

A study into the effects of personality type and methodology on cohesion in software engineering teams

J. S. KARN, S. SYED-ABDULLAH, A. J. COWLING and M. HOLCOMBE

Department of Computer Science, University of Sheffield, Regent Court,
211 Portobello Street, Sheffield, S1 4DP, UK

The aim of the research described in this paper is to gain a qualitative understanding of how cohesiveness relates to personality type, performance and adherence to a particular software engineering methodology (XP). A variety of research methods were employed including ethnographic methods, questionnaires and interviews. An online personality test based on the Myers-Briggs Type Indicator (MBTI) was used to ascertain the personality types, and questionnaires were used throughout the project to measure levels of cohesiveness. Examples of how the teams worked together throughout the project are described, and whether and how this relates to the personality types of the individual members. The results indicate that certain teams were found to work consistently well over the project due to homogeneity in personality type and others were found to be very cohesive due to a mixture of types.

Keywords: Software engineering; Methodology; Personality type; Teams; Myers-Briggs

1. Introduction

This paper describes research carried out on student teams taking part in large-scale software development projects; the aim of the research is to uncover combinations of personalities who can work in a cohesive manner throughout the project and the influence on this (if any) of the methodology. Software engineering (SE) activities are essentially cooperative and are performed by teams (Saeki 1995). In these cooperative situations, different roles of workers such as clients, users, analysts, designers and managers and so on, participate in activities and communicate with each other (Saeki 1995). The importance of social processes and human factors to SE was stressed in 1995 when Yourdon and Constantine issued the prophetic warning that any future, substantial SE productivity improvements are heavily contingent on much greater emphasis being paid to the organisational, social and human factors of SE (Yourdon and Constantine 1995).

Extreme programming (XP) is the methodology that is generally used in the Genesys projects (an MSc level

course from which the research subjects were drawn), the essence of this study has been to observe qualitatively all of the teams working on Genesys projects using the XP methodology, in order to identify the effect of different combinations of personalities on the overall cohesion of the team whilst using XP, and in turn how the methodology combined with the personalities of individual developers affected the overall performance of the team.

For any SE project manager the ideal is that the members of a team should cooperate smoothly and efficiently, by resolving issues amicably by constructive debate. In practice, various studies have provided evidence that this ideal is not achieved in other contexts and have proposed explanations for this.

Bass and Dunteman described a three-way classification of individual personality types and identified some ways in which these types affected an individual's approach to working in teams (Bass and Dunteman 1963). Belbin developed this approach by focusing on the different roles that need to be played by members of teams (Belbin 1981),

*Corresponding author. {J.Karn, S.Abdullah, A.Cowling, M.Holcombe}@dcs.shef.ac.uk

and Elam and Walz took it further by studying teams working in the requirements definition phase of a software project and observing the effects of the interpersonal conflicts within the team (Elam and Walz 1988). More recently, Henry and Stevens have investigated how Belbin's roles apply to the particular case of SE development teams (Stevens and Henry 1999).

Other researchers have also analysed the shift from traditional SE methodologies to XP in various industrial and educational contexts. Holcolombe, Macias *et al.* (2002) described the difficulties involved in introducing and teaching XP to undergraduate students, while Williams and Cockburn (2001) found that the introduction of pair programming increased many measures of code quality, although at a slight increase in programming time. Mugridge, MacDonald *et al.* (2003) found that the on-site customer and 40-hour weeks could not be instituted in student-based projects, and Lappo (2002) described a 12-week course taken by MSc level students in which only the planning game and 40-hour work practice were successfully established. Noll and Atkinson (2003) conducted an experiment in a 10-week SE course and found that teams using traditional methodologies, produced software with more functionality, while XP teams experienced problems with weak customer engagement, a lack of collective code ownership and weak batch integration.

The work described here attempts to build on this previous work by merging the issue of XP adaptation with that of team dynamics, particularly personality psychology. An area that seemed to offer further investigation to the authors was that of the interplay between different personalities in an SE team and the effect this had on the overall cohesion of the team, while they were attempting to follow the tenets of XP. Figure 1 shows the average cohesion after three readings.

Therefore the paper is laid out as follows: section 2 gives a description of XP, the relevant personality research and workgroup cohesion; section 3 describes the motivation behind this work and the major research questions; section 4 describes the organisational setting and how the observed behaviour was classified and analysed; section 6 describes possible threats to validity; section 7 summarises the main results for the five teams studied; section 8 then evaluates these results; finally, section 9 draws conclusions from the work done and proposes future work.

2. Background

2.1 XP

Extreme programming (XP) is perhaps the best known of a new breed of 'agile' SE methodologies. The founder of XP – Beck – states that there are 12 core practices related

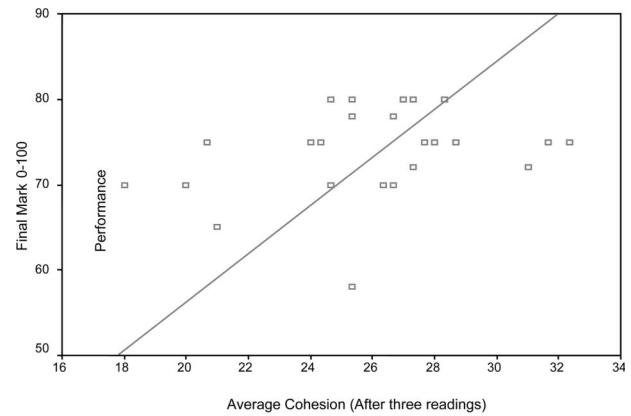


Figure 1. Average cohesion and final mark.

Note: Average cohesion level was used to study the relationship rather than the cohesion level at the third reading because this data represents the overall feeling of cohesiveness among the team members, as opposed to at a particular point in time i.e. when a reading was taken.

to this methodology, which he describes fully in *Extreme Programming Explained: Embrace Change* (2000). XP was created in response to problem domains whose requirements change and to address the problem of project risk. XP begins with four values: Communication, Simplicity, Feedback and Courage. It then builds up to the 12 practices that XP projects should follow. Many of the XP practices were created and tested as part of the Chrysler C3 project (Hendrickson 1999). Beck introduced XP as a solution to the problems encountered by more formal methods. XP focuses on four humanistic values: communication, simplicity, testing and aggression, and how each of them is interrelated.

Beck was always careful to make sure people were realistic when discussing XP; this becomes apparent in the quote: 'If you want to try XP, for goodness sake don't try to swallow it all at once. Pick the worse problem in your current process and try solving it the XP way'.

In 1999, Beck was forced to admit that XP was not suitable for all projects and sensibly distanced himself from people who were shouting from the rooftops that XP was indeed the 'Silver Bullet' to slay the mythical lycanthrope of problems with teams working on SE projects (Brooks 1987).

XP is aimed at small to medium-sized teams. The physical environment is also important; communication and coordination should be maintained at all times. Business culture is another focal issue in XP; if there is any resistance against XP principles on behalf of project members, management or customers, then it may be enough to fail the whole process.

2.2 Personality research

2.2.1 The work of C.G Jung. Jung is widely regarded as one of the greatest names in the history of empirical science. His own personal experiences and experiments formed the basis of most of his theories. Jung came up with many ideas and was a prolific writer – a full description of Jung’s promulgations is beyond the scope (and is indeed impractical) for a research paper. Interested readers should refer to the 18 Volumes of the ‘Collected Works of C.G Jung’.

It is Jung’s work on psychological types that is of direct relevance to this research. Jung began his career in psychiatry in December 1900 when he was appointed as an assistant physician at the Burghoeltzli Mental Hospital in Zurich under Bleuler, an expert on – and the man who devised the name of – schizophrenia (Storr 1998).

Jung first became interested in personality differences when he considered his own approach to problems of the day with those adopted by Freud and Adler. How was it, he asked himself, that each person could interpret the same material so differently? From this question Jung concluded that people did indeed belong to different psychological types and in 1921 he published his book, ‘Psychological Types’ (Jung 1976).

Jung went on to state: ‘I distinguish four functions: thinking, feeling, sensation and intuition. The essential function of sensing is to establish that something exists, thinking tells us what it means, feeling what its value is and intuition surmises whence it comes and whither it goes’ (Jung 1936).

2.2.2 The Myers–Briggs Type Indicator (MBTI). The MBTI is the brainchild of Mrs Isabel Myers and her mother Mrs Katherine Cook Briggs. Myers–Briggs dedicated their lives to understanding, interpreting, expanding and adapting Jung’s theory of psychological type. They searched in vain for a test that would indicate a person’s Jungian type and finally decided to create one of their own. The test used in this research is not the official paper version of the MBTI as devised originally by Myers and Briggs; rather it is an online test based on the MBTI developed by Human Metrics, a consortium of Israeli psychologists (www.humanmetrics.com/cgi-win/JTypes2.asp). Israeli statisticians and psychologists have found no significant statistical differences between this test and the official paper version of the MBTI.

So how did Myers–Briggs extend Jung’s theory? A brief overview of the descriptions they applied to Jungian functions and attitudes is necessary at this point:

- Sensing: Depend on their five senses for perception. Whatever comes directly from the senses is part of the sensing types own experience and is therefore

trustworthy. Anyone preferring sensing to intuition is primarily interested in actualities.

- Intuition: The intuitive listens for the intuitions that come up from their unconscious with enticing visions of possibilities. They are mainly interested in possibilities. The only fields that interest intuitives are those that give inspiration some play. They abhor the routine as it leaves nothing for inspiration to accomplish.
- Thinking: Is essentially impersonal. Its goal is objective truth, independent of the personality and wishes of the thinker or anyone else. So long as the problems are impersonal and can be judged from the ‘true-false’ standpoint then thinking is the better instrument
- Feeling: Human motives are notably personal. Therefore in the sympathetic handling of people where personal values are important, feeling is the most effective instrument (Myers and Myers 1997).

Myers–Briggs decided that one more preference was relevant to the identification of type. This is the choice, between a perceptive attitude and a judging attitude. There is a fundamental opposition between the types as described below:

- Judging: When using this attitude to come to a conclusion people will shut off perception for the time being. All the evidence is in and anything else is irrelevant; the time has come to arrive at a verdict.
- Perceiving: In this attitude people shut off judgement. Not all of the evidence is in; new developments will occur. It is much too soon to do anything irrevocable (Myers and Myers 1997).

Both of these types need each other to some degree. An adequately developed judging process should support a perceptive type; otherwise they will simply drift without making any decisive decisions. A judging type with insufficient perception will have no ‘give’ or ‘cooperation’ in them. They will be narrow, rigid and incapable of seeing any point of view except their own. This characteristic of the relentlessly judging individual is recognised in the word, prejudice – a pre-judgement impervious to perception (Myers and Myers 1997).

The description of the introvert–extravert dichotomy is similar to Jung’s. A conduct of an extravert is based on the outer situation whereas an introvert starts with the inner ideas – the mental concepts derived from what Jung termed the ‘archetypes’.

One process – sensing, intuition, thinking or feeling – must have clear sovereignty, with opportunity to reach its full development, if a person is to be really effective. For

people to be balanced they need adequate (not equal) development of a second process; not to rival the dominant process, but to act as a welcome auxiliary.

The fact that Jung didn't describe the use of auxiliary processes was the main criticism levelled at him from Myers–Briggs. Jung in contrast to Myers–Briggs described each process in its sharpest focus and with maximum contrast between its extraverted and introverted forms. Describing these rare theoretically 'pure' types with no development of the auxiliary was not applicable to mankind in general.

The MBTI has four dichotomous scales, resulting in a classification of one of four letters indicating the preference in each dichotomy. The four scales are as follows:

Extraversion – Introversion (E–I)

Intuition – Sensing (I–S)

Feeling – Thinking (F–T)

Judging – Perceiving (J–P)

There are 16 possible types that can result from the test. An example is INTJ (Introvert iNtuition Thinking Judging). On completion of the test the student will also know their clarity of preference; for example the result could be: I 52% N 33% T 22% J 62%. An indicator for clarity of preference is:

- 40 or higher (30 for T/F) – Very clear preference
- 21–39 (21–29 T/F) – Clear preference
- 11–20 – Moderate preference
- 1–10 – Slight preference

2.3 Workgroup cohesion

Workgroup cohesion refers to the degree to which team members have close friendships with others in their immediate work unit and their personal attraction to members of the group. According to Klein, increased task performance by cohesive groups is due to more frequent, less-inhibited, task-related communication. Cohesion has been proposed to be one of the important determinants of workgroup performance (Mullen and Copper 1994, Klein and Mulvey 1995).

Research into job-related diversity has found out that factors such as age and gender have less impact than personality. Much of the research on diversity in workgroups has concluded that diversity can be a 'double-edged sword', and specifically that it can result in a high-quality product while at the same time decreasing the cohesion (Webber and Donahue 2001).

Examinations of the teams in this study show that there is group diversity in the form of demographic, personality

and functionality. Functional diversity refers to the differences in skill and knowledge (Pelled 1996).

The questionnaire used to ascertain levels of workgroup cohesion in this project can be viewed in Appendix A.

3. Motivation and research questions

What this work aims to uncover is whether certain personalities contribute more in terms of ideas, speaking in public, being a foil for ideas, or working hard behind the scenes, and how they effect cohesion and adherence to a methodology (XP). Our research questions are therefore as follows:

- RQ1: Do certain combinations of personality type (as measured by the online MBTI test) promote harmonious cohesive teamworking practice?
- RQ2: Are certain personalities better suited to XP?

4. The experimental context

4.1 The software engineering observatory

The context for this study is the Software Engineering Observatory at the University of Sheffield, a research facility that is run by the Verification and Testing (VT) research group. The objective of this observatory is twofold: firstly, it is to create an environment for training and developing skills that are associated with the successful construction of a software solution with a real commercial client and secondly, it is for the carrying out of research work that would be impossible in the real software industry. The observatory consists of three student projects: The Software Hut, which is a second-year undergraduate project; Genesys Solutions, an MSc level project; and the Maxi project, another MSc level project. Each project involves teams of students working on projects with the aim of producing a software system for an industrial client.

The observatory allows empirical researchers to observe, question or interview software developers working on real industrial projects. Fourth-year and MSc students in the Department of Computer Science run a professional software house known as Genesys Solutions; it is these students who represent the research subjects in this paper.

4.2 Genesys Solutions

Genesys was set up as a response to the challenges of trying to introduce the entrepreneurial dimension as part of the higher learning curriculum. The concept is to allow the students to run the software house with minimum intervention from the lecturers, whose function is more as consultants (Holcombe 2001). Everyone connected with Genesys Solutions – students, staff and researchers – agreed

to sign a non-disclosure agreement at the start of the project. As a result of this agreement, only the names of a selection of recent projects can be given. In the period 2001–2004, Genesys' students worked on nine new and six maintenance projects.

4.3 Subjects

Genesys' projects start with a briefing by the project managers. Shortly afterwards the students are divided into; a marketing team, whose main functions are to promote Genesys to businesses, liaise with clients, attract new clients and maintain the company website; a research and development (R&D) team, whose main task is to ensure that all hardware and software applications required by the development teams are available; and finally the software development teams, which make up the main part of the company. Students are generally encouraged to volunteer for projects that interest them. The main income for this company comes from projects that are developed by these teams.

The Genesys project requires students to spend two-thirds of their time during the academic year working on projects for external clients from the requirements analysis through to the final implementation phase. The teams were selected on the basis of personality type, nationality and previous skills/experience. The project managers were keen to have a mixture of home and overseas students in each team.

Five teams were studied throughout the project; they were all software development teams. A decision was made not to study the marketing or research and development teams, because they were not working on any projects; therefore, it would be difficult to explain their results in the context of an SE project.

4.4 Procedure

The students were observed in projects spanning the entire academic year – from the end of September to the end of May. There was an examination period in January and no timetabled sessions for Genesys during this time. However, during this and other such periods, many students still worked on aspects of the Genesys project. However, it must be stressed that no observations took place outside of official university hours.

4.4.1 Week 1. The first week of the project was devoted to group formation; the students were introduced to their new teammates and took part in a team-building exercise. During this week the students were also asked to complete the online personality test and return the result to the project managers. Personality results for all of the teams are included in Appendix B. Communication and interaction between team members were promoted and improved

at an early stage. This aspect of team formation is important to observe, for the study.

4.4.2 Week 2. During the second week, projects and tasks were assigned to the individual teams. The teams were also asked to arrange to meet their external client; production of monthly reports also started in this week. From this point on there was also a weekly board meeting, when the project managers asked questions about the overall Genesys situation. From this point on, the students were observed in client, manager and team meetings. Ethnographic methods were employed to observe selected student teams. Social anthropologists originally founded these methods to aid them in their understanding of different cultures and environments. The students were observed within their natural setting; so this aspect of the study constituted field research, with a focus on collecting qualitative data.

Initial focus group (meaning the team under study) interviews were also conducted to get a glimpse of the students' understanding of the XP methodology. Focus group interviews were chosen because they are the most appropriate method for studying attitudes and experiences, for exploring exactly how opinions are constructed (Kitzinger 1995), and understanding behaviours, values and feelings (Patton 2002).

The difficulties encountered during the early interview sessions resulted in the development of an XP Activities/Product Table (Holcombe, Gheorge *et al.* 2001). The first workgroup cohesion questionnaire was also administered during this time.

Questionnaires were employed in this research to gauge the level of workgroup cohesion; they facilitated the collection and analysis of mass data quickly. The scale used for measuring workgroup cohesion was developed by Price and Mueller (1986). The validity and reliability of this scale has been demonstrated in other studies (Martin and Hunt 1980, Agho, Price *et al.* 1992). The reliability of a scale indicates how free it is from random error. Two frequently used indicators of a scale's reliability are test–retest reliability and internal consistency. For this research, internal reliability indicator was assessed because the scale is designed to measure the fluctuation of mood over a period of time during which the software was developed. Internal consistency refers to the degree to which the items that make up the scale are all measuring the same underlying attribute. Internal consistency can be measured in several ways, but for this research, the most common Cronbach coefficient alpha was used (Pallant 2001). The Cronbach alpha in this study is 0.71.

4.4.3 Weeks 3–11. The students carried on working on the projects. The second data collection relating to workgroup cohesion was carried out in week 11. Two types of documents were used during this research: project documents

and assessment documents. Software project documents were an important source of data in this research. The availability of the documents such as commentary reports, plans, requirement documents, project manuals and minutes of the meetings conducted by each team helped the researchers gain a better understanding of the interrelationship between the members, the problems faced and solutions agreed by the groups in adopting XP practices.

Assessment documents were also analysed during this period. There were two types of documents used for calculating the performance level of each member in a development team. The first was past grades achieved by each members and the second was grades given by the clients and the managers. The teams' performance levels for the past and current semester were derived from these documents.

4.4.4 The second semester. Observations continued until the end of the project. A couple of weeks before the end the final collection of data was carried out. The interview sessions were scheduled to suit the student's timetable. During these sessions, the final questionnaire on workgroup cohesion was administered.

5. Threats to validity

There are a number of threats to the validity of the experiment that need to be discussed before considering the results that were obtained. The subjects were students, not practitioners, although some of them did have experience in SE. Therefore we must be careful about generalising from

students to practitioners. Although it must be stressed that some students may have had relevant experience prior to their studies, since we were dealing with a masters level course, this was true for a number of individuals, and the students without industrial experience had obviously gained more SE experience than undergraduates.

The second threat is that the sample size is relatively small. This constitutes a threat to the external reliability of the study. The third threat was related to the collection of XP related data. The means to which the XP practices were adjusted could only be measured qualitatively, through interview sessions, a self-report from the development teams, progressive reports and the verifying report from the team, which inherited a project. During the course of this study, activities that represented specific XP practices were identified. There was, though, a need to identify precise experimental metrics to capture qualitatively the attributes of interest – that is the number of practices and the extent to which they were used.

6. Analysis of results

A scatter plot (figure 2) was plotted to examine the relationship between average workgroup cohesion (after getting the average of the three readings) and performance. The relationship between the two variables (average cohesion level and performance) was then investigated using Pearson product-moment correlation coefficient. The analysis shows a small positive correlation between the two variables [$r = 0.278$, $n = 27$, $p < 0.005$], with high level of workgroup cohesion associated with high level of

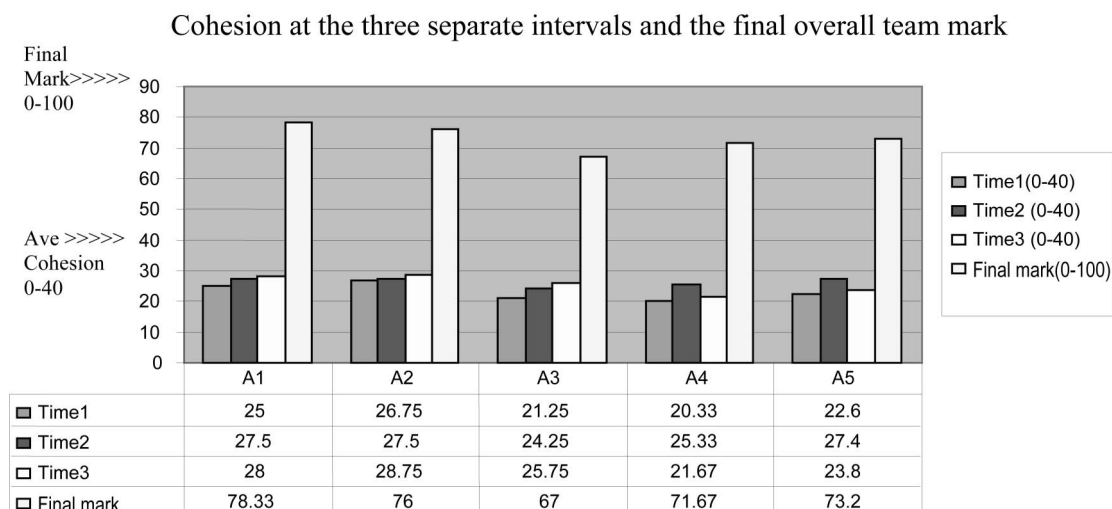


Figure 2. Work cohesion and final mark.

Note: Time1, Time2 and Time3 represent the average level of cohesion at the three separate time intervals (week 2 semester 1, week 11 semester 1 and week 10 semester 2). A semester at the University of Sheffield runs for 12 weeks. The term 'Final Mark' represents the mark given to the team by the client and project manager at the end of the academic year.

performance. This is in accordance with the conflicting findings in group research, regarding the relationship between cohesion and performance (Klein and Mulvey 1995). A review by Stogdill (1948) on this relationship reveals substantial inconsistency suggesting that performance is indirectly related to group cohesion. Further studies in group research suggest other contributing factors such as goal setting (Latham and Locke 1990), group tasks (Klein and Mulvey 1995), group experience (Littlepage, Robinson *et al.* 1997), personality (Thomson and Martinko 1995), organisational culture and group process (Chuang, Church *et al.* 2004), thus mediating the effect of cohesion on performances.

In this study, emphasis on collectivism was made to the teams at the beginning of the term. The importance of teamwork was stressed through the early 'icebreaking' game, continuous pair programming coaching, the use of CVS for frequent incremental integration and the constant client's review. The aim of these approaches is to integrate members who were diverse in terms of race, knowledge and skill.

It was observed that there were two patterns of work-group cohesion experienced by the Genesys teams:

Pattern 1: A team experiencing this pattern would have increased cohesion levels as they progressed along the project. This pattern is consistent with previous studies that illustrate cohesiveness among the team members increases with time due to increased communication and improved coordination (Teasley, Covi *et al.* 2002) during the group development process.

Pattern 2: The team experienced an increase in group cohesion level during the first semester, but the cohesion level started to decrease during the second semester. Overall, the cohesion level is higher at the end of the year than when they started the project. Possible factors to be considered for this pattern are group diversity (Webber and Donahue 2001) and team potency (Pearce, Gallagher *et al.* 2002).

6.1 Results of Team A1

This team in particular worked consistently well over the entire academic year; they worked as a cohesive unit and had a mixture of personalities that ostensibly cooperated and worked well together. Five of the six members of this team had an NT (intuition thinking MBTI preference), a typical profile for engineers.

Each member of the team also had a preference for judging over perceiving. This proved vital, as there is the danger that clear intuitive types can lack persistence unless balance is attained through the development of a judging process. Five of the six members also had a preference for thinking over feeling. These people placed a higher value on

logic than on sentiment and were on the whole impersonal and businesslike in their relations with each other. They were always truthful with opinions on a piece of work and this went down well with teammates. This businesslike, efficient approach to work was not seen as lacking in sociability; it was viewed as a logical sequence of stating the subject, making the necessary points, coming to a conclusion and avoiding repetition.

This team experienced a gradual increase in the group cohesion as they progressed along in their project. Even though there was a demographic diversity, there was no tendency for the members to have a preference for members of their own ethnic group and this helped the team to work harmoniously. The ability of the team members to coordinate themselves according to their expertise and non-existence of gap in skills enabled this team to work harmoniously.

Even though this team scored the highest in term of performance, there were glimpses of task-related conflict during the focus group discussion. Task-related conflict refers to disagreement among the team members about task issues, including the nature and importance of the task goal. Studies on task conflicts have shown both positive (Jehn 1995, Lovelace, Shapiro *et al.* 2001) and negative (Liang, Moreland *et al.* 1995) results. In the case of team A1, the conflict is more on practising pair programming rather than the project-related task. The experience of using XP in previous projects enabled some of the members to recognise the strength and weakness of its practices. These are the members who appointed themselves as XP enforcers, creating a slightly uncomfortable environment to those who were reluctant to work in pairs all the time. Nevertheless, the enforcers managed to make the pair programming practices compulsory, albeit reluctantly, hence this could have contributed to the slight increase in the group cohesion level. The task conflict in this case seems to have more favourable effect on the performance of the team.

Another factor that might contribute to this pattern of cohesiveness is that the group recognised some of their member's expertise. Prior group experience allowed group members to understand the ways in which other members might be able to contribute and this has the advantage of increasing the group cohesiveness and performance. Studies have found that team experience led to more trust in members expertise and the group that were trained together were better able to recognise other members expertise (Liang, Moreland *et al.* 1995).

6.2 Results of Team A2

There were four NTs (intuition thinking on the MBTI) out of five people in this group. In comparison with A1, A2 had a broader range of personalities including one who

preferred sensing and one who preferred feeling. This range of personalities was a source of strength rather than discord in this team because the different types came up with different ways of approaching the same problem. The main players in this group were an ENTJ (Extravert Intuitive Thinking Judging) and an ISTP (Introvert Sensing Thinking Perceiving).

This group had one perceiving type among all of the judging types. This meant that decisions were not made on a hasty basis as the perceiving types liked to keep a decision open for as long as possible before doing anything irrevocable. These people are typically flexible, adaptable and tolerant, and provide a good foil for judging types who aim to settle matters as quickly as possible. They can work together providing the clarity of preference for the judging types is not very clear and vice versa for the perceiving types.

The feeling type in the team ensured that personal human values were not totally neglected in the quest for a logical conclusion. This is one of the reasons for the high levels of cohesiveness within the team. They had one person who was naturally friendly and very interested in taking human feelings into account and this may have contributed to the group experiencing the first pattern of cohesion.

The group diversity for this team is more demographic than functional. The cohesiveness among the team members was very high from the start of the project. During the two focus group discussions, it could be sensed that the members were very comfortable with each other. The maturity among some of the members helped in fostering better interpersonal relationships because these members were known to occasionally break the 'developers block' during the project meeting by taking a break at the nearest pub or discussing topics other than task-related.

Examinations of the XP practices used by this team revealed that they were the team that had applied the most practices. The presence of a member who had used some of these practices before helped the team to adhere to the selected XP practices as closely as possible. To practice pair programming needs effort on the part of all the team members and in the case of this team, members reported that pair programming activities had been an enjoyable way of working. The members found that the changing of partners among pairs could be easily accommodated. This may be due to the improvising approach exercised by the team, whereby if the time did not permit them to change partners at certain times, then the members exchanged part of the system. The use of CVS, a version management tool also helped the team to foster collective ownership among the members. The interpersonal relationship between the members helped the team to appreciate the flexibility of the methodology by modifying it to suit the team and the project.

6.3 Results of Team A3

This team had only two NTs, the combination of functions usually associated with scientists and engineers. They were also the worst performing team. They had two feeling types and two sensing and a 3–2 ratio with extraversion predominant over introversion. This team experienced many 'political' problems with the client as opposed to SE problems; these problems were caused in the main by the intransigence of the client. It would be reasonable to expect any team to struggle in the face of this obduracy, whether they would struggle to the same degree as A3 is open to speculation; one is tempted to think that this would depend on the combination of personalities. The feeling types worked hard to create a harmonious situation in this team. They had found a way to work out which issues were important when looked at from a personal angle. In teaching, acting, and the other arts, in oratory, in the relations of the clergy to their congregations, in family life, in social contacts and in any sort of counselling it is feeling that serves as a bridge between one human being and another.

Similarly the sensing–intuition divide in this group was interesting. The problem was the intuitives had a very clear preference for this over sensing and it was therefore difficult for them to reconcile their ideas with the sensing types.

This team's diversity is in term of demographics, functionality and MBTI type. The task-related diversity that is present in this team should have been able to propel it to perform better, but the tendency of task-avoidance with a specific member of the team also had an effect on their effectiveness. The added political problem faced with the client, in addition to the SE problems seems to be the worst-case scenario for any developers to face. Nevertheless the group managed to progress with their project, albeit slowly. The presence of a sensing–feeling personality in this team was the driving force in keeping them going. This behaviour was typical of a sensing–feeling type as they managed to act as a peace broker in interteam understanding in the spirit of reconciliation with other members of the team. This sensing–feeling personality along with another member who had previously experienced working with the required methodology even though not adhering to it properly before, helped the team to use as many XP practices as possible. Previous research has shown that task experience can lead to ability-based enhancement when conditions promote transfer of specific knowledge or strategy (Littlepage, Robinson *et al.* 1997).

This team experienced the first pattern of group cohesiveness. The factor that might contribute to this pattern is the effort made by the team members to adhere to the practices that are easier to use such as pair programming and collective ownership. Due to the composition of the team and very few timetable clashes, members managed

to meet regularly in order to apply the activities that improved cohesiveness among them. Previous experience in working in pairs helped the members to progress to the next stage of pair programming, i.e. the changing of partners between the pair. In addition to the method of doing pair programming, the team also practiced having different test partners. These interactions are possible in reducing any emotional conflict that might arise in the team.

6.4 Results of Team A4

This team was diverse in terms of demographics, personality, skill and knowledge. They made earnest efforts to apply as many XP practices as possible. Attempts to apply all of the pair programming activities, which included changing of pairs regularly or swapping of the software modules were not successful due to timetable clashes and differences in skill. This team had two members less than the other project teams; this also contributed to the decision of doing 'selective pair programming' towards the end of the project. Selective pair programming refers to the cases of developers reverting to this practice when there are major or more difficult tasks to be completed.

Team A4 was an interesting one that contained two perceiving types, and only one NT (Intuitive Thinking). One of the perceivers was of ESTP (Extraverted Sensing Thinking Perceiving) type and they worked well as an organiser and kept the team moving. These are very useful attributes in an industrial context when time is of the essence. The ESFP (Extraverted Sensing Feeling Perceiving) focused on the facts but handled these with personal warmth. SFs (Sensing Feeling) have a tendency to be sympathetic and friendly and enjoy occupations that provide practical help and service for people. In a US study, SFs (Sensing Feeling) made up only 5% of the sample of science students (Laney 1949).

During the focus group discussion, the group explained that even though three of them had some limited working experience before joining the company, those experiences did not include the challenge of working as a team as demanded by the XP approach. Cohesiveness among the members started to diminish as the project became more complex and the members reverted to working individually in order to complete the project on time. The above factors might have contributed to the group experiencing the second pattern of cohesiveness.

6.5 Results of Team A5

On the whole this group worked well, but sometimes there was almost too much emphasis on not offending people. This meant that work was sometimes not done and there were no recriminations. In some cases more discipline was needed as well as a clear plan. Team A5 did not suffer from

continuous serious disruptions, but they did encounter communication problems and occasional personality clashes.

Team A5 had a majority of feeling types. They tended to weigh how deeply they cared about things that could be gained or lost by each of the alternatives. The prominent members of the group ensured that each final decision had a sound basis because they took into account facts, possibilities and human values.

One member of this group was an ENTJ (Extraverted Intuitive Thinking Judging) with very clear preferences for his functions and attitudes. His no-nonsense, efficient, often boisterous approach to decision-making did not go down too well with certain introverted feeling types. Potentially volatile situations were usually defused by excellent negotiation and people management skills from senior members of the team.

The team members were diverse in terms of demographics and skills. The differences in their experience on task and group work contributed to the team experiencing the second pattern in group cohesiveness. The focus group interviews conducted with the team revealed that they did not function as well as the other teams because of the loafing behaviour and effort-avoidance of particular team members. This in turn led to the feeling of resentment towards some of the group members. Research has discovered that teamwork does not always increase the participants mindful engagement in learning and thus improve its outcomes, and the existence of social-psychological effects such as understanding and empathy can also debilitate team performance (Solomon and Globerson 1989).

The discussion on XP applications identifies that the team did not truly adhere to most of the XP practices when developing their project. Pair programming was not fully practised, whereby the team members were supposed to pair as much as possible and in addition to changing partners. The nearly non-existence of this practice among some of its members is a possible contribution to the second cohesive pattern experienced by this team.

7. Initial findings

The initial findings provide some tentative answers to our research questions.

7.1 RQ1: Do certain combinations of personality type (as measured by the online MBTI test) promote harmonious cohesive teamworking practice?

Evidence was found to lend support to the view that certain combinations of personality type are likely to work better during an SE project. The two most consistent teams throughout the academic year (A1 and A2) had an

overwhelming majority of types routinely associated with engineering and science. These teams demonstrated that agreement is straightforward if one has a team of almost homogenous personality types. Because the teams worked well, the level of cohesiveness increased as the year went on. Although this kind of relationship is useful for efficiency and efficacy it does not help to foster a spirit of camaraderie between team members. Nonetheless, the ostensible success of these teams had the effect of increasing cohesion as the projects reached their climax.

An interesting finding relating to this question was that the team with the highest cohesion rating (A2) did not have the highest performance rating; admittedly, they came second by a narrow margin, but this finding suggests that a high level of cohesiveness is not the only factor that can contribute to the overall success of a team. The relationship between the members of A2 was decidedly personal and they all felt comfortable in each other's company. The fact that they also had a majority of typical engineering personality types, coupled with experience of using XP meant that they were always likely to be a high-performing team.

Another interesting finding was that the worst performing team was the furthest away from having a typical personality profile for an SE team. This group had a broad mixture of types and also suffered from 'political' problems with clients and team members not contributing to the desired level. It is reasonable to state that such political problems could have affected other teams, but whether they would have been allowed to reach such a critical stage is open to debate.

One of the teams (A4) performed reasonably well, yet they had the lowest levels of cohesion throughout the project. The main factor here was the personality type of the people involved with this team. One member had a boisterous and noisy motivational style and this alienated him from certain other members of the team who were quiet, retiring types. Serious disruption was averted by good people-management skills; this was impressive when one considers the diversity of this team and the potentially explosive clashes that could have occurred. In spite of this the team lacked a high level of cohesion (in comparison with other teams) because of intolerable differences and obduracy from the persons concerned.

These initial results lend support to the theory that different combinations of personality types can and do have an effect on the performance of SE teams. The results show that teams with a more traditional science/engineering personality profile tend to outperform those with more multifarious types. The initial results also show that the two teams with the highest levels of cohesion throughout the project were also the two top teams in terms of performance. This adds support to the theory that high cohesion equals a successful team.

7.2 RQ2: Are certain personalities more susceptible to XP?

While there was some evidence to support the argument that certain personality types are more open to the adaptation of XP, another reliable predictor was if one or more team members had prior experience of the methodology. In Team A1, a majority of the members had prior experience, and they appointed themselves as XP enforcers. This ensured that XP was compulsory in this team.

Team A2 applied the most XP practices throughout the project. This was largely due to one member (a sensing/feeling type) with prior experience who was keen to share his knowledge with others. The patient nature of the members with XP experience coupled with the interpersonal relationship between the members helped the team to appreciate the flexibility of the methodology by modifying it to suit the team and the project.

Team A3 had only two NTs, the combination of functions usually associated with scientists and engineers and were also the worst performing team. However, the presence of a sensing/feeling personality encouraged the team to follow as many of the practices as possible, albeit without adhering to them in the correct manner.

Team A4 made determined efforts to apply XP practices. These efforts were largely unsuccessful as this team was very diverse in terms of demographics, personality, skill and knowledge. Due to this diversity, working as a cohesive team and following XP guidelines proved to be very difficult for this team. As the project progressed the team became less cohesive and less likely to follow XP guidelines.

The discussion on XP applications identifies that the team do not truly adhere to most of the XP practices when developing their projects. Pair programming was not fully practised, whereby the team members were supposed to pair as much as possible and in addition, to change partners. The nearly non-existence of this practice among some of its members is a possible contribution to the second cohesive pattern experienced by this team.

For team A5, the caring considerate nature of the prominent group members meant that enforcing XP guidelines was never going to be easy. No action was taken against people who refused to adhere to these guidelines, and it was clear that in some cases a more direct and disciplined approach was needed. Due to the lack of cohesion in the team, some XP practices, particularly pair programming were not fully practised.

8. Conclusions

Three emerging conclusions were found during this research. The first is that combinations of personality types are important and they can have an effect both on performance and cohesion. The second is that teams with a high level of cohesion tend to outperform other teams

with lower levels. The third is that the highest measure of cohesion does not equal the most successful team in terms of performance. A fourth emerging conclusion is that it could be said that there are two types of cohesion, which might be termed social and technical. These different types of cohesion could be distinguished by looking at the kind of relationship(s) between the different team members and whether purely project-related or more personal human-centred issues dominate the agenda of a particular team. What this shows is that a team whose members do not have very close friendly relationships with each other, can still be cohesive and outperform others that enjoy more open relaxed relationships. Such a team achieves this by virtue of concentrating on the task in hand and devoting all energies to completing a given project.

This paper gives an insight into how the collectivistic culture of an organisation benefits the group diversity. Early studies by Chatman *et al.* (1998) and a theoretical study by Chuang *et al.* (2004) discussed how the organisational culture moderates group diversity and intragroup conflict. This paper has shown an insight into the group process taking place when the organisation emphasises collectivism. The decision of the consultants to insist on the use of an agile methodology and to specifically stress the collective behaviour practices such as pair programming, collective ownership, continuous integration and continuous review have some effect on the behaviour of the software development teams.

Many people in different subjects have stated that differences between people can either help a team to grow and be strong or they can help to destroy it from within. What is lacking in such statements is an explanation of how and why the people are different and what it is about them that makes them either a cooperative or intransigent person. By analysing personality differences and measuring levels of cohesion in the different teams this study has gone some way towards explaining how these differences and the XP methodology can contribute to the overall cohesion of a team.

References

- AGHO, A.O., PRICE, J.L. and MUELLER, C.W., 1992, Discriminant validity of measures of job satisfaction, positive affectivity and negative affectivity. *Journal of Occupational and Organizational Psychology*, **65**, pp. 185–196.
- BASS, B.M. and DUNTEMAN, D.G., 1963, Behaviour in groups as a function of self, interaction and task orientation. *Journal of Abnormal Social Psychology*, **66**, pp. 19–28.
- BECK, K., 2000, *Extreme Programming Explained: Embrace Change* (Reading, Mass, USA: Addison Wesley, Longman), pp. 37–43.
- BELBIN, R.M., 1981, *Management Teams: Why they succeed or fail* (Oxford: Butterworth-Heinemann).
- BROOKS, F.P., 1987, No silver bullet: essence and accidents of software engineering. In *IEEE Computer*, **20**, pp. 10–19.
- CHATMAN, J.A., POLZER, J.T., BARSADE, S.G. and NEALE, M.A., 1998, Being different yet feeling similar: the influence of demographic composition and organizational culture on work processes and outcomes. In *Academy of Management Journal*, **42**, pp. 273–287.
- CHUANG, Y.T., CHURCH, R., ZIKIC, J., *et al.* 2004, Organizational culture, group diversity and intra-group conflict. In *Team Performance Management*, **10**, pp. 26–34.
- ELAM, J.J. and WALZ, W.D., 1988, A study of conflict in group design activities: implications for computer supported cooperative environments. *Proceedings of the 21st Annual Hawaii International Conference on Decision Support and Knowledge Based Systems Track*, pp. 247–254.
- HENDRICKSON, C., 1999, DaimlerChrysler: The best team in the world. In *IEEE Computer*, **32**, pp. 60–69.
- HOLCOMBE, M., 2001, Real software engineering projects at the University of Sheffield. *Forum for Advancing Software Engineering Education (FASE)*, **11**, pp.
- HOLCOMBE, M., GHEORGE, M., MACIAS, F., *et al.*, 2001, Teaching XP for real: Some initial observations and plans. *Proceedings of the 2nd International Conference on Extreme Programming and Flexible Processes in Software Engineering (XP2001)*, pp. 14–17.
- www.humanmetrics.com/cgi-win/JTypes2.asp.
- JEHN, K.A., 1995, A multimethod examination of the benefits and detriments of intragroup conflict. In *Administrative Science Quarterly*, **40**, pp. 256–282.
- JUNG, C.G., 1936, Psychologische Typologie. In *Sueddeutsche Monatshefte*, **5**, pp. 264–272.
- JUNG, C.G., 1976, *Psychological Types*, Routledge and Kegan Paul, London.
- KITZINGER, J., 1995, Qualitative Research: Introducing Focus Group. In *British Media Journal*, **3**, pp. 299–302.
- KLEIN, H.J. and MULVEY, P.W., 1995, Two investigations of the relationships among group goals, goal commitment and performance. In *Organizational Behavior and Human Decision Processes*, **61**, pp. 44–53.
- LANEY, A.R., 1949, *Occupation implications of the Jungian personality function types as identified by the Briggs-Myers type indicator* (Washington D.C.: George Washington University).
- LAPPO, P., 2002, No Pain, No XP: Observations on teaching and mentoring extreme programming to university students. In *Proceedings of the Third International Conference on Extreme Programming and Agile Processes in Software Engineering (XP2002)*, pp. 35–38.
- LATHAM, G.P. and LOCKE, E.A., 1990, *A theory of goal setting and performance* (Englewood Cliffs, NJ, US: Prentice Hall).
- LIANG, D.W., MORELAND, R., ARGOTE, L., *et al.*, 1995, Group versus individual training and group performance: the mediating role of transactive memory. In *Personality and Social Psychology Bulletin*, **21**, pp. 384–393.
- LITTLEPAGE, G., ROBINSON, W., REDDINGTON, K., *et al.*, 1997, Effects of task experience and group experience on group performance, member ability and recognition of expertise. In *Organizational Behavior and Human Decision Processes*, **69**, pp. 133–147.
- LOVELACE, K., SHAPIRO, D.L., WEINGART, L.R., *et al.*, 2001, Maximising cross-functional new product teams' innovativeness and constraints adherence: a conflict communications perspective. In *Academy of Management Journal*, **44**, pp. 779–794.
- MACIAS, F., HOLCOMBE, M., MACIAS, F., *et al.*, 2002, Empirical experiments with XP. In *Proceedings of XP 2002*, pp. 225–228.
- MARTIN, T.N. and HUNT, J.G., 1980, Social influence and intent to leave: A path-analytic process model. In *Personnel Psychology*, **33**, pp. 505–528.
- MUGRIDGE, R., MACDONALD, B., TEMPERO, F., *et al.*, 2003, Five Challenges in Teaching XP. In *Proceedings of the Fourth International Conference on Extreme Programming and Agile Processes in Software Engineering (XP2003)*, *Lecture Notes in Computer Science* (Genoa, Italy: Springer-Verlag), pp. 406–409.

- MULLEN, B. and COPPER, C., 1994, The relation between group cohesiveness and performance: An integration. In *Psychological Bulletin*, **2**, pp. 210–227.
- MYERS, I.B. and MYERS, P.B., 1997, *Gifts Differing Understanding Personality Type* (California, USA: Davis Black), pp. 1–45.
- NOLL, J. and ATKINSON, D.C., 2003, Comparing Extreme Programming to Traditional Development for Student Projects: A Case Study. In *Proceedings of the Fourth International on Extreme Programming and Agile Processes in Software Engineering (XP2003)*, *Lecture Notes in Computer Science* (Genoa, Italy: Springer-Verlag), pp. 372–374.
- PALLANT, J., 2001, *SPSS Survival Manual* (Buckingham, Philadelphia, USA: Open University Press), p. 85.
- PATTON, M.Q., 2002, *Qualitative Research & Evaluation Methods* (London: Sage Publications).
- PEARCE, C.L., GALLAGHER, C.A., ENSLEY, M.D., et al., 2002, Confidence at the group level of analysis: A longitudinal investigation of the relationship between potency and team effectiveness. In *Journal of Occupational and Organizational Psychology*, **75**, pp. 115–119.
- PELLED, L.H., 1996, Demographic diversity, conflict, and work group outcomes: an intervening process theory. In *Organization Science*, **7**, pp. 615–631.
- PRICE, J.L. and MUELLER, C.W., 1986, *Handbook for Organizational Measurement* (Marsfield, Mass., USA, Scranton: Harper-Collins).
- SAEKI, M., 1995, Communication, collaboration and cooperation in software development – how should we support group work in software development. In *Proceedings of the 1995 Asia Pacific Software Engineering Conference*, pp. 12–22.
- SOLOMON, G. and GLOBERSON, T., 1989, When teams do not function the way they ought to. In *International Journal of Education Research*, **13**, pp. 89–99.
- STEVENS, K.T. and HENRY, S.M., 1999, Using Belbin's leadership role to improve team effectiveness: an empirical investigation. In *Journal of Systems and Software*, **44**, pp. 241–250.
- STOGDILL, R.M., 1948, Personal Factors Associated with Leadership: A Survey of the Literature. In *Journal of Psychology*, **25**, pp. 35–71.
- STORR, A., 1998, *The Essential Jung: Selected Writings*, (London, UK: Fontana Press), pp. 130–147.
- TEASLEY, S.D., COVI, L.A., KRISHNAN, M.S., OLSON, J.S., et al., 2002, Rapid software development through team collocation. In *IEEE Transactions on Software Engineering*, **28**, pp. 217–233.
- THOMSON, N.F. and MARTINKO, M.J., 1995, The relationship between MBTI types and attributional style. In *Journal of Psychological Types*, **35**, pp. 22–30.
- Webber, S.S. and DONAHUE, L.M., 2001, Impact of highly and less job-related diversity on work group cohesion and performance: a meta-analysis. In *Journal of Management*, **27**, pp. 141–162.
- WILLIAMS, L. and COCKBURN, A., 2001, The Costs and Benefits of Extreme Programming. In *Extreme Programming Explained* (Boston, MA, USA: Addison-Wesley), pp. 223–244.
- YOURDON, E.L. and CONSTANTINE, L.L., 1995, Re-Inventing IS/IT. *Australian Computing Society, Seminar Proceedings*,

This questionnaire consists of a number of words that describe different situations, feeling and emotions. For each item indicate to what extent you have felt this way during this week.

Circle a number that represents your feeling.

1 Not at all or slightly 2 A little 3 Moderately
4 Quite a bit 5 Very or Extremely

1. To what extent are individuals in your project team friendly?

1 2 3 4 5

2. How often do you do things socially with individuals in your project team outside of the project?

1 2 3 4 5

3. How often do you discuss important personal problems with individuals in your project group?

1 2 3 4 5

4. To what extent are individuals in your project team helpful to you in getting your work done?

1 2 3 4 5

5. To what extent do you trust individuals in your project team?

1 2 3 4 5

6. To what extent do individuals in your project team take an interest in you?

1 2 3 4 5

7. To what extent do individuals in your project do favours for you at considerable cost to themselves?

1 2 3 4 5

8. How much do you know about individuals in your project team?

1 2 3 4 5

Appendix A Workgroup Cohesion Questionnaire

Name: _____ Team: _____

Date (answering the questionnaire) : _____

Appendix B MBTI Scores for Individual Teams

Table 1. MBTI results for Team A1 with % clarity of preference.

Team member	E-I	N-S	F-T	J-P
1	E 33	N 44	T 22	J 11
2	E 44	N 1	T 33	J 44
3	E 44	N 33	T 11	J 56
4	E 11	N 11	T 44	J 50
5	I 44	N 22	T 56	J 33
6	I 54	N 48	F 38	J 37

Table 2. MBTI results for Team A2 with % clarity of preference.

Team member	E-I	N-S	F-T	J-P
1	E 28	N 67	T 44	J 50
2	I 11	S 22	T 33	P 22
3	I 33	N 11	T 11	J 33
4	I 33	N 44	F 11	J 22
5	I 1	N 67	T 22	J 56

Table 3. MBTI results for Team A3 with % clarity of preference.

Team member	E-I	N-S	F-T	J-P
1	E 22	S 11	F 33	J 67
2	E 11	N 67	F 11	J 44
3	E 26	N 37	T 46	J 11
4	I 33	N 44	T 1	J 44
5	I 22	S 1	T 22	J 44

Table 4. MBTI results for Team A4 with % clarity of preference.

Team member	E-I	N-S	F-T	J-P
1	E 78	N 22	T 22	J 67
2	E 11	S 24	T 33	P 24
3	E 33	S 22	F 11	P 33

Table 5. MBTI results for Team A5 with % clarity of preference.

Team member	E-I	N-S	F-T	J-P
1	E 88	N 67	T 74	J 56
2	E 22	N 78	F 11	J 33
3	I 34	S 44	F 24	J 22
4	I 67	N 11	T 22	J 33
5	I 33	N 22	F 28	J 36
6	I 33	N 22	F 44	J 44

Copyright of Behaviour & Information Technology is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.