Journal of Universal Computer Science, vol. 19, no. 10 (2013), 1474-1495
submitted: 25/3/13, accepted: 27/5/13, appeared: 28/5/13 © *J.UCS*

A Team Formation and Project-based Learning Support Service for Social Learning Networks

Howard Spoelstra, Peter van Rosmalen
(Centre for Learning Sciences and Technologies, Open University of the Netherlands
Heerlen, The Netherlands
{howard.spoelstra, peter.vanrosmalen}@ou.nl)

Evert van de Vrie
(School of Computer Science, Open University of the Netherlands
Heerlen, The Netherlands
evert.vandevrie@ou.nl)

Matija Obreza, Peter Sloep
(Centre for Learning Sciences and Technologies, Open University of the Netherlands
Heerlen, The Netherlands
matija.obreza@gmail.com, peter.sloep@ou.nl)

Abstract: The Internet affords new approaches to learning. Geographically dispersed self-directed learners can learn in computer-supported communities, forming social learning networks. However, self-directed learners can suffer from a lack of continuous motivation. And surprisingly, social learning networks do not readily support effective, coherence-creating and motivating learning settings. It is argued that providing project-based learning opportunities and team formation services can help overcome these shortcomings. A review of existing team formation tools evidences that a new design for team formation and the initiation of project-based learning is required before these can be supported in social learning networks. A design is proposed which identifies *'knowledge'*, *'personality'* and *'preferences'* as categories in which data is needed to form teams, and it specifies how the required data are gathered and assessed. The design defines rules deduced from team formation principles from prior team formation research to optimise team formations towards *increased productivity*, *creative solutions* or *higher learning outcomes*. The rules are implemented in three team formation expressions each calculating one of the desired team formations. The expressions are deployed on a set of test data, demonstrating the effectiveness of the team formation service design. The article includes a discussion of the results and provides indications for future research.

Keywords: Social learning networks, project-based learning, project team formation, team formation service, team formation rules, team formation expressions, self-directed learning
Categories: L.6.0, L.6.1, L.6.2, L.3.6

1 Introduction

The 21st century requires new approaches to innovation and learning. More and more, learning takes place in geographically dispersed networks, which we call social learning networks (SLNs). SLNs are defined as computer-supported networks of informal (non-formal) learners. In these networks, people can learn, share and develop knowledge and technology helps them to do so [Sloep 2009]. They aim at supporting potentially large groups of distributed self-directed learners, who can, in their efforts to acquire competences, work and learn collaboratively (e.g., for innovation, research or assignments) or set up working groups, communities, discussions or conferences [Koper and Sloep 2002, Koper 2009, Sloep and Berlanga 2011]. However, some of the characteristics of these groups of self-directed learners are that there are only weakly linked (they initially have limited knowledge about other learners) [Jones, Ferreday and Hodgson 2008] and that they may find it difficult to remain motivated [Kim 2009].

There are various ways to improve the coherence and the motivation of learners in a network, ranging from recommending resources to each other [Drachsler, Hummel and Koper 2008], doing small activities together [Van Rosmalen et al. 2008], to actively working together [Goodyear 2005]. For SLNs in particular, the introduction of project-based learning (PBL) opportunities should fit very well. It would enable self-directed learners to engage in focused and motivating learning activities in close collaboration with other learners. The benefits of PBL are found in that it improves the learners' motivation, so that learners are more inclined to deal with harder problems and spend more time studying [Johnson, Johnson, Stanne and Garibaldi 1990, Marin-Garcia and Lloret 2008]. Furthermore, it blends learning and working and thus creates a realistic (inter-professional) learning experience [Westera and Sloep 1998, Springer, Stanne and Donovan 1999, Felder, Felder and Dietz 1999]. Recent research by [Hsiung 2010] shows that collaborative learning leads to an increase in learning outcomes, when compared to individual learning.

However, introducing PBL in a SLN is not a straightforward operation. In traditional, formal educational settings, teachers have the expertise to define projects that fit in a formal educational curriculum and are responsible for the formation of the project teams. Teachers might rely on personal knowledge about the learners and/or data sources (grades, prior courses taken) from e.g., a Learning Management System (LMS) to form teams. In traditional educational settings, the learners learn in cohorts (with respect to place, time and collective progress in the curriculum) and commit themselves to the formal educational regime. Such an educational context stands in stark contrast with a SLN learning context. In a SLN there is no teacher with curriculum knowledge and team formation expertise. Furthermore, the data as mentioned above required to form teams are not readily available, while the learners exhibit self-directing and self-organising behaviour. The learners most probably do not know each other. When designing PBL and team formation support for SLNs, we therefore have to consider that in SLNs, projects will be started by a learner (or a stakeholder connected to the network), that these projects are not necessarily positioned in a well-defined curriculum and that prospective team members can have a wide variety of knowledge backgrounds, personalities and project-related preferences.

In earlier work we introduced a team formation process model [see Figure 1] [Spoelstra, Van Rosmalen and Sloep, 2012] for use in SLN contexts. The model describes the assessment of learner *knowledge*, *personality* and *preferences*, in order to determine a *fit*-value for a team of learners for a specific project. We demonstrated that there is support for this approach to PBL and team formation from the educational field. By allowing variations in the strength and weight of the learners' *knowledge* and *personality* in the suggested teams, the model also introduced the ability to direct the team formation process towards specific project outcomes (such as facilitating learning from other team members while solving a project problem, coming up with creative solutions for the project problem, or expertly and productively solving a project problem). These variations in knowledge and personality are defined in *team formation rules*. The collection of learner *preferences*, however, denote 'condiciones sine qua non' for collaboration and thus determine whether a project can take place at all with a particular team of learners. So the preferences serve as constraints on the application of the team formation rules.

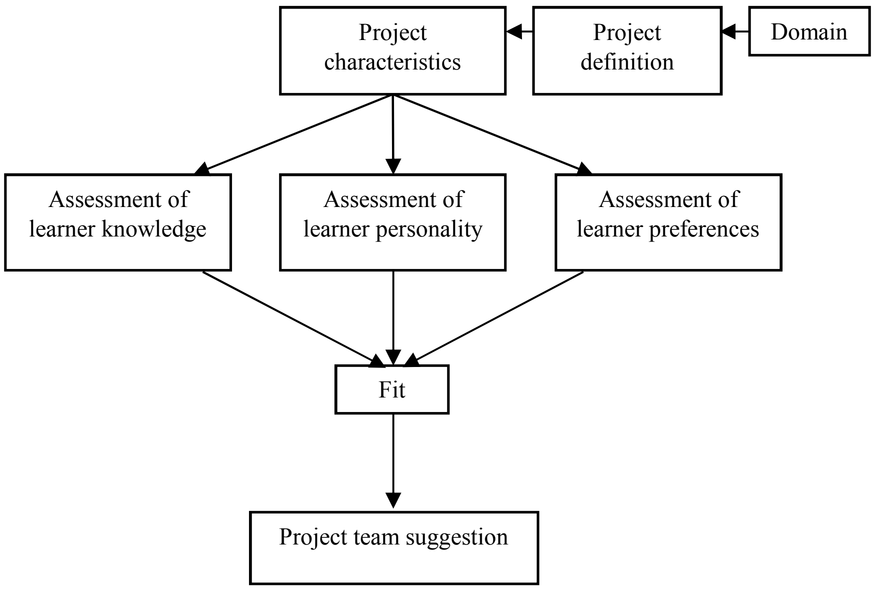

Figure 1: The model for the team formation process.

In the design of a PBL and team formation service for use in SLNs we take into account the differences between traditional educational settings and SLN settings as introduced above. As indicated, the most important differences are that there is no team formation expert (teacher) available, that the learners themselves should be enabled to start projects, and that the data used to start PBL and team formation has to be derived from different sources than in traditional learning settings. Therefore, in this article we address the following the question: *How can one design a team formation service for project-based learning in social learning networks that optimises either learning outcomes, creative outcomes or productive team performance outcomes?*

The remainder of the article is devoted to answering this question. It is organised as follows: [Section 2] provides an overview of prior research on team formation systems, the approaches to team formation they take, their aims and the data they rely on for forming teams. The section concludes with our assessment of their usefulness in SLN contexts. [Section 3] presents the design of our PBL and team formation service as well as the principles through which the data that feed into it are gathered. It also presents the definitions of the team formation rules and the formalisation of these rules into formal expressions. The expressions allow the calculation of team compositions from the data gathered. [Section 4] reports on the outcomes of a team formation exercise using the expressions on a set of simulated data. In [section 5] we discuss the outcomes, draw conclusions and indicate directions for future research.

2 Existing team formation approaches and systems

Team formation is a very active research area. Initially, this research was started in the human resource management (HRM) domain. However, as learning in teams also is considered to be preparatory for real life working conditions, team formation has also become an important topic in the educational research field. More recently, team formation is also being researched in the social network domain, using social network analysis (SNA) techniques. Team formation can be studied from different perspectives, such as competence, cultural, or personality perspectives. It can be performed for different aims and can be based on a multitude of different kinds of data. It can be studied as a separate entity, but also as being embedded in e.g., the management of international teams. An example of the latter is e.g., the People-Capability-Maturity-Model as applied to the area of Global Software Development [Colomo-Palacios, Casado-Lumbreras, Soto-Acosta, Misra and Garcia-Penalvo 2012].

The research outlined above resulted in a variety of team formation systems that are currently available. They use different data and various teaming criteria, support different aims and contexts and sometimes use multiple technologies to team up people. In the subsections 2.1 through 2.3, we provide a review of which data these systems use to form teams, sorted by the application domains: Human Resource Management, Social Networks and Education. The question we aim to answer from this review is whether these systems and the data sources they use to form teams for the goals they support can also be used in social learning networks. In subsection 2.4 we will argue that these systems all have drawbacks, prohibiting their use in the context of SLNs, thereby further strengthening our case for the design of a new team formation service.

2.1 Systems developed for use in Human Resource Management

1. Knowledge and collaboration habits [Wi, Mun, Oh and Jung 2009]. The system suggested provides grouping based on project keywords and data in knowledge repositories, keyword search in reports, paper, patents, and books. It also uses SNA techniques for finding co-authors of publications.

2. Competences mined from employee publications [Rodrigues, Oliveira and de Souza 2005]. The proposed system aims at facilitating collaboration and knowledge sharing, dissemination and creation in scientific organizations. Terms from user publications have to be manually connected to competences (and level of mastering) by a user in the role of 'knowledge manager'. After a project manager creates a project model, the system can mine the best suited project members through the required competences.

3. Knowledge, personality and working relationships [Chen and Lin 2004]. From a representation of knowledge, teamwork capability (experience, communication skills, and flexibility in job assignment) and collegiality (using the Myers-Briggs type indicator test), teams are suggested.

2.2 Systems developed for use in Social Networks

4. Type of relationship, subject, institution, geographic location, time [Monclar, Oliveira, de Faria, Ventura, de Souza and Campos 2011]. The analysis aims at discovering emerging groups in Social Networks.

5. Co-authors and related research papers [Cheatham and Cleereman 2006]. The research uses co-authorship information to create a network of relations in combination with user concept maps to enable ad-hoc team formation.

6. *References in scientific papers* [Sie, Drachsler, Bitter-Rijpkema and Sloep 2012]. The proposed system creates a network from user publications, using the measures of "betweenness" and keyword similarity. The system can either recommend authors for future publications, including prior co-authors (to strengthen current bonds between authors and strive for acceptance of a certain research topic), or recommend new co-authors (to foster creativity).

2.3 Systems developed for use in Education

7. Gender, nationality, age, previous marks, team role, and learning style [Ounnas, Davis and Millard 2009]. The authors suggest a system in which the grouping constraints and their strengths are ranked by an instructor, who also sets the project to be staffed. The system aims to increase the satisfaction of grouping constraints and to overcome the orphans' problem (learners not assigned to a team after the team formation process has ended).

8. Learner knowledge related to a task knowledge model represented in learning objects [Pollalis and Mavrommatis 2008]. The system proposed keeps track of learner knowledge and aims to group learners with comparable knowledge backgrounds to the knowledge required to perform a defined task. It is aimed at distance learners but disregards grouping criteria outside 'knowledge' as the authors suggest group formation in distance learning has less use for criteria such as gender, age, nationality, or religion.

9. Creativity score and rating of ideas [Ardaiz-Villanueva, Nicuesa-Chacón, Brene-Artazcoz, Sanz de Acedo Lizarraga and Sanz de Acedo Baquedano 2011]. The system calculates a creativity value for a user, based on the number and the length of the user provided responses to a generated idea. It uses user ratings given to ideas gathered in a brainstorm, combined with the creativity value, to suggest teams. An instructor can change the team formations. The project topics are already set, as the system works inside a PBL setting.

10. Thinking styles [Wang, Lin and Sun 2007]. The system proposed is a teacher-based tool, called DIANA. It uses data on psychological variables from questionnaires on thinking styles. It can form heterogeneous groups with respect to these styles.

11. Learner characteristics [Tobar and de Freitas 2007]. The system uses data as defined in IMS LIP (which defines both set data, such as ID, name, address, phone, email, web-address, physical, technical and cognitive characteristics, and variable data, such as goals, learning plans, learning preferences). These data are contained in a learner database and can be used by a teacher to form groups.

12. Knowledge and learning styles [Christodoulopoulos and Papanikolaou 2007]. An instructor can form heterogeneous and homogenous groups from enrolled students, based on 3 criteria (knowledge, and two axis of learning style test results). Learners take a test to determine their learning style. Unfortunately, we could not determine how the authors derived the score on knowledge.

13. Performance in previous work, activity in collaboration [Soh, Khandaker and Jiang 2008]. A system called I-MINDS can form buddy groups for unstructured collaborations and teams for structured cooperative learning activities. It uses computer-based agents to model the learners or the groups. The user model is gradually filled, based on learner activities. Structured cooperative learning follows a model with a teacher-predefined set of activities.

14. Performance and personality traits [Graf and Bekele 2006]. This research is aimed exclusively at forming heterogeneous groups, based on group work attitude, interest for the subject, achievements motivation, self-confidence, shyness, level of performance in the subject, and fluency in the language of instruction. The data on the users is represented in a vector space.

2.4 Assessment of the usability of existing systems and approaches for team formation in SLNs

The above overview of systems, aims and contexts for forming teams also describes what data these systems and approaches use to form teams. It might suggest there is a considerable overlap with the data our approach suggests to use to form teams. There are, however, distinct differences between the aims and implementation contexts in which these systems can be used and the SLN aims and implementation context:

- The systems for use in human resource management (systems 1, 2 and 3) rely on the availability of data in enterprise repositories
- The systems for use in social networks (systems 4,5 and 6), while not relying on e.g., users filling out questionnaires and taking interviews, do expect the availability of detailed logs of interactions between users

- The systems for use in education are sometimes constrained to learning situations where specific team formations are required (Systems 10 and 14), or are sometimes based on data contained in, e.g., a LMS (Systems 7, 8, 11, and partly, 12).
- Often the systems reviewed require users (administrators, teachers, tutors or instructors) to define projects, to start the team formation process or to solve team formation problems (Systems 7, 8, 9, 10, 11, 12, and 13).

However, as explained in [Section 1], a SLN does not necessarily provide the data on which these existing systems can operate. Therefore, alternative approaches have to be explored, such as asking the learners to submit specific evidence on the required knowledge or having them point to relevant entries in their e-portfolio [Penalvo et al. 2012]. SLNs also have no users in the specific roles required to run these systems. And while most of the systems examined from the educational domain only support curriculum-based activities, SLNs support self-directing learners in potentially wider knowledge domains. These differences, combined with the fact that SLN learners currently cannot easily benefit from focussed and motivating collaborative learning opportunities, warrant that we design a new approach to forming teams for project-based learning in these social learning networks.

3 PBL and team formation service design for use in SLNs

The team formation model presented in [Section 1] might readily be recognized as belonging to traditional educational settings. PBL theory and team formation theory [Oakley, Felder, Brent and Elhajj 2004, Obaya 1999] suggest that in such settings a team formation expert (e.g., a teacher) should initiate projects, while using knowledge about the curriculum to define an appropriate task. This expert uses knowledge (which can be both implicit and explicit) about the learners to form teams. However, as explained above, in SLN settings, these data nor teachers, are available. As SLN learners self-direct and self-organise we need to design a support service that enables learners themselves to perform the chain of activities required to initiate PBL and team formation.

Following the model introduced earlier, our service is designed to gather three categories of data for initiating PBL and team formation:

I) *Knowledge,* contained in: a) the collective learning materials available in the SLN, which make up the *domain*, b) *projects* and their *characteristics* (such as preferred team size, duration etc.) as defined by learners or other stakeholders in the network, c) *knowledge* available from possible team members, as evidenced by learners submitting materials for that purpose

II) *Personality*: data on the learners' personalities

III) *Preferences*: data on the learners' preferences with respect to project activities.

In order to perform the assessments depicted in the model, these data are handled by different *experts-by-proxy*. We differentiate between a *knowledge proxy*, a *personality proxy* and a *preferences proxy*.

3.1 The proxy designs

The aim of the *knowledge proxy* is three-fold: 1) to create a representation of the knowledge contained in all the topics in the learning materials present in the SLN, 2) to deduce which of these topics are addressed in the project task and 3) to assess whether and how much knowledge learners have available on the topics addressed.
The knowledge proxy *operates* on a) the collective learning materials available in the SLN that make up the domain, b) descriptions of projects by learners (or other stakeholders in the network), c) knowledge available from possible team members, as evidenced by learners submitting materials for that purpose. It is important to notice that we assume that these sources are all explicitly available in a textual form.

The assessment of knowledge through the analysis and comparison of data in a textual form is a complex task. However, prior research demonstrated the successful application of a textual analysis method, called Latent Semantic Analysis (LSA), to match *people* to *jobs* and *learning materials* [Laham, Bennett and Landauer 2000, Landauer, Foltz and Laham 1998, Landauer 2007]. In our knowledge proxy design, these entities translate to *learners*, *projects* and the *collective learning materials* in the SLN domain. [Figure 2] depicts an example of a simplified version of the process the *knowledge proxy* performs: It creates a representation of the knowledge in the domain (containing topics 1 through 6) and it analyses a project description, which is shown to relate to 3 topics in the domain (Topics 1, 3, and 5). After learners submit knowledge evidence on these topics the proxy analyses the degree to which the learner' knowledge overlaps the knowledge in the domain topics by using the domain topics as reference points. In [Figure 2] the results of these analyses are depicted as percentages.

The *personality proxy* takes a different approach in that it uses data on learner personality, which is gathered through a personality test. We specifically chose to assess learners on the personality construct "conscientiousness" (which measures learner *carefulness, thoroughness, sense of responsibility, level of organization, preparedness, inclination to work hard, orientation on achievement,* and *perseverance*) because it predicts a person's future performance in a team [Goldberg 1990, Jackson et al. 2010]. The learner conscientiousness score is established by using the Big Five personality test [Barrick and Mount 1991].

The *preferences proxy* establishes a learner preferences profile, in which learners enter data on such variables as availability, time zone, possible collaboration languages and preferred tools. The proxy then determines the overlap with respect to the project characteristics mentioned above and the learners' project work related preferences. When preferences do not overlap at all, this fully blocks user inclusion in a team. (E.g., when one learner indicates to be available only on Mondays, while another learner indicates to never be available on Mondays, their calendars are mutually exclusive and thus these two learners will never be matched in a team). We currently envision the learner to enter this data in the profile.

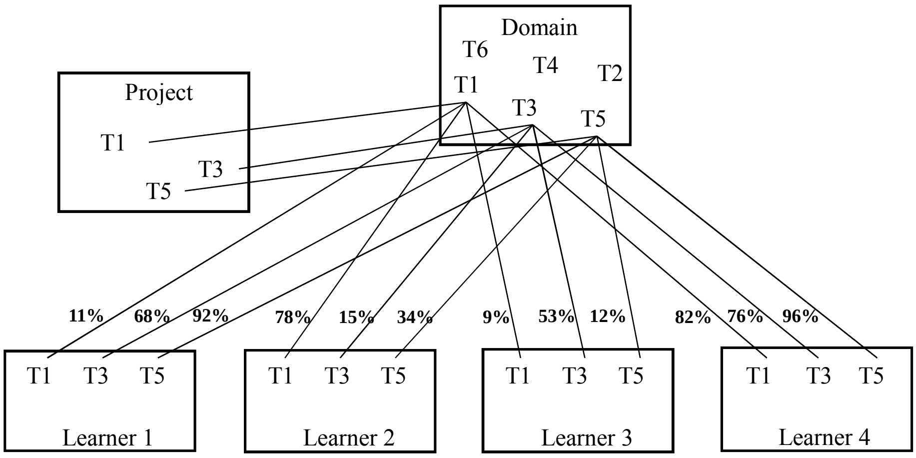

Figure 2: An example of the knowledge proxy process: A project refers to 3 topics in the domain (T1, T3, and T5). Learner submitted knowledge evidence (from Learners 1 to 4) for these topics is compared with the topic materials in the domain. The percentages indicate the degree of the knowledge overlap.

From this it follows that the *first step* in the team formation process is finding overlapping sets of preferences by comparing the project characteristics and learner preferences. By doing this, the proxy's result limits the number of learners from which teams can be formed. The team formation process then continues with the data on knowledge and personality.

It is important to notice that the data gathered on learners is *not of a static nature*, but can be refreshed every time a learner re-enters knowledge evidence for a project, retakes the personality test, or updates preferences. Furthermore, future iterations of the team formation service might be enabled to connect to user data already available in such e-portfolios as described in e.g., the TRAILER project [Penalvo et al. 2012].

For the remainder of this article we assume that the results of the assessments are available.

3.2 Definition of the team formation service and rules for targeting productive, creative or learning outcomes

The proxies' data gathering designs presented above provide the data to the team formation service. The service combines the two separate sets of data by following *team formation rules*. We discern three possible teamwork target outcomes and indicate three sets of rules, one for each outcome. The rules are based on existing research:

(1) Productive problem solving:

- Forming teams from learners who have different conscientiousness scores impedes their task negotiations after the project team has been formed, which would then hinder the team task execution [Gevers and Peeters 2009]
- Members of productive teams should be capable and conscientious and must have domain knowledge [Isaksen and Lauer 2002]

The general team formation rule we deduce is: *Productivity is fostered when team members have high scores on knowledge of the project topics and show high levels of conscientiousness*.

(2) Creative solutions:

- Too much complementary fit in knowledge can lead to a loss of creativity and to group thinking [West 2002]
- People with high conscientiousness scores tend to be less creative [George and Zhou 2001, Wolfradt and Pretz 2001]
- Groups with members that possess different knowledge backgrounds will be more innovative because they contribute from different perspectives [Paulus 2000]
- Successful research teams are heterogeneous [Dunbar 1997]

We deduce as general team formation rule: *Team creativity is fostered when team members have highly differentiated scores on knowledge of the project topics and show low levels of conscientiousness.*

(3) Facilitating learning:

- Learning is fostered when team members provide a complementary fit in knowledge backgrounds and show a supplementary fit in personalities [Werbel and Johnson 2001].
- Mutual teaching and learning are among the most important activities in defining and solving problems [Paulus 2000].
- There is a maximum 'distance in knowledge' (the zone of proximal development) that can be bridged when learning with more capable peers [Vigotsky 1978]

From these findings we deduce as general team formation rule: *Learning in a team is facilitated when knowledge on the project topics is distributed over the members (allowing each member to learn and teach) and the differences in the levels of project topic knowledge between the members are not too high and the members' conscientiousness levels all are high*.

Table 1 provides an overview of the team formation rules (with respect to learner knowledge and conscientiousness) and the target outcomes. In the table, the terms "supplementary" and "complementary" are used to denote "sharing knowledge with other members” and “providing knowledge to the team which other members lack”, respectively.

Research basis	Kind and level of knowledge rule	Conscientiousness rule	Target outcome
Gevers and Peeters, Isaksen and Lauer	Supplementary and high	All high	Productive problem solving
George and Zhou, Wolfradt and Pretz, West, Paulus, Dunbar	Complementary and high	All low	Creative solutions
Werbel and Johnson, Vigotsky, Paulus	Complementary and high, but within limits	All high	Facilitating learning

Table 1: Research basis, the kind and level of knowledge rule and conscientiousness rule for specific target outcomes.

3.3 Team formation expressions

Based on the target outcomes defined in [Table 1], we devised three mathematical team formation expressions that can be applied to the data gathered. They suggest formations of productive, creative, or learning teams, respectively. Applying the expressions results in measures of fitness calculated for all possible teams of a chosen size, recruited from a given set of learners. For each possible team, the team fitness value is represented in a value between “0” and “1”, with “1” indicating the highest possible fit for that outcome. This allows for comparing teams with respect to fitness over their different target outcomes. Weights can be used to indicate the importance of e.g., knowledge over conscientiousness in the team formation process. In the expressions below all weights are equal and sum up to 1. Other weight distributions are likely of relevance but have not been systematically explored. In all expressions, for demonstration purposes, the maximum score on knowledge (*Max_K*) is set to 10 and the maximum score on conscientiousness (*Max_C*) is set to 5. Both the desired team size (*n*) and number of topics (*k*) the project refers to are arbitrarily set to 4.

3.3.1 Productive teams

The team formation expression for the outcome “productive problem solving” [see Table 1 and Figure 3] describes teams whose members have the highest average score on knowledge and the highest average score on conscientiousness.

$$FitP_i = W_K * \frac{Avg\_K_i}{Max\_K} + W_C * \frac{Avg\_C_i}{Max\_C}$$

Figure 3: Team formation expression for productive teams.

Explanation of [Figure 3]: In the first part, the average score on knowledge of all members of team *i* over all topics is calculated (Avg_K_i) and divided by the maximum knowledge score (*Max_K)*. In the second part the average score on conscientiousness over all members is calculated (Avg_C_i) and divided by the maximum conscientiousness score (*Max_C*). These two scores are multiplied by their weights (W_K, W_C) separately and then summed. As the two parts each result in a value between 0 and 1 and the sum of the weights always is 1, this results in a measure of fitness ($FitP_i$) for each team considered between 0 and 1. In [Table 2] we present an example of a score set leading to a *FitP* of 1.

Member	**Topic 1**	**Topic 2**	**Topic 3**	**Topic 4**	**Cons**
L01	10	10	10	10	5
L02	10	10	10	10	5
L03	10	10	10	10	5
L04	10	10	10	10	5

Table 2: Example of scores on topic knowledge and conscientiousness leading to a FitP of 1.

3.3.2 Creative teams

The mathematical team formation expression for the outcome "creative solutions" [See Table 1 and Figure 4] maximises when team members have a maximum difference in knowledge between their best score and their second best score *over their own topic scores*, and when there is a maximum difference in knowledge between the best score and the second best score *inside a topic*. It minimises the average conscientiousness score in the team.

$$FitC_i = W_K * \frac{\sum_j DifK_j}{TeamSize * Max\_K} + W_E * \frac{\sum_t DifK_t}{NumTop * Max\_K} + W_C * \frac{Max\_C - Avg\_C_i}{Max\_C}$$

Figure 4: Team formation expression for creative teams

Explanation of [Figure 4]: In the first part the expression calculates the differences for each team member *j* between their highest score on a topic and the next best score on a topic ($DifK_j$) and sums these differences up over all team members. The result is divided by the product of the team size (*TeamSize*) and maximum score on knowledge (*Max_K*). In the second part, the differences for each topic *t* between the highest score on that topic and the next best score on that topic ($DifK_t$) are summed up. The result is divided by the product of the number of topics (*NumTop*) and the maximum score on knowledge (*Max_K*). Finally, in the third part, from the maximum conscientiousness score (*Max_C*) the all-member average conscientiousness score (Avg_C_i) is subtracted. The result is divided by the maximum conscientiousness score (Max_C).

All three scores are multiplied by their weights (W_K, W_E, W_C) separately and then summed. As all three parts each result in a value between 0 and 1 and the sum of the weights always is 1, this results in a measure of fitness ($FitC_i$) for each team considered between 0 and 1. In [Table 3] we present an example of a score set leading to a *FitC* of 1.

Member	Topic 1	Topic 2	Topic 3	Topic 4	Cons
L01	10	0	0	0	0
L02	0	10	0	0	0
L03	0	0	10	0	0
L04	0	0	0	10	0

Table 3: Example of scores on topic knowledge and conscientiousness leading to a FitC of 1.

3.3.3 Learning teams

The team formation expression for the outcome "facilitating learning" [see Table 1 and Figure 5] mathematically describes teams whose members can teach and learn to and from each other inside each knowledge topic, while having a high score on Conscientiousness. It optimises the match between possible teachers and learners in the team by using Vygotsky's 'zone of proximal development' [Vygotsky 1978] as a parameter (*zpd*) to calculate teaching and learning effectiveness for the team over all project topics.

$$FitL_i = W_K * \frac{\sum_t \sum_l \sum_j \left| score_{t,j} - score_{t,l} \right|}{d_{jt} \cdot zpd \cdot n \cdot k} + W_C * \frac{Avg\_C_i}{Max\_C}$$

Figure 5: Team formation expression for learning teams.

Explanation of [Figure 5]: In the first part, every topic score of a member is compared to the other member's topic scores ($|score_{t,j}$, $score_{t,l}|$). When there is no difference between the scores, the members cannot teach to each other, nor learn from each other. If the difference is inside the parameter *zpd* (currently set to be between 0 and 3), then that member becomes a teacher to the other member. The member' teaching effectiveness depends on the difference from the set *zpd*. For example when member 1 scores 8 on topic 1 while member 2 scores 6 on topic 1, then the difference is 2. With a *zdp* set to 3, the teaching effectiveness between these members is calculated as 2/3. In the same manner, learning effectiveness is calculated. This is repeated for all other members. For each member the teaching and learning effectiveness scores are summed up and then divided by that member's summed number of times being a teacher and number of times being a learner in the topic (d_{jt}). We define the result as that member's effectiveness in the team. This process is repeated for all members (n) inside the topic. Finally, all teaching scores are added, all learning scores are added and all effectiveness scores are summed. With the multiplication of the sum of the effectiveness scores with the sum of the sum of all learning scores and the sum of all teaching scores, we arrive at a score for that topic, which is then normalised. This process is repeated over all topics (k), and all topic scores are summed. This final sum represents the teams learning capability.

In the second part, the average team conscientiousness score (Avg_C_i), divided by the maximum conscientiousness score (Max_C) is calculated. The two scores are multiplied by their weights (W_K and W_C) separately and then summed. As the two scores each result in a value between 0 and 1 and the sum of the weights always is 1, this results in a measure of fit for each team considered ($FitL_i$) between 0 and 1.

There are two exemptions to the rule: If the difference between two topic scores is higher than the parameter zdp, or when a teacher has a score on a topic lower than a set minimum score (currently set to 6), teaching and learning effectiveness for that teacher/learner pair is set to be ~0. In [Table 4] we present an example of a score set leading to a *FitL* of 1 when the zone of proximal development is set to 3 and the minimum teacher topic score is set to 6.

Member	**Topic 1**	**Topic 2**	**Topic 3**	**Topic 4**	**Cons**
L01	10	9	10	9	5
L02	7	6	10	6	5
L03	10	6	7	9	5
L04	7	9	7	6	5

Table 4: Example of scores on topic knowledge and conscientiousness leading to a FitL of 1.

We anticipate that the application of the three expressions to the same data set will result in differentiated team formation suggestions for each of the three outcomes, and that the results indicate which outcome fits best to any of the teams possible.

4 Results of the application of the team formation expressions on a test data set

For the simulation we used a set of test data on 10 learners [see Table 5]. The test data set presupposes that the project description had already been analysed and was found to refer to knowledge on 4 topics in the domain. It further presupposes that the analysis of knowledge evidence on these 4 topics, as submitted by 10 learners, had already been performed. This is reflected in the numerical scores under the topics 1 through 4 (ranging from 1 to 10, where 10 indicates the highest possible score on a topic). The scores on Conscientiousness (Cons) in [Table 5] are the simulated results of a personality test (ranging from 1 to 5, where 5 indicates the highest level). The team size of the teams to be formed was arbitrarily set to 4 learners per team.

Member	Topic1	Topic2	Topic3	Topic4	Cons
L01	9	8	8	9	5
L02	4	6	4	5	4
L03	4	3	4	9	1
L04	5	4	6	8	5
L05	3	4	10	2	1
L06	8	9	8	5	4
L07	4	9	5	3	2
L08	8	9	8	7	3
L09	5	8	7	8	3
L10	4	5	3	4	1

Table 5: The test data set.

4.1 Application of the expressions

When the expressions above are applied to the test data set, all fitness values for the 210 unique combinations [Number_of_learners! / ((Number_of_learners – team_size)! * team_size!)] of 4 learners are calculated. The output we receive lists all possible teams and their scores on FitP, FitC and FitL, totalling to 630 values. In [Table 6] we present only the 3 highest scores per outcome, and the lowest score (all results are truncated to 3 decimals). In the three columns FitP, FitC and FitL the scores are sorted from high to low.

Team members	FitP	Team members	FitC	Team members	FitL
L01,L04,L06,L08	0.797	L03,L05,L07,L10	0.500	L02,L04,L06,L09	0.660
L01,L04,L06,L09	0.784	L03,L05,L06,L10	0.442	L04,L06,L08,L09	0.609
L01,L02,L04,L06	0.781	L03,L05,L08,L10	0.442	L02,L04,L06,L08	0.598
~	~	~	~	~	~
L03,L05,L07,L10	0.363	L01,L02,L04,L06	0.092	L03,L05,L07,L10	0.126

Table 6: Team formations for 4 teams of 4 learners, sorted by FitP, FitC or FitL.

The individual team members and their scores on Topics 1 to 4 and conscientiousness for the teams with the highest scores on FitP, FitC and FitL are presented in [Table 7].

	Team with highest FitP					Team with highest FitC					Team with highest FitL				
	T1	**T2**	**T3**	**T4**	**Cons**	**T1**	**T2**	**T3**	**T4**	**Cons**	**T1**	**T2**	**T3**	**T4**	**Cons**
M1	9	8	8	9	5	4	3	4	9	1	4	6	4	5	4
M2	5	4	6	8	5	3	4	10	2	1	5	4	6	8	5
M3	8	9	8	5	4	4	9	5	3	2	8	9	8	5	4
M4	8	9	8	7	3	4	5	3	4	1	5	8	7	8	3

Table 7: Individual learner scores (M1 to M4) on topic knowledge (T1 to T4) and conscientiousness (Cons) for the teams with the highest fit values on FitP, FitC and FitL from Table 6.

For FitP, a team comprised of learners L01, L04, L06, and L08 receives the highest score (0.797), while the lowest score (0.363) is for a team comprised of learners L03, L05, L07, and L10. For FitC, a team formed from learners L03, L05, L07, and L10 receives the highest score (0.500), while a team of learners L01, L02, L04, and L06 receives the lowest score (0.092). As for FitL, a team with learners L02, L04, L06, and L09 scores highest (0.660). A team with learners L03, L05, L07, and L10 scores lowest (0.126).

4.2 Differentiations in team formation suggestions

When sorted for FitC, the highest scoring team on FitP is found on position 208 and when sorted for FitL that team is found on position 6. Both when sorted for FitP and for FitL, the highest scoring team on FitC is found on position 210. When sorted for FitP, the highest scoring team on FitL is found on position 16 and when sorted for FitC, it is found on position 196. The differentiation is not only relevant with respect to rank in the results, but also with respect to actual fitness value calculated. [Table 8] allows for comparing the teams with the highest fitness values on a particular outcome (these fitness values are highlighted in the table) with how well they fit to any of the other outcomes.

Team of members	FitP	FitC	FitL
L01,L04,L06,L08	0.797	0.100	0.569
L03,L05,L06,L10	0.363	0.500	0.126
L02,L04,L06,L09	0.713	0.142	0.660

Table 8: Team fitness values on FitP, FitC and FitL for the highest scoring teams on FitP, FitC and FitL, respectively.

The results also indicate which kind of team could preferably be formed from these learners: the highest overall fitness-value (0.797) is received for a team (consisting of the learners L01, L04, L06, and L08) that aims at the outcome “productive problem solving”. An interesting find is that a team consisting of learners

L02, L04, L06, and L09, while receiving the highest fitness value for the outcome "facilitating learning" (FitL = 0.660), would likely do better if it were to aim for "productive problem solving" as an outcome (FitP = 0.713).

5 Discussion and conclusion

We set out to answer the question: *How can one design a team formation service for project-based learning in social learning networks that optimises either learning outcomes, creative outcomes or productive team performance outcomes?*

Our perspective was that social learning networks currently do not readily support effective, coherence-creating and motivating learning settings. We therefore suggested to provide these learners with a project-based learning and team formation service. As a starting point we took our team formation model [Spoelstra, Van Rosmalen and Sloep 2012]. A survey of existing team formation tools and techniques revealed that these are not easily applicable in a "team formation for project-based learning in social learning networks"-approach. They assume data and user roles that are not available in SLNs. For this reason we proposed a design which allows project-based learning and team formation to be based on data that can be acquired directly from the SLN and its learners. The design puts learners in control of the process of defining and staffing projects, thus honouring these learners' self-directing and self-organising behaviour. The design uses the data categories 'knowledge', 'personality', and 'preferences' (as defined in the team formation model) and describes the ways in which the data can be gathered and processed to suggest team formations. A benefit of the design is that it is also based on personality characteristics, which is rarely the case in existing tools, but which – according to literature [Roberts, Kuncel, Shiner, Caspi and Goldberg 2007] – is highly relevant.

The team formation and project-based learning service deploys three different proxies to gather and assess data: 1) To assess both required and available knowledge, the *knowledge proxy* analyses textual data; 2) To assess learner personality, the *personality proxy* determines a learner's conscientiousness by using a personality test; 3) To determine project work preferences, the *preferences proxy* determines whether collaborative project work can happen at all.

In order to determine how learners should be teamed up based on knowledge and personality we analysed existing research on team formation principles. The outcomes led to the definition of team formation rules for forming productive, creative, or learning teams, respectively. These rules were formalised in team formation expressions.

The application of the expressions to a set of simulated test data demonstrates their ability to form teams and to suggest different teams based on the desired teamwork outcomes. The results provide both team rank on all three possible outcomes and the absolute fitness values for those outcomes. The results further allow us to suggest which outcome would fit best to any of the teams that could be formed. We believe these results clearly show the ability of the expressions to differentiate between teams fit for any of the proposed teamwork outcomes.

Future research can introduce further differentiation in the results: when one primary outcome is selected for a team, its fitness scores on the other outcomes might act as qualifiers for that outcome. This would provide a method for closer selection of teams, based on how the primary outcome will likely be achieved.

With its strong base in PBL and team formation research, we believe our approach addresses important issues in team formation. However, one could argue that knowledge might also be contained in other forms of evidence currently not taken into account, and that even though 'conscientiousness' is very important predictor of a learner' success in future project work, it is not the only personality aspect playing a role in team work. Furthermore, research by e.g., [Kirton 2003] indicates that the more diverse a team is, the greater its potential for problem solving will be, but the more difficult it becomes to manage. This might be of particular interest in the case of creative teams, where the favoured low average conscientiousness, combined with highly diverse knowledge could lead to teams having difficulty working together. Future research will determine whether the introduction into the expression for creative teams of additional personality factors such as 'Extravertness' [Barrick and Mount 1991] are necessary to mitigate this effect.

However, our premise was that social learning network learners only have limited knowledge of other learners, and that these networks do not have historic data on learner performance. We therefore believe that our team formation service offers an import first step in supporting project-based learning and team formation in such networks. Nevertheless, (parts of) the team formation service can have a wider application in settings where the required data is already partly available. When data on prior knowledge and preferences are available (e.g., in a classroom setting), the service only requires the addition of personality data to be usable. The preliminary analyses of the results from a survey about whether and under which conditions teachers would accept team formation suggestions from an automated system based on the proposed design indicate that of 11 responses, 5 express acceptance of automated team suggestions, while 5 responses express acceptance with some reservations. These reservations are mostly concerned with aspects such as who has the final say in team formation. As our tool delivers team formation suggestions from which users can deviate, we feel convinced that a team formation tool based on the principles outlined above will be welcomed.

Another area of application might be found in the context of Massive Open Online Courses (MOOCs), where the use of team formation tools could be a way to enhance the currently rather limited interaction between students.

In our future research we will report on an implementation of the knowledge, personality and preferences proxies using real student reported data in a large-scale experiment. A next step will then be to further implement the knowledge proxy, for which we suggest to use the LSA textual analysis method to match knowledge from learners to knowledge required by a project.

References

[Ardaiz-Villanueva, Nicuesa-Chacón, Brene-Artazcoz, Sanz de Acedo Lizarraga and Sanz de Acedo Baquedano, 11] Ardaiz-Villanueva, O., Nicuesa-Chacón, X., Brene-Artazcoz, O., Sanz de Acedo Lizarraga, M., Sanz de Acedo Baquedano, M.: Evaluation of computer tools for idea generation and team formation in project-based learning, Computers & Education, 56, 3 (2011), 700-711

[Barrick and Mount, 91] Barrick, M. R., Mount, M. K.: The big five personality dimensions and job performance: A meta-analysis, Personnel Psychology, 44, (1991), 1–26.

[Christodoulopoulos and Papanikolaou, 07] Christodoulopoulos, C.E., Papanikolaou, K.A.: A Group Formation Tool for e-Learning Environments, Proc.19th IEEE International Conference on Tools with Artificial Intelligence, 2, (2007), 117–123.

[Colomo-Palacios, Casado-Lumbreras, Soto-Acosta, Misra and Garcia-Penalvo, 12] Colomo-Palacios, R., Casado-Lumbreras, C., Soto-Acosta, P., Misra, S., Garcia-Penalvo, F.: Analyzing Human Resource Management Practices Within the GSD Context, Journal of Global Information Technology Management, 15, 3 (2012), 30–54.

[Drachsler, Hummel and Koper, 08] Drachsler, H., Hummel, H., Koper, R.: Personal recommender systems for learners in lifelong learning: requirements, techniques and model, International Journal of Learning Technology, 3, 4 (2008), 404–423.

[Dunbar, 97] Dunbar, K. (1997): How scientists think: On-line creativity and conceptual change in science, In T. B. Ward, S. M. Smith, and J. Vaid (Eds.), Creative thought: An investigation of conceptual structures and processes, 461–493. Washington, DC: American Psychological Association.

[Felder, Felder and Dietz, 99] Felder, R. M., Felder, G. N., Dietz, E. J. (1999): A longitudinal study of engineering student performance and retention v. comparisons with traditionally-taught students, Journal of Engineering Education, 87, 4 (1999), 469–480.

[George and Zhou, 01] George, J. M., Zhou, J. (2001): When openness to experience and conscientiousness are related to creative behavior: An interactional approach, Journal of Applied Psychology, 86, 3 (2001), 513–524.

[Gevers and Peeters, 09] Gevers, J. M. P., Peeters, M.: A pleasure working together? The effects of dissimilarity in team member conscientiousness on team temporal processes and individual satisfaction, Journal of Organizational Behavior, 30, (2009), 379–400.

[Goldberg, 90] Goldberg, L. R.: An alternative description of personality: The big-five factor structure, Journal of Personality and Social Psychology, 59, 6 (1990), 1216–1229.

[Goodyear, 05] Goodyear, P.: Technology and the articulation of vocational and academic interests: reflections on time, space and e-learning, Studies in Continuing Education, 28, 2, (2006), 83–98.

[Graf and Bekele, 06] Graf, S., Bekele, R.: Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization, Proc. 8th International Conference on Intelligent Tutoring Systems (ITS 2006), Springer Berlin / Heidelberg, (2006), 217–226.

[Hsiung, 10] Hsiung, C. M.: An experimental investigation into the efficiency of cooperative learning with consideration of multiple grouping criteria, European Journal of Engineering Education, 35, 6 (2010), 670–692.

[Isaksen and Lauer, 02] Isaksen, S. G., Lauer, K. J.: The Climate for Creativity and Change in Teams, Creativity and Innovation Management, 11 (2002), 74–86.

[Jackson, Wood, Bogg, Walton Harms and Roberts, 10] Jackson, J., Wood, D., Bogg, T., Walton, K., Harms, P., Roberts, B.: What do conscientious people do? Development and validation of the Behavioral Indicators of Conscientiousness (BIC), Journal of Research in Personality, 44, 4 (2010), 501–511.

[Johnson, Johnson, Stanne and Garibaldi, 90] Johnson, D.W., Johnson, R., Stanne, M.B. Garibaldi, A. (1990): Impact of group processing on achievement in cooperative groups, The Journal of Social Psychology, 130, 4 (1990), 507–516.

[Jones, Ferreday and Hodgson, 08] Jones, C.R., Ferreday, D. Hodgson, V.: Networked learning a relational approach: weak and strong ties, Journal of Computer Assisted Learning, 24 (2008), 90–102.

[Kim, 09] Kim, K.J.: Motivational Challenges of Adult Learners in Self-Directed e-Learning, Journal of Interactive Learning Research, 20, 3 (2009), 317–335.

[Kirton, 03] Kirton, M.: Adaption—Innovation in the context of diversity and change, London: Routledge (2003).

[Koper, 09] Koper, R. (Ed.): Learning Network Services for Professional Development, Berlin and Heidelberg: Springer (2009).

[Koper and Sloep, 02] Koper, R., Sloep, P. B.: Learning Networks connecting people, organizations, autonomous agents and learning resources to establish the emergence of effective lifelong learning, RTD Programme into Learning Technologies 2003-2008. More is different. Heerlen, The Netherlands: Open Universiteit Nederland (2002).

[Laham, Bennett and Landauer, 00] Laham, D, Bennett, W. Jr, Landauer, T.K.: An LSA-Based Software Tool for Matching Jobs, People, and Instruction, Interactive Learning Environments, 8, 3 (2000), 171–185.

[Landauer, 07] Landauer, T. K.: LSA as a theory of meaning; Chapter in the Handbook of Latent Semantic Analysis. 3–34. Edited by Thomas K. Landauer, Danielle S. McNamara, Simon Dennis, Walter Kintsch. Psychology Press. Lawrence Erlbaum; 1st edition (2007).

[Landauer, Foltz and Laham, 98] Landauer, T. K., Foltz, P. W., Laham, D.: Introduction to Latent Semantic Analysis, Discourse Processes, 25 (1998), 259–284.

[Marin-Garcia and Lloret, 08] Marin-Garcia, J.A., Lloret, J.: Improving teamwork with university engineering students: the effect of an assessment method to prevent shrinking, WSEAS Transactions on Advances in Engineering Education, 5, 1 (2008), 1–11.

[Oakley, Felder, Brent and Elhaij, 04] Oakley, B., Felder, R.M., Brent, R., Elhajj, I.: Turning student groups into effective teams, Journal of Student Centered Learning, 2, 1 (2004), 10–34.

[Obaya, 99] Obaya, A.: Getting cooperative learning, Science Education International, 10, 2 (1999), 25–27.

[Ounnas, Davis and Millard, 09] Ounnas, A., Davis, H. C., Millard, D. E.: A Framework for Semantic Group Formation in Education, Educational Technology & Society, 12, 4 (2009), 43–55.

[Paulus, 00] Paulus, P.: Groups, Teams, and Creativity: The Creative Potential of Idea-generating Groups, Applied Psychology, 49 (2000), 237–262.

[Penalvo et al., 12] Penalvo, F. J. G., Zangrando, V., Holgado, A. G., Gonzalez, M. A. C., Pardo, A. M. S., Griffiths, D., Forment, M. A., Mykowska, A., Janssen, J., Alves, G. R., Minović, M.: TRAILER project overview: Tagging, recognition and acknowledgment of informal learning experiences, Proc. International Symposium on Computers in Education (SIIE), (2012), 1–6.

[Pollalis and Mavrommatis, 08] Pollalis, Y.A., Mavrommatis, G.: Using similarity measures for collaborating groups formation: A model for distance learning environments, European Journal of Operational Research. 193, 2 (2009), 626–636.

[Roberts, Kuncel, Shiner, Caspi and Goldberg, 07] Roberts, B. W., Kuncel, N., Shiner, R., Caspi, A., & Goldberg, L.: The power of personality: A comparative analysis of the predictive validity of personality traits, SES, and IQ, Perspectives in Psychological Science, 2 (2007), 313 – 345.

[Rodrigues, Oliveira and de Souza, 05] Rodrigues, S., Oliveira, J., de Souza, J. M.: Competence Mining for Team Formation and Virtual Community Recommendation, Proc. Ninth International Conference on Computer Supported Cooperative Work in Design, 1 (2005), 44 – 49.

[Sie, Drachsler, Bitter-Rijpkema and Sloep, 12] Sie, R. L. L., Drachsler, H., Bitter-Rijpkema, M., Sloep, P. B.: To whom and why should I connect? Co-author Recommendation based on Powerful and Similar Peers, International Journal of Technology Enhanced Learning (IJTEL), 4, 1 (2012), 121–137.

[Sloep, 09] Sloep, P.B.: Fostering Sociability in Learning Networks through Ad-Hoc Transient Communities, In M. Purvis, B. T. R. Savarimuthu (Eds.). Computer-Mediated Social Networking. Proceedings of the First International Conference, ICCMSN 2008. Dunedin, New Zealand, June 2008, revised selected papers. (62–75). Heidelberg, Germany: Springer, (2009).

[Sloep and Berlanga, 11] Sloep, P. B., Berlanga, A. J.: Learning Networks, Networked Learning [Redes de Aprendizaje, Aprendizaje en Red], Comunicar, XIX, 37 (2011), 55–63.

[Soh, Khandaker and Jiang, 08] Soh, L., Khandaker, N., Jiang, H.: I-MINDS: A Multiagent System for Intelligent Computer-Supported Collaborative Learning and Classroom Management, International Journal of Artificial Intelligence in Education. 18, 2 (2008), 119–151.

[Spoelstra, Van Rosmalen and Sloep, 12] Spoelstra, H., Van Rosmalen, P., Sloep, P.: Project team formation support for self-directed learners in social learning networks, Proc. IADIS International Conference on Web Based Communities and Social Media (ICWBC & SM 2012), (2012), 89–96.

[Springer, Stanne and Donovan, 99] Springer, L, Stanne, M.E., Donovan, S.S.: Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology, Review of Educational Research, 69, 1(1999), 21–51.

[Tobar and Luís de Freitas, 07] Tobar, C.M., Luís de Freitas, R.: A Support Tool for Student Group Definition, Proc. Frontiers In Education Conference - Global Engineering: Knowledge Without Borders, Opportunities Without Passports. FIE '07. 37th Annual 10-13 Oct. 2007. T3J-7 - T3J-8.

[Vygotsky, 78] Vygotsky, L. S.: Mind in society: the development of higher psychological processes, edited by Michael Cole [et al.] Harvard University Press, Cambridge, (1978).

[Wang, Lin and Sun, 07] Wang, D.Y., Lin, S.S., Sun, C.T.: Diana: A computer supported heterogeneous grouping system for teachers to conduct successful small learning groups, Computers in Human Behavior, 23, (2007), 1997–2010.

[Werbel and Johnson, 01] Werbel, J.D., Johnson, D.J.: The use of person-group fit for employment selection: A missing link in person-environment fit, Human Resource Management, 40, (2001), 227–240.

[West, 02] West, M.: Sparkling fountains or stagnant ponds: An integrative model of creativity and innovation implementation in work groups, Applied Psychology: An International Review, 51, (2002), 355–424.

[Westera and Sloep, 98] Westera, W., Sloep, P. B.: The Virtual Company: Toward a Self-Directed, Competence-Based Learning Environment in Distance Education, Educational Technology, 38, 1 (1998), 32–37.

[Wi, Oh, Mun and Jung, 09] Wi, H., Oh, S., Mun, J. Jung, M.: A team formation model based on knowledge and collaboration, Expert Systems with Applications, 36, 5 (2009), 9121–9134.

[Wolfradt and Pretz, 01] Wolfradt, U., Pretz, J. E.: Individual differences in creativity: personality, story writing, and hobbies, European Journal of Personality, 15, (2001), 297–310.