Personality-Based Group Formation

Amir Mujkanovic, Andreas Bollin

▶ **To cite this version:**

Amir Mujkanovic, Andreas Bollin. Personality-Based Group Formation. Open Conference on Computers in Education (OCCE), Jun 2018, Linz, Austria. pp.207-217, 10.1007/978-3-030-23513-0_21 . hal-02370910

HAL Id: hal-02370910
https://inria.hal.science/hal-02370910

Submitted on 19 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Personality-based Group Formation

A Large-Scale Study on the Role of Skills and Personality in Software Engineering Education

Amir Mujkanovic, Andreas Bollin
Alpen-Adria-Universität Klagenfurt, Department of Informatics-Didactics,
Klagenfurt, Austria, amujkano@gmail.com, Andreas.Bollin@aau.at

Abstract. Extensive research confirms the benefits of group work in various educational and business domains. There has, however, been little consideration to rigorous formation of groups, especially project teams, in software engineering disciplines to improve the outcomes of these groups. Previous studies show that the outcome of groups will be affected by a number of different factors, such as the context in which these groups interact, the characteristics and the behaviour of each individual and the group composition. This research evaluates the extent to which it is possible to enhance the group outcomes by systematically reconstructing the groups of students and hence improve the performances and raise the overall outcome level of a software engineering lecture at two universities, the Alpen-Adria University of Klagenfurt and the Technical University of Košice. An empirical experiment has been carried out involving 69 groups and 140 individuals. The results of this experiment were then compared with historical data of 961 groups (approximately 2,400 students) on group outcomes over a period of 12 years. The findings show statistically significant improvements of the outcomes for those groups that were systematically constructed. These results could enable business leaders and educators to systematically form their groups for improving the outcomes of these groups.

Keywords. Software engineering, systematic group formation, improving group outcomes

1 Introduction

A winning team is required in almost any business and engineering discipline to achieve quality results such as, for example, the development of a product or the delivery of a service to clients. One can argue that every team within a company has to be successful in achieving their goals and many experts confirm that the composition of the team is the key to success. Examples of group types include business teams which might exist to generate financial profit, project teams which might exist to achieve certain project goals, club teams which might exist to have fun, families which might exist for reproduction purposes or educational groups which might exist to achieve certain learning outcomes. Panitz [1] points out a number of benefits that result from collaborative work in educational settings. These benefits include academic, social and psychological aspects which have been discussed in detail by Mujkanovic and Bollin [2] and by Mujkanovic et al. [3]. Groups generally exist for a particular reason and they typically target one or multiple outcomes [4].

A systematic and thorough construction of groups is a very demanding challenge, especially in composing the group members in a way that the intended group outcomes can be improved [5]. Many factors will have an impact on these outcomes, including the context in which the group activity takes place, the individual characteristics, the individual behaviour and the group composition. The meaning of these factors will be fully discussed in section 2.1.

The composition of groups plays a significant role in achieving the outcomes, but we still do not know much about the strength of this influence. Therefore, this research assesses the extent to which a systematic formation of software engineering groups affects the outcomes of these groups. Specifically, we aim at assessing the extent to which group outcomes can be improved through systematic reformation of groups during an ongoing lecture. Additionally, we are interested in the extent to which the outcomes of a lecture can be improved through a systematic reformation of groups.

This paper presents the results of an examination of data that have been collected over more than 10 years, involving randomly formed and systematically formed groups. The systematic formation is based on personality types and skills of each participants. The results show an improvement in the group outcomes when groups were systematically formed.

We commence with reviewing and studying the existing literature in section 2. Section 3 discusses the approach that has been taken to address the hypothesis and the subsequent research questions. Section 4 presents the findings of this research. Threats to validity will be discussed in section 5, and conclusions and further work will be addressed in section 6.

2 Background

There are important terms that will be used throughout this paper. It is therefore important to clarify the meaning of these terms. *Individual characteristics* are observable traits that can be used to differentiate between individuals. These traits exist independently of a human's behaviour and include cognitive and physical abilities, cultural values, personality traits, etc. Examples of individual characteristics include the age, the level of knowledge or the intelligence quotient of an individual. *Individual behaviours* are the actions of an individual within the context established by a particular task occurring within a particular environment. Examples include the level and diversity of chat dialogue that occurs between group members or potentially the number and nature of requests for assistance, etc. *Context* is the environment in which group activities take place. *Group outcomes* is the evaluation criterion that is defined to assess the group results. The *group composition* is a systematic arrangement of group members with certain personalities and skills which will contribute towards achieving the group outcomes.

In the following, we now briefly summarise related work in respect to group formation, project-based learning and the five-factor model, before presenting the research objectives that are covered by our study.

Many researchers [6 – 8] have studied the *group formation* problem. An overview is given by Magnisalis [9], in which the approaches to group formation have been

summarised and clustered by the methods used to form these groups. Various approaches have used clustering techniques [8], fuzzy and genetic algorithms [7] and hidden Markov models, as well as approaches that used learning styles of students.

Graf and Bekele [10] point out the importance of collaborative learning and the group formation process. They address the formation of heterogeneous groups that is defined as the level of diversity of achievements within the groups. The heterogeneity is measured by the Euclidean distance between the attributes of group members. Ant colony optimisation is applied to improve the "goodness of the heterogeneity" of groups. Their research addresses the problem of the famous travelling salesman that is often discussed in literature on optimisation problems. Students are represented as nodes and the travelling salesman optimisation is applied to find the closest students and create groups. The evaluation of this approach has been made through a study that involved 512 students. The authors show also the scalability of their proposed method and the application to the real world. While their work uses sophisticated artificial intelligence methods to address an important problem and improve the group formation process, it does not consider the formation of homogeneous groups. The focus of their research seems to be the quality of the group formation process itself [10].

We are not aware of any existing research that has used personality types and skills for the systematic reformation of groups during an ongoing software project management course and a project-based learning scenario. Our approach uses a simple and scalable group formation model with focus on systematic group reformation within a well-established simulation environment called AMEISE. The reasons for using the AMEISE framework in our research include a high standard of the lecture's content that has not changed much over the past few years, as well as the nature of the lecture in which student groups perform two simulations. This provides a perfect context to test our hypothesis with random groups (1st simulation) and systematically formed groups (2nd simulation). Another reason that make the AMEISE a perfect environment for our research is the very stable assessment scheme of the entire lecture. More information on the AMEISE simulation framework and a justification for using this framework for this research can be foundin a previous paper [2].

Concerning *project-based learning*, there are countless reasons for its importance for student careers. Krajcik and Blumenfeld [11] give an overview on the key elements that should be considered in project-based learning environments. These elements include: (1) a formulation of the key questions and hence a problem that has to be solved; (2) students tackle the problem by engaging in real problem-solving processes that are essential to expert work in the field; (3) teaching staff and fellow students including the community begin to engage in collaborative activities and support the project team; and (4) students develop a set of outcomes that represent the learning outcomes of the lecture. In project-based learning, students solve real-world problems and gain knowledge and skills and they also reflect on their skills and their personality [12]. To test our hypothesis, it is required to categorise individuals into different personalities. One way is using the Five-Factor Model.

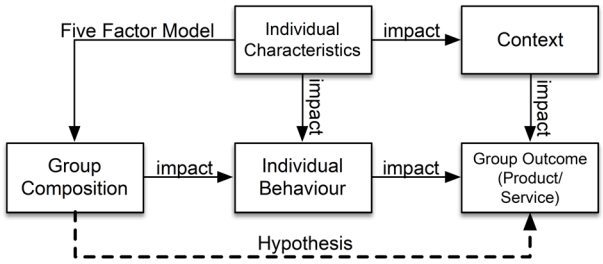


Figure 1. Group outcome model: Factors that have an impact on group outcomes [2]

Salleh et al. [13] and Yamada et al. [14] confirm that the personality has an impact on students' performances. They used the *Five Factor Model* (Extroversion, Agreeableness, Consciousness, Neuroticism, Openness) to assess the impact of personality on the outcomes. Salleh et al. found out that consciousness and openness had an impact on the performances of students. Yamada et al. suggest constructing the groups with members with different individual characteristics. Another study by Alfonseca et al. [6] found that certain learning styles impact on student performances and that collaborative learning might be improved through systematic formation of groups. Systematic group formation is exactly the core of this research, which aims at improving group outcomes.

In an initial study [2], we assessed the impact of the findings by other researchers [13, 14] in our context. As the results of our initial study provided promising results that supported our hypothesis, we felt strengthened to use these results for further studies. The learnings from existing work and our own examinations were applied to our group outcomes model. Initial studies [2] introduced the group outcome model where various factors that impact on these group outcomes had been discussed. This group outcome model had been further developed and includes now also the Five-Factor-Model (as represented in the top left in Figure 1), using individual characteristics to obtain the personality types of each individual. These personality types are then used to compose the groups in a way that the desired group outcomes will be more likely.

To our knowledge, there has been no consideration of how groups might be reorganised during a lecture to improve learning outcomes. We have therefore formulated the following hypothesis and research questions that will address this research opportunity:

Hypothesis: By systematically reconstructing groups of students it is possible to enhance their outcomes and improve the individual performances and therefore raise the outcome level of the software engineering lecture.
Research question RQ1: To what extent is it possible to improve the group outcomes by systematically reorganising the student groups?
Research question RQ2: To what extent is it possible to improve the outcomes and raise the total outcomes of the lecture by reorganising the student groups?

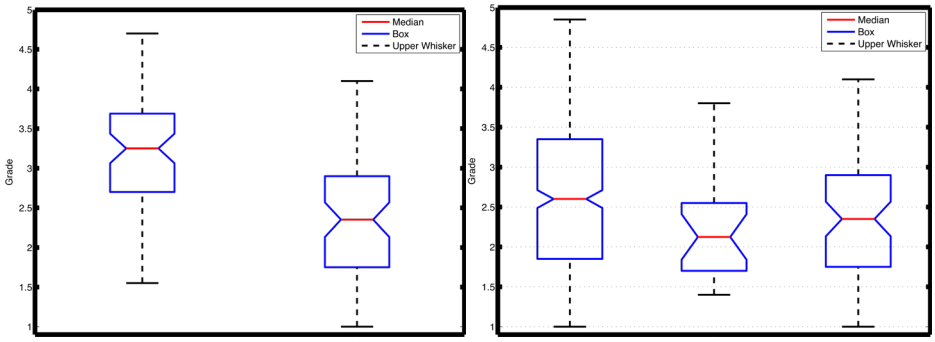
3 Methodological approach

The methodological approach was adapted from an initial pilot study in 2015 and was further developed through feedback of the pilot study [2]. Before we start discussing the details, it is useful to explain the structure of the experiment and the overall approach, which has been conducted in its two major phases. In the first phase, participants of the study performed an assignment during a software project management course using the AMEISE simulation framework that provided empirically validated and quantitative data for grading. In the second phase, participants of the study were systematically placed into groups of four different cohorts aiming at addressing our hypothesis and improving their outcomes through rigorous and systematic reformation of the project groups. After both phases, the atmosphere was observed through a pre-test (after phase 1 means prior to systematic group formation) and a post-test (after phase 2 means after systematic group formation) to capture any circumstances (e.g. conflicts between group members) which might have had an impact on the outcomes. The participants were split into four different student cohorts that were systematically constructed depending on their personality and results of the first simulation run. These cohorts included a random cohort (RG), a cohort (MC) that included at least one manager or coach per group as recommended by Sunaga et al. [15], a cohort of students that achieved best results in phase 1 (UC), and a cohort that included at least one analyst or renovator per group (AR). No roles were assigned to the students (so, only personality traits were used to form the groups). Results from both phases were then compared to examine whether the results could be improved through systematic formation of groups.

The students at both institutions were used to working together in different team constellations (even though they preferred to work with colleagues they knew), and then, from a student's perspective, they were randomly assigned, as explained before. All students were informed that they were taking part in an experiment, and surveys at the end of the course showed that they were satisfied with their group re-formation.

RQ1 was addressed by comparing the results of phase one (randomly formed groups) and phase 2 (systematically formed groups). These results are fully discussed in section 4. RQ 2 was assessed by analysing the grades and historical data that was available for all the courses that used the AMEISE environment to teach software project management between 2006 and 2016. The group outcome (grade) is a weighted composite of a number of factors that were kept the same in both phases (for more details, see a previous paper [2]). Both research questions are addressed in the large-scale study reported here, designed and conducted at the University of Klagenfurt and the Technical University of Košice in 2016.

To address both research questions, two separate examinations were carried out. Research question one was addressed by analysing the results of the experiment as described above. During both phases, participants conducted a full AMEISE simulation which was assessed at the end of the course. A total number of 69 groups completed a software project management assignment and each group received grades on a scale



(a) **(b)**

Figure 2. (a) Box plot of the group outcomes of randomly formed groups in phase 1 (to the left, n = 69 groups) and systematically formed groups in phase 2 (to the right, n = 69 groups) and 140 individuals. (b) Box plot of grades achieved in software project management lectures at the University of Klagenfurt using AMEISE between 2006 and 2016 (left, n=666) and 2015 (middle, n=22) and in 2016 (right, n=69).

between 1.0 and 5.0 (1=excellent, 2=good, 3=passed, 4=satisfactory, 5=fail). The data collection process and determination of individual characteristics (skills and personality) remained the same as in our pilot study [2].

RQ2 was addressed by examining historical data of 2,397 software engineering students (961 groups) enrolled in a software engineering course at the Alpen-Adria University of Klagenfurt (59 groups) and the Technical University of Košice in Slovakia (532 groups) over an observation period of 12 years. The remaining 370 groups were enrolled in the same course at other institutions. The students worked on their assignment mostly in *pairs* and *triads*, and in Klagenfurt some of the students had, due to their software engineering focus, a slightly higher previous knowledge in project management. During the experiments, when systematic group formation had been applied, students only knew that they were being part of a scientific experiment, but they did not know any details about the experiment.

4 Results

RQ1 examines the extent to which the groups' outcomes can be improved by systematically reorganising the student groups during a software project management course. Figure 2(a) depicts at the left the group grades in phase one, in which groups were randomly formed, as well as the grades of all groups in phase two (the right boxplot) that were systematically formed. The AMEISE framework determines several performance measures which are automatically transformed into grades. This scheme is a well-established assessment method that has been used for ten years. As a result of our large-scale study, the average group outcomes improved from 3.2 to 2.32, which is an improvement of approximately 27.5 % on average. A comparison of the means through a t-test using the MATLAB function *ttest2* returns additional insight. The *h*

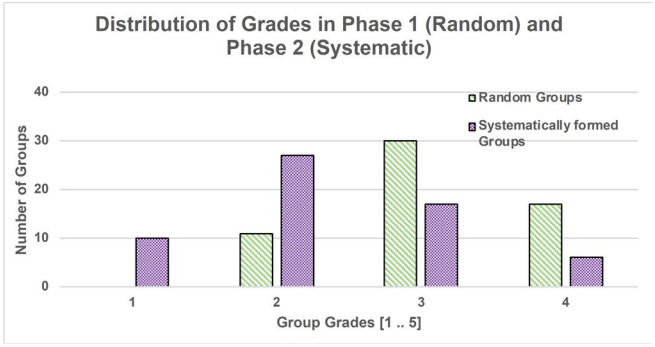


Figure 3. Distribution of the outcomes of randomly formed groups from the different phases.

and *p* values (representing the test for the null hypothesis) returns h=1 and p<0.0001 which tells that our null hypothesis can be rejected at a significance level of 0.0001.

The median of achieved grades in phase one is 3.25. The 75% percentile is 3.7 and the upper adjacent (lowest grade) is 4.7. The 25% quartile is 2.7. The best grade achieved is 1.6. Half of the data (inter-quartile range) lie between grades 2.7 and 3.25.

In phase two, where systematic formation of groups was applied, the grades were improved. Figure 2 (a) shows at the right the grades achieved in phase two. The median lies at 2.35. The 75% quartile is at 2.9 and the maximum at 4.1, which is the lowest grade achieved by a group. The 25% quartile is at 1.75 and the lowest value (best grade) is exactly 1. Half of the data lie between 1.75 and 2.7. From these two box plots, it can be seen that in phase two, when the systematic formation of groups has been applied, the notches do not overlap with the notches of the results in phase 1. Krzywinski and Altman [16] confirm the medians differ significantly when notches do not overlap – supporting our hypothesis.

Figure 3 presents the distribution of the group grades in both. These grades were gathered from the AMEISE simulation framework with possible values between 1.0 and 5.0. Considering the histogram of random groups, it is obvious that this graph represents a unimodal (one peak) distribution with no outliers that is skewed left. We have a concentration of the grades among the lower grades, with a small number of good grades (2, meaning "good" on the Austrian grade scale) and no excellent grade. The centre of the distribution is around the average grade 3.2. The minimum value, which is the best grade achieved, is 1.6 and the highest value that represents the lowest grade is at 4.7. This represents a range of grades from 1.6 to 4.7 which is a range of 3.1.

Now considering the distribution of the group grades when we applied our methodology of systematic formation of groups, we can report an improvement of the grades. The shape of the distribution is still a unimodal distribution that has changed the skew towards the right. The centre of the distribution is located around the average group grade 2.3. The minimum value is 1.00, which is the highest possible grade, and the maximum value is at 4.1. The range of grades remained the same (as in phase one where groups had been randomly formed); however, as the minimum and maximum values represent, we can report a shift of the mean to the right.

RQ2 examines the extent to which the results of the software project management course can be improved by systematically constructing the groups. To assess this issue, it is useful to consider historical data from the past years of the same course.

Figure 2(b) presents the achievements of groups between 2006 and 2016, when random group and self-assigned formation had been applied, as well as the results of studies when systematic group formation had been applied in 2015 [2] and 2016.

The lower adjacent of the results between 2006 and 2016, and therefore the best grade achieved, is a 1, which corresponds to an excellent grade. The upper adjacent, and therefore the lowest grade achieved, is a 5, which corresponds to a fail. The 75% percentile is at 3.65 and the median is at 2.75. The lower and upper limits of the notch are about 2.62 and 2.87. The 25% percentile is at 1.95. The inter-quartile range lies between the grades 1.9 and 3.65.

The lower adjacent of the results in 2015, and therefore the best grade achieved, is a 1.4, which corresponds to an excellent grade. The upper adjacent, and therefore the lowest grade achieved, is 3.8, which corresponds to a satisfactory grade. The 75% percentile is at 2.55 and the median is at 2.13. The lower and upper limits of the notch are about 1.84 and 2.4. The 25% percentile is at 1.7. The inter-quartile range lies between the grades 1.7 and 2.55.

The lower adjacent of the results in 2016, and therefore the best grade achieved, is 1, which corresponds to an excellent grade. The upper adjacent, and therefore the lowest grade achieved, is 4.1, which corresponds to a satisfactory grade. The 75% percentile is at 2.9 and the median is at 2.35. The lower and upper limits of the notch are about 2.13 and 2.56. The 25% percentile is at 1.75. The inter-quartile range lies between the grades 1.75 and 2.9. It is worthwhile mentioning that during the study in 2016 a new teaching staff (therefore with little experience) prepared the simulations in AMEISE, which might have had an impact on the overall results.

Comparing both results from our initial study in 2015 and the large-scale study in 2016 with the data of the past 10 years, there has been an improvement of the outcomes when systematic group formation has been applied. A comparison of the historical data of 12 years (the left box plot in Figure 2 (b)) and the results of the study in 2016 (the right box plot in Figure 2 (b)) through a t-test returns $h=1$ and $p=0.0028$, which is evidence that our results can be claimed as statistically significant.

5 Threats to validity

Validity considers the entire scientific experiment and examines whether the findings meet the requirements of the scientific method. Before we discuss the details of validity issues, it is worthwhile mentioning that we kept everything the same between the pilot study in 2015 [2] and the study that examined research question one in this paper. The subject was taught by the same academic staff, with an additional teacher in 2016. Also, the course material including the assignments remained the same.

Internal validity focuses on the examination if each and every step of the experiment follows the scientific method and whether other factors that have not been considered could have an impact on the results. External validity focuses on the generalisation of the results to other settings and to other populations.

Internal validity might be affected by the nature of the experiment, as involving humans in research studies is a known challenge and we are aware that capturing individual characteristics through a survey may not accurately represent the skills and

personality types of each participant, especially when the characteristics are self-perceived. An idea that might provide more accurate individual characteristics could be a system that collects data of how people perceive others when they interact with each other. Such a system would enable additional individual characteristics that are not self-perceived. However, as we have approximately 2,400 students that were included in our study, we are confident that most students respond carefully and honestly to the personality tests.

The experiment has been set up in two different phases; there is a possibility that the improvement of grades has been achieved through a learning effect. However, if the results have been biased by a learning effect, then this learning effect has influenced all participating groups on average. Therefore, a possible learning effect can be seen as irrelevant. An additional issue might be given by the diversity of students' previous knowledge. Even if they undergo the same curricula, their previous background and therefore their skills might have camouflaged impact on our findings.

External validity is certainly an issue of these finding as it cannot be assumed that these findings can be applied to other settings with a guarantee to achieve the same results. The settings that have been chosen for our work include two different cultural environments, one at the Alpen-Adria University of Klagenfurt in Austria and the second at the Technical University of Košice in Slovakia.

6 Conclusion and further work

The work presented in this paper intended to test our hypothesis and to assess the extent to which group outcomes can be improved by systematically re-organising the student groups during a software project management course. The hypothesis has been decomposed into two core research questions which have been addressed separately. The findings of both research questions provide results in favour of our hypothesis and therefore contribute to the body of knowledge.

Research question one considered the improvements of grades by systematically reorganising the student groups. The findings suggest that there is a statistically significant improvement of group outcomes by 27.5% when they are rigorously and systematically constructed. Research question two considered the improvements of the software project management course through a systematic formation of project groups. A comparison of data over ten years showed that the results were significantly improved on average by 14.6% when systematic formation of groups was applied. The performance increase is based on simple methods and two central questions in the Five Factor Model, rather than on complicated artificial intelligence methods.

These findings are promising, as they provide evidence that a systematic formation of groups might enable business leaders and educators to systematically form their teams, especially in highly technical environments, and therefore improve the key performance indicators of their business.

Teaching staff could systematically form groups of students across different school levels and therefore increase the learning outcomes of students. A transfer of these results and further studies in various schools, as well as in semiconductor industry, will be subject to further work.

References

1. Panitz, T.: The case for student centered instruction via collaborative learning paradigms. Retrieved from: http://home.capecod.net/~tpanitz/tedsarticles/coopbenefits.htm (1999)
2. Mujkanovic, A., Bollin, A.: Improving Learning Outcomes Through Systematic Group Reformation: The Role of Skills and Personality in Software Engineering Education. In Proceedings of the 9th International Workshop on Cooperative and Human Aspects of Software Engineering, ACM, Austin, TX, 97–103 (2016)
3. Mujkanovic, A., Lowe, D., Willey, K.: Adaptive group formation to promote desired behaviours. Australasian Association for Engineering Education, Melbourne, VC, 850-858 (2012)
4. Donelson, R.F.: Group Dynamics (2nd ed.). Brooks/Cole Publishing Company, Pacific Grove, CA (1983)
5. Cole, G.A.: Organisational Behaviour. Continuum, London (2001)
6. Alfonseca, E., Carro R.M., Martín, E., Ortigosa, A., Paredes, P.: The impact of learning styles on student grouping for collaborative learning: a case study. User Modelling and User-Adapted Interaction, 16(3–4), 377–401 (2006)
7. Crespo, R.M, Pardo, A., Pérez, J.P.S., Kloos, C.D.: An algorithm for peer review matching using student profiles based on fuzzy classification and genetic algorithms. Innovations in Applied Artificial Intelligence, 3533, 685–694 (2005)
8. Zakrzewska, D.: Cluster Analysis in Personalized E-Learning. In Nguyen, N.T., Szczerbicki, E. (Eds.) Intelligent Systems for Knowledge Management. Studies in Computational Intelligence, 252, 229–250 (2009)
9. Magnisalis, I., Demetriadis, S., Karakostas, A.: Adaptive and Intelligent Systems for Collaborative Learning Support: A Review of the Field. IEEE Transactions on Learning Technologies, 4(99), 5-20 (2011)
10. 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.) Intelligent Tutoring Systems, 4053, 217–226 (2006)
11. Krajcik, J. S., Blumenfeld, P.C.: Project-Based Learning. The Cambridge Handbook of the Learning Sciences, 317-333 (2006)
12. Graham, R.: UK Approaches to Engineering Project-Based Learning. Bernard M. Gordon-MIT Engineering Leadership Program, 1–48 (2010)
13. Salleh, N., Mendes, E., Grundy, J., Burch, G.S.J.: An empirical study of the effects of personality in pair programming using the five-factor model. Presented in the 3rd International Symposium on Empirical Software Engineering and Measurement, 214–225 (2009)
14. Yamada, Y., Inaga, S., Washizaki, H., Kakehi, K., Fukazawa, Y., Yamato, S., Okubo, M., Kume, T., Tamaki, M.: The impacts of personal characteristic on educational effectiveness in controlled-project based learning on software intensive systems development. Presented in the IEEE 27th Conference on Software Engineering Education and Training, 119–128 (2014)
15. Sunaga, Y., Washizaki, H., Kakehi, K., Fukazawa, Y., Yamato, S., Okubo, M.: Relation between Combinations of Personal Characteristic Types and Educational Effectiveness for Controlled Project-based Learning Course. IEEE Transactions on Emerging Topics in Computing, 99, 1–9 (2016)
16. Krzywinski, M., Altman, N.: Points of Significance: Visualizing samples with box plots. Nature Methods, 11(2), 119–120 (2014)