Improving Learning Outcomes Through Systematic Group Reformation

The Role of Skills and Personality in Software Engineering Education

Amir Mujkanovic
Carinthia University of Applied Sciences
Europastraße 4
9524, Villach, Austria
A.Mujkanovic@cuas.at

Andreas Bollin
Alpen-Adria University of Klagenfurt
Universitätsstraße 65-67
9020, Klagenfurt, Austria
Andreas.Bollin@aau.at

ABSTRACT

Group or team formation has been a well-studied field in numerous contexts, such as business teams, project teams, and educational teams. There has, however, been little consideration of how groups or particularly project teams might have to be reorganized during a lecture in order to yield an optimal learning outcome. Empirical studies show that the outcome of groups will be affected by a number of factors, including the individual behavior of each group member, their skills and inclinations. In this research, the authors aim at finding correlations between the individual characteristics and learning outcomes during a running lecture so that teaching and learning can be improved. A study has been carried out that involved software engineering students over the past three years. The results show an improvement in the learning outcomes for groups that were systematically formed, which, for future settings, could enable educators as well as project leaders to systematically form groups and improve the outcomes of these groups in various domains.

Keywords

Group Formation, Learning Outcomes, Individual Performance, Software Engineering Education

1. INTRODUCTION

A large body of literature confirms the advantages of collaborative group work. These advantages may however vary upon circumstances [9] [26]. Panitz [21] points out that there are academic, social, and psychological advantages that might arise from collaborative work. As part of academic advantages, when learning collaboratively, Panitz sees the development of critical thinking, improved engagement of learners in the learning process and improved problem-solving skills. Social advantages include the development of conflict resolution skills, and learners, engaged in multi ethnic groups, experience how different people from different cultures approach their work. One of the major advantage of collaborative learning is however that students can support each other when teaching staff is not available and they therefore gain important skills for their future career [26]. We assume that these advantages are also valid in our domain of interest, software project management courses.

When we reflect on our personal situation we find ourselves as members of various groups throughout our lives [5]. We might be members of business groups, project teams, learning groups, club teams, etc. Each group member will have different characteristics including the age, skills, motivation, personality, etc. Regardless of the variety of group members characteristics, groups will typically exist for a particular reason and will target at one or multiple outcomes [10]. A general distinction can however been made between task-oriented groups (business teams to make financial profit, project teams to achieve various project related outcomes, sports team to win a match or a competition) and groups that are natural (families) [1].

A systematic and rigorous formation of groups is a very complex and challenging task when it comes to particularly designing the group composition in such a way that desired outcomes can be improved [8]. There are also other factors that might affect the outcomes of groups. These factors include the context in which the group activity takes place, the individual characteristics, the individual behavior and the group composition. As we will use these terms throughout the paper it is useful to briefly clarify their meaning:

- **Individual characteristics** are observable traits that can be used to differentiate between individuals. These traits exist independently of human's behavior 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.
- **Group metrics** are measures of innate group properties that are solely a consequence of the group membership. These measures might potentially be derived from the available individual characteristics (once a group has been formed). Examples include the average age of group members or the range of ages within a group.
- **Individual behavior** are the actions of an individual within the context established by a particular task

CHASE'16 May 16, 2016, Austin, Texas, USA

ISBN 978-1-4503-4155-4/16/05…$15.00
DOI: http://dx.doi.org/10.1145/2897586.2897615

occurring within a particular environment. Examples include the level and diversity of a chat dialogue that occurs between group members or potentially the number and nature of requests for assistance.

In this paper we are strongly interested in the reformation of learning groups to enhance their outcomes and therefore improve their individual performances. We begin by reviewing and analyzing related work in chapter 2. Chapter 3 discusses our approach to systemic formation of groups, the experiment design, the setting where the experiment took place, research questions, the methodology and the study including the results. Threats to validity will be discussed in chapter 4. Chapter 5 discusses the results and concludes.

2. BACKGROUND

2.1 Related Work

This section presents the review and analysis of the literature that is related to understanding project-based learning, Five-Factor Model (FFM), Five-Factor and Stress Theory (FFS), personality types, group effectiveness and the importance of the group composition. A previously developed group outcome model will be introduced that was developed as a result of an extensive literature review on collaborative work, group dynamics and technology enhanced learning [20].

The interest in project-based learning in engineering education disciplines has been increasing and is very often seen as a key mechanism to prepare students for future jobs. In good project-based learning situations, such as for example introduced by Graham [12], students solve real world problems through which they gain knowledge and skills by tackling complex tasks. They also get a chance to reflect on their individual skills and their personality.

Research on personality has a long history and also very big potential for future research. Researchers have identified five dimensions of personality that provide a beginning of understanding what psychologist are interested in. These five dimensions are *Extraversion*, *Agreeableness*, *Conscientiousness*, *Neuroticism* and *Openness*. People who fall under the *Extraversion* category could be described as active, assertive, enthusiastic, outgoing, talkative, etc. Someone who falls in the category of *Agreeableness* could be described as appreciative, forgiving, kind, trusting, etc. *Conscientiousness* describes people who are efficient, organized, reliable, responsible, thorough, etc. People who tend to be anxious, self-pitying, tense, unstable, worrying fall under the *Neuroticism* category and artistic, curious, imaginative and original people fall in the category of *Openness* [19].

Studies such as the one by Salleh, Mendes and Grundy [27] investigated the extent to which the personality based on the Five-Factor Model affects the outcomes of pair programming exercises. Their study particularly examines the impact of *Conscientiousness* on the educational success of undergraduate students. Their experiment has been conducted at The University of Auckland and involved 218 students. It turned out that *Openness* has a stronger impact on the academic performances of students than *Conscientiousness*. Our study looks at a similar problem just from a quite different perspective. In our research we aim at finding be best group composition of students to improve their learning outcomes.

Similarly, the pilot study by Peslak [24] investigates the relationships between the personality and group outcomes in the context of Information Technology. The focus of their study is at finding impacts of personality traits and diversity of personalities within teams on the general team processes and the project success. Findings of their study did not show significant impact of personality on the outcomes. However, their results have to be carefully interpreted due to a very small sample size of 18 students. This threat to validity has also been addressed by the author of the study.

Parrington and Harries [23] introduced a study that focuses on the investigation whether the balance of team roles has an effect on the total team performances. They conducted their study with 43 teams of students, but did not find any significant relationship between the balances of certain team roles. They conclude that some of these roles *might* have an impact on the total performances.

Yamada et al. [31] present research based on FFS Theory that has been carried out to find characteristics that high performing teams have in common. In addition, their study investigates the learning process and the experience perceived by the students. Yamada et al. found out that heterogeneous teams perform better than homogeneous teams. Members of heterogeneous teams have different strength and skills and they can learn from each other as they work on complex tasks.

Chong [6] targets at finding relations between the group member roles, the total number of existing roles, the role development stage and the total team performance. To address this issue, 342 management students (organized into 33 teams) participated in a study where they planned a production of a custom made paper bag. As a result of the study, correlations could be found between the team role and total team performance. Although the examination of the impact of roles on the total performances seems to be an interesting research issue, it is not the focus of our work.

Many researchers have already considered the group formation problem. A rich consideration of literature that focuses on the formation of groups in various domains (where a number of different approaches and techniques for the formation of groups has been used) is given in [16].

The example introduced by Tourtoglou and Virvou [30] supports students with learning the Unified Modelling Language (UML). The formation of groups for this purpose is based on the students personality, performances type and the level of expertise. More detailed information on the group formation process is however missing.

Another approach by Zakrzewska [32] shows research that uses different learning styles and unsupervised classification methods to form groups. Similarly, the work by Papadimitriou, Grigoriadou and Gyftodimos [22] uses the learning style but also the total performance to form student groups.

A web-based tool for group formation has been introduced by Christodoulopoulos and Papanikolaou [7] in which up to three characteristics can be used to form homogeneous and heterogeneous groups. Heterogeneous groups are formed by randomly assigning students to groups based on the uniform distribution and homogeneous groups are created by using fuzzy classification methods.

Alfonseca et al. [2] have carried out a study to find out whether the learning style of students has an impact on the learning outcomes achieved by self-selected groups. Their intention is also to use this knowledge to automatically form educational groups in adaptive e-Learning systems. Their findings suggest that certain learning styles have an impact on the quality of students' results and that collaborative learning might be improved by systematically forming groups of students.

Martin and Paredes [17] introduce a merge between learning styles and collaborative learning. They primary used learning styles to form groups of students to meet their "learning preferences". For this purpose the authors have used *TANGOW* which provides opportunities for adaptation of online courses. The authors have strongly focused on a group formation that is based on the students learning style. Their system generates by default diverse group regarding their learning style. Once there are enough students available to complete a collaborative task an appropriate workspace will be dynamically created. Students with a dominant visual learning style will perform the task in a workspace with more graphical user interface and students with a dominant verbal learning style will complete their task in an environment with more textual interface.

It is obvious that these relationships between the characteristics of groups and the intended outcomes are complex. As a complement to other approaches, our research now focuses on a systematic re-formation of groups that will enhance the learning outcomes of groups and their individuals.

2.2 Group Outcome Model

A significant research area that considers groups is the scientific field of *groups dynamics*. This research field became very popular after the 1940s and was coined by Kurt Lewin [14] who studied the individual interaction within small groups.

It seems to be clear how a group is defined until one begins to analyze this concept. Is a number of people on the bus, train or aircraft a group? How about a crowd at a concert or a number of people having dinner, or is a group something else? Apparently these questions are not easy to answer and it is therefore challenging to identify factors that define a group. The etymology of the word *group* suggests that this word was used to express that there is more than a single person or multiple items behind this concept. Many researchers looked deeper into this area and formulated definitions of a group. Two definitions that are very often found in the literature are the definitions by Bass and Schein [1].

Schein's understanding is that groups generally consist of any number of people who interact with one another in some way, who are physiologically aware of one another and who perceive themselves as a group. Bass however regards a group as a number of individuals whose existence as a whole group is beneficial for each individual. According to Bass, group members do not necessarily perceive that they are member of a group nor they share common goals. Even if interaction and the diverse roles of group members are included in many other definitions they are not part of Bass' understanding of a group [1] [15].

These results in a significant disagreement in Schein's and Bass' definition which suggests that a group is a general concept with significant variations in their characteristics that are assumed to define the existence of a group [1] [20].

There is an interesting comparison between teams and work groups in [13] which broadly relates to the collective actions and the collective outcomes. From this it can be concluded that there will be different types of groups which will exist for various reasons and outcomes and therefore they will have different characteristics such as different structures and compositions that will affect the overall group outcome. Extensive literature considers the effectiveness of a group. One example is Hackmann's model of group effectiveness that provides an insight into the components that affect the effectiveness of a group. These components include the organizational context, material resources, group synergy, process criteria and, particularly, the design of groups [10, 20].

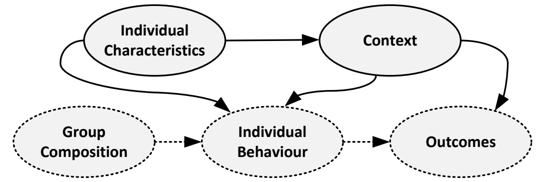

Figure 1: Group Outcome Model [20]

Figure 1 presents a group outcome model that has been developed by analyzing the literature on group dynamics, team effectiveness, collaborative work and technology enhanced learning. This model shows the key components that might affect the outcomes with the focus of this research with dashed lines. By considering this model it can be concluded that the outcome of groups will be directly affected by the context in which the group activities take place and by the individual performances. These individual performances might in turn be affected by the individual characteristics and by the composition of the group.

There are many ways how the outcome of a group might be improved. Looking at Figure 1, one could try to control the context in some way. However, another option is to keep the context stable and to change the group attributes by controlling the composition of the group [20]. This is exactly what we did when we tried to improve the learning outcomes of a software project management lecture that has been part of the curriculum in Klagenfurt since 2002.

3. SYSTEMATIC GROUP FORMATION

3.1 AMEISE

Experience-dominated subjects like software project management cannot be learned by merely attending lectures [18, 29]. Additional labs, however, even with only modest real-life projects, call for substantial effort to be spent by the instructors as well as by the partaking students. With these issues in mind, the concept for AMEISE (A Media Education Initiative for Software Engineering) has been developed. It is a simulation framework for practicing management of software engineering projects and is based on Stuttgart University's SESAM (Software Engineering Simulation by Animated Models) [11]. The AMEISE tool-set allows for repeatedly experiencing the complexity of software project management within a game-like simulation environment [3] [4].

Among other Universities in Europe, AMEISE is used in a couple of Software Engineering lectures in Klagenfurt and we selected these lectures for very good reasons:

- First, the lectures' content did not change a lot during the past couple of years. In every lecture two simulation runs are done by groups of students, with assessment and reflection units in-between and after the simulation runs. And, standardized questionnaires (pre-, intermediate and post-tests) show the learning progress.

- Secondly, the assessment schemes (criteria) of the whole lectures and the simulation runs did not change. Furthermore, all AMEISE results are still stored in a database for our analysis.
- Finally, AMEISE lectures have constantly been improved (adopting also brain-based teaching recommendations in 2013), so that we thought that a further improvement of the lecture is hard to achieve, setting a high bar for our research questions.

The traceability of all the results of the lectures and the constancy of the teaching model behind AMEISE-related lecture units make them an ideal candidate for studying effects in adjusting the group outcome model in Figure 1.

The outcome is measured by several dimensions. First, the achievements in the AMEISE simulation runs are taken into account. Typical measures include the duration of the project, the project cost, and product (and intermediate products) quality. The AMEISE web-site[1] provides more information about the data that is gathered. It allows for a very fine-granular assessment of the goal achievements.

Another set of measures for the performance of students is rooted in the results of pre-, intermediate- and post-tests related to previous knowledge and software project management skills. Here, again, we use a set of questionnaires that has been unchanged for years now. The AMEISE web-site[2] contains the latest versions of these documents, too.

3.2 Research Questions

Based on the previous discussion (and different to other approaches), it is reasonable to assess the extent to which the outcome of AMEISE groups might be improved by rigorously constructing these development groups and to see whether this also improved the individual performances. To address this issue the following hypothesis is proposed:

Hypothesis 1 *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.*

This hypothesis can be broken down into two core research questions:

1. To what extent is it possible to improve the group outcomes by systematically reorganizing the student groups?
2. To what extent is it possible to improve the outcomes and raise the total performances of each individual and therefore achieve better total outcomes of the lecture by reorganizing the student groups?

The following methodology has been applied to address our core research questions.

[1]Typical AMEISE evaluation criteria can be found at http://ameise.aau.at and following the link to Tutorial → Tutorial Documents → Classification of the Evaluation Criteria.

[2]AMEISE questionnaires can be found at the web-site by following the link to Documents → Set of Questionnaires.

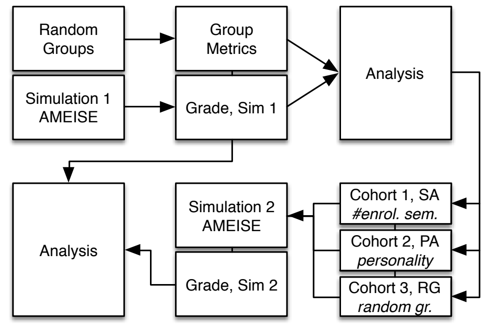

Figure 2: Experiment Design for the Examination of the Group Outcomes

3.3 Methodology

Prior to drilling down into details it is worthwhile to mention that our study has been conducted in two major stages (see Figure 2 for an overview). The first stage was the random formation of groups that performed a software project management exercise. The outcomes of the first stage were then analyzed, and the analysis results were used for the second stage, where we tried to reconstruct the cohorts in a way that an improvement of the outcomes became more likely, which is exactly the focus of this research.

Through the questionnaires (pre-tests as mentioned above), individual characteristics of each student were gathered. Before we started, we looked at the previous knowledge and skills of the students (e.g. the number of enrolled semesters or passed lectures) and their personality type according to the aforementioned FFS model. Once these individual characteristics have been collected, each student was randomly placed into groups (pairs of students), and each group performed an AMEISE simulation run which was then assessed as described above (simulation results and results from the intermediate test), yielding the outcome for each group. The outcome and the students' metrics were analyzed to find dependencies between the characteristics of the groups and the actual outcome. The two metrics that showed the highest correlation factors (enrolled semester and personality) were selected as an optimization criteria for the second step, consisting of another simulation run and a post-test. But, this time we changed the composition of the groups. One third of the groups then consisted of those students with the highest number of enrolled semesters and skills (highest previous knowledge, Cohort SA), one third was grouped according to their personality types (Cohort PA), and one third of the groups stayed unchanged (Cohort RG). At the end, we looked at their performances to answer research question 1.

Research question 2 was examined by comparing the individual grades (following a standard grading scheme in these lectures, taking the collaboration during the semester, the results of the simulation runs and the results of the final examination into account) and historical data that is available for all the courses which used the AMEISE environment to teach software project management.

Both research questions will be addressed in the following experiment that has been designed and conducted at the Alpen-Adria University of Klagenfurt in June 2015.

		Group Nr.		Questionnaires [..]				Simulation Results		
		Sim 1	Sim 2			FFS Informal/Formal	FFS Dominant/Easy-Goin	Sim 1	Sim 2	
Stud. [ID]	Cohort [ID]	Group [ID]	Group [ID]	Sex [m,w]	Sem. [Num]	[-4..4]	[-4..4]	Grade [1..5]	Grade [1..5]	Delta
1	RG	aau-20	aau-20	m	12	-1	2	1.45	1.90	0.45
2	RG	aau-16	aau-16	m	11	0	2	1.00	1.30	0.30
3	SA	aau-12	aau-05	m	10	2	-2	2.30	1.45	-0.85
4	PA	aau-07	aau-11	w	4	1	-1	1.95	2.10	0.15
5	PA	aau-07	aau-12	w	16	1	2	1.95	1.30	-0.65
6	PA	aau-18	aau-17	m	6	-1	-1	1.65	1.50	-0.15
7	PA	aau-02	aau-12	m	1	2	0	2.10	1.30	-0.80
8	PA	aau-05	aau-09	m	6	2	-1	3.35	2.30	-1.05
9	SA	aau-02	aau-02	w	14	2	-1	2.10	1.90	-0.20
10	SA	aau-15	aau-01	m	16	4	2	1.90	1.65	-0.25
11	SA	aau-11	aau-07	m	8	-2	0	2.90	1.75	-1.15
12	RG	aau-14	aau-14	m	4	2	-3	1.40	1.20	-0.20
13	SA	aau-06	aau-05	m	10	2	0	1.80	1.45	-0.35
14	RG	aau-14	aau-14	w	10	2	2	1.40	1.20	-0.20
15	SA	aau-18	aau-07	m	10	2	-1	1.65	1.75	0.10
16	SA	aau-17	aau-01	m	15	1	1	1.75	1.65	-0.10
17	PA	aau-01	aau-15	m	4	1	2	2.65	1.30	-1.35
18	PA	aau-05	aau-19	m	18	-2	2	3.35	1.90	-1.45
19	PA	aau-09	aau-18	w	10	2	2	1.65	1.75	0.10
20	SA	aau-06	aau-03	w	14	-1	1	1.80	1.90	0.10

Figure 3: Data gathered for students 1 to 20. Out of 48 factors, the table summarizes the student ID, the assigned cohort, the assigned simulation groups for both runs, the sex, and the number of semester enrolled. (Continued in Figure 4).

		Group Nr.		Questionnaires [..]				Simulation Results		
		Sim 1	Sim 2			FFS Informal/Formal	FFS Dominant/Easy-Goin	Sim 1	Sim 2	
Stud. [ID]	Cohort [ID]	Group [ID]	Group [ID]	Sex [m,w]	Sem. [Num]	[-4..4]	[-4..4]	Grade [1..5]	Grade [1..5]	Delta
21	PA	aau-12	aau-19	m	8	2	2	2.30	1.90	-0.40
22	RG	aau-10	aau-10	w	4	-4	0	1.15	1.50	0.35
23	SA	aau-01	aau-08	w	8	1	1	2.65	2.15	-0.50
24	PA	aau-19	aau-15	m	4	-4	-2	2.25	1.30	-0.95
25	SA	aau-08	aau-08	m	8	2	2	2.50	2.15	-0.35
26	RG	aau-10	aau-10	m	8	1	1	1.15	1.50	0.35
27	RG	aau-04	aau-04	m	10	3	-2	1.15	1.45	0.30
28	PA	aau-11	aau-22	m	10	-2	2	2.90	1.60	-1.30
29	RG	aau-13	aau-13	m	10	0	-1	1.15	1.60	0.45
30	PA	aau-03	aau-17	m	8	1	2	2.60	1.50	-1.10
31	RG	aau-04	aau-04	m	10	-2	-2	1.15	1.45	0.30
32	SA	aau-17	aau-06	m	10	-1	-3	1.75	2.00	0.25
33	SA	aau-19	aau-06	w	10	-2	0	2.25	2.00	-0.25
34	RG	aau-16	aau-16	m	6	-1	1	1.00	1.30	0.30
35	PA	aau-08	aau-09	m	6	0	1	2.50	2.30	-0.20
36	PA	aau-15	aau-18	m	6	-2	1	1.90	1.75	-0.15
37	PA	aau-03	aau-11	m	4	-1	2	2.60	2.10	-0.50
38	RG	aau-13	aau-13	w	6	3	-4	1.15	1.60	0.45
39	SA	aau-22	aau-02	m	14	-2	-3	1.45	1.90	0.45
40	SA	aau-22	aau-03	m	14	-2	3	1.45	1.90	0.45

Figure 4: Data gathered for students 21 to 40. The table also includes the two factors from the FFS-model (informal/formal and dominant/easy-going dimensions), the grades for the first and second run, and the delta between the two grades.

3.4 The Study

At the beginning, the study involved 42 software engineering students enrolled in 3 software engineering subjects in summer 2015. The students knew that they are taking part in a study, but did not know what the study was about. To measure the effect of group reorganization, the study has been split into two steps. During both steps, student groups (pairs of students) conducted a full AMEISE simulation run [4] and were assessed as usual at the end of the lecture. The lecturing staff, the exercises and the structure of the lecture remained the same for the entire experiment, extending over 13 weeks in total.

As mentioned in Section 3.3, Figure 2 represents each processing step of the study which began by randomly selecting 42 students to groups. During the courses, two students performed the exercise on their own and they needed to be removed from the study, yielding 20 groups at the end.

Twenty groups performed a software project management exercise and each group was assessed on a scale between 1 and 5 (1=excellent, 2=good, 3=passed, 4=satisfactory, 5=fail). Group metrics that represent the individual characteristics of the group members were correlated with the simulation results and, additionally, two factors that represent the personality (according to the FFS model) of each student have been calculated. For this, the students had to fill out a short questionnaire determining their level of formality (on a range between $[-4..4]$) and dominance (on a range between $[-4..4]$). According to FFS this helps in estimating the degree of being a Renovator, an Analyst, a Coach, or a Manager (most of us are mixed-types). Figures 3 and 4 summarize parts of the data that has been gathered (due to limitations of space, a lot of other factors we collected have been excluded from the tables).

To find dependencies between group metrics and the simulation results, a correlation analysis has been performed. According to an Anderson-Darling test, the grades of the simulation runs are normally distributed (with a given significance of 0.05). The data set describing the enrolled semester is not normally distributed, the personality types Renovator, Analyst, and Coach are normally distributed, the manager type is not normally distributed.

Table 1: Outcome correlation R_P, R_S, R_K (n=40) including p-value in brackets.

Measure	R_P (p)	R_S (p)	R_K (p)
Semester	0.793 (0.04)	0.437 (0.12)	0.521 (0.07)
Formality	0.853 (0.03)	0.818 (0.03)	0.784 (0.03)
Dominance	0.090 (0.26)	0.106 (0.26)	0.113 (0.19)

So, as we can not assume normal distribution, three different statistical tests were used to assess the data: the Pearson's Correlation Coefficient, the Spearman's Rank Correlation Coefficient, and the Kendall's Tau Correlation Coefficient. The Pearson's correlation coefficient (R_P) measures the degree of association between the variables, assuming normal distribution of the values ([25, p.212]). Although this test might not necessarily fail when the data is not normally distributed, the Pearson's test only looks for a linear correlation. It might indicate no correlation even if the data is correlated in a nonlinear manner. The Spearman's rank correlation coefficient (R_S) ([25, p.219]) is a nonparametric test of correlation and assesses how well a monotonic function describes the association between the variables. Finally, the Kendall's robust correlation coefficient (R_K) can be used as an alternative to the Spearman's test ([25, p.200]). It is also nonparametric and investigates the relationship among pairs of data. However, it ranks the data relatively and is able to identify partial correlations.

As shown in Table 1, there is medium to high positive correlation between the number of the enrolled semester, formality, and the results of the first simulation run. However, there is no correlation between the level of dominance, but as the p-value is higher than 0.05, this result is not significant. Other factors (not shown in Figures 3 and 4) did

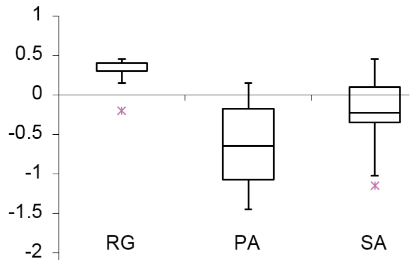

Figure 5: Box-plot of group outcome improvements (a value lower than 0 indicates an improvement) achieved by systematically formed students cohorts (RG, n=12; PA, n=14; SA, n=14).

not yield high correlations, thus we decided to focus on the number of enrolled semesters and personality only, and for testing our hypothesis we created three different cohorts.

Cohort 1 (SA) involved students of the highest possible semester. For the second group we followed the recommendations in literature [28, p.7], stating that one should have at least one manager (meaning high informality and low dominance in the FFG model) or coach in a team (meaning high informality and high dominance in the FFG model). This yielded cohort 2 (PA), where one member of the group was either more of a manager or coach type. Cohort 3 (RG) then consisted of the remaining randomly formed groups.

3.4.1 Results - Research question 1

All three cohorts performed the same exercise again using the AMEISE simulation environment. As a result of this study, the average simulation outcomes across all cohorts improved from grade 1.86 (on average) to 1.68 (on average). This was expected, as there is, of course, also a learning effect between the first and the second simulation run. One can assume that this learning effect is equally distributed over all groups, but in our case the improvement of the groups was different.

The greatest improvement has been made by the groups that were constructed according to the personality type. They raised their average grades from 2.38 to 1.73 (which is 27%). The second best performance has been achieved by the semester optimized cohort. This cohort raised the average grade from 2.02 to 1.83 which is an improvement by 9.3% on average. Interestingly, the student cohort that had excellent grades in the first simulation run could not improve their average simulation results. Their average simulation result dropped from 1.20 to 1.45 (on the Austrian scale 1..5). These results on the changes of the average simulation outcome are visualized in Figure 5.

3.4.2 Results - Research question 2

There has also been an improvement of performances of individuals. Figure 6 presents the points of individuals achieved in software project management lectures between 2013 and 2015. The plot on the left side represent an aggregation of results in the years 2013 and 2014 where the lower extreme lies at 70 and the upper extreme at 100 points. The lower quartile is at approximately 85 and the upper quartile at 95 points. The median is at around 91 points. In comparison with the year 2015 when the systematic group formation has been performed as previously discussed, a significant improvement of individual learning outcomes has been achieved. The plot on the right side shows the lower extreme at 80 points, the lower quartile at roughly 91 and the upper quartile at almost 100 points. The median lies at 96 points which clearly shows that the performances of individuals were raised comparing to previous years.

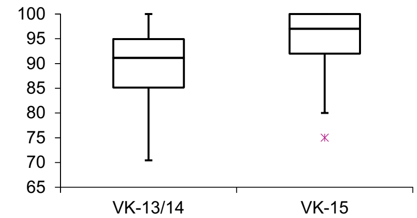

Figure 6: Box-plot of individual points achieved in software project management lectures at AAU Klagenfurt using AMEISE in 2013 and 2014 (to the left, n=82) and 2015 (to the right, n=40).

Based on previous discussions and the presented results there is evidence that supports our hypothesis that group and individual outcomes can, to some extent, be improved by systematically construction groups.

4. THREADS TO VALIDITY

The individual characteristics have been gathered through a survey. The survey is based on self-perception and the responses may therefore not accurately represent the skills and personality of each student. Though not very likely, there is also a chance that the different improvement of the outcomes was based only on an learning effect between the first and second simulation run. Another issue is that even if students undergo the same curricula, their previous knowledge had a camouflaging impact on our study. These are possible threats to internal validity. Threats to external validity are that we can not guarantee that our results can be transferred to other settings, as the number of participating groups and lecture types/settings is too low.

5. DISCUSSION AND CONCLUSION

This paper presented a group reformation approach that is based on the results of a correlation analysis and previous studies with the aim to improve the learning outcomes of groups as a whole and also the performances of each individual within a group. Our results provide evidence that support our hypothesis. The average group outcomes could be improved by reorganizing the students cohorts based on the results of our correlation analysis and on the insights of previous studies. Our plan is to conduct further studies in the future (currently we have a lecture at TU Kosice with 170 students) and collect more evidence that either supports or rejects our hypothesis.

6. ACKNOWLEDGMENTS

The author wish to acknowledge previous work done by Professor David Lowe from The University of Sydney and Professor Keith Willey from The University of Technology Sydney as supervisors of Amir's PhD research. Many thanks also to the students who participated in this study and the reviewers of the SEET-track of ICSE 2016.

7. REFERENCES

[1] J. Adair. *Effective teambuilding: How to make a winning team*. Pan Books, London, 1986.

[2] E. Alfonseca, R. M. Carro, E. Martín, A. Ortigosa, and P. Paredes. 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, September 2006.

[3] A. Bollin, E. Hochmüller, R. Mittermeir, and L. Samuelis. Experiences with integrating simulation into a software engineering curriculum. In *Software Engineering Education and Training (CSEE&T), 2012 IEEE 25th Conference on*, pages 62–71, 2012.

[4] A. Bollin, E. Hochmüller, and R. T. Mittermeir. Teaching software project management using simulations. *2011 24th IEEE-CS Conference on Software Engineering Education and Training (CSEE&T)*, pages 81–90, 2011.

[5] D. Cartwright. *Group Dynamics*. Concrete series. Harper & Row, 1960.

[6] E. Chong. Role balance and team development: A study of team role characteristics underlying high and low performing teams. volume 8(3). Journal of Behavioral & Applied Management, 2007.

[7] C. E. Christodoulopoulos and K. A. Papanikolaou. A Group Formation Tool in an E-Learning Context. In *19th IEEE International Conference on Tools with Artificial Intelligence(ICTAI 2007)*, volume 2, pages 117–123. IEEE, October 2007.

[8] G. A. Cole. *Organisational Behaviour*. Letts Higher Education List Series. Continuum, 2001.

[9] P. Dillenbourg. What do you mean by collaborative learning? In P. Dillenbourg, editor, *Collaborative learning Cognitive and computational approaches*, volume 1, chapter 1, pages 1–15. Elsevier, 1999.

[10] R. F. Donelson. *Group Dynamics*. Brooks/Cole Publishing Company, Pacific Grove, California, 2nd edition, 1983.

[11] A. Drappa and J. Ludewig. Simulation in Software Engineering Training. In *Proc. 23rd International Conference on Software Engineering*, pages 199–208. IEEE-CS and ACM, 2001.

[12] R. Graham. UK Approaches to Engineering Project-Based Learning. In *Bernard M. Gordon - MIT Engineering Leadership*, pages 1–48. Cambridge, 2010.

[13] J. R. Katzenbach and D. K. Smith. The discipline of teams. *Harvard business review*, 71(2):111–20, 1993.

[14] K. Lewin. Field Theory and Experiment in Social Psychology: Concepts and Methods. *American Journal of Sociology*, 44(6):868–896, May 1939.

[15] J. Luft. *Group processes: an introduction to group dynamics*. Mayfield Publishing Co., 3rd edition, 1984.

[16] I. Magnisalis, S. Demetriadis, and A. Karakostas. Adaptive and Intelligent Systems for Collaborative Learning Support: A Review of the Field. *Learning Technologies, IEEE Transactions on*, 4(99):1–1, 2011.

[17] E. Martin and P. Paredes. Using learning styles for dynamic group formation in adaptive collaborative hypermedia systems. In *Proceedings of the First International Workshop on Adaptive Hypermedia and Collaborative Web-based Systems (AHCW 2004)*, pages 188–198. Citeseer, 2004.

[18] Mary Shaw. *Future of Software Engineering 2000*, chapter Software Engineering Education: A Roadmap, pages 373–380. ACM, 2000.

[19] R. McCrae and O. John. An Introduction to the Five-Factor model and its application. volume 60(2), pages 175–215. Journal of Personality, 1992.

[20] A. Mujkanovic, D. Lowe, C. Guetl, and T. Kostulski. An architecture for automated group formation within remote laboratories. In *Remote Engineering & Virtual Instrumentation*, pages 91–100, Brasov, 2011.

[21] T. Panitz. The case for student centered instruction via collaborative learning paradigms. Technical report, 1999.

[22] A. Papadimitriou, M. Grigoriadou, and G. Gyftodimos. Interactive Problem Solving Support in the Adaptive Educational Hypermedia System MATHEMA. *Learning Technologies, IEEE Transactions on*, 2(2):93–106, 2009.

[23] D. Parington and H. Harris. Team role balance and team performance: an empirical study. volume 18(8), pages 694–705. Journal of Management Development, 1999.

[24] A. R. Peslak. The Impact of Personality on Information Technology Team Projects. pages 273–279. Proceedings of the 2006 ACM SIGMIS CPR: Forty four years of computer personnel research: achievements, challenges & the future, ACM New York, 2006.

[25] D. Rees. *Essential Statistics (4th Edition)*. Chapman & Hall/CRC: Boca Raton, Florida, 2003.

[26] T. Roberts. *Computer-supported collaborative learning in higher education*. Idea Group Publishing, Hershey, October 2005.

[27] N. Salleh, E. Mendes, J. Grundy, and G. St. J Burch. An Empirical Study of Effects of Personality in Pair Programming using the Five-Factor Model. pages 214–225. ESEM, 2010.

[28] Y. Sunaga, H. Washizaki, K. Kakehi, Y. Fukazawa, S. Yamato, and M. Okubo. Relation between Combinations of Personal Characteristic Types and Educational Effectiveness for Controlled Project-based Learning Course. *IEEE Transactions on Emerging Topics in Computing, PrePrints, doi:10.1109/TETC.2016.2526664*, 1(1):1, 2016.

[29] J. B. Thompson. Software Engineering Practice and Education: An International View. In *Proc. SEESE'08*, pages 95–102, 2008.

[30] K. Tourtoglou and M. Virvou. User Modelling in a Collaborative Learning Environment for UML. *Fifth Int. Conference on Information Technology: New Generations (itng 2008)*, pages 1257–1258, April 2008.

[31] Y. Yamada, S. Inaga, H. Washizaki, K. Kakehi, Y. Fukazawa, S. Yamato, M. Okubo, T. Kume, and M. Tamaki. The impacts of personal characteristic on educational effectiveness in controlled-project based learning on software intensive systems development. In *2014 IEEE 27th Conference on Software Engineering Education and Training (CSEE&T)*, pages 119–128. IEEE, April 2014.

[32] D. Zakrzewska. Cluster Analysis in Personalized E-Learning. *Intelligent Systems for Knowledge Management*, 252:229–250, 2009.