We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX INDEXED

CLARIVATE ANALYTICS

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

A Case Study of Using Big Data Processing in Education: Method of Matching Members by Optimizing Collaborative Learning Environment

Keiko Tsujioka

Abstract

The purpose of this paper is to optimize the combination of members for collaborative learning that utilized learning management system (LMS), a kind of social media. It is considered that there is a problem of this combinatorial optimization because of various discrete elements in education and it is difficult to find exact solution. Then, we have solved this problem by the method of approximate solution in nursing science class with big data processing, for instance, individual traits, learning outcome, and so on. The result shows continuously learning effects. We will report in this fundamental research how to gather learners' various data and optimize matching members of team by local searching. It might be explained how to solve problems of combinatorial optimization by AI.

Keywords: combinatorial optimization, matching members of team, method of approximate solution, big data processing, collaborative learning, feedforward control, feedback control

1. Introduction

Effective collaborative learning is required in nursing science class because of the shortage of numbers in an aging society with a declining birthrate in Japan. LMS, a kind of social media, which any course members are allowed to connect with all information uploaded by them, such as movies, documents, message and so on, likes a social network system, so called computer supported collaborative learning (CSCL) [1]. This system has been brought into nursing science class in order to prepare for practical training with team members. It seems effective for collaborative learning; however, there happened problems which team members had difficulties interacting with each other. It is supposed that there were problems in their relationship among them. From this reason, we have addressed to find out the method of combinatorial optimization for team members so that students can interact with each other through LMS.

CSCL has been studied by many researchers [2], because collaborative learning is expected to have learning effects through interactive communication among

group members [3]. Along with developments of social network, the problems of the relationships between individual and individual and individual and groups or community have become revealed [4]. Koschmann [5] pointed out that it is difficult to find out the solution of CSCL problems because the educational system is related to lots of elements and then complicated. From this point of view, it is considered that finding out the exact solution of combinatorial optimization is not easy because of computational complexity, but approximate solution might be possible to reach better solution by local search [6, 7].

In the field of learning sciences, however, Sawyer said that the innovated reform in education, like scaling up from systems approach, is difficult to succeed [8]. The method of scaling up, for instance, a case of server, improves the whole function of the system not by reforming the sever system but by raising the level of CPU. Dede [9] mentions, although the reform of fast food may transform easily to any restaurants affiliating with a certain franchise, it is difficult to prevail a new type of strategy for instructions even in the same school and obtain general acceptances. Therefore, it is predicted to be unsuccessful cases in Education by scaling up of traditional method, because of its definition which were determined by how much becoming widespread and the level of the reliability.

Upon this, he recognizes the value of successful cases which are given priority to a criterion in a certain context and an adjustment with practical research [10, 11]. The reason of his viewpoint is that it is important to be customized to each educational field because it is difficult to adjust with rapid progress or reform, like vegetation and animals are not able to adjust with their rapid habitant changes. Consequently, Dede [12] said "Examining scalability in the context of his subset of powerful conditions may yield a workable index, but only investigation its feasibility by using real data can determine the potential validity and value of such a measure."

For instance, studies of educational data mining (EDM) have been increasing in quantity and analyzed learners' behavior from various aspects [13] and predicted student performance [14]. The research group of Márquez-Vera [15] has found out the method for predicting dropout students as soon as possible by different data mining approaches with high dimensionality and imbalanced data [16]. Similarly, studies of social network analysis (SNA) have gathered learners' data related to not only behavior but also relationships between small groups and individuals. In the field of social psychology, researchers have been studying about small groups. Guetzkow [17] reported that the conflicts of relationships might be influenced by personality or a sense of value and those of problem solving for projects might be influenced by traits of perception or cognition. In Japan, research groups have continued studying and found out that if conflicts of relationship within groups are lower and those of projects within them are higher, learning effects would be higher. And moreover, the motivation of members would become higher to the next subject [18]. Those results of experiments were outcomes within laboratories and not practical ones; however, it must be a good example of a successful combination of team members.

Sawyer [7] introduced that "scale-up researchers successful strive to improve the implementation with each successive interaction of design-implement-evaluate-redesign cycle." Then, we suppose that it is important to choose successful cases in education when we optimize matching members of team with scale-up method of design-implement-evaluate-redesign cycle.

2. Design

In order to optimize combination of team members, supporters (author) gather various progressing learners' information concerning with practices and analyze

them, in other words, big data processing and analysis [19]. Before gathering data, which information and how to gather them will be discussed and planned. Results are returned to the instructors. Supporters explain the results to instructors and learners. Then next, how to improve instructions and learning will be discussed.

2.1 Structure of big data processing system

Big data processing system consisted of the measuring system, the data analysis system, and the results of analysis output system (**Figure 1**). Big data processing system provides instructors students' data which are gathered by the measuring system and analyzed by the analysis system so that they can predict students' behavior as feedforward control.

2.2 Concept of big data processing system

After measuring learner's response (2) (**Figure 2**), those data are uploaded to data analysis system (3). The results of analysis are processed by output system (4). Instructors are able to access to the system of output system anytime in order to check the results of learner's assessment (5).
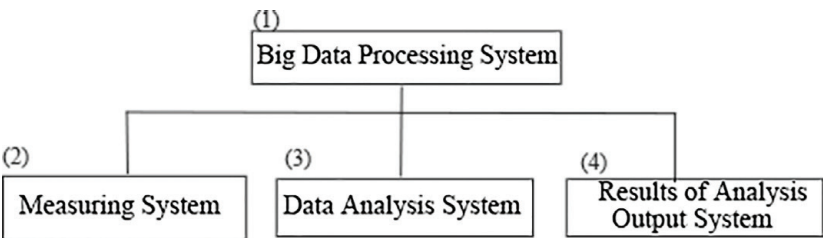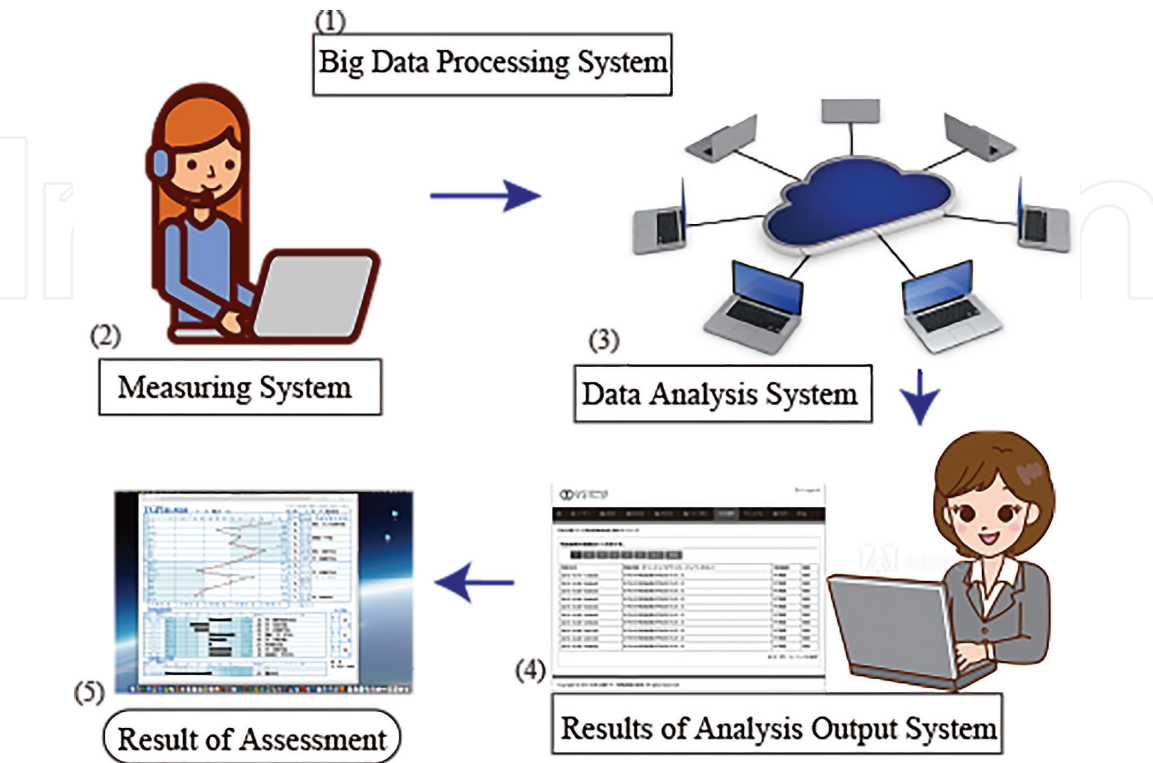
Figure 1.
Big data processing system.

Figure 2.
Concept of big data processing system.

2.2.1 Measuring system

At the stage of the measuring system (2), questionnaires are concerning with personality and presented to participants as a task. Participants are required to solve problems and make decision whether their daily behavior or attitude are similar or not to content of sentences one by one. There are 120 questionnaires which consisted of 20 traits. Each trait has 10 questionnaires which have similar contents to each other. Participants respond to them by selecting answers from yes, no, or neither.

The measuring system will gather information of participants, for instance, attributes, responses, decision time, and so on. Decision time will be measured from the beginning of presenting questionnaires to output information of decisions in which participants selected their response.

There are two ways for presenting questionnaires, sound voice and letters.

Each media presented sentences separately to participants, a total of 240 questionnaires, and gathered information of responses and decision time of each.

2.2.2 Data analysis system

After gathering information by the measuring system, those data will be processed by the data analysis system so that we are able to analyze them, for instance, clustering, categorizing, correlating, and so on, depending on purposes in order to predict behavior and attitude.

2.2.3 Assessment system

In one of the analyses, the results will supply assessments for personality of participants to observers. They will be able to obtain profiles of each participant (Appendix 1). Those profiles show 12 kinds of traits and help the observer to find out characteristics of participants, for example, social introversion, depression, nervousness, and so on. Moreover, from the curved line of the profile, we can categorize personality types, for instance, A to E type.
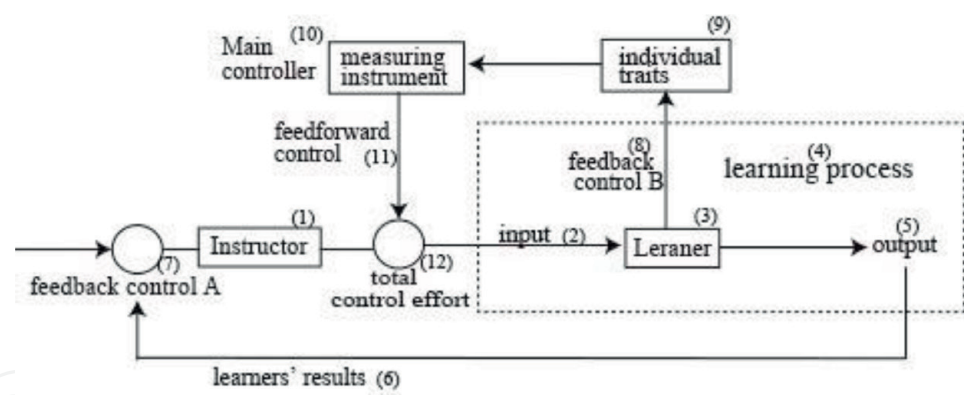
2.3 Personalized education and learning support system

Instructors (1) (**Figure 3**) need to make an instructional design including grouping and teaming before classes. In this case, they are required to consider learners' individual traits (9) concerning with learning process (4); however, if learners are freshmen (3), instructors have not had enough information about students (2, 5–7) [feedback control A]. Then, individual traits are measured by PELS (10) beforehand (11) [feedforward control] so that instructors can predict learners' behavior and design instructions (12). Because learners are continuously learning (8) [feedback control B] through classes, instructors are required to gather learners' information and redesign instructions. PELS supports them with scale-up method (**Table 1**) [20].

2.4 Hypothesis

1. If relationships among (a–d) members (**Figure 4**) are good and individual traits are different, learning outcomes improve through interactive communication.

2. Instructors are able to find out constrictive conditions of a successful combination of team members through their empirical knowledge about interaction in class (**Figure 5**). They are able to improve combination supported by PELS and continuously obtain learning effects.
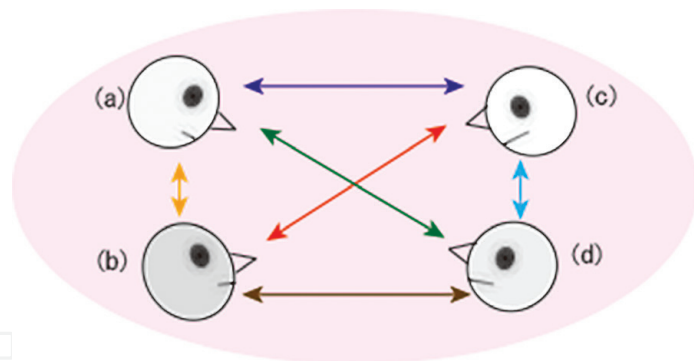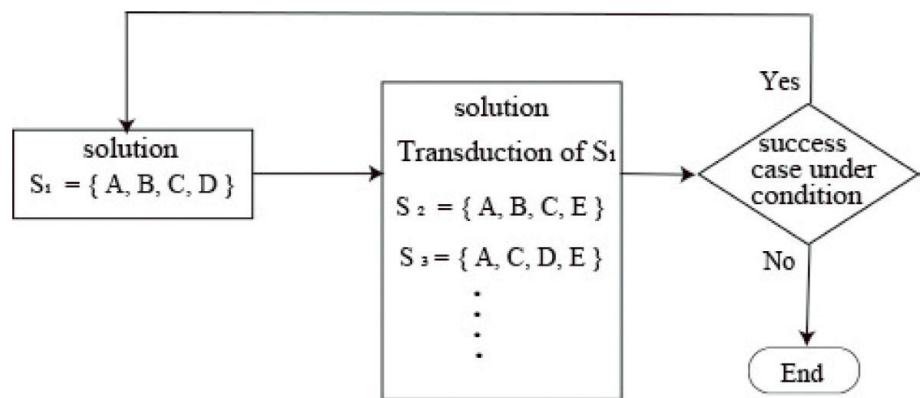
Figure 3.
Model of personalized education and learning support system (PELS).

(1)	Instructor	(7)	feedback control A
(2)	input	(8)	feedback control B
(3)	Learner	(9)	individual traits
(4)	leaning processing	(10)	Main controller (measuring instrument
(5)	output	(11)	feedforward control
(6)	learners' results	(12)	total control effort

Table 1.
Elements of personalized education and learning support system (PELS).

Figure 4.
Model of interaction among team members during collaborative learning.

Figure 5.
Local search for solution of combinatorial optimization.

3. Method

3.1 Teacher training

3.1.1 Participants

Participants are 35 female freshmen and 21 sophomores, a total of 56 students. Two instructors participated in the training. Students are divided into teams of four members: freshman, nine teams, and sophomore, six teams.

3.1.2 Duration

Practical research was implemented from April in 2015 to March in 2016. The first semester: 15 classes (90 min each). The second semester: 30 classes (90 min each).

3.1.3 Aims of training

The purpose of the teacher training is to find out the examples of combination of team members.

3.1.4 Procedure

1. Preparing for instructions of prototype practices in nursing science class Deciding how to evaluate students' performance

2. Dividing participants into teams consisting of four members each Deciding restricted conditions (e.g., avoiding close friends)

3. Implementing practices with team members and evaluating their performance

4. Measuring students' traits by Yatabe-Guilford Personality Inventory (YGPI)

5. Comparing between the results of performance and team combination by personality types

6. Deciding restricted conditions for optimizing combination of team members

3.2 Practical research

3.2.1 Participants

Participants are 98 female freshmen and divided into 25 teams.

3.2.2 Duration

Practical research was implemented from April in 2015 to March in 2016. The first semester: 15 classes (90 min each). The second semester: 30 classes (90 min each).

3.2.3 Procedure

1. Designing instructions with supporters reflecting the results of prototype practices in teacher training

2. Measuring students' traits by YGPI

3. Dividing participants into 25 teams with four members each (except 2 teams) under restricted conditions which are decided in teacher training

4. Implementing pre-/posttest (low-stakes assessments) which is concerning with conceptual reconstruction related to nursing sciences, before class and end of class

5. Explaining about individual differences to students and instructors by supporters before the first and the second semester

6. Reporting observations in class from instructors to supporters

3.2.4 Data gathering

1. High-stakes assessments: students' performance practiced by traditional method of combination in 2014

2. High-stakes assessments: students' performance practiced by optimizing method of combination in 2015 (e.g., low-stakes assessment; LMS, video, documents, interactive communication, outcomes, reports, questionnaires, and so on).

3.3 Investigation

Duration: from April in 2016 to September in 2018.

Data gathering: high-stakes assessments: students' performance practiced by optimizing method of combination.

Interview: three instructors; one is an expert (a chief instructor); two new members (one is from 2015; the other is from 2017) were asked some questions about an optimization of the combination of team members by an interviewer (an author).

4. Method of analysis

Visualization: comparing between successful and unsuccessful teams by categorization of personality and other factors

Qualitative analysis: comparing between traditional and optimizing methods by analyzing from interview

Quantitative analysis: comparing students' performance (average score) among the passing of years

5. Results

5.1 Results of teacher training

After a prototype practical experiment, students' personality had been measured. Two of nine teams have completed their presentation for freshman, and four of nine teams dropped out. All teams of sophomore have completed and succeeded their presentation. **Figure 6** shows the examples of the relation between

performance and combination of team members' personality type. Instructors were required to report their analyses of those relations which teams were success or not from the aspect of not only outcome but also interactive communication during practice. Supporters have explained to instructors about how to understand the results of measurement and help them to predict students' behavior and attitude beforehand [feedforward control].

Then next, they were required to decide on restricted conditions for combinatorial optimization. They have reported:

(1) A type ≥2 or 1; (2) B type <2; (3) C–E type ≤2.

5.2 Results of practical research

Ideally speaking, the method of combinatorial optimization is like **Figure 7**. According to the results of teacher training, types of students' personality were not distributed equally but discrete.

Therefore, we have decided to locally search for a solution of combinatorial optimization along with restricted conditions in which the instructors had found out the rules during teacher training. The results were succeeded, for instance,
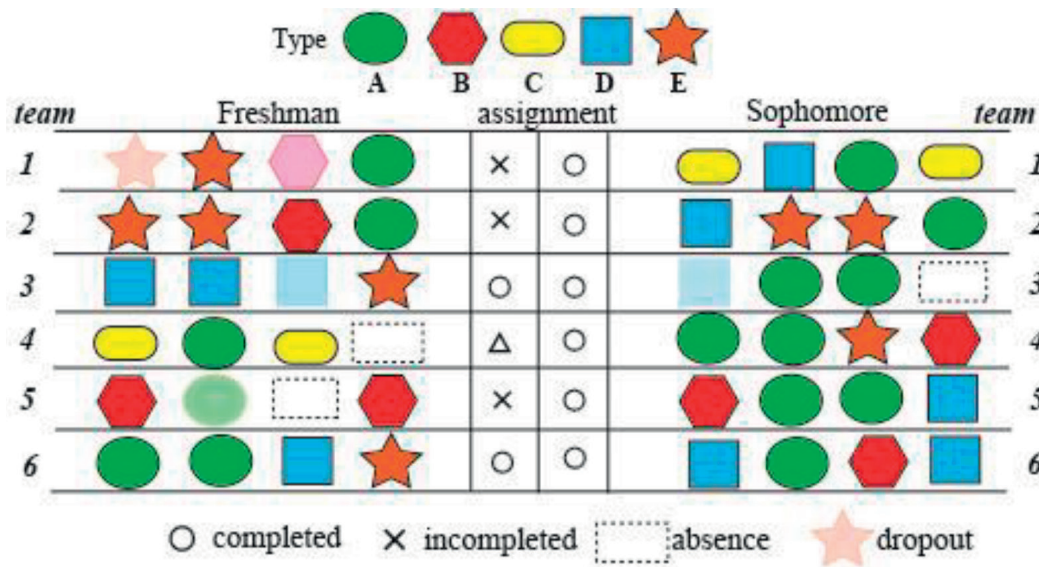

Figure 6.
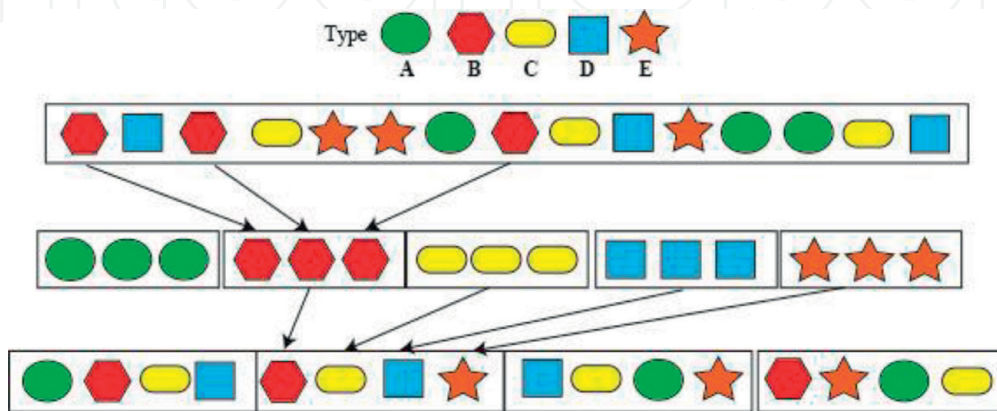*Comparison of combination among team members by prototype method.*

Figure 7.
Model of optimization method of combination under restricted conditions.

all teams have taken out their assignments using LMS, and their average learning outcome (83.95) has become significantly better than those of the previous year (58.94) (df.193, t = −14.1, p < 0.001). Especially, instructors have reported that their interactive communication had become smooth.

Comparing performance depending on teams in 2015, however, some of them were succeeded but some of them were not. Accordingly, comparing both high- and low-stakes assessments among teams, we have chosen successful and unsuccessful teams, Team B and Team E (**Figure 8**). The combination of members of both teams was satisfied by restricted conditions which instructors had decided.
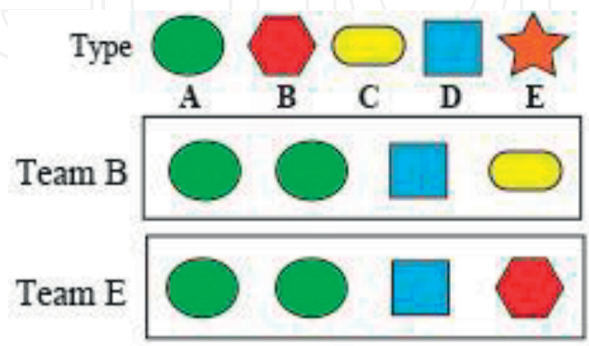
Figure 8.
Comparison of combination between successful and unsuccessful teams: personality types.
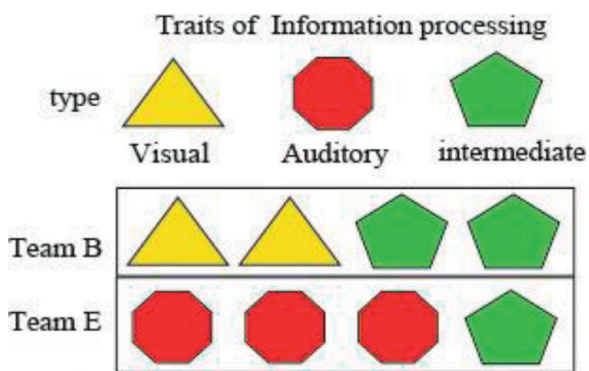
Figure 9.
*Comparison of combination between successful and unsuccessful teams: cognitive types.*
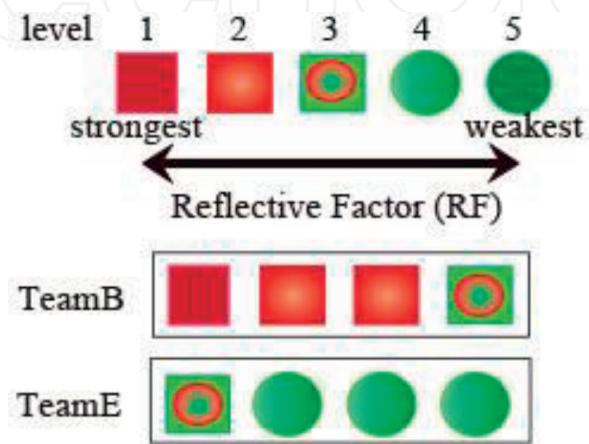
Figure 10.
Comparison of combination between successful and unsuccessful teams: reflective types.

And next, we have analyzed the structure of both team members from the other factors (**Figures 9** and **10**). In both factors, there were unbalances in the combination of Team E. On the other hand, there was a balance in cognitive types for Team B but in reflective factor. In this case, three of them were good at reflection which had effects on their performance.

5.3 Results of investigation

We have carried out a follow-up survey on combinatorial optimization in nursing science class from 2014 to 2018. **Figure 11** shows the results of average scores of high-stakes assessment which were evaluated by the criteria of credits which are required to obtain the qualification of nursing national examination. In the first semester, the average scores have been gradually increased. In contrast, in the second semester, they have been decreasing (**Figure 12**).
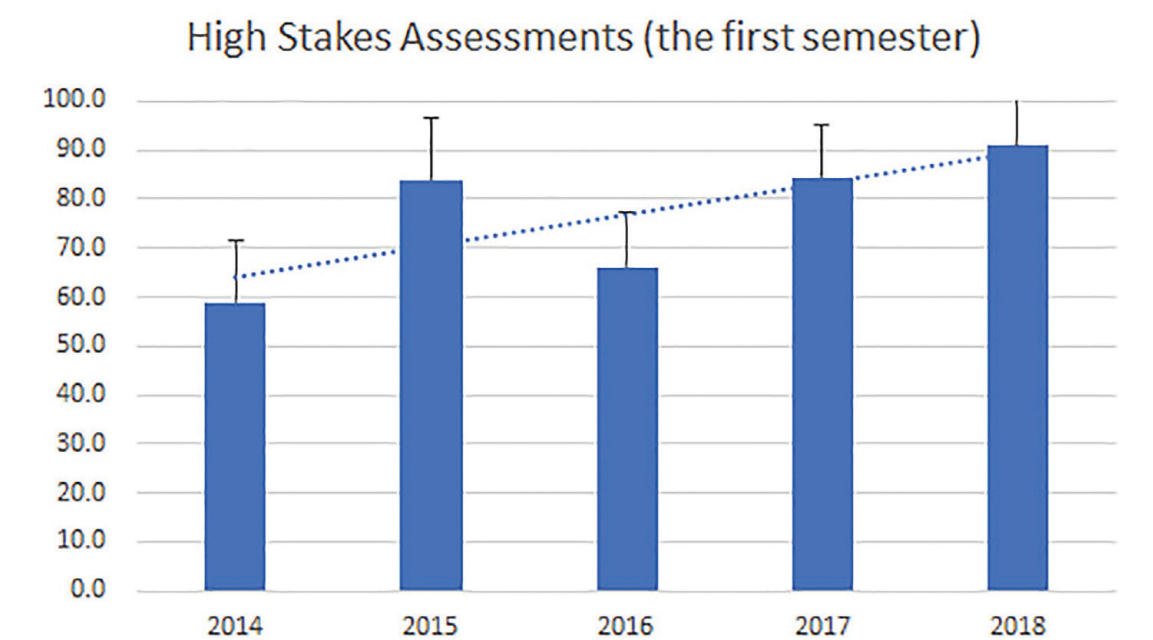
Figure 11.
Changing over the years (from 2014 to 2018).
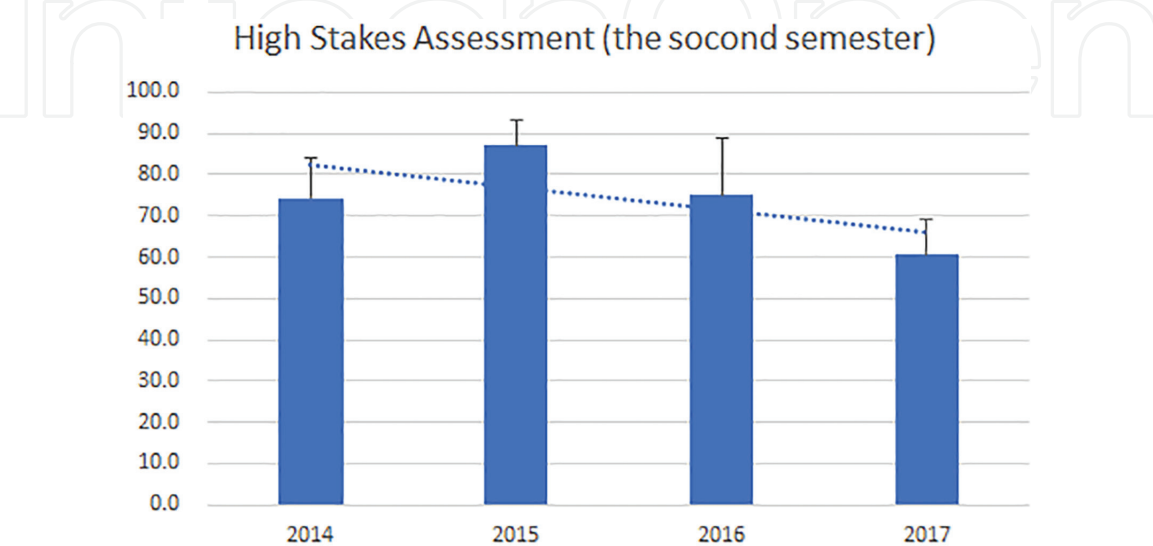
Figure 12.
*Changing over the years (from 2014 to 2017).*

Three instructors have been interviewed in 2018. The chief instructor has said that she had obtained the method of combination during teacher training. Until then, students had not been able to interact with each other and behaved passive attitude to practice. Other two new members said that they have been referring to the results of measurement of personality while they are teaching. It seems to be a well progression; however, they have not observed the results of personality in detail, for instance, reflective factor and so on.

6. Discussion

We have conducted teacher training and a practical research along with our design (**Figure 3**) in order to examine two hypotheses. The former one, whether learning outcomes improved through interactive communication among team members who were combined by different traits or not, has been examined statistically. The results showed that the average of the traditional method (n = 97) in 2014 was 58.9 and those by the optimizing support model (PELS) (n = 98) in 2015 was 83.9. The disparity was 25 points, and apparently the results in 2015 were significantly better than those of 2014 (df. = 97, t = −11.7, p < 0.001) [20].

Moreover, both instructors and supporters have observed that students' behavior and attitude in 2015 were favorable and they have built an excellent relationship. Especially, The members of team B had their interactive communication with each other, even by social network (LMS), and their documents were written out significantly excellently, comparing with those other teams. We have also found that outcomes of Team B had been observed by members of other teams, using LMS. In other words, many of the students had visited to see the documents and conversations of Team B through the network. That is, our optimizing support method might have the synergistic effect, not only within team but also between teams. Although many researchers pointed out the problems about interaction though social network [21], the results of our fundamental research seem fruitful. Then, whether this method might be able to be applied to other cases and how to do it should be discussed.

Then next, the latter hypothesis that an empirical knowledge of instructors helps them to find out the systemic rules of a favorable combination of team members through teacher training should be examined (**Figures 4** and **5**). Taken all together,
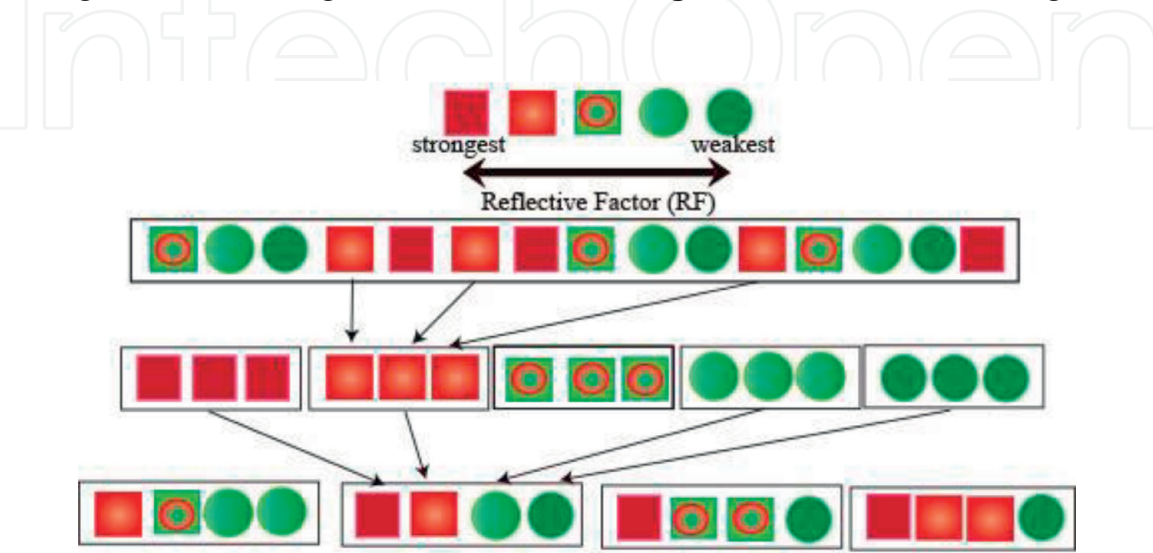
Figure 13.
Model of combinational optimization by reflective factor.

the examination of previous hypothesis has proved an effectiveness of matching members by their optimization. Moreover, the average of students' outcomes in the first semester over the passing of years (from 2015 to 2018) is slightly increasing (**Figure 11**). From this point of view, it is expected learning effectiveness continuously with this method. In contrast, however, there is a slight decrease in the results of the second semester. In addition to this, from the comparison of the results of individual teams, there are successful teams and unsuccessful ones (**Figure 8**). This means, in some extent, their empirical knowledge is recognized by the examination, but some problems of the methodology of optimization remain. The results of comparisons between Team B and Team E (**Figures 9** and **10**) might give us hints of solution. In the case of the second semester, explanations to new instructors and students about categorical visual and auditory types had not been provided in 2017. Concerning with reflective factor, which is one of the evaluations in YGPI, new instructors were also not explained in detail. From those points of view, the problems are caused by insufficient supports.

Look at those issues from different point of views, such as feedforward control and feedback control B (**Figure 3**), in the first semester, the model of optimization by personality types might be an example of a success case, on the other hand, those by other factors might be unsuccessful cases. This means that combinatorial optimization should be supported continuously for instructors and students. This, however, might be an ideal solution; it is supposed that feasibility and machine learning for AI might help us solve this problem with other factors (**Figure 13**). From this aspect, examples of successful and unsuccessful cases might help us to establish algorithm for solution of combinatorial optimization by local searching [22].

7. Conclusions

Computer-supported collaborative learning (CSCL) has begun to be paid attention after progress of social network. Because learners are always able to connect with each other, then learning effect by social interaction is expected. In contrast, many researchers reported the problems concerning with distance communication. In this paper, it is supposed that the problem might be caused by ill combination of team members. Therefore, we have begun to support instructors and students so that they can interact with each other smoothly by using the strategy of approximate solution with the method of scaling up.

We have designed teacher training and practical experiments that utilized personalized education and learning support system (PELS) which is structured by feedforward and feedback control, so that instructors can find out a concrete combinatorial optimization step by step. Consequently, they might have been able to find the method for combination of team members and students' performance had been significantly better than those by traditional method. On the other hand, problems concerning with discrepancies among teams and the example of combinatorial optimization by local search remained, finding transduction of successful team members continuously from a variety of factors. It seems, difficult to practice, however, it is important to develop the method with machine learning by AI.

Acknowledgements

Appendix 1
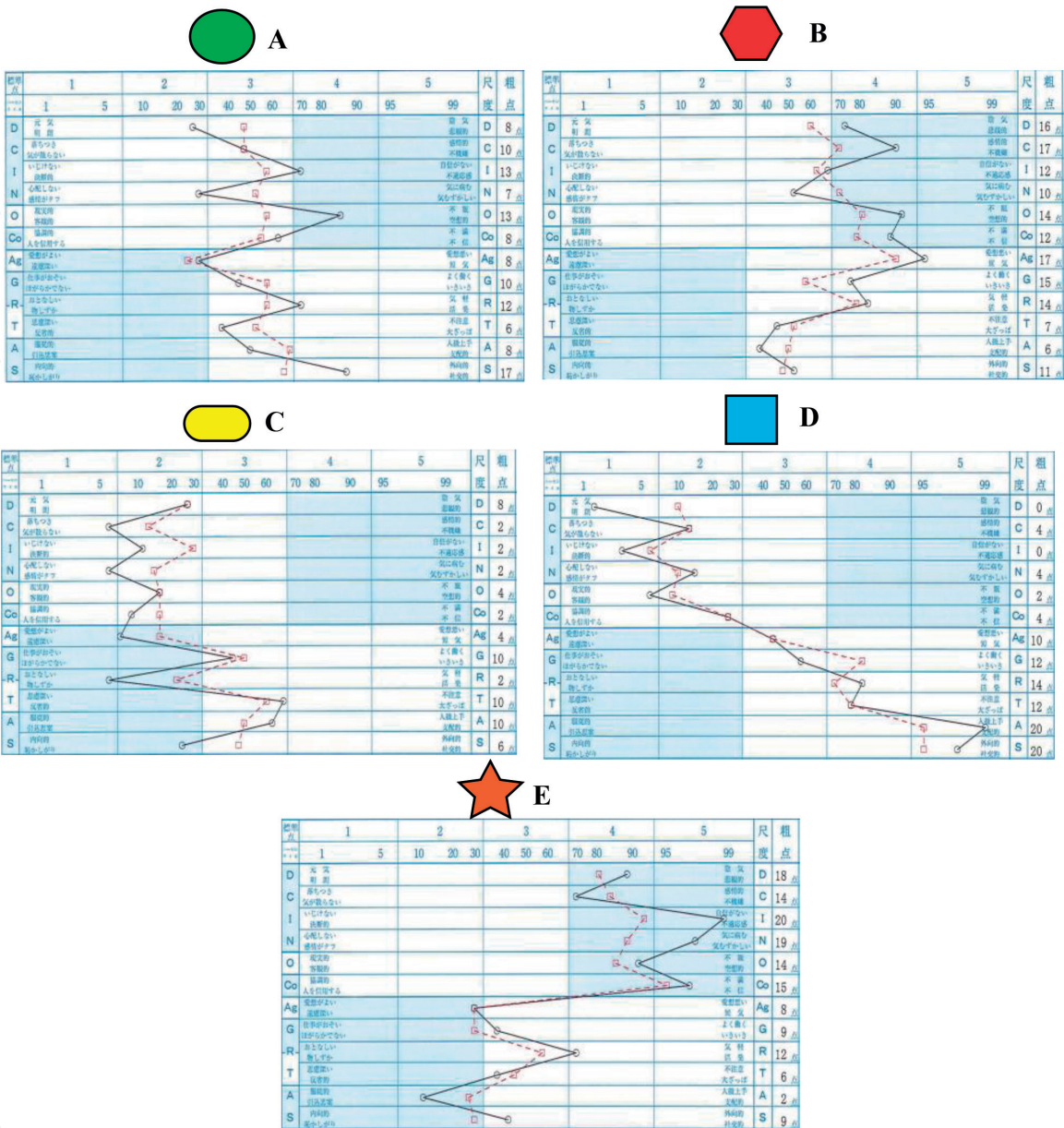
Author details

Keiko Tsujioka
Institute for Psychological Testing, Osaka, Japan

*Address all correspondence to: keiko_tsujioka@sinri.co.jp

IntechOpen

References

[1] Stahl G, Koschmann T, Southers DD. Computer supported collaborative learning. In: Sawyer RK, editor. The Cambridge Handbook of the Learning Sciences. Cambridge University Press; 2006. pp. 409-426. ISBN: 100-521-60777-9. paperback

[2] Koschmann T. Paradigm shifts and instructional technology. In: Koschmann T, editor. CSCL: Theory and Practice of an Emerging Paradigm. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 1996. pp. 1-23

[3] Fransen J, Weinberger A, Kirschner PA. Team effectiveness and team development in CSCL. Educational Psychologist. 2013;**48**(1):9-24. DOI: 10.1080/00461520.2012.747947

[4] Kreijns K, Kirschner PA, Vermeulen M. Social aspects of CSCL environments: A research framework. Educational Psychologist. 2013;**48**(4):229-242. DOI: 10.1080/00461520.2012.750225

[5] Koschmann T. CSCL: Theory and Practice of an Emerging Paradigm. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 1996

[6] Johnson D. Approximation algorithms for combinatorial problems. Journal of Computer and System Sciences. 1974;**9**:256-278. DOI: 10.1016/S0022-0000(74)80044-9

[7] Crescenzi P, Kann V. Approximation on the web: A compendium of NP optimization problems. In: Rolim J, editor. Randomization and Approximation Techniques in Computer Science. RANDOM 1997. Lecture Notes in Computer Science. Vol. 1269. Berlin, Heidelberg: Springer; 1997

[8] Nathan MJ, Sawyer RK. Foundations of the learning sciences. In: Sawyer RK, editor. The Cambridge Handbook of the Learning Sciences (Second Edition). Cambridge University Press; 2014. pp. 21-42

[9] Clarke J, Dede C, Ketelhut DJ, Nelson B, Bowman C. A design-based research strategy to promote scalability for educational innovations. Educational Technology. 2006;**46**(3):27-36

[10] Nelson BC, Ketelhut DJ, Clark J, Dieterle E, Dede C, Elandson B. Robust design strategies for scaling educational innovations; the river city case study. In: Shelton BE, Wiley D, editors. The Educational Design and Use of Computer Simulation Games. Rotterdam, The Netherlands: Sense Press; 2007. pp. 224-246

[11] Clarke J, Dede C. Design for scalability: A case study of the river city curriculum. Journal of Science Education and Technology. 2009;**18**:353-365. DOI: 10.1007/s10956-009-9156-4

[12] Dede C. Scaling up: Evolving innovations beyond ideal settings to challenging contexts of practice. In: Sawyer RK, editor. The Cambridge Handbook of the Learning Sciences. Cambridge University Press; 2006. pp. 551-565. ISBN: 100-521-60777-9. paperback

[13] Baker R, Simens G. Educational data mining and learning analytics. In: Sawyer RK, editor. The Cambridge Handbook of the Learning Sciences. 2nd ed. Cambridge University Press; 2014. pp. 253-271. ISBN: 100-521-60777-9. paperback

[14] Dutt A, Isamil MA, Herawan T. A Systematic Review on Educational Data Mining, IEEE Access. Vol. 52017. pp. 15991-16005. DOI: 10.1109/ACCESS.2017.2654247. Electronic ISSN: 2169-3536

[15] Márquez-Vera C, Cano A, Romero C, Noaman AYM, Fardoun HM, Ventura S. Early dropout prediction

using data mining: A case study with high school students. Expert Systems. 2016;**3**(1):107-124. DOI: 10.1111/ exsy.12135

[16] Márquez-Vera C, Cano A, Romero C, Ventura S. Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data. Applied Intelligence. 2013;**38**(3):315-330. DOI: 10.1007/ s10489-012-0374-8

[17] Guetzkow H, Gyr J. An analysis of conflict in decision making groups. Human Relations. 1954;**7**:367-381

[18] Murayama A, Miura A. Intragroup conflict and subjective performance within group discussion—A multiphasic examination using a hierarchical linear model [In Japanese]. The Japanese Journal of Experimental Social Psychology. 2014;**53**(2):81-92. DOI: 10.2130/jjesp.1203

[19] Tsujioka K. A Case Study of ICT Used by Big Data Processing in Education: Discuss on Visualization of RE Research Paper; ICIET, Association for Computing Machinery; 2018. In printing. ISBN: 978-1-4503-4791

[20] Tsujioka K. Development of Support System Modeled on Robot Suit HAL for Personalized Education and Learning; EITT, Society of International Chinese and Education Technology, IEEE2017. pp. 337-338

[21] Katz N, Lazer D, Arrow H, Contractor N. Network theory and small groups. Small Group Research. 2004;**35**(3):307-332. DOI: 10.1177/1046496404264941

[22] Skansi S. Introduction to Deep Learning: From Logical Calculus to Artificial Intelligence, Computer Science. Springer International Publishing; 2018. ISBN: 978-3-319-73004-2