, *Social Network Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1994.

[8] F. Riquelme, R. Munoz, R. Mac Lean, R. Villarroel, T. S. Barcelos, and V. H. C. de Albuquerque, "Using multimodal learning analytics to study collaboration on discussion groups," *Universal Access Inf. Soc.*, vol. 18, no. 3, pp. 633–643, Aug. 2019.

[9] J. A. R. Guasch, "Análisis de roles de trabajo en equipo: Un enfoque centrado en comportamientos," Ph.D. dissertation, Departamento de Psicología Social, Universitat Autònoma de Barcelona, 2006.

[10] (2022). *Belbin Team Roles*. [Online]. Available: https://www.belbin.com/about/belbin-team-roles

[11] R. M. Belbin, *Management Teams: Why They Succeed or Fail*, 3rd ed. Oxford, U.K.: Butterworth, 2010.

[12] B. Watkins and M. Gibson-Sweet, "Sailing with Belbin," *Educ. Training*, vol. 39, no. 3, pp. 105–110, Apr. 1997.

[13] S. Scherer, N. Weibel, L.-P. Morency, and S. Oviatt, "Multimodal prediction of expertise and leadership in learning groups," in *Proc. 1st Int. Workshop Multimodal Learn. Anal. (MLA)*, Oct. 2012, pp. 1–8.

[14] M. Worsley, D. Abrahamson, P. Blikstein, S. Grover, B. Schneider, and M. Tissenbaum, "Situating multimodal learning analytics," in *Proc. ICLS*, 2016, pp. 1346–1349.

[15] S. Scherer, M. Worsley, and L.-P. Morency, "1st international workshop on multimodal learning analytics," in *Proc. 14th ACM Int. Conf. Multimodal Interact.* New York, NY, USA: ACM Press, Oct. 2012, pp. 609–610.

[16] H. Cornide-Reyes, F. Riquelme, D. Monsalves, R. Noel, C. Cechinel, R. Villarroel, F. Ponce, and R. Munoz, "A multimodal real-time feedback platform based on spoken interactions for remote active learning support," *Sensors*, vol. 20, no. 21, p. 6337, Nov. 2020.

[17] K. Schwaber and J. Sutherland. (2021). *The Scrum Guide 2020*. [Online]. Available: https://scrumguides.org

[18] I. Myers and K. Briggs, *The Myers-Briggs Type Indicator: Manual*. Palo Alto, CA, USA: Consulting Psychologists Press, 1962.

[19] I. Myers, *MBTI Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator*. Palo Alto, CA, USA: Consulting Psychologists Press, 1998.

[20] R. Stein and A. B. Swan, "Evaluating the validity of Myers–Briggs type indicator theory: A teaching tool and window into intuitive psychology," *Social Pers. Psychol. Compass*, vol. 13, no. 2, Feb. 2019, Art. no. e12434.

[21] C. J. Lake, J. Carlson, A. Rose, and C. Chlevin-Thiele, "Trust in name brand assessments: The case of the Myers–Briggs type indicator," *Psychol.-Manager J.*, vol. 22, no. 2, pp. 91–107, May 2019.

[22] H. Zhu, "Role-based collaboration," in *E-CARGO and Role-Based Collaboration: Modeling and Solving Problems in the Complex World*. Hoboken, NJ, USA: Wiley, 2021, ch. 3, pp. 69–102.

[23] H. Zhu, "Predict team performance with E-CARGO," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2018, pp. 3559–3564.

[24] H. Zhu, "Computational social simulation with E-CARGO: Comparison between collectivism and individualism," *IEEE Trans. Comput. Soc. Syst.*, vol. 7, no. 6, pp. 1345–1357, Mar. 2020.

[25] H. Zhu, "E-CARGO and role-based collaboration," in *Proc. IEEE 24th Int. Conf. Comput. Supported Cooperat. Work Design (CSCWD)*, May 2021, p. 3.

[26] R. M. Belbin, *Team Roles at Work*, 2nd ed. London, U.K.: Elsevier, 2010.

[27] O. Salinas, F. Riquelme, R. Munoz, C. Cechinel, R. Martinez, and D. Monsalves, "Can analytics of speaking time serve as indicators of effective team communication and collaboration?" in *Proc. X Latin Amer. Conf. Hum. Comput. Interact.*, Nov. 2021, pp. 1–4.

[28] F. Riquelme, R. Noel, H. Cornide-Reyes, G. Geldes, C. Cechinel, D. Miranda, R. Villarroel, and R. Munoz, "Where are you? Exploring micro-location in indoor learning environments," *IEEE Access*, vol. 8, pp. 125776–125785, 2020.

[29] R. Noël, D. Miranda, C. Cechinel, F. Riquelme, T. T. Primo, and R. Munoz, "Visualizing collaboration in teamwork: A multimodal learning analytics platform for non-verbal communication," *Appl. Sci.*, vol. 12, no. 15, p. 7499, Jul. 2022.

[30] Y. Song, L.-P. Morency, and R. Davis, "Learning a sparse codebook of facial and body microexpressions for emotion recognition," in *Proc. 15th ACM Int. Conf. Multimodal Interact.* New York, NY, USA: ACM Press, Dec. 2013, pp. 237–244.

[31] F. Vieira, C. Cechinel, V. Ramos, F. Riquelme, R. Noel, R. Villarroel, H. Cornide-Reyes, and R. Munoz, "A learning analytics framework to analyze corporal postures in students presentations," *Sensors*, vol. 21, no. 4, p. 1525, Feb. 2021.

[32] R. Noel, F. Riquelme, R. M. Lean, E. Merino, C. Cechinel, T. S. Barcelos, R. Villarroel, and R. Munoz, "Exploring collaborative writing of user stories with multimodal learning analytics: A case study on a software engineering course," *IEEE Access*, vol. 6, pp. 67783–67798, 2018.

[33] A. Datta, J. Tan Teck Yong, and A. Ventresque, "T-RecS: Team recommendation system through expertise and cohesiveness," in *Proc. 20th Int. Conf. companion World Wide Web*. New York, NY, USA: ACM Press, Mar. 2011, pp. 201–204.

[34] V. S. Sharma and V. Kaulgud, "Studying team evolution during software testing," in *Proc. 4th Int. Workshop Cooperat. Hum. Aspects Softw. Eng.*, May 2011, pp. 72–75.

[35] N. A. Batista, M. A. Brandão, G. B. Alves, A. P. C. da Silva, and M. M. Moro, "Collaboration strength metrics and analyses on GitHub," in *Proc. Int. Conf. Web Intell.*, Aug. 2017, pp. 170–178.

[36] C. Amrit, "Coordination in software development," *ACM SIGSOFT Softw. Eng. Notes*, vol. 30, no. 4, pp. 1–7, 2005.

[37] K. Ehrlich and M. Cataldo, "The communication patterns of technical leaders," in *Proc. 17th ACM Conf. Comput. Supported Cooperat. Work Social Comput.*, Feb. 2014, pp. 733–744.

[38] A. Boone, M. Roelants, K. Hoppenbrouwers, C. Vandermeulen, M. Dubois, and L. Godderis, "Teamwork in medical education: Are medical students team workers?" *Saf. Health Work*, vol. 13, p. S103, Jan. 2022.

[39] D. Tamayo Avila, W. Van Petegem, and M. Snoeck, "Improving teamwork in agile software engineering education: The ASEST+ framework," *IEEE Trans. Educ.*, vol. 65, no. 1, pp. 18–29, Feb. 2022.

[40] A. Aranzabal, E. Epelde, and M. Artetxe, "Team formation on the basis of Belbin's roles to enhance students' performance in project based learning," *Educ. Chem. Engineers*, vol. 38, pp. 22–37, Jan. 2022.

[41] R. Aguilar-Vera, J. Díaz-Mendoza, and J. Ucán-Pech, "Influencia de la teoría de roles de belbin en la medición de software: Un estudio exploratorio," *RISTI—Revista Ibérica de Sistemas e Tecnologias de Informação*, no. 31, pp. 50–65, 2019.

[42] M. Godskesen, "Improving teamwork abilities across cultural differences: Belbin group role theory applied," in *Proc. 9th Int. Workshop Act. Learn. Eng. Educ.*, Barcelona, Spain, 2009, pp. 1–10. [Online]. Available: http://upcommons.upc.edu/revistes/handle/2099/7810

[43] L. Gutiérrez, V. Flores, B. Keith, and A. Quelopana, "Using the Belbin method and models for predicting the academic performance of engineering students," *Comput. Appl. Eng. Educ.*, vol. 27, no. 2, pp. 500–509, Mar. 2019.

[44] F. Gander, I. Gaitzsch, and W. Ruch, "The relationships of team role- and character strengths-balance with individual and team-level satisfaction and performance," *Frontiers Psychol.*, vol. 11, Nov. 2020, Art. no. 566222.

[45] J.-M. Flores-Parra, M. Castafon-Puga, R. D. Evans, R. Rosales-Cisneros, and C. Gaxiola-Pacheco, "Towards team formation using belbin role types and a social networks analysis approach," in *Proc. IEEE Technol. Eng. Manage. Conf. (TEMSCON)*, Jun. 2018, pp. 1–6.

[46] (2022). *Belbin GetSet*. [Online]. Available: https://www.belbin.es/getset/

[47] W. Huang, S.-H. Hong, and P. Eades, "How people read sociograms: A questionnaire study," in *Proc. Asia–Pacific Symp. Inf. Vis.*, vol. 60, 2006, pp. 199–206.

[48] X. Molinero, F. Riquelme, and M. Serna, "Cooperation through social influence," *Eur. J. Oper. Res.*, vol. 242, no. 3, pp. 960–974, May 2015.

[49] P. Runeson and M. Höst, "Guidelines for conducting and reporting case study research in software engineering," *Empirical Softw. Eng.*, vol. 14, no. 2, pp. 131–164, Apr. 2009.

[50] M. Romero and O. T. Gebera, "Serious games para el desarrollo de las competencias del siglo XXI," *RED—Revista de Educación a Distancia*, no. 34, pp. 1–22, 2012.

[51] J. Roos, B. Victor, and M. Statler, "Playing seriously with strategy," *Long Range Planning*, vol. 37, no. 6, pp. 549–568, Dec. 2004.

[52] A. Ponce-Sandoval, D. Monsalves, F. Riquelme, and H. Cornide-Reyes, "MMLA approach to analyze collaborative work in lego serious play activities," in *Proc. 40th Int. Conf. Chilean Comput. Sci. Soc. (SCCC)*, Nov. 2021, pp. 1–8.

DIEGO MONSALVES is currently pursuing the Ph.D. degree in applied computer engineering with the Universidad de Valparaíso. He is also a Computer Engineer with the Universidad de Valparaíso. His research interests include data analysis and multimodal learning analytics.

HÉCTOR CORNIDE-REYES received the master's and Ph.D. degrees in informatics engineering from the Pontificia Universidad Católica de Valparaíso. He is currently an Assistant Professor with the School of Computer Engineering, Universidad de Atacama. His research interests include software engineering education, learning analytics, and agile methods.

FABIÁN RIQUELME received the M.Sc. degree from the Universidad de Concepción, Chile, and the Ph.D. degree in computing from the Universitat Politècnica de Catalunya, Spain. He is currently the Research and Innovation Coordinator with the Faculty of Engineering, Universidad de Valparaíso, Chile. His research interests include social network analysis, social computing, decision systems, and data analysis.