

# Software Engineering Team Diversity and Performance

VREDA PIETERSE AND DERRICK G. KOURIE

University of Pretoria

and

INGE P. SONNEKUS

University of South Africa

---

Reflection on software engineering as a sub-discipline within computer science reveals that it is as much about people and teamwork as it is about technical expertise. It is therefore important that a software engineer is not only competent in software development but also able to work effectively in a team. Based on the scholarly literature we derived a simple model for team composition and used it to guide, without coercion, our students during team formation. This research reports on an investigation into the role of personality diversity within teams of tertiary students taking a course in software engineering. Data collected confirms that both personality diversity in teams as well as competence of teams impacts positively on team performance. In contrast to other studies, our research data does not seek to couple personality with appropriate role allocation in the team; rather, our measurement is in respect of "raw" team diversity alone. In this study the correlation between personality diversity and the success of teams were captured at different stages during the development of a software engineering project tackled by the teams. The same was done with correlation between competence and team success. The personality diversity of the teams showed a strong correlation with team success during the inception phase. This correlation however weakened during the course of the year while the correlation between competence and success started slightly weaker than personality diversity during the inception phase, but grew very strong towards the completion of the team projects.

Categories and Subject Descriptors: D.2.9 [Software Engineering]: Management – *Programming teams*; K.3.2 [Computers and Education]: Computer and Information Science education – *Computer Science Education*; K.6.1 [Management of Computing and Information Systems]: Project and People Management – *Staffing, training*

General Terms: Management, Human Factors

Additional Key Words and Phrases: Personality, Team diversity, Team ability, Team performance, Teaching teamwork.

---

## 1. INTRODUCTION

It is commonly recognised that various factors influence software engineering team performance. Factors which are believed to play a role are team skill, managerial involvement as well as software development tools and methods [Guinan et al. 1998]. Other factors that have been researched are team size [Biffel and Halling 2003], team role allocation according to personality [Gorla and Lam 2004], and skill diversity [Smith et al. 2001]. This research reports on an investigation into the role of personality diversity within teams of tertiary students taking a course in software engineering.

In the next section, prominent models for team formation are surveyed. We then describe our sample, indicating how the teams were formed. Subsequently, the metrics used to measure team performance and ability are described, as well as the formula we formulated to measure team diversity.

From the literature findings, it was known that higher personality diversity links to better team performance. The results of this study, presented in the penultimate section, confirm this and indicate that this relationship compares well with the relationship between academic ability and the performance of a team especially during the initial stages of team development.

## 2. BACKGROUND

Paulk [2002] states:

The work we do in software projects is human-centric and design-intensive, and the process is what we do.

Issues related to teamwork are essential in Software Engineering. Hazzan and Tomayko [2004] state:

Even though it is technical, people issues are central.

Nevertheless, Denning [1992] observes that employers consider that new personnel do not know how to communicate and that they tend to have insufficient experience and preparation for working as part of a team. The general premise articulated by these and other authors is clear: namely, that human factors play a pivotal role in the

---

### Author Addresses:

V. Pieterse, Department of Computer Science, University of Pretoria, Lynnwood Road, Pretoria, South Africa; vpieterse@cs.up.ac.za

I.P. Sonnekus, Department of Further Teacher Education, University of South Africa, P O Box 392, UNISA, South Africa; sonneip@unisa.ac.za.

D.G. Kourie, Department of Computer Science, University of Pretoria, Lynnwood Road, Pretoria, South Africa; dgkourie@cs.up.ac.za

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage, that the copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than SAICSIT or the ACM must be honoured. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee.

© 2006 SAICSIT

success of software development in teams. To advance the discipline of software engineering, it is therefore important not only to expose students to teamwork, but also to engage in research to uncover critical success factors in undertaking teamwork.

A number of researchers have discussed the influence of team heterogeneity on successful team performance. Bradley and Hebert [1997] summarises the general conclusion of a number of studies as follows:

For complex problem solving, such as information systems system development, teams made up of different types of individuals with a variety of skills, knowledge, abilities and perspectives are more affective than groups that are more homogeneous. In other words, diversity in skills and knowledge combined with a balance of personality types is desirable for effective teams.

Bradley and Hebert [1997] emphasises the influence of personality type. They maintain that each personality type has a positive contribution to make to the overall effectiveness of the team and that a balance of personality types in a team should be sought. In an experiment involving two teams they found that personality types appear to be an important explanatory variable for differences in team performance. They conclude that teams that are heterogeneous in terms of personality type are more likely to create innovative solutions.

In an extensive study spanning several years, Belbin [1981] defined eight team roles based on personality types. Belbin identified team structures in terms of his team roles and argues that it is possible to create a winning team by carefully selecting the right mixture of members in terms of their personalities and mental abilities. An experiment by Rutherford [2001] was based on the formation of software engineering project teams in a tertiary institution by grouping members according to their personality types as identified by the application of the Keirsey Temperament Sorter [Keirsey and Bates 1984]. Heterogeneous teams showed a more open, communicative and varied approach to problem solving and enjoyed the teamwork more than homogenous teams. This was consistent with the one of the prominent themes throughout the studies conducted by Belbin [1981]. Gorla and Lam [2004] consider the four dimensions of personality as identified by the application of the Keirsey Temperament Sorter [Keirsey and Bates 1984] to assess the personalities of the members of the software project teams in their study. They identified ideal personality types for three team roles (team leader, systems analyst, and programmer, respectively) and concluded that the presence of the identified personality traits in each of the roles impacts positively on team success. They also linked team performance to the heterogeneity of personality dimensions identified between team leader and the rest of the team. Essentially these models emphasise the need for diverse personality traits in a team.

### 3. PRESENT STUDY

The Software Engineering module presented at third year level at the Computer Science Department of the University of Pretoria is an obligatory part of the curriculum for Computer Science majors. An important objective of the course is to teach software engineering concepts and methodologies. However, in recognition of the research findings already mentioned, an important sub-goal of the module is to teach teamwork. The course thus emphasizes both hard and soft skills. Students are required to design and develop a fairly large object oriented software product in a team.

The teams must have four or five members. An investigation into what is considered the norm for team size reveals that five members seem to be ideal. Koch [2003] advises to keep team smaller than six and Boehm et al [1998] remark that teams with five members were more efficient than teams with six members. For practical reasons we had to allow smaller teams but did not allow teams of less than four. We considered that larger teams were necessary to generate enough conflict and challenge students sufficiently to build their inter-personal skills.

Projects are intended to simulate real world contexts as closely as possible. This includes not only dealing with technical uncertainties but also dealing with difficult team members who could range from being aggressive and assertive, to shy and withdrawn. In addition, students have to contend with the prospect that their efforts will be laid bare for public exposure as would be the case in a real world context. Thus, the course culminates in an end-of-the-year exhibition at a project fair. The exhibits are judged not only by staff members but also by industry representatives, who contribute to prizes in various categories.

It was not possible to constitute student teams in conformance with the specific type of diversity suggested by studies mentioned above as the student body not contain enough members of the required personalities to populate all the teams adequately. Thus, instead of trying to enforce team assembly in accordance with recommendations prescribed by prior research, our model allowed initial teams to be formed spontaneously.

Prior to the formation of these teams, students were informed about Keirsey personality types, about the Belbin team roles and about research results that show the advantages of diversity in teams. The students were strongly advised to determine both their Keirsey personality type and their Belbin team roles. Their Keirsey type was determined using an online test [Boeree, 2003] while their Belbin team roles were determined using a printed version of the Self-Perception Inventory [Belbin 1981]. They were furthermore advised to use these results to form teams with as high as possible diversity according to their Keirsey personality types, race and gender while forming a team with a winning composition according to Belbin [1981]. Students were granted a week to group themselves into teams. No further checks were made as to whether or not our advice was followed.

Team Number	1	2	3	4	5	6	7	8	9	10	11	12
Team Ability	51	55	75	71	61	52	65	70	71	67	44	58

*Table 1. Average Ability of members of the teams*

Team Number	1	2	3	4	5	6	7	8	9	10	11	12
Team Diversity	4	6	5	7	6	4	3	3	8	4	3	5

*Table 2. Calculated personality diversity of the teams*

Team	1	2	3	4	5	6	7	8	9	10	11	12
Performance Term 1	62	74	83	90	72	68	61	52	87	85	58	78
Performance Term 2	55	73	81	87	79	69	70	72	86	83	63	78
Performance Term 3	55	62	83	81	82	62	70	78	83	80	52	70
Performance Term 4	52	60	83	84	81	59	65	80	78	78	50	72

*Table 3. Performance of the teams at the end of each term*

A simple team diversity metric (discussed below) was used to express the diversity of teams in terms of personality information with regard to the Keirsey temperaments only. Team diversity can be significantly improved in terms of this metric by removing, replacing or adding a single member. In the class of 122 students, 82 completed the test to determine their Keirsey personality type. Only teams for which the Keirsey personality type of all members were known were used in this study. There were twelve out of twenty-four such teams. In the teams that are part of our sample, there was sufficient variation in this metric across teams to provide some interesting information, as is discussed later.

#### 4. METRICS

For the purposes of this study, three metrics were at issue: team ability; team diversity and team performance. They were constituted as follows.

##### 4.1 Team ability

A value to represent the potential ability of a team has to be derived from a measure of the abilities of the individuals in the team. At first thought using a standardised scholastic ability test for this purpose seem to be most reliable. We however decided that the grades achieved for theoretical tests in our course are the best measure for individual ability in our context for the following reasons:

- Limited availability of Standardised Ability Test data
- Standardised ability tests are not always the most reliable predictors of success in specific courses. A study conducted by Baron and Norman [1992] revealed that high-school class rank added significantly to the prediction of academic success of university students whereas scholastic aptitude test did not. Similarly Edge and Friedberg [1984] found that an in-house pre-test combined with high-school class rank were more reliable than many of the available more generalised tests to predict success in a first course in calculus.
- All students in this study were subjected to the same tests at the same time. This contributes greatly to the ranking value that these marks provide. It can therefore be seen as a valid substitute for high-school class rank that was identified as a reliable predictor of success by various authors [Edge and Friedberg 1984; Baron and Norman 1992]
- House [2000] refers to several studies that shows positive correlations between Scholastic Aptitude Test (SAT) scores and grades in college courses as well as between American College Testing (ACT) scores and student achievement in college courses. This expected strong correlation can be interpreted in the reverse direction rendering that the grades awarded for tests at university, in particular the tests taken in this course, can be seen as an indication of the ability of the student.
- The tests written by the students in this study were open-book tests and consisted of questions that required personal insight in the concepts at hand and ability to apply their insight to new problems.

An individual evaluation mark was calculated for each student based on marks received for three tests. For the purpose of this study, the ability of a team was taken to be the average individual evaluation mark of the team's members. The calculated ability of the teams included in the study is shown in table 1.

##### 4.2 Team diversity

The team diversity is calculated using information about an established measurement of individual personality called the Keirsey-Bates Temperament Sorter [Keirsey and Bates 1984]. It is one of several instruments used to measure personality type preference. It is modelled after the Myers-Briggs Type Indicator (MBTI) [Myers and Briggs 1975].

The following discussion of the four dimensions of the Keirsey-Bates Temperament Sorter describes the key attributes of persons leaning towards each of the two preferences of each dimension:

- Dimension 1 ( $K_1$ )  
The first dimension is known as the social interaction dimension. It distinguishes between introverts and extroverts. Introverts are comfortable with long silences and prefer to work alone. Extroverts tend to talk about anything just to break the silence and enjoy interacting with others. Meetings tend to drain introverts while extroverts are energized by social interaction.
- Dimension 2 ( $K_2$ )  
The second dimension refers to how a person prefers to gather information. A person is considered to be either observing or intuitive. An observing person prefers tangible results, pays attention to detail and values information based on concrete facts. An intuitive person has a more holistic view, sees facts as contributors to concepts and tends to be more imaginative.
- Dimension 3 ( $K_3$ )  
This dimension is about how individuals make decisions. Some are considered thinkers and others are considered feelers. A thinking person is orderly and critical and bases decisions on logic and objective consideration. A feeling person is driven by personal values and bases judgments on subjective and personal considerations. Feelers are good team builders because they focus on implications for individuals.
- Dimension 4 ( $K_4$ )  
The fourth dimension describes the work style of the person. Work style is defined to be judging or sensing. Judgers are organised individuals. They plan ahead and prefer closure to avoid last minute stresses. They are inclined to establish deadlines and expect others to follow. Sensors are flexible people. They prefer to be spontaneous and to leave things open in anticipation of a better opportunity that might arise. They view deadlines only as guidelines and are energized by pressure.

The students completed this test before they formed their teams and submitted the results when they registered their teams. As the completion of the personality test was optional, only students who completed the test were included in this study.

$$f(K_i) = \begin{cases} 0 & \text{if all team members have the same preferences in dimension } i. \\ 1 & \text{if all but one team member has the same preference in dimension } i. \\ 2 & \text{otherwise.} \end{cases} \quad (1)$$

$$\text{Team Diversity} = \sum_{i=1}^4 f(K_i) \quad (2)$$

To calculate the team diversity, firstly the spread of each of the four Keirsey-Bates dimensions across the team was considered in turn. For each dimension, a value was assigned to the team as defined by the discrete function to calculate the team diversity with regard to one personality dimension shown in equation 1. For example considering Dimension 3; if there is a fair spread of thinkers and feelers  $f(K_3) = 2$ , if the team is predominantly feeling or predominantly thinking  $f(K_3) = 1$ , while if all members were preferred thinkers  $f(K_3) = 0$ , likewise  $f(K_3) = 0$  if all members were preferred feelers.

The overall diversity of the team in terms of these dimensions is expressed with the function in equation 2. The four different personality dimensions are treated equally. As a result, the maximum diversity of a team in terms of the formula is eight and the minimum diversity is zero. The calculated personality diversity of the teams included in the study is shown in table 2.

### 4.3 Team performance

The teams were required to follow the Unified Software Development Process for their chosen projects, handing in a series of deliverables as specified by this process. These documents were assessed by teaching assistants using rubrics that included correctness, comprehensiveness as well as other quality measures such as clarity, neatness of layout, and overall impression.

The teams were also evaluated on the implementation of their projects in a series of demonstrations where the teaching assistants assessed their projects using rubrics that include the scope of the project, the professional conduct of the team, compliance with their development plan, as well as the quality of the code in terms of readability, re-usability and efficiency.

The projects culminate in a final exhibition of the software that was developed at a project fair. The exhibitions are judged by staff members of the department and by sponsors from industry. The rubrics used at the fair include items like innovation, technical difficulty, customer satisfaction, usefulness to industry and marketability.

	Term 1	Term 2	Term 3	Term 4
Correlation Coefficient	0.8186	0.7261	0.5743	0.4849
Level of significance (P)	0.0563%	0.3746%	2.5348%	5.5341%

Table 4. Correlation coefficients between team diversity and team performance.

	Term 1	Term 2	Term 3	Term 4
Correlation Coefficient	0.5709	0.8091	0.9067	0.8825
Level of significance(P)	2.6221%	0.0722%	0.0024%	0.0073%

Table 5. Correlation coefficients between team ability and team performance.

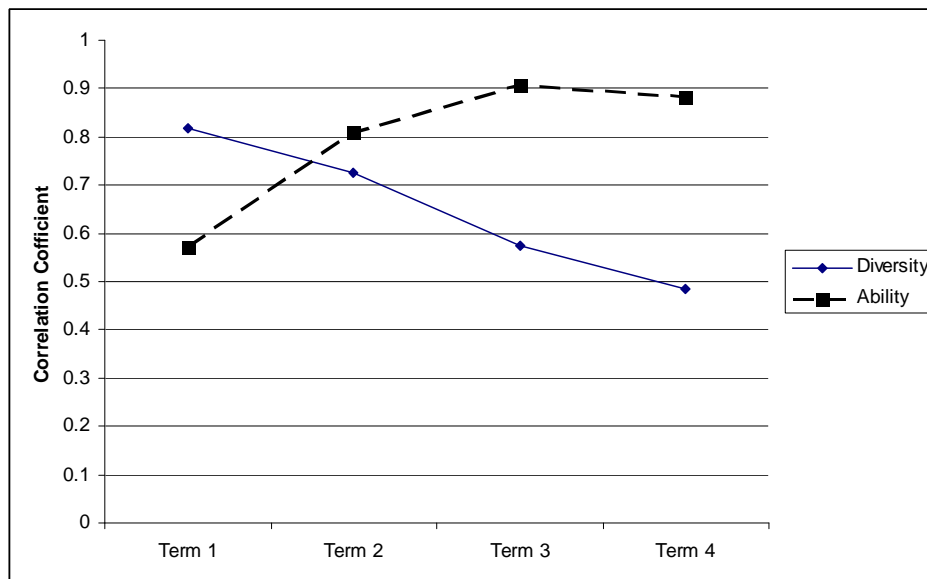


Figure 1. Correlation of team diversity and team ability with team performance.

For the purpose of this report, the success of a team was measured at the end of each term. Their performance is expressed in terms of the marks allocated to the deliverables and demonstrations of their projects to date. For the first term the evaluation included documents produced during the inception and elaboration phases of the projects while the second term evaluation is for all the design deliverables required for the projects. The third term evaluation includes progress demonstrations of the implementation. The last and final evaluation of team performance incorporated the marks for all the design deliverables, marks for the all the demonstrations of the implementation as well as the marks allocated by the adjudicators at the project fair. The calculated performance of the teams included in the study is shown in table 3.

## 5. ANALYSIS

Since the data is clearly not normally distributed, and therefore not amenable to conventional parametric correlation tests, it was decided to rely on non-parametric statistical tests. The well-known Spearman's rank correlation test was therefore applied to the data. The correlation between team diversity and team performance, as well as the correlation between team ability and team performance was determined at the end of each term. The results are shown in Table 4 and Table 5.

In all cases, these coefficients lead to the rejection of a null hypothesis of no correlation at a very high level of significance as the highest P value (one-tailed rejection) is only 5.5341%. It is therefore accepted that both personality diversity and average team ability impacts positively on team performance. At the end of the first term the correlation between team diversity and team performance is noticeably higher than that between team ability and team performance, while the opposite is true at the end of the year. As can be seen in Figure 1, the correlation between team diversity and the performance of the team was initially very strong but declined over time, while the correlation between team ability and the performance of the team grew stronger over time.

At the end of the first term it seemed as if personality diversity had a greater impact on the performance of the teams. The teams that provided positive examples of high team performance at that stage, not only scored well on average ability; they were also highly diverse in terms of the diversity metric. In contrast, teams that scored quite respectably on average ability but were homogenous in terms of the diversity metric, emerge as two of the three worst team performers. Another poor team performer, was not only poor in ability but was, once again, highly homogenous. These observations support the notion that personality diversity in a team can help the team to get productive sooner.

A team that appears to refute the trend the most was a team with a fairly high academic ability, though not exceptionally high, and relatively homogenous. Their team performance was the third highest in the sample at the end of the first term. This is a healthy caution that diversity alone should not be taken as an infallible predictor of initial performance and that other factors such as effective leadership, intra-team communication and group cohesion [Bradley and Hebert, 1997] may play a pivotal role during the initial stages of working in a team.

Relying on answers to qualitative questions that were included in regular peer evaluation forms, the reasons for the anomalous nature of this team were sought. Students complete these forms anonymously, except for indicating to which team one belonged. The question "To what do you contribute your success?" was asked. It is interesting to note that the four members of the team at hand attributed their success mainly to the fact that they 'enjoy working together' and 'function well as an egoless team'. It would therefore seem that the success of this team can at least in part be attributed to a healthy relationship among the members, resulting in a meaningful participation of all members. This is consistent with the conclusion of Lingard and Berry [2002] who note that team success is likely to increase if better participation among the team members is achieved.

## 6. CONCLUSION

There is no doubt that teamwork is an important aspect of software engineering that needs to be included in the curriculum. A simplified model for team composition that is based on the incorporation of diverse personalities formed the basis of this study. This consideration of personality diversity stemmed from an important common factor found in various other models for high-performance team formation reported in the literature. The reality of a given body from which team members must be selected, necessitated a departure from these models. This was accommodated by not specifying specific personality traits for specified team roles as was done in the existing models.

This pilot study suggests that personality diversity is a strong predictor of success that cannot be ignored, especially during the initial phases of team growth. The results also provides *prima facie* evidence that our model for team composition can be used, with a high probability of success, to assist in the formation of teams that are more likely to reach their productive stage sooner.

Thus, while already statistically significant, these results require both finer-grained and more extensive investigation – not only in pursuit of additional corroborative evidence, but also in order to identify threshold values at which sufficient diversity has been attained for sustainable impact on team performance. There is also a need to gain more insight about the relative importance of diversity in relation to other team attributes. In fact it would also be interesting to know, from this further research, if the different personality preferences had an independent role - e.g., if teams with many extroverts performed better than teams with many introverts, or vice versa - or whether it is sheer diversity that matters in all four dimensions

## 7. REFERENCES

- BARON, J. AND NORMAN, M. F. 1992. SATs, achievement tests, and high school class rank as predictors of college performance. *Educational and Psychological Measurement*, 52 (2), 1047–1055
- BELBIN, R.M. 1981. *Management Teams*. John Wiley & Sons: New York.
- BIFFL, S. AND HALLING, M. 2003. Investigating the Defect Detection Effectiveness and Cost Benefit of Nominal Inspection. *IEEE Transactions on Software Engineering* 29 (5), 385–397
- BOEREE, C.G., 2003, Jungian Types Test. URL: <http://www.ship.edu/~cgboeree/jungiantypetest.html>
- BOEHM, B., EGYED, A. KWAN, J. PORT, D. SHAH, A. AND MADACHY, R. 1998, Using the WinWin Spiral Model: A Case Study. *Computer* 31(7), 33–44.
- BRADLEY, J.H. AND HEBERT, F.J. 1997. The effect of personality type on team performance. *Journal of Management Development* 16 (5), 337–353
- DENNING, P.J. 1992. Educating the New Engineer, *Communications of the ACM*, 35 (12), 83–97.
- EDGE, O.P. AND FRIEDBERG S.H. 1984. Factors Affecting Achievement in the First Course of Calculus, *Journal of Experimental Education*, 52(3), 136–140
- GORLA N. AND LAM Y.W. 2004. Who should work with whom? *Communications of the ACM*, 46 (6), 79–82
- GUINAN, P.J, COOPRIDER, J.G. AND FARAJ, S. 1998. Enabling Software Development Team Performance During Requirements Definition: A Behavioral Versus Technical Approach. *Information Systems Research* 9(2), 101–125
- HAZZAN, O. AND TOMAYKO, J.E. 2004. Reflection Processes in the Teaching and Learning of Human Aspects of Software Engineering. *Proceedings of the 17th Conference on Software Engineering Education and Training (CSEET'04)*
- HOUSE, J.D. 2000. Academic Background and Self-Beliefs as Predictors of Student Grade Performance in Science, Engineering and Mathematics. *International Journal of Instructional Media*, 27 (2) 207–213
- KEIRSEY, D. AND BATES, M. 1984. *Please Understand Me*, Prometheus Nemesis Book Company, Del Mar, California.
- KOCH, P-P. 2003. The Ideal Web Team (part 1). *Digital Web Magazine*. [http://www.digital-web.com/articles/the\\_ideal\\_web\\_team\\_part1/](http://www.digital-web.com/articles/the_ideal_web_team_part1/)

- LINGARD, R. AND BERRY, E. 2002. Teaching Teamwork in Software Engineering based on an understanding of factors affecting group performance. *Proceedings of the 32th ASEE/IEEE Frontiers in Education Conference, Nov 6–9, Boston, MA.*
- MYERS, I. B. AND BRIGGS, K. 1975. *The Myers-Briggs type indicator (Form G)*. Palo Alto, CA: Consulting Psychologists Press.
- PAULK, M.C. 2002. Agile Methodologies and Process Discipline, *CrossTalk*, Oct 2002, 15–18.
- RUTHERFOORD, R.H. 2001. Using Personality Inventories to Help Form Teams for Software Engineering Class Projects. *ITiCSE 2001 6/01 Canterbury, UK.*, 73 – 76
- SMITH D.C., BECKER M., BURNS-HOWELL J. AND KYRIAKIDES, J. 2001. Creating High performance IS Teams, *SAICSIT, Pretoria 25 – 28 September 2001*, 172–181