Group Formation for Collaborative Learning
A Systematic Literature Review

Chinasa Odo[1(✉)], Judith Masthoff[2], and Nigel Beacham[1]

[1] University of Aberdeen, Aberdeen, UK
{r01cro17,n.beacham}@abdn.ac.uk
[2] Utrecht University, Utrecht, The Netherlands
j.f.m.masthoff@uu.nl

Abstract. This paper presents a systematic literature review (SLR) that investigates group formation, as a first step towards automated group formation for collaborative learning. Out of 105 papers selected for review, after using specific selection and a quality assessment method, a final list of 21 relevant studies was selected for analysis. The review revealed the current state of the art in group formation.

Keywords: Systematic review · Group formation · Collaboration

1 Introduction

A group is a structure that creates boundaries and sustains interaction amongst members. Grouping is essential in human nature specifically in learning. It is central to human lives, and it is hard to imagine human existence without a group [17]. Being in a group brings about collaboration in the achievement of set goals. Collaboration is rooted in the theory of [36] which explains the Zone of Proximal Development. Vygotsky's theory believes in the construction of knowledge through social interactions among peers in a community. Research has shown that learners feel more engaged when they are given the opportunity to be part of the learning process [7,17] because it is student-centered learning.

Groups, are formed when two or more people interact and influence each other's discussion for learning and understanding learning contents more completely [10,12]. In the case of online collaboration, it is based on the concept of knowledge construction and gradual building of knowledge through asynchronous online discussion among learners and the instructor. The instructor, acts as a facilitator who provides appropriate resources, learning activities and utilizes knowledge of learners' personality profiles.

This paper investigates the current state of the art in the ways collaborative learning groups are formed, with a particular interest in group types (homogeneous or heterogeneous), learner characteristics, group sizes and algorithms which should inform the automatic formation of learning groups.

2   Review Strategy

This research presents a review on group formation following the guideline by [32], which outlined a practical guide on how to carry out a systematic literature review in computer science, based on the work in [18]. Whilst we are interested in automated group formation, this review will be wider, given the limited work in this area, to inspire work on automated group formation.

Research Questions. The following research questions will be answered:

1. Do teachers consider homogeneity or heterogeneity when forming groups?
2. What learner characteristics are important criteria for forming groups?
3. What size is recommended as an ideal learning group for collaboration?
4. What are the various techniques in use for automated group formation?

Literature Sources and Data Gathering. The study included papers from online databases (EEE, Springer, ACM, Inspire, Crossref, ArXiv, GVK DBLP, Pubmed, PLOS, DOAJ) and search engines (Google Scholar, CiteSeerx) supported by Jabref. A search string was constructed using the method in [32], using synonyms for keywords. The search strings used were: *(a)* (group OR grouping OR team OR teaming) AND (forming OR formation); *(b)* (group OR grouping OR team) AND (collaboration OR collaborate); *(c)* (peer OR peering) AND (recommending OR recommendation OR recommender OR recommend); *(d)* (group OR grouping OR team Or teaming) AND (size OR sizing).

   A search on the titles, abstract and keywords resulted in the first set of 105 papers published from 2002–2017. The identified papers included some that did not address the purposes of this study, stored in multiple databases, or published in many sources. For selection, we applied inclusion and exclusion criteria.

Inclusion and Exclusion Criteria. A paper is included only if: (1) it contains online/conventional collaboration, or group/team formation; (2) it is published between 2002 and 2017; (3) it is duplicated or stored in multiple sources, only one copy is selected; (4) it has multiple publications, the most recent one or full version is selected; (5) it has both a conference and journal version, the journal version is selected. A paper is excluded if it: (1) is not related to education; (2) is presented in a language other than English; (3) it is only available in the form of a presentation; (4) it does not address the problem of group collaboration. The application of these criteria reduced the study papers to 48.

Quality Assessment. To assess and analyze the selected papers, a 9 item quality assessment checklist was developed and assessed as follows:

1. Venue was evaluated depending on the paper source: (i) For conference and workshop papers, the Computing Research and Education rankings (CORE) were used [8], with values assigned as A = 1.5, B = 1, C = 0.5, No ranking = 0.

(ii) For journal articles, the Journal Citation Report (JCR) was used which reports citation data [16]. Journals are ranked as Q1-Q4, with values assigned as Q1 = 2, Q2 = 1.5, Q3 = 1, Q4 = 0.5, No JCR = 0.

2. Other items such as if paper had been cited and whether it was relevant to each research question were rated as Yes = 1, No = 0.

Papers with an overall quality assessment score greater than 5 were included. At the end of this stage, a list of 21 papers were selected: S01 [3], S02 [5], S03 [6], S04 [9], S05 [20], S06 [21], S07 [25], S08 [29], S09 [31], S10 [34], S11 [37],S12 [38], S13 [1], S14 [2], S15 [4], S16 [13] S17 [15], S18 [22], S19 [23], S20 [26], S21 [30].

3   Results and Discussion

Research Question 1 - Which Group Type Is Considered (homogeneous/heterogeneous) by Teachers/instructors When Forming Learning Groups? S01, S02, S06, S13 S04, S05 S17 and S18 advocated the formation of heterogeneous groups which is in line with the study of [15,39]. Only S01 discussed homogeneous group formation which focused on language preference. Study by [33] also noted that homogeneous groups are less stigmatized.

Research Question 2 - What Learner Characteristics Are Considered as Important Criteria for Forming Collaborative Learning Groups? Most reviewed papers focused on only one learner characteristic for group formation. The exceptions are S01, which considered gender and language preferences, S20 which considered interest and background for group formation and S17 which considered personality traits and performance feedback. While [27] advocated for group collaboration with interest in combination different learner characteristics. S02, S13 and S16 considered complementary skills of strong and weak learners. S02 suggested that: (i) All members should be expert in one of the identified complementary skills; (ii) Only one member can lead the team. This is like S07, which proposed to combine learners who are more knowledgeable with those who are less knowledgeable but did not mention how this knowledge will be determined. S04, S05 and S18 considered learning style as a measure to bring learners together. However, the use of learning styles is controversial as reported in [11]. S10 considered feedback as a criterion for group formation, maintaining that the quality of previous collaboration is important when forming a group. S11 and S12 mentioned diversity but were not specific on the area of diversification. S17 and S19 proposed personality traits as criteria for forming a group. Research by [24] shows that personality trait is an important factor in collaboration. S02, S03 and S13 proposed getting learners' team work profile using a questionnaire.

Research Question 3 - What Group Size Is Recommended as an Ideal for Learning Collaboration? S05, S11, S12 and S15 suggested small groups

without stating any number to constitute a group. S06 suggested that the size should depend on the tutor's choice. only S15 mentioned specific group sizes of 3, 5 and 7. This is supported by the Ringelmann effect (as mentioned by [19]) who noted that members become less productive as group size increases.

Research Question 4 - What Are the Various Techniques in Use for Automatically Assigning Learners to Groups? In S01 and S21, a binary integer technique which takes the values of 0 or 1 was proposed. Data mining was proposed by S04. There are many types of data mining techniques which are used to explore and analyze large data set in order to discover meaningful patterns as noted by [28,35], but S04 did not specify the type they used for grouping. S05 and S11 used genetic algorithms in group formation problem. This consists a set of students S and a set of groups G. The goal is to allocate all learners in S to a group in G, such that the groups are as heterogeneous as possible. The genetic approach in [14] shows that learners are drawn from the population to produce the fittest groups by changing individuals to form better groups which takes different learner characteristics into account.

S06, S08, S09 used an approximation algorithm. In S06, group allocation is made by finding individuals to act as leaders for each group by minimizing a leadership cost function, and then adding individuals to the groups by minimizing the communication cost function (using Greedy Search). The user provides feedback on the resulting groups in terms of which learners to keep in the groups. The algorithm is run again, till the user is satisfied. In S08, a group allocation is made using learners activity ratings; learners with similar preferences are put together. This approximation is evolved into a final group allocation with the desired number of groups. S09 uses backtracking.

S07 used a semantic algorithm, which maximized the diversity in knowledge in the groups. Artifacts (such as essays) produced by learners were analyzed to extract knowledge of each learner. The learners concepts were aggregated into a unified data model, and used to calculate diversity. S10 is based on a group technological approach where similar characteristics are identified and grouped together to take advantage of the similarities. The input data is composed of two matrices: (1) learner characteristics compatibility and (2) assignments of the characteristics to learners. A clustering approach is then used to form groups based on these matrices. S14 used a Bayesian network. Initially learners were divided into disjoint teams. After every activity, learners evaluates their peers by stating the most predominant role of each teammates. At each iteration, Bayesian learning was employed to update the probability for a learner given the evaluation history, these probabilities are then used to form the next teams.

Finally, S17 and S20 used Ant colony optimization and Particle swarm optimization respectively. The first is inspired by the collective foraging behaviour of specific ant species. The objective of the algorithm is to maximize the heterogeneity of all groups based on the Goodness Heterogeneous values of all groups. In S20, the particle swarm optimization technique is used. In this technique, each particle has: (1) a current position in the search space, (2) a current velocity, and

a personal best position in the search space. During each iteration, each particle in the swarm is updated using (1) and (2). In S20, each particle represents a distribution of learners over groups.

4   Conclusions and Future Work

This paper provided a systematic literature review on group formation for collaborative learning, as a first step towards automated group formation by a computer agent for group collaborative learning. In the light of the findings, the research was able to identify that (1) The reviewed papers have not specifically considered which of the learner characteristics are considered important when forming a group but tended to focus on a particular characteristic. (2) The reviewed papers did not mention an ideal size to consider when forming a group. (3) The reviewed papers used a wide variety of algorithms with no studies to compare the relative effectiveness of such algorithms. Our future studies will use a mixed method research method with triangulation to determine which learner characteristics to combine to achieve effective collaborative learning groups.

References

1. Al-Adrousy, W.M., Ali, H.A., Hamza, T.T.: A recommender system for team formation in MANET. J. King Saud Univ. Comput. Inf. Sci. **27**(2), 147–159 (2015)
2. Alberola, J.M., del Val, E., Sanchez-Anguix, V., Julian, V.: Simulating a collective intelligence approach to student team formation. In: Pan, J.-S., Polycarpou, M.M., Woźniak, M., de Carvalho, A.C.P.L.F., Quintián, H., Corchado, E. (eds.) HAIS 2013. LNCS (LNAI), vol. 8073, pp. 161–170. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40846-5_17
3. Amarasinghe, I., Leo, D.H., Jonsson, A.: Intelligent group formation in computer supported collaborative learning scripts. In: Chang, M., Chen, N., Huang, R., Kinshuk, Sampson, D.G., Vasiu, R. (eds.) 17th IEEE International Conference on Advanced Learning Technologies, ICALT 2017, Timisoara, Romania, 3–7 July 2017, pp. 201–203. IEEE Computer Society (2017)
4. Babar, M.A., Kitchenham, B.: The impact of group size on software architecture evaluation: a controlled experiment. In: Proceedings of First International Symposium on Empirical Software Engineering and Measurement (ESEM 2007), pp. 420–429, September 2007
5. Battur, S., et al.: Enhancing the students project with team based learning approach: a case study. In: 2016 IEEE 4th International Conference on MOOCs Proceedings of the Innovation and Technology in Education (MITE), pp. 275–280, December 2016
6. Borges, J., Dias, T.G., Cunha, J.F.E.: A new group-formation method for student projects. Eur. J. Eng. Educ. **34**(6), 573–585 (2009)
7. Burke, A.: Group work: how to use groups effectively. J. Eff. Teach. **11**(2), 87–95 (2011)
8. Computer Research and Education: CORE Rankings Portal - Computing Research and Education (2017)

9. Costaguta, R., de los Angeles Menini, M.: An assistant agent for group formation in CSCL based on student learning styles. In: Proceedings of the 7th Euro American Conference on Telematics and Information Systems, EATIS 2014, pp. 24:1–24:4. ACM, New York (2014)
10. Dillenbourg, P.: What do you mean by 'collaborative learning'?, no. 1, pp. 1–19 (1999)
11. Coffield, F., Moseley, D., Hall, E., Ecclestone, K.: Should we be using learning styles? What research has to say to practice. Technical report, The Learning and Skills Research Centre (2014)
12. Galanes, G.J., Adams, K.K.L.: Effective Group Discussion: Theory and Practice. McGraw-Hill, Boston (2013)
13. Garcia, P., Balmaceda, J.M., Schiaffino, S., Amandi, A.: Automatic detection of team roles in computer supported collaborative work. IEEE Latin Am. Trans. **11**(4), 1066–1074 (2013)
14. Gogoulou, A., Gouli, E., Boas, G., Liakou, E., Grigoriadou, M.: Forming Homogeneous, Heterogeneous and Mixed Groups of Learners, pp. 33–40 (2007)
15. Graf, S., Bekele, R.: Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization. In: Ikeda, M., Ashley, K.D., Chan, T.-W. (eds.) ITS 2006. LNCS, vol. 4053, pp. 217–226. Springer, Heidelberg (2006). https://doi.org/10.1007/11774303_22
16. JCR: Journal Citation Reports (2017)
17. Johnson, D.W., Johnson, R.T.: Assessing Students in Groups: Promoting Group Responsibility and Individual Accountability. Corwin Press, Thousand Oaks (2004)
18. Kitchenham, B., et al.: The impact of limited search procedures for systematic literature reviews: a participant-observer case study. In: Proceedings of the 3rd International Symposium on Empirical Software Engineering and Measurement, pp. 336–345 (2009)
19. Kravitz, D.A., Martin, B.: Ringelmann rediscovered: the original article (1986)
20. Lescano, G., Costaguta, R., Amandi, A.: Genetic algorithm for automatic group formation considering student's learning styles. In: 8th Euro American Conference on Telematics and Information Systems, EATIS 2016, Cartagena, Colombia, 28–29 April 2016, pp. 1–8. IEEE (2016)
21. Li, C.T., Shan, M.K.: Composing activity groups in social networks. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management, CIKM 2012, pp. 2375–2378. ACM, New York (2012)
22. Liu, S., Joy, M., Griffiths, N.: An exploratory study on group formation based on learning styles. In: Proceedings of the IEEE 13th International Conference on Advanced Learning Technologies, pp. 95–99, July 2013
23. Lykourentzou, I., Antoniou, A., Naudet, Y.: Matching or crashing? Personality-based team formation in crowdsourcing environments (2015)
24. Lykourentzou, I., Antoniou, A., Naudet, Y., Dow, S.P.: Personality matters: balancing for personality types leads to better outcomes for crowd teams. In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, CSCW 2016, pp. 260–273. ACM, New York (2016)
25. Manske, S., Hoppe, H.U.: Managing knowledge diversity: towards automatic semantic group formation. In: Chang, M., Chen, N., Huang, R., Kinshuk, Sampson, D.G., Vasiu, R. (eds.) 17th IEEE International Conference on Advanced Learning Technologies, ICALT 2017, Timisoara, Romania, 3–7 July 2017, pp. 330–332. IEEE Computer Society (2017)

26. Maria-Iuliana, D., Constanta-Nicoleta, B., Alexandru, B.: Platform for creating collaborative e-learning communities based on automated composition of learning groups. In: Proceedings of the 3rd Eastern European Regional Conference on the Engineering of Computer Based Systems, pp. 103–112, August 2013
27. Odo, C.R., Masthoff, J.F.M., Beacham, N.A., Alhathli, M.A.E.: Affective State for Learning Activities Selection, May 2018
28. Pujari, A.K.: Data Mining Techniques. Universities Press (2001)
29. Roy, S.B., Lakshmanan, L.V.S., Liu, R.: From group recommendations to group formation (2015)
30. Sadeghi, H., Kardan, A.A.: Toward effective group formation in computer-supported collaborative learning. Interact. Learn. Environ. **24**(3), 382–395 (2016)
31. Sancho-Asensio, A., Sol, X., Montero, J., Navarro, J., Canaleta, X., Vernet, D.: Support tool for the formation of working groups in collaborative learning environments. In: Proceedings of the 9th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–6, June 2014
32. Silva, R.L.S., Neiva, F.W.: Systematic literature review in computer science - a practical guide (2016)
33. So, A., Agbayewa, J.O.: Effect of homogenous and heterogeneous ability grouping class teaching on student's interest, attitude and achievement in integrated science. Int. J. Psychol. Couns. **3**(3), 48–54 (2011)
34. Srba, I., Bielikova, M.: Dynamic group formation as an approach to collaborative learning support. IEEE Trans. Learn. Technol. **8**(2), 173–186 (2015)
35. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2009)
36. Vygotskii, L.S.L.S.: Thought and Language. M.I.T. Press, Massachusetts Institute of Technology (1962)
37. Wang, D.Y., Lin, S.S.J., Sun, C.T.: DIANA: a computer-supported heterogeneous grouping system for teachers to conduct successful small learning groups. Comput. Hum. Behav. **23**, 1997–2010 (2007)
38. Wichmann, A., Hecking, T., Elson, M., Christmann, N., Herrmann, T., Hoppe, H.U.: Group formation for small-group learning: are heterogeneous groups more productive? In: Proceedings of the 12th International Symposium on Open Collaboration, OpenSym 2016, pp. 14:1–14:4. ACM, New York (2016)
39. Zamani, M.: Cooperative learning: homogeneous and heterogeneous grouping of Iranian EFL learners in a writing context. Cogent Educ. **3**(1), 1149959 (2016)