University of Reading

Human resource allocation to multiple projects based on members' expertise, group heterogeneity and social cohesion

Article

Accepted Version

Ballesteros-Pérez, P., Phua, F. T. T. and Mora-Melià, D. (2019) Human resource allocation to multiple projects based on members' expertise, group heterogeneity and social cohesion. Journal of Construction Engineering and Management, 145 (2). ISSN 0733-9364 doi: https://doi.org/10.1061/(ASCE)CO.1943-7862.0001612 Available at http://centaur.reading.ac.uk/78762/

To link to this article DOI: http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0001612

Publisher: American Society of Civil Engineers

Human resource allocation to multiple projects based on members' expertise, group heterogeneity and social cohesion

Pablo Ballesteros-Pérez, Ph.D.[1*]; Florence Ting Ting Phua, Ph.D.[2]; Daniel Mora-Melià, Ph.D.[3]

Abstract

Project managers regularly allocate human resources to construction projects. This critical task is usually executed by fulfilling the minimum project staffing requirements normally based around the quantity and competence of project members. However, research has shown that team performance can increase by up to 10% and 18%, respectively, as a consequence of the group members' heterogeneity and social cohesion. Also, there is currently no practical quantitative tool which incorporates these aspects to allow project managers to achieve this task efficiently and objectively.

A new quantitative model for the effective allocation of human resources to multiple projects, which takes into account group heterogeneity and social cohesion is proposed. This model is easy to build, update and use in real project environments with the use of a spreadsheet and a basic optimization engine (e.g. Excel Solver). A case study is proposed and solved with a Genetic Algorithm to illustrate the model implementation. Finally, a validation example is provided to exemplify how group heterogeneity and social cohesion condition academic achievement in an academic setting.

[1] Senior Lecturer. School of Architecture, Building and Civil Eng., Loughborough University, Loughborough, Leicestershire LE11 3TU, UK. *Corresponding author: +44 (0) 150 922 3771, p.ballesteros-perez@lboro.ac.uk

[2] Associate Professor. School of the Built Environment. University of Reading, Whiteknights, Reading RG6 6DF, UK. Phone: +44 (0) 118 378 8194, f.phua@reading.ac.uk

[3] Associate Professor. Departamento de Ingeniería y Gestión de la Construcción, Universidad de Talca, Facultad de Ingeniería, Curicó, camino Los Niches km.1 3340000, Chile. Phone: +56 951 265 044, damora@utalca.cl

Keywords

Human resource management; human resource allocation; team performance; heterogeneity; faultlines; cohesion; sociometry

Introduction

It is a well-known saying that ‘people are the lifeblood of organizations’. Indeed, despite living in an era of constant technological advancement, most of our tasks are still done, handled or supervised by human beings. In organization life, the size and/or complexity of many undertakings nowadays demand the involvement of many people (sometimes from different organizations) working together to achieve a common goal. This goal can be anything, but many times involves creating deliverables (products, services) to enhance a company’s internal performance, to make profit, or both. However, people (employees, workers) who take part in these undertakings are normally subject to constraints. For example, they are qualified to do certain jobs and not others; they have different levels of competence in different domains; they cannot be present in multiple locations; and, certainly, they have physical constraints in terms of how long they can work for (Hendriks et al. 1999).

Therefore, when there are several, sometimes concurrent projects that require the participation of people to be completed, a project manager faces a practical dilemma: how to best allocate his/her human resources on-hand to deliver his/her projects successfully. ‘Successfully’ can mean completing the projects on time, on budget and within an agreed (or shared) quality threshold, or just meeting the key stakeholders’ expectations (Xia et al. 2017). In any case, as long as there are ongoing projects, the project manager will require competent human resources to engage in certain tasks for a period of time before they are freed and able to join other ongoing or upcoming projects. An essential part of the project manager's role in the allocation of optimum human resources is to ensure, as much as possible, that the individuals within the projects can work cooperatively with each other (Anvuur and Kumaraswamy 2016).

However, collaboration between project members does not happen by chance. There are indeed many factors that prevent this from happening. These factors can be communication-related for instance, and/or have to do with the project member's demographic attributes such as (differences in) nationality, education, religion, experience, to cite a few (Al-Bayati et al. 2017). Sometimes, there are people who do not like working with certain individuals, and this can also be really detrimental to the project progress and its eventual success (Chen et al. 2017b). In this regard, Phua (2004) and Phua and Rowlinson (2004) have found that cooperative behavior between project members is influenced, to a certain extent by individual members' intrinsic social and psychological factors which have to do with many more factors other than just their extrinsic demographic profile such as age, sex, education, work experience and roles. For this reason, we will consider both cohesion and heterogeneity factors later when aiming to build high-performing teams.

Given our existing understanding of the various factors that affect team performance, there is however, a scarcity of quantitative and objective tools that enable the effective allocation of human resources in terms of where and when they are to be allocated to projects (Ahmadian Fard Fini et al. 2017). Conventionally, this type of allocation issues largely fall within mainstream Human Resource Management (HRM) application which has its roots in social sciences.

A different, maybe opposite, scenario can be found within Operational Research (OR), which deals with the modeling and application of advanced analytical methods to make better decisions. The problem of allocating multiple human resources to a single project is relatively recent in OR, but it has been well studied and is known nowadays as the 'Team Formation Problem' (TFP) (Tseng et al. 2004). When there are multiple simultaneous projects, the TFP becomes the 'Multiple Team Formation Problem' (MTFP). Particularly, the grouping of individuals to create teams have been made by attending to multiple factors: the resources' temporal availability, current workload, individuals' skills, level of competence, geographical distance, seniority, number of contacts, among many others (Gutiérrez et al., 2016). In this line of research, it is not common to find theoretically-grounded sociological considerations in the composition of teams. This means that, whereas it is relatively easy to come across OR models that allocate resources that meet some functional (e.g. skills, competence) project members' requirements, it is very rare to find models that try to optimize other socially-based group traits like intra-group social preferences and group cohesion (Ballesteros-Pérez et al. 2012). This piece of research proposes to take a step forward in bridging this gap.

In this paper, a new human resource allocation model that takes into account, not just basic project staff requirements and employees' profiles, but also group heterogeneity (diversity) and social cohesion, is developed. This is a worthwhile contribution because, as discussed earlier, team performance has been demonstrated to be significantly influenced by these two factors. Hence, it seems logical to incorporate this knowledge when creating high-functioning teams which comprise the 'right' individuals working together. To this end, the rest of the paper will be structured as follows. The *literature review* section will go over the major contributions published in the areas of the MTFP, group heterogeneity and social cohesion. The *materials and methods* section will formulate the model, define its major variables and explain how these are interrelated under mathematical expressions for measuring team performance. A *case study* will exemplify the model implementation in a fictitious company environment with twenty people and three simultaneous projects. A short *validation* section will implement the model in a real academic setting where a cohort of 15 MSc students worked in groups to deliver three projects. The *discussions* will provide some insight and further analysis on the implications and limitations of the model. Finally, the *conclusions* will summarize the paper and convey why the proposed tool is relevant to the wider project management community.

Literature review

The proposed model draws from research developed in two very different areas – operational research (OR) and applied psychology (AP) –, but it is applied on a third one: Human Resource Management (HRM). The amount of works published in connection with HRM within both OR and AP is endless, so it is necessary to narrow down significantly the works to be presented here. In this regard, only three very relevant topics will be reviewed: the MTFP, group heterogeneity and faultlines, and group cohesion and sociometry.

The Multiple Team Formation Problem (MTFP)

The MTFP involves the distribution of people with different skillsets to a series of teams (projects) that usually require more than a single area of expertise while optimizing other criteria (e.g. profits, execution time, number of people). This problem is known to be NP-hard (Non-deterministic Polynomial-time Hard) even for instances with a single project (Gutiérrez et al., 2016). This means the MTFP belongs to the set of OR problems that are harder to solve.

The first attempt to model and compute a solution to the TFP is relatively recent and was developed by Lappas et al. (2009) when trying to create teams of experts from professional profiles posted on social networks. Just a year later, Dorn and Dustdar (2010) proposed solving the TFP with a first heuristic approach, whereas Li and Shan (2010) improved the Enhanced-Steiner algorithm that was one of the two original algorithms used to solve the TFP.

A year later, Yin et al., (2011) were the first to consider social influence among the teams of experts. Additionally, Farhadi et al. (2011) allowed for the possibility of different competence levels among the human resources, a generalization that will also be considered in our model.

In 2012, the number of works published on the TFP grew exponentially. Among the most relevant: Sorkhi et al. (2012) proposed a game theoretic approach to form and rank project teams; Farhadi et al. (2012a, 2012b) extended the second original algorithm that had proven to be very effective when dealing with the TFP – the Rarest First algorithm –; whereas Gajewar and Sarma (2012) proposed three new optimization algorithms and successfully applied them to the MTFP for the very first time.

Next, Shi and Hao (2013) formulated the MTFP with a multi-criteria decision-making ranking approach involving the individuals’ social networks. Then, Teixeira and Huzita (2014) approached the MTFP considering the human resources' contextual information (culture, idiom, temporal distance and previous experience), besides task requirements and the interpersonal relationships among human resources. Our proposed model will also take advantage of similar constructs in order to create a multi-dimensional model. Also, Agrawal et al. (2014) focused on educational settings allowing the MTFP to be implemented without allowing overlaps between the different student teams, a feature that will also be considered in our model. Still in the same year, Awal and Bharadwaj (2014) tried to capture the synergy produced among team members by means of a new ad-hoc concept named ‘Collective Intelligence’ and also used a Genetic algorithm to solve their problem formulation. In this paper, the solution of the case study proposed later will also make use of a genetic algorithm approach as the way this ad-hoc index was defined share some similarities with our objective function.

Although there have been many other recent works published on the MTFP, these will not be recounted here as they are not directly germane to this study. However, one that is perhaps worth highlighting is the work from Gutiérrez et al. (2016) which formally included sociometric preferences among individuals in the MTFP. Our proposed model also shares a similar approach for modeling group cohesion. However, the algorithmic approach will be totally different to Gutiérrez et al.'s as our model includes other dimensions, which makes our model no longer quadratic.

Group heterogeneity and faultlines

Research on how team effectiveness is influenced by the team composition has been abundant too. Most of this research has focused precisely on measuring and analyzing the effects of group heterogeneity on team performance. Group heterogeneity (homogeneity) refers to a measurement of how different (similar) the members' demographic attributes (age, sex, ethnicity, etc.) are with each other. There are many reviews on group heterogeneity (see Earley and Gibson (2002) for a comprehensive one) but they will not be recounted here either. In this piece of research, we are focusing on the quantitative aspects of how heterogeneity is measured and what are its effects on team performance, rather than the mechanisms or factors that cause it.

With this in mind, the first indices that captured quantitatively how diverse (homogeneous/heterogeneous) a group can be were defined by Blau (1977) and Allison (1978). Generally, these and other later indices involved measuring group homogeneity as the members' demographic attribute overlaps. With those indices, heterogeneity was also generally defined as the inverse of homogeneity, that is heterogeneity=1/homogeneity.

Additionally, for a long time, it was believed that the presence of faultlines (demographic features that divide a bigger group into two or more relatively homogeneous subgroups) was detrimental to group performance (Lau and Murnighan 2005). It was not until the work of Gibson and Vermeulen (2003), who proposed a new metric for measuring group heterogeneity – the Subgroup Strength – , that it was understood that the presence of subgroups (faultlines) could indeed promote team learning behavior and improve their performance. The Subgroup Strength (SS) has many advantages over previous homogeneity metrics (indices) as it allowed researchers to identify group faultlines much more effectively. Indeed, it was shown recently by Meyer and Glenz (2013) in a comprehensive comparative study that the SS is one of the simpler, yet more powerful metrics for measuring group heterogeneity in the presence of two or more subgroups. For these reasons, SS will also be used in our model later to describe subgroups' heterogeneity.

Finally, Gibson and Vermeulen (2003) also showed that a team's performance seemed to vary by up to 10% depending on the SS. Again, this was supported recently by another study by Chen et al. (2017). This study also confirmed another speculation of Gibson and Vermeulen's: that the relationship between SS and team performance was an inverted U-shape whose minima (lower performance) were to be expected for extremely homogeneous and heterogeneous groups. Finally, many other works have been published on the effects of group heterogeneity on intra- and cross-subgroups demographic faultlines (Lau and Murnighan 2005), but only some related to group cohesion will be reviewed later in the *Discussions* to clarify the effect of possible collinearities between both variables.

Group cohesion and sociometry

Group cohesion is a desirable attribute because research has proven it to be positively related to team performance, as well as a wide range of other positive behavioral outcomes (better individuals' attitude, well-being, lower absenteeism, etc.) (Chang and Bordia 2001; Chen et al. 2017b). However, very few pieces of research have actually quantified the extent to which team performance is influenced by group cohesion or dissociation.

One exception is a recent and comprehensive review performed by Evans and Dion (2012). These authors, beyond concluding that there is a positive relationship between cohesion and performance, recounted that cohesive groups seem to perform around 18 percentile points on average above the average (uncohesive) groups. This figure will be used later in our model as other research has also corroborated the cohesion-performance relationship even when different settings (e.g. business, education, research) or group sizes are considered (Castaño et al. 2013). Furthermore, because existing research on cohesion and performance has operationalized cohesion almost completely in terms of interpersonal attraction (see evidence from Lott and Lott (1965) to Beal et al. (2003) for instance), it makes theoretical sense for our model to adopt sociometry to model group cohesion.

Sociometry was devised by Jacob Levy Moreno (Moreno 1941) and is a method that can be used for estimating the quality of group dynamics. It is one of the few methods that allows the gathering of quantitative information about the informal structure of a group that is difficult to obtain in other ways. Sociometry was extensively used between the 40s and 60s at schools, companies and research settings to examine social interrelations and communication patterns within groups (Salo 2006). In sociometry, interpersonal relations are measured by asking group members to express their preferences and rejections for particular companions in a certain situation or activity (Festinger et al. 1950). Hence, the advantageous simplicity of sociometry is, at the same time, its major limitation: it requires that group members are truthful and open in stating who they prefer and not prefer to work with. A reasonable question then is whether group cohesion can be adequately represented by sociometric choices and if these choices r can be eventually captured by means of questionnaires that request group members to state their preferences and rejections towards other group members. In fact, both aspects have been subjected to multiple research studies in many varied settings. An example of a brief but reassuring and confirmatory review can be found in Salo (2006).

Finally, there is one question that needs to be addressed before formulating the model. As stated earlier, the proposed model will group individuals under different projects that have some minimum staff (areas of expertise and levels of competence) requirements. According to a recent piece of research (Mathieu et al. 2015), when people with the right combination of expertise work together, as expected, this is positively related with team performance. However, this same piece of research also showed that this is unrelated to team cohesion. With this in mind, we will allow our model to effectively separate the effect of the constraints (i.e. minimum project staffing requirements) from the group performance variables (i.e. group heterogeneity and cohesion metrics).

Materials and methods

Model outline

In this section, an OR model that allocates a pool of skilled individuals to a series of simultaneous projects with specific staffing (expertise and competence) requirements is proposed and mathematically described in detail. This model will take into account how similar (homogeneous) these individuals are and how they get along with each other (group cohesion).

Mathematical notation

Let us assume two individuals *i* and *j* where *i, j* belong to a set of *n* people (workers) who are available to be allocated into teams. Let us assume that these individuals can be combined into a number of non-overlapping teams (subgroups) where each team is noted by the letter *k* and whose size is noted as n_k (number of members of team *k*).

For every individual *i* (or *j*) it is assumed that the following information is known as illustrated in the following examples:

- Professional level of competence l_i where $l_i \in L$ and L={junior, intermediate, senior}
- Functional department d_i where $d_i \in D$ and D={architecture, civil, mechanical, electrical}
- Age a_i where a_i = positive integer.
- Gender g_i where $g_i \in G$ and G={Male, Female}
- Ethnicity e_i where $e_i \in E$, and where *E*, for simplicity, will be assumed here as the continent of origin, that is E={African, Antarctican, Asian, Australian, European, North American, South American}
- Team tenure (seniority in the same group or company) t_i where t_i = positive integer.
- Sociometric preference of individual *i* towards individual *j*, that is s_{ij} where $i \neq j$, $s_{ij} \in S$ and $S=\{-1, 0, +1\}$. Particularly, $s_{ij}= -1$ means *i* dislikes working with *j*, $s_{ij}= 0$ means *i* is neutral towards (or has never worked with) *j*, and $s_{ij}= +1$ means *i* likes working with *j*. The set of all values s_{ij} correspond to a non-symmetrical matrix of size $n \times n$.

Sociometric preferences aside, these individuals' attributes have been selected here as they were the ones adopted by Gibson and Vermeulen (2003) in their seminal work on group faultlines. This set of attributes has been widely tested in subsequent research (e.g. Chen et al. 2017a; Meyer and Glenz 2013) and it is still largely accepted that they provide a robust representation of group diversity.

Hence, given n people available from whom we know their l_i, d_i, a_i, g_i, e_i, t_i and s_{ij}, we will create subsets (subgroups/teams) of n_k individuals, each of which will be working on a different project k. Individuals can only be allocated to either a single group k or no subgroup at all (those unallocated individuals will be idle resources). This implies that no individual can be present in two or more subgroups, even if they could only work part-time in several projects. We use this simplified assumption to make this model more accessible from the point of view of its first mathematical formulation.

Therefore, as implied above, every subgroup k will be allocated to a single project and we will note projects and subgroups (teams) with the same subscript k from now on. Each project k will have specific staffing requirements (p_k). For instance, p_k={1 senior Architect, 1 intermediate civil engineer, 1 junior civil engineer, 2 intermediate electrical engineers}. Any subgroup of workers n_k that matches or exceeds (both in number and/or competence) these requirements will be considered a feasible subgroup that can potentially be allocated to project k.

Team performance measurement

In order to determine which feasible allocation of subgroups is most desirable, it is necessary to anticipate how much better each possible alternative allocation of subgroups would perform if eventually chosen. Additionally, it is worth emphasizing that each feasible allocation might encompass multiple subgroups as each subgroup will be allocated to one project. Therefore, it is necessary to create an index that captures, not just how efficient each subgroup is, but also how efficient all groups are on average; that is, how efficient the allocation is altogether. This index will be named ‘Global Efficiency (E)’ and will correspond to a weighted average calculated from the subgroup Efficiencies of each subgroup k (noted as E_k), that is:

$$E = \sum E_k \cdot w_k \tag{1}$$

In this expression, w_k corresponds to the weight of each subgroup k. This way w_k can be calculated, for instance, proportionally to each project k’s budget (b_k). Alternatively, w_k can also be calculated proportionally to the number of people n_k from each project, divided by the total people available n (allocated or not) or the total number of allocated people only (Σw_k). These alternatives are expressed in equations (2) and (3), respectively:

$$w_k = \frac{b_k}{\sum b_k} \tag{2}$$

$$w_k = \frac{n_k}{\sum n_k} \quad \text{or} \quad w_k = \frac{n_k}{n} \tag{3}$$

With the global (allocation) Efficiency E defined in (1) as a function of each subgroup’s E_k and w_k values, now it is necessary to detail how E_k values can be calculated.

E_k is a composite efficiency index obtained as the product of two other indices that represent the expected performance of that subgroup k in terms of its homogeneity (P_k^{SS}) and social cohesion (P_k^{S}). Namely,

$$E_k = P_k^{SS} \cdot P_k^{S} \tag{4}$$

Particularly, P_k^{SS} estimates the Performance of a subgroup k based on the Subgroup Strength (SS) as defined by Gibson and Vermeulen (2003). To calculate the SS value of a subgroup k (noted as SS_k), it will be necessary to calculate the subgroup k’s homogeneity value h_k first, as well as the individuals’ degree of overlaps in terms of different diversity factors (we will use: functional department, age, gender, ethnicity, and team tenure as justified later).

P_k^S is an index that measures the differential level of performance expected for subgroup *k* given a particular level of cohesion, which is measured by sociometric indices. In this case, S_k will be calculated as the interpersonal social preferences and rejections stated by all members belonging to subgroup *k*.

What follows are the details on how P_k^{SS} and P_k^S are calculated. Once these two values are known for each potential subgroup *k*, obtaining E_k will be straightforward with (4).

Let us start with P_k^{SS}, the performance metric coming from the Subgroup Strength metric. Conventionally, a subgroup *k*'s homogeneity h_k has been defined as:

$$h_k = \frac{\sum_{i<j} \left\{ O_{ij}^d + O_{ij}^a + O_{ij}^g + O_{ij}^e + O_{ij}^t \right\}}{\frac{n_k(n_k - 1)}{2}} = \frac{\sum_{i<j} O_{ij}}{\frac{n_k(n_k - 1)}{2}} \quad (5)$$

Where O_{ij} is the total overlap between individuals *i* and *j*, and which is computed as the sum of O_{ij}^d, O_{ij}^a, O_{ij}^g, O_{ij}^e and O_{ij}^t which, in turn, represent the overlaps between two individuals *i* and *j* on functional department, age, gender, ethnicity and team tenure, respectively. The sum in the numerator is restrained to $i<j$ (but it could have also been $i>j$ indistinctly) to avoid the cases where $i=j$ (individuals' self-overlaps) as well as to prevent the symmetrical O_{ij} values (that is $O_{ij}=O_{ji}$) from being counted twice.

Also in the same vein, the factor $n_k(n_k-1)/2$ in the denominator of (5) corresponds to the total number of pairs analyzed (all possible combinations of *i* and *j*, excluding those cases where $i \geq j$).

With all this in mind, and according to Gibson and Vermeulen (2003), the different overlaps between a group of individuals can be calculated as follows:

Functional department overlap: $O_{ij}^{d} = 1$ if $d_i = d_j$, else 0 (6)

Age overlap: $$O_{ij}^{a} = \frac{\min(a_i, a_j) - 18}{\max(a_i, a_j) - 18} \tag{7}$$

Gender overlap: $O_{ij}^{g} = 1$ if $g_i = g_j$, else 0 (8)

Ethnicity overlap: $O_{ij}^{e} = 1$ if $e_i = e_j$, else 0 (9)

Team tenure overlap: $$O_{ij}^{t} = \frac{\min(t_i, t_j)}{\max(t_i, t_j)} \tag{10}$$

Overlap values can vary between [0, 1]. Hence, values of h_k will vary between [0, 5]. We are aware that other diversity factors could have also been included in the definition of h_k such as for example, language, education, experience. However, in the interest of keeping to the model's simplicity and for illustrative purpose in the case study which follows, we deemed it reasonable to stick to the diversity factors in the definition of h_k as proposed by Blau (1977) and Allison (1978).

And now that the overlaps of all individuals O_{ij} and the subgroup k's homogeneity value h_k have been detailed, the Subgroup Strength of a subgroup k (SS_k) is defined as the population standard deviation of the O_{ij} values from all n_k members belonging to subgroup k, that is:

$$SS_k = \sqrt{\frac{\sum_{i<j} \left(O_{ij} - h_k\right)^2}{\frac{n_k(n_k - 1)}{2}}} = Std.Dev.O_{ij} \tag{11}$$

As defined, SS_k will vary from 0 to 1.25 (since h_k domain was restricted to [0,5]). Additionally, Gibson and Vermeulen (2003) proved that team diversity (represented by means of SS_k) and group performance were quadratically related (inverted U-shape) approximately as described in Figure 1a.

<Insert Figure 1 here>

Also, a recent study by Chen et al. (2017) suggested that this quadratic expression is quasi-symmetrical and that the value of δ seems generally close to 10% on average. Therefore, the subgroup k's performance P_k^{SS} can be calculated from the subgroup strength SS_k value as:

$$P_k^{SS} = -2.56\,\delta\, SS_k^2 + 3.2\,\delta\, SS_k + 1 - \delta \qquad \text{with } \delta \approx 0.10 \tag{12}$$

Expression (12) can vary between [1–δ, 1] and is obtained from a quadratic polynomial which is forced to cross the points: (0,1–δ), (1.25/2, 1) and (1.25,1–δ).

On the other hand, the subgroup k's cohesion-related performance index P_k^S is calculated from the subgroup k members' sociometric preferences s_{ij} towards each other (i.e. preferences and rejections to work with a particular individual). These preferences and rejections do not have to be symmetrical (that is, $S_{ij} \neq S_{ji}$ or $S_{ij} = S_{ji}$). Hence, we define a subgroup k's cohesion S_k as:

$$S_k = \frac{\sum_{i \neq j} s_{ij}}{n_k (n_k - 1)} \tag{13}$$

Similarly, the term $n_k(n_k - 1)$ corresponds to the total number of pairs analyzed excluding the choices of individuals with themselves. So, as s_{ij} can be equal to –1 (meaning i dislikes j), 0 (i is neutral or have not met j), or +1 (i likes j), S_k actually represents how well (or badly) all subgroup k's members get along with each other on average. Analogously, S_k can take on values within the range [-1, 1].

Finally, previous researchers' results suggest that the average cohesive group seems to perform around 18% better than average (non-cohesive or non-uncohesive) groups (Evans and Dion 2012). For the purpose of this paper, this performance differential will be called φ . However, it is worth pointing out that in those previous pieces of research it is not always clear how group cohesion is measured or quantified. Also, there is a total absence of studies clarifying whether the cohesion-performance relationship is linear or if it indeed follows a different pattern. In light of this, it seems prudent to take the simplest alternative and assume that group cohesion (represented now by S_k) and performance (P_k^S) will just be linearly related as represented in Figure 1b. Hence:

$$P_k^S = 1 + \varphi S_k \qquad \text{with } \varphi \approx 0.18 \tag{14}$$

After defining expression (13), all variables involved have been presented and related to each other. We are now able to calculate the global group efficiency (E) from the different simultaneous subgroups' efficiencies E_k and their respective weights w_k. This is summarized at the bottom of Figure 1. Hence, from now on, every possible subgroups' allocation can be measured in relative performance terms and each feasible complete group allocation can be compared against each other. The following is an example to illustrate how we can apply the model based on a fictitious case study which reflects as much as possible, a real project environment.

Application

The case study follows approximately the same order of calculations that was presented in the previous section. For the interested reader, the complete step-by-step calculations can be found in an Excel file accessible from the *Supplemental Online Material*.

Particularly, this case study comprises a group of 20 individuals with different levels of professional competence and who belong to four functional departments. The entire professional, demographic (homogeneity-related) and sociometric (cohesion-related) information from the 20 individuals is described in Figure 2.

<Insert Figure 2 here>

These 20 individuals are to be allocated to three simultaneous projects, each of which has different staffing (professional level and functional department) requirements as well as budgets. This information is detailed in Figure 3. Also, at the bottom of Figure 3, the weights of each project have been calculated as a function of the project budgets according to expression (2).

<Insert Figure 3 here>

Any subgroup allocation that meets or exceeds the staffing requirements described in Figure 3 will be a feasible solution. What is necessary now is to be able to calculate the global efficiency E from any feasible grouping solution. With this aim in mind, the first step will be to obtain all individuals' overlaps concerning functional department, age, gender, ethnicity and team tenure, so that these values can be reused anytime when two individuals are put together in the same subgroup. Due to its length (5 tables, one per type of overlap), these calculations have been included as *Supplemental Online Material*. Figure 4 just summarizes the total overlaps, which have been obtained from the simple addition of all the individuals' overlap values.

<Insert Figure 4 here>

What remains is calculating, for any potential and feasible subgroup k, its homogeneity h_k (with expression (5)), subgroup strength SS_k (with expression (11)), and cohesion S_k (with expression (13)) values. Then, with SS_k and S_k, known, calculating the homogeneity-related P_k^{SS} performance metric (with expression (12)) and the cohesion-related P_k^{S} performance metric (with expression (14)) can be performed. Next, with P_k^{SS} and P_k^{S} known, we can calculate E_k (by means of expression (4)). Once the values of E_k are all known for all the simultaneous subgroups (three in our example, as there are three projects), and by knowing the weight of each subgroup w_k (with expressions (2) or (3), and as detailed at the bottom of Figure 3), it is possible to obtain the global efficiency E of that group configuration by means of expression (1). This series of calculations are represented vertically from top to bottom in the lower half of Figure 5.

<Insert Figure 5 here>

In Figure 5, a random solution directly allocating the individuals available to meet the project staffing requirement, but without any further (homogeneity, nor cohesion) considerations, is presented. At the top of Figure 5, one can find the allocation of each individual to each subgroup/project. The column to the right sums each individual's allocations and verifies that no individual is allocated more than 100%, that is, to more than one project. These are necessary but not sufficient problem constraints which need to be met to qualify any allocation as feasible.

However, every time there is any change (for instance a member is allocated to a different project/subgroup), all values need to be recalculated. Therefore, the only way of finding good solutions is by iterating these calculations multiple times while testing as many feasible solutions as possible. Hence, for the model to be practically useful, this task must be automated by an optimization algorithm.

The global efficiency of the random solution depicted in Figure 5 corresponds to 0.981. By definition, the value of the global efficiency E will vary between [1-δ-φ, 1+φ]. Hence, the closer the value of E to 1+ φ, the higher the expected groups' performance. Now, in looking for that optimum solution, there is one last point to be discussed.

As discussed earlier, the proposed model falls within a particular case of the MTFP. The MTFP is NP-hard, and that opens the door to the use of metaheuristics when looking for (near) optimum solutions. Among common metaheuristics, Genetic Algorithms (GA) are one of the quicker and simpler, but also effective, options. On top of that, they are usually available within commercial solvers included by default in spreadsheet software like Microsoft Excel ®. Moreover, GA have been in use in resource-constrained allocation for a long time and they have been considered as one of the most efficient metaheuristics when dealing with the MTFP in recent studies (e.g. Ahmadian Fard Fini et al. 2016; Awal and Bharadwaj 2014). For all these reasons, we will use GA for solving multiple instances of our case study and find a quick and good (despite maybe less than optimal) solution.

Therefore, once the model is implemented in a spreadsheet where every time a new allocation is proposed all the subgroup calculations can be automatically and instantly updated, the GA can start looking for new solutions. The problem constraints are the ones specified in grey cells (one per individual's maximum allocation time plus one per project staffing requirements check). The objective function corresponds to E, which on this occasion is to be maximized. On resorting to Excel Solver, the best solution was found in seconds by our GA as shown in Figure 6.

<Insert Figure 6 here>

The solution shown in Figure 6 (with E=1.087) is comparatively much more efficient than the one from Figure 5 (with E=0.981). If the same GA would have been aimed at minimizing E, the worst solution found (not included) would have had an E=0.911. Within the [1-δ-φ, 1+φ] interval, the three solutions correspond to the following percentiles: 56.7% (the random solution), 79.8% (the best solution) and 41.5% (the worst solution). As can be seen, based on the results, there seems to be valid reasons to try to optimize the model outputs with the help of an optimization algorithm. Mostly, when doing this manually, would have been an unsurmountable task.

Validation

In the previous section it was shown how the model can be implemented to fulfill its most common purpose: finding the optimum (or near optimum) allocation of a set of available human resources into a series of projects, each with not necessarily equal staffing requirements. However, before accepting that the model outputs constitute a fair description of reality, it is necessary to verify whether its parameters actually influence different levels of team performance. Particularly, the most relevant model parameters are the ones proposed in equations (12) and (14), that is, the group diversity-related performance (P_k^{SS}) and the social cohesion-related performance (P_k^S). Hence, if higher values of these two parameters exhibit correlation with higher values of team performance, the model will be of some value. Conversely, if there is no such correlation, the model, at least as currently formulated, would render useless.

With this purpose in mind, a first exploratory and validation study was conducted comprising the academic performance of fifteen MSc Civil Engineering students at the Universidad de Talca (Chile). This group of students were enrolled in a module named

‘Projects’ which is a transversal integration module and its purpose is to determine how well students can apply knowledge and understanding from previous related modules. The module was led by one of the authors in the second semester of 2017. It required that fifteen students submitted three assignments (projects) each. The three projects which will be named Project 1, Project 2 and Project 3, had progressive submission dates every two months. Students worked in groups of five to deliver these three projects. After each project was completed, the groups were reshuffled so that most students had to work with different team mates in the next project.

In short, for the first assignment (Project 1), there were three groups with 5 students (named here as groups A, B and C) each submitting a different project. The same happened for Project 2 and 3, but with groups whose member composition was different from Project 1. Each of these three projects was assessed and given a mark between 0 and 100. In total, there were 9 different marks: one per assignment (Project 1, 2 and 3) and group (A, B and C). However, each student only received three marks (one from each Project) whose average resulted in the module’s final mark for him/her.

The demographic attributes of the fifteen students can be found in Figure 7. By columns, the five individuals’ attributes had a close equivalence with the five attributes described in our model: background (akin to functional department), age, gender, ethnicity and work experience (akin to team tenure). However, as expected from a group of students, the sample was also relatively more homogeneous than other real-life projects (most individuals had similar ages, similar experience, and a less varied set of backgrounds/degrees).

<Insert Figure 7 here>

The group demographic attributes were directly retrieved from their registrations information. This MSc programme required a minimum work experience of 2 years as an admission criterion. This is the reason why most students exhibit similar years of experience, but also similar ages. Additionally, most were local students, which meant most of them were South American.

The Sociometric individuals' preferences matrix was populated using the registration of the same students for a series of lab sessions which ran in parallel to this module. Particularly, before allocating the fifteen students to those lab sessions, they were asked with whom they would like to carry out the lab sessions and whom they would prefer to avoid. Although it was not compulsory for the instructors to implement those preferences, those registers proved useful later for allocating the students to the Project groups, and also to populate the sociometric matrix.

Finally, the last three columns in Figure 7 correspond to the three marks that each student was awarded at the end of the module.

The information from Figure 7 is enough to develop a first, simple, and representative correlation analysis between the model parameters and team performance. For this purpose, the groups' homogeneity, subgroup strength and social cohesion values were calculated for the nine five-student groups that submitted the three projects. With these values, calculating the group diversity- and cohesion-related performances was straightforward. These calculations are all presented, along with the students' allocations, in Figure 8.

<Insert Figure 8 here>

Values highlighted in blue, red and green correspond to Projects 1, 2 and 3, which were also identified in the last three columns of Figure 7. For reference purposes, the project marks were also shown at the bottom row of Figure 8.

The last step consists of showing how the regression plots of the variables included in the model align with the expected performance outputs. For this purpose, several plots are shown in Figure 9 from the numerical data represented in the lower rows of Figure 8.

<Insert Figure 9 here>

Figures 9a) and 9b) show how subgroup strength and group social cohesion are related to group diversity-related and group cohesion-related performance, respectively, based on equations (12) and (14). Additionally, in Figures 9d) and 9e), we can see how group diversity-related performance and group-cohesion related performance are both related to team performance as well (team performance is represented by the project marks). The very fact that these two graphs exhibit approximately linear trends between the X and Y variables indicate that the regression expressions represented in Figures 9a) and 9b) were supported. This is because expressions (12) and (14) are actually transforming the group diversity-related (from Figure 9a) and group cohesion-related performance (from Figure 9b) into a series of points with a linear correlation with Team Performance. If that had not been the case, then there would be no linear relationship, and quite possibly, no trend would be found at all.

Finally, Figure 9c) represents the Y-axis variables from Figures 9a) and 9b) and barely shows any trend. This proves that the level of correlation between subgroup strength and group cohesion is very low. This relationship was also hypothesized earlier and this is proven numerically and graphically here.

Therefore, a quick observation of the three plots at the bottom of Figure 9 show evidence of a moderate/strong correlation between the model parameters (independent variables) and the project marks (dependent variable). As the project marks can be considered as a good proxy for group performance, we can conclude that the model formulation seems to be fairly representative and is correctly indicating that certain heterogeneity-related and cohesion-related group attributes can ultimately lead to higher (or lower) team performances.

Of course, the conclusions of this validation case study have to be taken with some caution too. The analysis is based on an academic environment, rather than a real project. Real life projects tend to consist of a more diverse group of professionals (higher dispersion of the demographic attributes) with generally many more variables which may be difficult (but necessary) to control. Notwithstanding this, we acknolwedge further validation using real projects is needed in order to improve the validity of the model. However, resorting to an academic environment also has numerous advantages. First, the outcome of 'project' performance can be known (under some simplifying assumptions) as all assignments are graded and awarded a mark. And second, these project cycles are usually faster which also allows data retrieval to be generated faster than in real-life projects.

Other limitation of our validation case study is the reduced number of points (only nine) and the reduced variation of some of the performance measurements. In connection with the latter, the cohesion-related performance values of Projects 2 and 3 are very close to each other, obscuring the type of relationship that more dispersed values could have shown. Also, although the rest of the cases show clearer trends, it is necessary to point out that these might not be necessarily linear. This, despite us resorting to three points, two of them still remain too close in the cohesion-related performance graph to infer properly potentially non-liner trends.

Finally, it is clear from Figure 9 that Project 1 seemed to be more challenging to the students as they all got lower marks (probably because it was the first assignment), whereas the other two seemed easier (they received higher marks). Similarly, for future validation studies, it will be advisable to gather individual marks from each student (by means of individual exams, for example). Only with this additional piece of information, will it be possible to better compare different levels of group members' performance (as groups made up of bright people usually perform better than ones with mediocre students).

Discussion

In this paper, a new model for allocating human resources that considers team functional requirements, group heterogeneity and social cohesion has been proposed and validated. In formulating the model, a few simplifications and constraints were assumed. These will be now be reviewed and discussed in detail.

First of all, as stated earlier, this model has necessarily oversimplified the nature of real life work collaboration issues. Real life team work is complex and dynamic. Certainly, it cannot be reduced to two variables –group diversity and social cohesion– without neglecting aspects that make from group collaboration something rich and distinct from other engineering and technical challenges. In real life, group members' exhibit behaviors and possess attributes that have not been included in this model (e.g. how introvert/extrovert group members are; their dedication, devotion, preferences or just personal or professional interests or goals; their soft skills or motivation to work in groups; the asymmetrical personal relationships as a consequence of the lines of command, etc.). However, the intention here was not to include an exhaustive list of group attributes but to present a simple and self-contained model. And despite all the necessary simplifications, the model still seems to be robust, at least, based on the preliminary validation results shown here. The inclusion of further variables will be something that, no doubt, will be considered in future versions of the model when it is applied in real project settings.

Secondly, one might raise the question that the way group heterogeneity and cohesion have been defined in this paper might lead to some collinearity or, at least, covariance between the two constructs. This is because there is a possibility that both may be capturing some common aspects of a group configuration. Our model, however, has instrumented both constructs in a multiplicative way, that is, $P_k^{SS} \times P_k^{S}$, not additive, because they do not substitute for one another when contributing to subgroup performance E_k. All the same, we agree that this might be an over simplification, but existing research so far does not seem to have reached an agreement on whether this is an untenable assumption.

For example, Festinger et al. (1950) in an early attempt found contradicting results in two experiments analyzing the group heterogeneity-cohesion relationship. Much more recently, Dion (2000), on performing analysis in two houses of war veterans found again inconclusive findings indicating that in one house cohesion was related with homogeneity, whereas in the other it was not. And even more recently, in a study conducted by Chiocchio and Essiembre (2009), it was shown that a group's homogeneity or heterogeneity does not appear to affect the social cohesion-behavioral performance correlations in either academic or organizational settings. In line with this, probably the most enlightening stance has been the one taken by Sturgis et al. (2014), who claimed that the relationships between the different subcomponents of group heterogeneity and cohesion might be very different from each other, even cancelling out each other's effects. They also emphasized that more research is necessary to validate this.

However, and fortunately, because our model only tries to relatively (not absolutely) generate the most desirable subgroups allocations from the same pool of human resources, the effect of potential (if existing) collinearities between subcomponents of heterogeneity and cohesion will not be that critical so as to invalidate the model. This, as despite correlation between both variables might cause some scale distortion, the relative rank (order) of solutions should not have altered much.

Additionally, other simplifications have been assumed along the model formulation. Probably the two most relevant have been limiting the allocations of individuals to be in full-time working arrangement and not part-time. Also, the relationship between group cohesion and group performance has been assumed to be linear.

The first simplification is relatively easy to address but it would complicate the mathematical expressions to a point where they are no longer that intuitive. In this paper, we have tried to encourage understanding of the model's utility and to avoid distractions by complicating it too much. However, allowing part-time allocations might make finding better solutions somewhat easier for an optimization algorithm. This, as the objective functions of many OR models are generally easier to optimize when the decision variables are closer to being continuous (Gutiérrez et al., 2016).

Finally, the second simplification cannot be satisfactorily addressed until there is more research to determine the nature of the cohesion-performance relationship. This might be a critical aspect for further model development. It may indeed lead to some adjustments in some of the equations (probably in expression (4) and surely in expression (14)), but for now there is no point in us speculating how it might impact the model formulation, or indeed if there is such an impact at all.

Conclusions

A model that allocates human resources to multiple projects with specific staffing requirements while also considering group homogeneity and cohesion has been proposed. This model constitutes a powerful and practical tool for any project manager who needs to efficiently allocate human resources and who wants to maximize the expected productivity of his/her group members. The mathematical expressions are, in general, quite straightforward and can be easily implemented by means of a spreadsheet. The optimization algorithm for finding near-optimal solutions can also be implemented with the aid of a very simple commercial solver like Excel Solver (currently a free, despite capped, version of Frontline Solvers®).

Human resources are a key component of project success, but there is a lack of practical, quantitative tools that allow project managers to efficiently allocate these resources and build high performing teams. There are many reasons that can keep a team from functioning effectively. In this paper, two factors that are found to strongly and consistently influence group performance – group homogeneity and group cohesion – have been incorporated within the model. This model allows the measuring and comparing of any set of feasible subgroup allocations to several projects simultaneously.

Namely, group homogeneity has been defined by the subgroup strength metric and the sum of overlaps between subgroup members on five different demographic sub-factors (functional department, age, sex, ethnicity and team tenure). Group cohesion has been defined as the degree of acceptance (or rejection) that all members have with each other. The information on the five sub-factors in the group cohesion construct is generally very easy to obtain from the group members' professional profiles. In terms of the degree of acceptance/rejection that each group member has toward the rest of their group members, these can generally be known by using sociometric questionnaires. The latter, despite its limitations, have also been proven in previous research to be quite representative and relatively easy to use and update. Basically, these questionnaires require asking all group members who have finished a project: "Who would you like to keep working with?" and "Who would you prefer not to work with? From the group members' answers it is possible to populate (and keep updating) a sociometric matrix that is eventually useful for measuring how cohesive each potential subgroup is or can be.

Furthermore, previous research has proven that group homogeneity can reduce/increase group performance by up to 10% on average. Similarly, group cohesion is responsible for average increases (or decreases) of group performance by up to 18%. Both figures have been included in the proposed model and allow the objective measuring of the relative group performance differences between multiple feasible subgroups. Feasible subgroups are those who fulfill the minimum project staffing requirements stated by some simultaneous projects.

With all this, the proposed mathematical model has been detailed concerning all its components and variable relationships. A fictitious case study involving twenty workers who are allocated to three projects have been proposed and solved by means of a simple Genetic Algorithm. Finally, a validation case study based on an academic setting has also been included which involved fifteen MSc students who were allocated to three groups and were required to complete three sequential projects.

The proposed model is a simple and yet powerful way of addressing the commonplace challenges of a typical project manager in efficiently allocating human resources in projects. Despite some intentional simplifications, the model shows promise in helping project managers to make more objective and efficient decisions about their human resource allocations. However, more validating studies will be required in the future to test the actual utility of the model in real project contexts.

Although validation with real projects is necessary, this will also increase the complexity of the model's application due to the number of variables to be considered, as well as generally bigger team sizes. Indeed, this wider range of variables will have some human resource implications in terms of the sturcture of social cohesion of individuals within the projects. For example, the interactions amongst members in real projects may be underpinned by career-related imperatives, and hence, are likely to be more dynamic and nuanced, when compared with students'. In this vein, a potentially fruitful avenue for future research would be to use real life projects in conjunction with using academic projects as controlled experiment to enable researchers to study the nature and structure of social cohesion more precisely.

Data availability

All data generated or analyzed during the study are included in the submitted article or supplemental materials files.

References

Agrawal, R., Golshan, B., and Terzi, E. (2014). "Grouping students in educational settings." *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '14*, ACM Press, New York, New York, USA, 1017–1026.

Ahmadian Fard Fini, A., Akbarnezhad, A., Rashidi, T. H., and Waller, S. T. (2017). "Job Assignment Based on Brain Demands and Human Resource Strategies." *Journal of Construction Engineering and Management*, 143(5), 4016123.

Ahmadian Fard Fini, A., Rashidi, T. H., Akbarnezhad, A., and Travis Waller, S. (2016). "Incorporating Multiskilling and Learning in the Optimization of Crew Composition." *Journal of Construction Engineering and Management*, 142(5), 4015106.

Al-Bayati, A. J., Abudayyeh, O., Fredericks, T., and Butt, S. E. (2017). "Managing Cultural Diversity at U.S. Construction Sites: Hispanic Workers' Perspectives." *Journal of Construction Engineering and Management*, 143(9), 4017064.

Allison, P. D. (1978). "Measure of inequality." *American Sociological Review*, 43, 865–880.

Anvuur, A. M., and Kumaraswamy, M. M. (2016). "Effects of Teamwork Climate on Cooperation in Crossfunctional Temporary Multi-Organization Workgroups." *Journal of Construction Engineering and Management*, 142(1), 4015054.

Awal, G. K., and Bharadwaj, K. K. (2014). "Team formation in social networks based on collective intelligence – an evolutionary approach." *Applied Intelligence*, Kluwer Academic Publishers, 41(2), 627–648.

Ballesteros-Pérez, P., González-Cruz, M. C., and Fernández-Diego, M. (2012). "Human resource allocation management in multiple projects using sociometric techniques." *International Journal of Project Management*, 30(8).

Beal, D. J., Cohen, R. R., Burke, M. J., and McLendon, C. L. (2003). "Cohesion and performance in groups: a meta-analytic clarification of construct relations." *The Journal of applied psychology*, 88(6), 989–1004.

Blau, P. M. (1977). *Inequality and Heterogeneity*. Free press, New York.

Castaño, N., Watts, T., and Tekleab, A. G. (2013). "A reexamination of the cohesion-performance relationship meta-analyses: A comprehensive approach." *Group Dynamics: Theory, Research and Practice*, 17(4), 207–231.

Chang, A., and Bordia, P. (2001). "A Multidimensional approach to the Group Cohesion – Group Performance relationship." *Small Group Research*, 32(4), 379–405.

Chen, S., Wang, D., Zhou, Y., Chen, Z., and Wu, D. (2017a). "When too little or too much hurts: Evidence for a curvilinear relationship between team faultlines and performance." *Asia Pacific Journal of Management*, Asia Pacific Journal of Management, 34(4), 931–950.

Chen, Y., McCabe, B., and Hyatt, D. (2017b). "Relationship between Individual Resilience, Interpersonal Conflicts at Work, and Safety Outcomes of Construction Workers." *Journal of Construction Engineering and Management*, 143(8), 4017042.

Chiocchio, F., and Essiembre, H. (2009). *Cohesion and Performance: A Meta-Analytic Review of Disparities Between Project Teams, Production Teams, and Service Teams*. *Small Group Research*.

Dion, K. L. (2000). "Group cohesion: From 'Field of Forces' to Multidimensional Construct." *Group Dynamics: Theory, Research, and Practice*, 4(1), 7–26.

Dorn, C., and Dustdar, S. (2010). *On the Move to Meaningful Internet Systems: OTM 2010*. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Lecture Notes in Computer Science, (R. Meersman, T. Dillon, and P. Herrero, eds.), Springer Berlin Heidelberg, Berlin, Heidelberg.

Earley, P. C., and Gibson, C. B. (2002). *Multinational Work Teams: A New Perspective*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Evans, C. R., and Dion, K. L. (2012). "Group Cohesion and Performance: A Meta-Analysis." *Small Group Research*, 43(6), 690–701.

Farhadi, F., Hoseini, E., Hashemi, S., and Hamzeh, A. (2012a). "TeamFinder: A Co-clustering based Framework for Finding an Effective Team of Experts in Social Networks." *2012 IEEE 12th International Conference on Data Mining Workshops*, IEEE, 107–114.

Farhadi, F., Sorkhi, M., Hashemi, S., and Hamzeh, A. (2011). "An Effective Expert Team Formation in Social Networks Based on Skill Grading." *2011 IEEE 11th International Conference on Data Mining Workshops*, IEEE, 366–372.

Farhadi, F., Sorkhi, M., Hashemi, S., and Hamzeh, A. (2012b). "An Effective Framework for Fast Expert Mining in Collaboration Networks: A Group-Oriented and Cost-Based Method." *Journal of Computer Science and Technology*, 27(3), 577–590.

Festinger, L., Schachter, S., and Back, K. (1950). *Social pressures in informal groups; a study of human factors in housing*. Harper, Oxford, England.

Gajewar, A., and Sarma, A. Das. (2012). "Multi-skill collaborative teams based on densest subgraphs." *Proceedings of the 12th SIAM International Conference on Data Mining, SDM 2012*, 165–176.

Gibson, C., and Vermeulen, F. (2003). "A Healthy Divide : Subgroups as a Stimulus for Team Learning Behavior." *Administrative Science Quarterly*, 48(2), 202–239.

Gutiérrez, J. H., Astudillo, C. A., Ballesteros-Pérez, P., Mora-Melià, D., and Candia-Véjar, A. (2016a). "The multiple team formation problem using sociometry." *Computers and Operations Research*, 75.

Gutiérrez, J. H., Astudillo, C. A., Ballesteros-Pérez, P., Mora-Melià, D., and Candia-Véjar, A. (2016b). "The multiple team formation problem using sociometry." *Computers and Operations Research*, Elsevier, 75, 150–162.

Hendriks, M. H. A., Voeten, B., and Kroep, L. (1999). "Human resource allocation in a multi-project R&D environment: Resource capacity allocation and project portfolio planning in practice." 17(3), 181–188.

Lappas, T., Liu, K., and Terzi, E. (2009). "Finding a team of experts in social networks." *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '09*, ACM Press, New York, New York, USA, 467.

Lau, D. C., and Murnighan, J. K. (2005). "Interactions within Groups and Subgroups: The Effects of Demographic Faultlines." *The Academy of Management Journal*, Academy of Management.

Li, C.-T., and Shan, M.-K. (2010). "Team Formation for Generalized Tasks in Expertise Social Networks." *2010 IEEE Second International Conference on Social Computing*, IEEE, 9–16.

Lott, A. J., and Lott, B. E. (1965). "Group cohesiveness as interpersonal attraction: A review of relationships with antecedent and consequent variables." *Psychological Bulletin*, 64(4), 259–309.

Mathieu, J. E., Kukenberger, M. R., D'Innocenzo, L., and Reilly, G. (2015). "Modeling Reciprocal Team Cohesion – Performance Relationships , as Impacted by Shared Leadership and Members ' Competence Modeling Reciprocal Team Cohesion – Performance Relationships , as Impacted by Shared Leadership and Me." *Journal of Applied Psychology*, 100(3), 713–734.

Meyer, B., and Glenz, A. (2013). "Team Faultline Measures: A Computational Comparison and a New Approach to Multiple Subgroups." *Organizational Research Methods*, 16(3), 393–424.

Moreno, J. L. (1941). "Foundations of Sociometry: An Introduction." *Sociometry*, 4(1), 15–35.

Phua, F. T. T. (2004). "The antecedents of co-operative behaviour among project team members: an alternative perspective on an old issue." *Construction Management and Economics*, Taylor & Francis Group , 22(10), 1033–1045.

Phua, F. T. T., and Rowlinson, S. (2004). "How important is cooperation to construction project success? A grounded empirical quantification." *Engineering, Construction and Architectural Management*, Emerald Group Publishing Limited, 11(1), 45–54.

Salo, M. (2006). *The Relation Between Sociometric Choices and Group Cohesion*. Arlington, Virginia.

Shi, Z., and Hao, F. (2013). "A strategy of multi-criteria decision-making task ranking in social-networks." *The Journal of Supercomputing*, 66(1), 556–571.

Sorkhi, M., Alvari, H., Hashemi, S., and Hamzeh, A. (2012). "A game-theoretic framework to identify top-k teams in social networks." *KDIR 2012 - Proceedings of the International Conference on Knowledge Discovery and Information Retrieval*, 252–257.

Sturgis, P., Brunton-Smith, I., Kuha, J., and Jackson, J. (2014). "Ethnic diversity, segregation and the social cohesion of neighbourhoods in London." *Ethnic and Racial Studies*, Taylor & Francis, 37(8), 1286–1309.

Teixeira, L. O., and Huzita, E. H. M. (2014). *DiSEN-AlocaHR: A Multi-Agent Mechanism for Human Resources Allocation in a Distributed Software Development Environment International Conference*. Advances in Intelligent Systems and Computing, (S. Omatu, H. Bersini, J. M. Corchado, S. Rodríguez, P. Pawlewski, and E. Bucciarelli, eds.), Springer International Publishing, Cham.

Tseng, T.-L. (Bill), Huang, C.-C., Chu, H.-W., and Gung, R. R. (2004). "Novel approach to multi-functional project team formation." *International Journal of Project Management*, 22(2), 147–159.

Xia, N., Zhong, R., Wu, C., Wang, X., and Wang, S. (2017). "Assessment of Stakeholder-Related Risks in Construction Projects: Integrated Analyses of Risk Attributes and Stakeholder Influences." *Journal of Construction Engineering and Management*, 143(8), 4017030.

Yin, H., Cui, B., and Huang, Y. (2011). *Advanced Data Mining and Applications*. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Lecture Notes in Computer Science, (J. Tang, I. King, L. Chen, and J. Wang, eds.), Springer Berlin Heidelberg, Berlin, Heidelberg.