Assigning Students in Groups:

Self-formed Groups versus Automatically-formed Groups

Razmerita, Liana[1] and Brun, Armelle[2]
[1] Copenhagen Business School, CBS, ISV,
Dalgas Have, 15, DK-2000 Frederiksberg, Denmark
lr.isv@cbs.dk
[2] LORIA-Nancy Université
615, rue du jardin botanique - 54506 Vandœuvre les Nancy
armelle.brun@loria.fr

RÉSUMÉ : Le travail collaboratif, et particulièrement le travail en groupe, sont des approches couramment utilisées dans les études visant à améliorer les processus d'apprentissage. De nombreux travaux dans la littérature se sont intéressés à l'étude de la performance du travail en groupe et notamment l'impact de la composition du groupe sur ses performances ainsi que sur les performances individuelles. Cependant, peu de travaux se sont intéressés à la formation automatique de groupes. Dans cet article nous proposons une méthodologie pour aider les enseignants à former automatiquement des groupes d'étudiants, en prenant en compte non seulement des critères de regroupement mais également le profil des étudiants. Cette méthodologie se base sur une étude préliminaire qui analyse la composition de groupes formés par des étudiants.

Mots clés : Apprentissage collaborative, travail en groupe, formation de groupes

ABSTRACT: Group work has been adopted as an important tool to support collaborative work in order to enhance learning processes. There is a wealth of literature related to group performance and the impact of group composition on group and individual performance. However, very few studies address the issue on how to automatically form groups. This article proposes a methodology that could be used by professors to form groups automatically taking into account different criteria as well as the students' profile. This methodology is based on a pilot study that analyzes group composition of self-formed student groups.

Keywords: group formation, group work, collaborative learning,

1 Introduction

Group work and collaborative work are important pedagogical tools for classroom assignments and have proved to have a strong impact on individual and cooperative learning [13]. Furthermore Web2.0 has created new possibilities for students to engage, interact and collaborate in various learning tasks that may enhance learning processes and the overall learning experience. In this context the didactic challenge for educators is to design and integrate a new set of tools based on specific didactic principles associated with a specific domain of learning [25]. Consider a professor who wants to assign group work to students during a course. One issue he has to handle is: How to form groups? He/she may either let the students form the groups by themselves or may form student groups randomly or based on their physical proximity (e.g. their position in the class,). Even though this is a very easy convenient approach, randomly formed groups may not be fair or may not be the best approach. Group work has many variables and different factors that can influence group performance. For example, literature related to motivation and group learning shows that performance is not only linked to the interest in the subject to learn but may also be related to relations to peers, gender differences, age, individual differences, cultural backgrounds within the group, personality traits, etc. [12, 14]. Furthermore, explaining to students why they form a group and why they are performing certain tasks may lead to better performance [14]. Moreover, a student has to feel comfortable in a group in order to communicate his/her ideas, to express his point of view with his/her group-mates. Thus, a professor may select a special criterion based on which he/she wants to create the groups. This criterion can be either a single criterion or a set of different criteria (e.g. heterogeneous in relation with their background or level of knowledge or homogeneous groups). However forming groups of students, especially for large size classes, is an intractable, time consuming task for the professor.

Previous studies related to collaboration and group work have emphasized the importance of heterogeneity for performance, creativity and learning [15]. A heterogeneous group is made of members that are balanced in terms of diversity based on some criteria. These criteria may be varied: e.g. culture, background, gender, personality, level of knowledge, etc. There is a wealth of literature related to group performance and their composition [10,18] but very

few literature focus on how to automatically form groups. Group work and collaborative learning open a number of challenging research questions related to performance such as: How heterogeneity or homogeneity of a group influence the performance and level of students' competence? Do homogenous groups perform better than heterogeneous groups? How students form groups? Which information can be useful and collected to form student's profile in the context of group formation? and how to form automatically groups taking into account the student's profile?

Let imagine a course in which assignments have to be handled in groups. These assignments may be of different types (guided problem solving exercises, mid or long-term knowledge discovery exercises) and may have different objectives; Literature suggests that groups should be formed differently according to the type of assignments. Homogeneous groups are better for achieving specific goals (e.g. short-term and guided problem solving) while heterogeneous groups are better for long-term knowledge discovery problems [14].

This article is a preliminary study that focuses on how students form groups and proposes a method to form groups automatically. In particular, this article proposes a methodology that could be used to help professors to form groups automatically taking into account different criteria as well as the students' profile. This paper relies on a study case investigating the way students form groups, in the framework of a course where they have to conduct research and present project related ideas and findings in group. Students make use of a Web 2.0–enabled platform to communicate and exchange their ideas on a topic of interest related to the course they have selected themselves.

The article is structured in five sections. The second section presents an overview of literature on collaborative learning, collaborative work, group composition and group formation. The third section presents a pilot study of group work within a heterogeneous classroom supported by a Web2.0 enabled learning environment. The fourth section proposes a methodology that can be used to create groups. The methodology is a tool that enables professors in very different teaching areas to form groups automatically either without a priori knowledge about their students or when groups need to be formed quickly based on certain criteria. The last section discusses conclusions and outlines future work plans.

2 Literature review

2.1 Web 2.0 and e-Learning 2.0

Web 2.0 and associated technologies have changed the way the web is used; Web is a dynamic, social space where participation, collaboration, on-line interaction are core elements. Web 2.0 or social software may be approached from different perspectives: as a new social media tool, a facilitator of new forms of interaction and knowledge sharing [16], an enabler of personal information and knowledge management tools and new didactic tools that facilitate interaction and social processes [2]. Web 2.0 has a large influence on learning approaches (e.g. using wikis, Wikipedia, blogs, micro blogs) and it offers new means to interact, socialize on-line, find information, and communicate using a wide range of new collaborative services. The growing popularity of Web 2.0 suggests that students' learning might be accomplished under circumstances different from those currently used in universities [17]. Learning is not anymore viewed as a unidirectional process, where teachers are in the same place at the same moment than learners, and knowledge is transferred from teachers to learners. Learners are now participants in the learning process thanks to the tools that enable and encourage them to participate, interact and collaborate more easily with other learners, teachers or professors, etc.

2.2 Collaborative learning and cooperative learning

Collaborative learning can be described as a situation in which two or more people learn or attempt to learn something together. Collaborative learning is different from cooperative learning. In cooperative learning peers split the work in tasks and tend to solve these tasks individually and then assemble their results into the final output. While in collaborative learning learners interact and do work together in order to complete their assignments. Despite the difference, many articles use these two terms interchangeable. For example, cooperative learning is defined as an instructional strategy in which students work actively and purposefully together in small groups to enhance both their own and their teammates' learning [27].
Literature emphasizes that collaborative learning is one of most the successful techniques to enhance student performance; several studies report that group work and cooperative learning enhance the learning of an individual compared to when he/she learns alone [2, 3]. Several tools have been proposed to support collaborative learning with the aim of facilitating information and resource sharing between students [4, 5]. Group work can be performed either

by students physically present at the same place and at the same time (synchronous work) or remotely (distant in place and time) through asynchronous work [12]. In this context some studies focused on the learning processes in groups and the appropriate group or factors for a learner to join a learning group.

2.3 Group composition

The quality of the learning process in the context of collaborative work highly depends on the characteristics of the group. One question that has been frequently asked is: what factors influence group performance? and in the e-learning context: what factors influence group work and lead to high quality learning? Previous studies suggest that groups should be formed randomly or by students themselves according to the criterion of their choice and groups should have a small size, made up to four-members and made up of students with different level of knowledge [6]. Related work emphasized the importance of personality attributes, gender, school background, ethnic background, motivation [7, 8] in group performance. The learning style is also an important criterion in group composition [11, 21].

It has been observed that the quality of learning in groups is influenced by their diversity. Heterogeneous groups may outperform homogeneous groups [9]. Some studies emphasized that heterogeneous groups may be more creative and innovative [26] and furthermore they may be more effective for individual learning. One of the first heterogeneity criteria that has been studied was in relation with the level of knowledge and skills [6, 19]. Diversity in background, opinions [1], ideas, personality, gender, are also criteria that can be considered in the heterogeneity [10].

Forming groups, based on these criteria is time-consuming for the professor, especially when a large number of students have to be grouped.

As a consequence, constructing groups automatically would lead to effective group formation, while the criteria the professor wants to use can be taken into account. In the literature, few studies have focused on the way to create groups automatically. Groups may be automatically formed based on thinking styles [21], competence, learning style and interactions [23], learning achievements [24]. The clustering approach proposed are either genetic algorithms [22, 23, 24] or ant colony optimization with Euclidean distance [14, 22].

3 The pilot study

3.1 The StudyBook context

This pilot study aims to investigate the use of a Web2.0 collaborative platform, called StudyBook, to support new ways of teaching and learning. The underlying hypothesis of the overall study is that collaborative groupwork creates natural opportunities for the learners to articulate their understanding, reflect on and justify actions and these activities may improve learning. Furthermore new Web2.0/Web3.0 tools integrated in a learning platform may support group work, collaboration and learning processes in new ways that are more natural for new generation of learners. In particular, the StudyBook project aims to investigate the use of social-collaborative services for collaborative learning and groupwork. StudyBook has already been used in several studies and courses [25]. A screenshot of StudyBook is presented in the Figure 1.

This study has been conducted within the Web Interaction Design and Communication course, an elective course for bachelor students at the Copenhagen Business School in Denmark. This course enrolled 46 bachelor students from seventeen different countries and from different study programs (Business Administration, European studies, International Business, Marketing, HRM, etc). Within this course, a project has to be conducted collaboratively, partly in group and partly individually. The students had to select a topic of interest, to form groups of three or four and work in group on this topic. Within their projects, the students define their own topic to work on in relation with their core areas of interest and the main course topics. As the course is an elective course, students are from various study programs and years, they do not know each other. The resulting groups are thus heterogeneous in knowledge background, in level, in culture.

The project work is divided in two steps:

1. Work in group: The group has to work on the project, by using StudyBook platform that facilitates interaction between students to work on the project. Indeed, students are not in the same study programs and thus do not have common free time-slots. Their project work had to be constructed collaboratively using wikis and afterwards makes a presentation.

2. Work individually: In the second step, students have to continue to work on the selected topic individually in order to write a ten page report that will be graded.

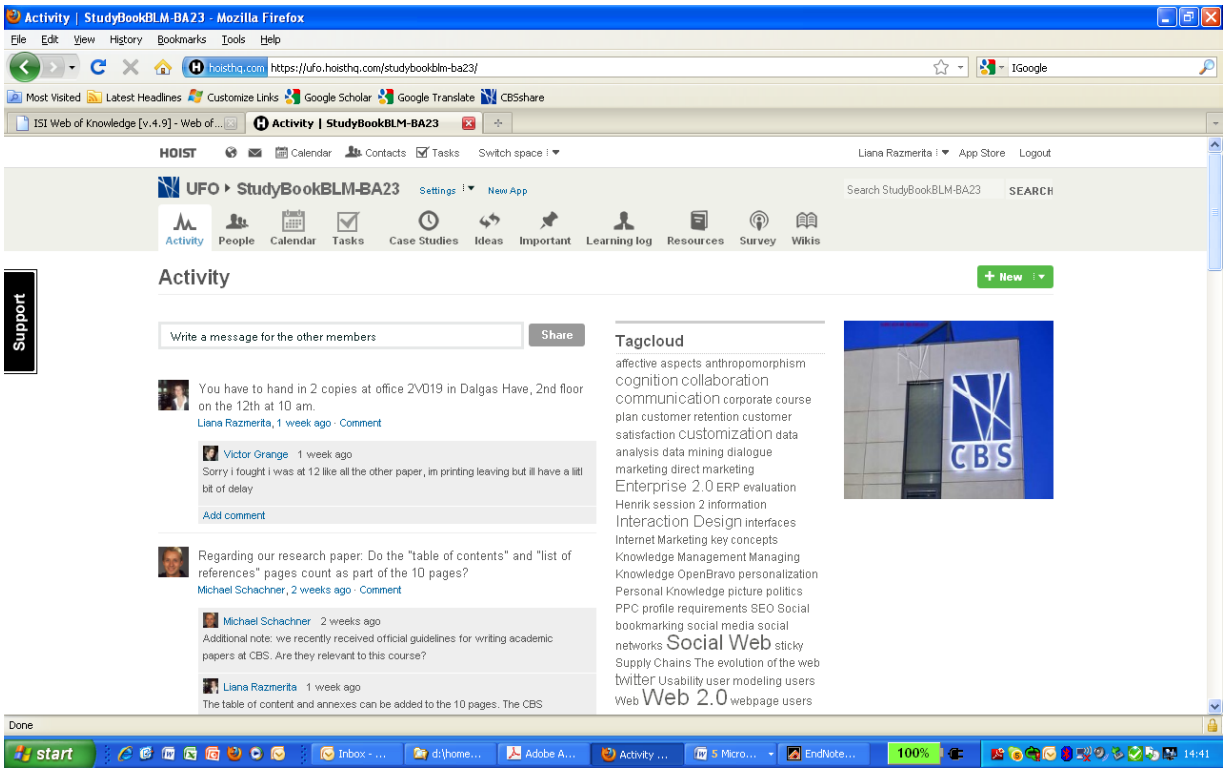

Figure 1 Screenshot of StudyBook platform

3.2 Data collection: the questionnaire

Two questionnaires have been designed with two objectives in mind: to assess the student's perception of StudyBook as a learning platform and collect data about students' and the way they form groups. Within the following section of the article we focus on the findings related to the analysis of the groups that have been formed and not on the evaluation of the StudyBook platform. In the questionnaire, data about students' include their background (study programs), culture (the country they come from), topics of interests and level of knowledge in relation with the following topics of interests related to the course:

- New technologies and new communication channels
- Interaction design underlying concepts
- Hands-on, experience, real-case studies and strategies
- Evolution of the Web: Web 2.0, Web 3.0
- User modeling and cognition
- Evaluation frameworks
- Social Networks
- Programming technical knowledge
- Cultural differences and web design
- Usability and user experience

Students had to answer about both their level of knowledge (beginner, intermediate, advanced) and research interests (not interested, interested, very interested) in relation with the above mentioned topics of interests. In addition, students were asked about the way they have formed their groups. The following possibilities were proposed:

- Common topics of interests
- Compatible working periods
- Similar background and culture
- Different background and culture
- You know or you like the other persons in the group
- You selected a group because you had to form a group

3.3 Data Analysis: evaluation of groups

This section presents group formation in heterogeneous classes. Among the 46 students enrolled in the course, 29 students have answered the questionnaire. All 29 students have filled in questions but in two groups not all the members of the groups answered the questionnaire. Eight groups in total have answered the questionnaire. First of all, students declare they have enjoyed working in groups and have perceived positively their workgroup collaboration.

A first analysis of the answers shows that the students who answered the questionnaire are from fifteen different countries all over the world. The most represented countries are Denmark, Spain and France (4 students from each country) and Hong Kong (3 students). Thus, the formed groups are probably heterogeneous in terms of cultural background.

In our overall study heterogeneity is related to students' background (different study programs), cultural (different countries), topics of interests and the level of knowledge. In this preliminary study, we specifically focus on level of knowledge and topics of interest. In order to study homogeneity and heterogeneity of groups we compute a metric distance between students taking into account their declared level of knowledge or/and topics of interests. Each student is represented by a vector and the normalized Euclidean distance was used to compute the distance between two vectors. The average distance in a group (also referred to as heterogeneity) is computed as the average of the distances between all the students in the group, and results will be discussed in the next subsections.

3.3.1 Group heterogeneity in relation with their level of knowledge

Studies suggest that if students organize themselves in groups, they usually tend to form homogeneous groups [12]. The first column of Table1 represents the average distance in terms of level of knowledge within groups and within the class. The average distance between all students in the class is 0.23 and standard deviation is 0.17. Less than 3% of the pairs of students have a distance greater than 0.8. This means that few students are very different in terms of their level of knowledge. Among the eight groups, the average distance is 0.18, the most homogeneous group has an average distance of 0.02 and the less homogeneous has an average distance of 0.48. Three groups have an average distance less than 0.05. The average distance between students in the groups is smaller than the average distance in the whole matrix. Thus groups formed by students tend to be similar in terms of knowledge. The findings of our study confirm the fact that students tend to form homogeneous groups.

Group	Average distance in the level of knowledge	Average distance in the topics of interest
Gr1	**0.02**	**0.27**
Gr2	**0.48**	0.10
Gr3	0.24	0.23
Gr4	0.16	0.15
Gr5	0.22	0.20
Gr6	**0.02**	**0.05**
Gr7	0.05	0.18
Gr8	0.24	0.22
All groups	0.18	0.18
Class	0.23	0.18

Table 1 Average distances within groups

3.3.2 Group heterogeneity in relation with their topics of interests

The second column of Table 1 represents the average distance in terms of topics of interest within groups and within the class. The average distance between all students in the class is 0.18 and standard deviation is 0.12. Less than 0.2% have a distance greater than 0.7. Despite their heterogeneity, students appear to be quite homogeneous in relation with their level of interest.
Among the eight groups, the average distance is also 0.18, which is similar to the average distance among groups in terms of level of knowledge. The most homogeneous group has an average distance of 0.05 and the less homogeneous one has an average distance of 0.27. Five groups out of the eight have similar average interest distance and average knowledge distance. Within the class, the average distance between students in terms of topics of interest (0.18) is lower than the one in terms of level of knowledge (0.23); students are more similar in terms of topics of interest than in level of knowledge. However, the average distance within groups is similar to the average distance within the class; thus we can deduce that students do not tend to form similar groups in terms of topics of interest.

3.3.3 How students formed groups

Based on the assignment requirements described in section 3.1, the results of the study in relation with how students form groups are presented in Figure 2. Based on the questionnaire's answers, a large majority of students form groups based on identified common topics of interests (68,97%), followed by affinity with the others members (10, 34%) and different backgrounds and culture (10, 34%).

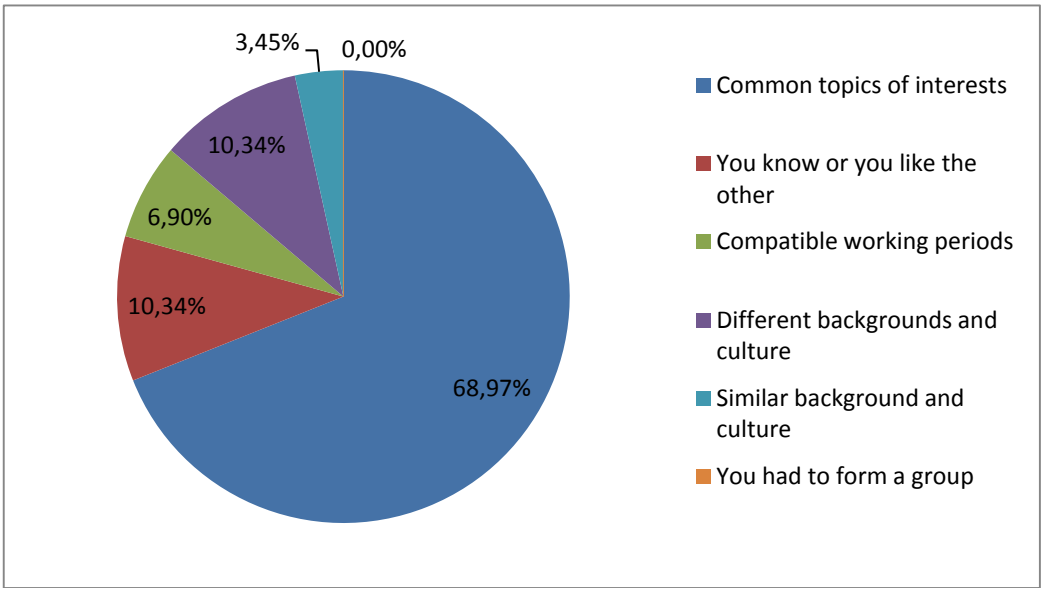

Figure 2 How students form groups

As a conclusion, the groups formed by students tend to be homogeneous in terms of the level of knowledge. However, these groups are not really homogeneous in terms of the topics of interests, whereas the students answered in the questionnaire they formed homogeneous groups in terms of common topics of interests. Thus, when leaving students to form groups by themselves, the resulting groups may not have the expected characteristics. Thus a tool that automatically forms groups may be highly useful to ensure that groups have given characteristics and lead to the expected type of learning and performance. In the following section, we propose a methodology to form groups automatically according to students' profile and specified citeria.

4. Methodology for forming groups automatically: Clustering students

The proposed methodology comprises four main steps: collecting data about students, initialize the vectors representing the students' characteristics, clustering students, evaluating the group performance.

As the data collected about students will be used by the clustering method (grouping method), the professor has to decide which criteria and which type of data is useful to form groups. Data can be collected through questionnaire or based on existing data about students.

Step 1:
Collect data about students. Questionnaires can provide an effective way to collect data about the students. However data can be provided by students or by the administration (e.g., the study program, the background) or by the professor (possible topic, learning concepts, learning objectives, students' skills, heterogeneous versus homogeneous groups). As described previously, in our pilot study the following data (topics of interest, level of knowledge, country of origin, study program) was collected with the purpose of automatic group formation and other purposes. The collected data needs to be preprocessed in order to be used by the algorithm. For example qualitative data need to be transformed into quantitative data.

Step 2:
Initialize the input vectors. Each student is represented by a vector with features/components that are made up of the attributes values associated with the student, initialized from the questionnaire.

Step 3:
Select and run the clustering algorithm (e.g. K-means, hierarchical clustering, etc) in order to generate the groups; depending on the algorithm select the number of clusters, the size of the cluster or the quality criterion. One distance measure that can be used to compute similarity between students in groups is for example the Euclidean distance:

$d(x,y) = \sum_i^n |x_i - y_i|$, where $x=[X_i]$ $i=1..n$ represents one student's profile and $y=[Y_i]$ $i=1..n$ represents another student's profile.

Once the grouping task is achieved, the students can work on their assignments in groups formed.

Step 4:
Evaluate the group performance in relation with the selected criterion or criteria. Depending on the assigned task and the learning objectives, the professor might decide to evaluate the performance of the groups and may decide to change or keep the type of clustering method for the following assignments. The performance may be evaluated at either group level or individual level.

5. Conclusion

This paper investigates group formation in the context of e-learning. This paper proposes a methodology to form groups of students which includes a preliminary study on self-formed groups within a course in a heterogeneous class. A pilot study has been conducted to study the way students form groups in the context of heterogeneous classes. According to questionnaires findings, students have enjoyed working in groups and have perceived positively their workgroup collaboration. However not all students in the class have formed a group.
Furthermore, the analysis of the groups reveals that students tend to form homogeneous groups in relation with the level of knowledge, which is in line with what other previous studies have suggested, for example[12]. In relation with the topics of interests, according to the students' answers, the groups were formed based on topics of interests but according to the analysis of groups as presented in section 3.3.2 groups are not as homogeneous as students declared. The proposed methodology using clustering algorithm tools for group formation is a useful tool to help professors form groups automatically using a certain criteria. By using this methodology, the professor decides the criterion or criteria to be used in order to form group in relation with type of assignments, type of course and learning objectives. Furthermore he/she can also decide if the groups should be homogeneous or heterogeneous and the metric used to compute the similarity distance between users, as well as, the type of clustering algorithm.
We plan to further test the methodology and the performance of automatically formed groups within different type of courses and assignments. In further studies we will assess the performance of automatically formed groups using different clustering methods.

Bibliography and webography

[1] Siemens, G., "Connectivism: A Learning Theory for Digital Age", *Elearnspace, http://www.elearnspace.org/Articles/connectivism.htm retrieved the 16th of May*

[2] Dansereau, D., Johnson, D. "Cooperative learning", *In: Druckman, D., Bjork, R.A.: Learning, Remembering, Believing: Enhancing Human Performance, National Academy Press, Washington, D.C. (1994), pp 83-111.*

[3] Slavin, R.E., "When Does Cooperative Learning Increase Achievement?", *Psychological Bulletin Vol. 94, 1983, pp 429-445.*

[4] Florea, A.M., "An Agent-based Collaborative Learning System: Advanced Research in Computers and Communications in Education", *In Proc. of the Int. Conf. on Computers in Education, IOS Press, 1999, pp. 161-164.*

[5] Krejins, K. and Kirschner, P.A., Jochems, W., "The Sociability of Computer-Supported Collaborative Learning Environments", *Educational Technology and Society, Vol. 5, No. 1 (2002), pp 26-37.*

[6] Slavin, R.E., "Developmental and motivational perspectives on cooperative learning: A reconciliation ", Child Development, *Vol. 58, (1987), pp 1161-1167.*

[7] Bradley, J.H., Herbert, F.J., "The effect of personality type on team performance", *Journal of Management Development, Vol. 16, (1997), pp 337-353.*

[8] Alfonseca, E., Carro, R., Martín, E., Ortigosa, A., Paredes, P., "The Impact of Learning Styles on Student Grouping for Collaborative Learning: A Case Study ", *User Modeling and User-Adapted Interaction Journal, Vol. 16, No. 3-4, (2006), pp 377-401.*

[9] Nijstad, B. A., Carsten, K. W. D., "Creativity and group innovation", *Applied Psychology, Vol. 51, NO. 3, (2002).*

[10] Slavin, R.E., " Cooperative learning: Theory, research, and Practice", *2nd ed. Boston: Allyn and Bacon. (1995)*

[11] Martin, E., Paredes, P., "Using Learning Styles for Dynamic Group Formation in Adaptive Collaborative Hypermedia Systems", *Proceeding of the conference on Adaptive Hypermedia and Collaborative Web-based Systems (AHCW'04), 2004.*

[12] Souren, P., Seetharaman, P., Samarah, I., Mykytyn, P., "Impact of Heterogeneity and collaborative conflict management style on the performance of synchronous global virtual teams", *Information and Management, Vol. 41, (2003), pp 303-321.*

[13] Johnson, D.W., Johnson, R.T., "What is cooperative Learning?", In *Perspectives on Small group Learning: Theory and practice*, ed. M. Brubacher.

[14] Bekele, R.,"Computer-Assisted Learner Group Formation Based on Personality Traits", *Phd Thesis, Hamburg University, 2005*

[15] Rich, M., "A Learning Community on the Internet: An Exercise with Masters Students", In *Proceedings of Americas Conference on Information Systems*, Indianapolis, 1997.

[16] Kirchner, K., Razmerita, L., Sudzina, F., "New Forms of Interaction and Knowledge Sharing on Web 2.0", in *Web 2.0: The Business Model*, E.D. M. Lytras, P. O. De Pablo, Ed. USA: Springer Science and Business Media, 2008, pp. 21-37.

[17] Mondahl, M., Rasmussen, J., Razmerita, L., "Web 2.0 Applications, Collaboration and Cognitive Processes in Case-based Foreign Language Learning," in *2nd World Summit On The Knowledge Society, WSKS 2009*. vol. LNAI 5736, M. Lytras, Ed. Creete, Grece: Lecture Notes in Computer Science, Springer-Verlag Berlin, 2009, pp. 98-107.

[18] Sharan, S.Handbook of cooperative learning methods (2nd ed.). Westport, CT: Praeger. 1999.

[19] Webb, N. "Peer interaction and learning in small groups". In *Intern. Journal of Educational Research*, 13, 21–39, 1989.

[20] Lin, S.S., Sun, C.T., "Team-forming recommendation for Web-based cooperative learning: Learning effect and partner preference". In *proceedings of the annual meeting of National Association for Research on Science and Teaching*. New Orleans, USA, (2000)

[21] Wang, D.Y. , Lin, S.J. , Sun, C.T., "DIANA: A computer-supported heterogeneous grouping system for teachers to conduct successful small learning groups", *Computers in Human Behavior*, Vol. 23, Issue 4, July 2007, pp 1997-201

[22] Graf, S., Bekele, R., "Forming Heterogeneous Groups for Intelligent Collaborative Learning Systems with Ant Colony Optimization", In *Intelligent Tutoring System Conference*, pp 217-226

[23] Ho, T.F., Shyu, S.J., Wang, F.H., Li, C.T., "Composing High-heterogeneous and High-interaction Groups in Collaborative Learning with Particle Swarm Optimization", In *2009 World Congress on Computer Science and Information Engineering*, 2009, pp 607-611

[24] Chan, T., Chen, C.M., Wu, Y.L., Jong, B.S., Hsia, Y.T., Lin, T.W., "Applying the genetic encoded conceptual graph to grouping Learning", In *Expert Systems with Applications*, 2010, Vol. 37, pp 4203-4118

[25] Mondahl, M., Rasmussen, J., Razmerita, L."Web 2.0 Applications, Collaboration and Cognitive Processes in Case-based Foreign Language Learning," in 2nd World Summit On The Knowledge Society, WSKS 2009. vol. LNAI 5736, M. Lytras, Ed. Creete, Grece:

[26] Paulus, P.B., Nijstad, B.A. Group Creativity: Innovation through Collaboration: Oxford University Press, 2003

[27] Abrami, P.C., Poulsen, C., Chambers, B. "Teacher motivation to implement an educational innovation: Factors differentiating users and non-users of cooperative learning," Educational Psychology, vol. 24, pp. 201-216, 2004.