ELSEVIER

The Journal of Systems and Software 44 (1999) 241–250

The Journal of Systems and Software

Using Belbin's leadership role to improve team effectiveness: An empirical investigation

Sallie M. Henry *, K. Todd Stevens

Department of Computer Science, Virginia Tech, 660 McBryde Hall, Blacksburg, VA 24061, USA

Received 16 October 1997; received in revised form 24 December 1997; accepted 29 January 1998

Abstract

This paper presents a controlled experiment conducted with senior software engineering students that demonstrates the utility of forming teams based on R. Meredith Belbin's set of team roles. The overall research effort focuses on the general utility of Belbin's roles in improving the effectiveness of teams, which can be viewed in two ways: performance and team viability. Performance effectiveness, which is the focus of this paper, clearly addresses a team's productivity. To address this problem, the first phase of the total research project consists of a controlled experiment that demonstrates that teams containing leadership roles perform better than teams that do not have this role filled. In a laboratory setting, a number of teams were formed that contained a single leader; others were formed that had no leader or multiple leaders. The results of this experiment are positive; they demonstrate that indeed Belbin's roles provide useful information to form teams. The specific conclusion of this controlled experiment is that a single leader on a team improves a team's performance over teams having multiple leaders or no leader. In other words, as one would expect, the mean time-to-completion for the leaderless teams was significantly greater than the teams with leaders. This means that Belbin's roles can be utilized in formation of new teams as well as in evaluation of extant teams, making certain that a team has a single leader. Both of these aspects, formation and evaluation, are extremely useful to managers of software programmers.

Keywords: Empirical studies; Software engineering; Team formation; Roles

1. Introduction

This paper presents an investigation of how to improve the effectiveness of software development teams in an innovative manner, namely by forming teams based on who can work effectively together. The *effectiveness* of a team is described in management literature in terms of two aspects: *performance and viability* (Sunstrom et al., 1990). Viability "entails members' satisfaction, participation, and willingness to continue working together" (Sunstrom et al., 1990). Performance and viability issues clearly correspond to the dual problems of productivity and human resources costs, which are described elsewhere (Cherlin, 1981; Boehm, 1981; DeMarco and Lister, 1987). This investigation focuses on the performance aspect of team effectiveness, measuring and comparing completion times to solve programming problems.

2. Background and motivation

This research project began with the following informal observation, "Why is it that a team of very gifted individual programmers doesn't necessarily make a *great* team?" This led us to investigate the software development process in non-traditional, non-computer science ways. Such interdisciplinary research in computer science has been investigated for decades (Weinberg, 1971; Shneiderman, 1980). In the 1980s, the realization that software could not keep pace with hardware improvements "focus(ed) attention on human factors in the process of system development as well as the performance of the end user of computer systems" (Nichols, 1982).

This investigation takes such an interdisciplinary approach to address the question of how to assess the effectiveness of software development teams by examining teams from the perspective of various disciplines, much the same way as Curtis has encouraged (Curtis et al., 1986). The approach taken for this investigation applies *psychological* approaches previously used in *management* research to the *computer science* discipline.

* Corresponding author. Tel.: +1 540 231 6931; fax: +1 540 231 6075; e-mail: henry@cs.vt.edu.

0164-1212/99/$ – see front matter
PII: S0164-1212(98)10060-2

2.1. Psychology background

This investigation began with an interest in personality characteristics and their effects on software teams, starting with an interest in the Myers-Briggs Type Indicator (MBTI) (Myers, 1987), which ultimately plays only a minor role in the overall research investigation. As the research progressed, interest shifted to a different approach that proved to be more effective, something more directly applicable to investigating *teams*. What was discovered in the literature on team evaluation was Belbin's roles (Belbin, 1981, 1993). Therefore, this section presents the basic concepts of roles from the psychology discipline, after which is an outline of Belbin's specific roles that are used in this study.

Role theory, functions of roles, and role complements are important components of this research. *Role theory* is defined in different ways by various authors; for example Biddle, from which much of this discussion is drawn, describes it as "behaviors characteristic of one or more persons in a context" (Biddle, 1979). Sarbin takes a particularly applicable view of role theory because he specifically includes personality in his description: role theory "is an interdisciplinary theory in that its variables are drawn from studies of culture, society, and personality" (Sarbin, 1954). One approach in role theory concerns the idea that "actions are patterned into coherent and ordered systems that govern both interpersonal interaction and societal functioning" (Knowles, 1982), which introduces the concept of functions. From these differing views of roles, one can see that a roles can therefore be viewed as a tripartite: a position, behaviors associated with the position, and a set of personality characteristics associated with those behaviors. *Positions* such as a manager, programmer, or system administrator are expected to engage in certain *behaviors* (for example, direct people, provide information, keep track of information), by virtue of being in that position (Jeffrey and Putman, 1994). In order to effectively carry out those behaviors, an individual must display certain *personality characteristics* (for example, introverted, extroverted, confrontational).

The characteristic behaviors mentioned above can be defined as *functions* associated with a given role. Biddle states that role functions are a "behavioral repertoire characteristic of a person or position" (Biddle, 1979). He further states that roles can be viewed in terms of their consequences, which are their "characteristic effects, or functions, within a societal system" (Biddle, 1979).

Two other concepts important to this research are *functional requisite* and *role complement*, which are essentially the foundation of Belbin's work (Belbin, 1981), presented in Section 2.2. Functional requisites are "functions that are necessary for the survival or maintenance of a social system" (Biddle, 1979). For this investigation we are interested in the necessity of these functions for the success or effectiveness of a team. A role complement is a set of roles defined for a given context, for example a programming team. These two concepts combine to form the functional requisites that are embodied by Belbin's set of roles, which is an example of a complement of roles.

2.2. Management background: Belbin's roles

The significance of this work lies in the applicability to the individuals who form teams: managers. Hence, management application to forming or evaluating teams is the focus of the work, and therefore it is necessary to present some background information from the management discipline. The following section presents an overview of Belbin's work, from the management literature, that is used in this empirical study.

R. Meredith Belbin conducted a series of experiments that produced results that are the very foundation of this research. His results consist of a model of *management teams* based on a set of roles that need to be present for a management team to be successful, in other words a role complement. He started with a simple idea that different types of people interact in different ways. Initially, he investigated teams composed of members who were all very similar. Then, he conceived of various other types through extensive observation of teams working together. For example, one such type was labeled a "Plant" because this type of team member appeared to "sit in the corner" and not interact a lot, like a house plant, but presented very innovative plans and ideas when he or she did interact. Eventually, through observation, hypothesis testing, and experimentation, Belbin defined eight roles based on these observed "types" that he felt were necessary for a team to be successful: Chairman, Shaper, Plant, Monitor-Evaluator, Resource Investigator, Team Worker, Company Worker, and Completer-Finisher. His work is described in Belbin (1981, 1993).

Belbin has a different view of roles than the traditional one described in Section 2.1. He believes that team members have two types of roles. The first is functional, as described above as a tripartite. The second is a team role, which are in fact the set of roles described in the previous paragraph. We are interested in investigating how these types of roles affect team performance. For a particular individual, the first type of role might be a typist on a programming team, whereas the second type might indicate that the individual is a Company Worker or a Team Worker. The team role describes how the individual fits into the team, not what particular function he or she performs.

Table 1 provides a brief description of all of the Belbin roles in terms of their team role behaviors and personality characteristics. Most of the table is directly from Belbin's original presentation of the material

Table 1
Descriptions of Belbin roles

Name	Symbol	Behavioral description	Typical features	Positive qualities	Allowable weaknesses
Chairman	CH	Guiding and controlling leader, knows the members' abilities well	Calm, self-confident, controlled	A capacity for treating and welcoming all potential contributors on their merits and without prejudice. Strong sense of objectiveness	No more than ordinary in terms of intellect or creative ability
Shaper	SH	Demanding, coercing, confrontational leader, pushes for members to excel	Highly strung	Drive and a readiness to challenge inertia, ineffectiveness, complacency or self-deception	Proneness to provocation, irritation and impatience
Plant	PL	Innovator and problem solver, the "idea" member	Individualistic, serious-minded, unorthodox	Genius, imagination, intellect, knowledge	Up in the clouds, inclined to disregard practical details or protocol
Resource Investigator	RI	Contact person for resources external to the team, brings resources into the team	Extroverted, en-thusiastic, curious, communicative	A capacity for contacting people and exploring anything new. An ability to respond to challenge	Liable to lose interest once the initial fascination has passed
Monitor-Evaluator	ME	Analyzes, evaluates proposed solutions and choices	Sober, unemotional, prudent	Judgement, discretion, hard-headedness	Lacks inspiration or the ability to motivate others
Company Worker	CW	Implements agreed upon plans	Conservative, dutiful, predictable	Organizing ability, practical common sense, hard-working, self-discipline	Lack of flexibility, un-responsiveness to unproven ideas
Team Worker	TW	Facilitates team functions, mediates issues within the team	Socially oriented, mild, sensitive	Ability to respond to people and to situations, and to promote team spirit	Indecisiveness at moments of crisis
Completer-Finisher	CF	Focuses on details and meeting deadlines	Painstaking, orderly, conscientious, anxious	A capacity for follow-through, perfectionism	A tendency to worry about small things, a reluctance to "let go"

(Belbin, 1981). The *Shaper* is central to this experiment, and it is described here; the remainder of the roles are described in Appendix A. The Shaper is one type of leader. A Shaper is a slave driver, questioning members to find the best approaches to problems. This role leads the team by stimulating the members to "challenge inertia, ineffectiveness, and complacency". They are typically confrontational. Shapers tend to be nervous, extroverted, competitive, and argumentative, just to name a few of their stronger characteristics.

While these roles may be interesting, the question arises "How can one determine the suitability of individuals for these roles?" Belbin developed a questionnaire as part of his investigations, which can be found in Belbin (1981). It is called the Belbin Self-Perception Inventory. This is the instrument used to gather data from individuals to analyze the teams. Similar to psychometrics such as the MBTI, this test provides indicators of an individual's natural propensity towards filling each role. One aspect of the test that affects how teams are set up concerns the fact that the numbers produced by the test are relative, not absolute. The test consists of several sections; for each section an individual distributes ten (10) points among eight (8) statements, based on how strongly they feel about each statement. Table 2 shows one section of the seven section test.

2.3. Using student programming teams

The subjects in the controlled experiments that form the core of this paper are senior college students. Other investigations have looked at student programming teams. Henry, for one, evaluates various methods used to set up programming teams for collegiate senior-level software engineering classes: random, letting them set

Table 2
Sample section of questionnaire for Belbin roles

Section 2: If I have a possible shortcoming in teamwork, it could be that:
1. I am not at ease unless meetings are well structured and controlled and generally well conducted.
2. I am inclined to be too generous towards others who have a valid viewpoint that has not been given proper airing.
3. I have a tendency to talk too much once the group gets on to new ideas.
4. My objective outlook makes it difficult for me to join in readily and enthusiastically with colleagues.
5. I am sometimes seen as forceful and authoritarian if there is a need to get something done.
6. I find it difficult to lead from the front, perhaps because I am over-responsive to group atmosphere.
7. I am apt to get caught up in ideas that occur to me and so lose track of what is happening.
8. My colleagues tend to see me as worrying unnecessarily over detail and the possibility that things may go wrong.

up their own, grade point average, etc. (Henry, 1983). She presents a heuristic for establishing teams that is based on amount of free time, schedule conflicts, and grade point average that achieves relatively equivalent teams. Scott and Cross also discuss issues in setting up student programming teams in an effort to make them relatively equivalent (Scott and Cross, 1995). Some of the issues that they present are academic performance (grades), team and project size, and, interestingly enough, psychological profiles including the MBTI. Unfortunately, their treatment is very superficial; they only consider psychological issues such as a team needing an introvert and an extrovert because the class requires written and oral presentations, and *clearly*, they conclude, "an Extrovert will be more comfortable with an oral presentation, while an Introvert may produce a better written report" (Scott and Cross, 1995).

Finally, the issue of using student programmers in experiments introduces the concern that such research does not directly apply to industry programmers. Holt et al. demonstrate that advanced students and professional programmers are statistically similar in terms of comparing their mental representation and various performance measures (Holt et al., 1987). This provides support for using students in studies, especially for investigations where an industry validation is to be done, which is true for this research.

2.4. Management team research applied to software teams

Numerous studies have been conducted on software development teams, e.g. Jeffery (1987), Mills (1983), Bostrom and Kaiser (1981), Shneiderman (1980), Trower and Straub (1991), von Mayrhauser (1984) and Walz et al. (1987) and Zahniser (1990). Some of this work has management overtones, others do not. One investigation by Rob Thomsett is of particular interest for this investigation because it applies Belbin's roles to software developers (Thomsett, 1990). His study presents a qualitative analysis of software teams in Australia using three measurement instruments for the three different models: the MBTI, the Belbin Self-Perception Inventory, and the Job Diagnostic Survey (JDS).

Thomsett's presentation includes the following in terms of results for his study: For one company that participated in Thomsett's study, "Immediate productivity increases of 200 percent (were) reported by the senior management of the computing group" (Thomsett, 1990). Thomsett describes some intuitive reasons that the Belbin roles, as well as the MBTI types, are distributed the way they are in his data: a 'cloning' effect that means that managers hire employees that are similar to themselves. Further, he observes that "at best because of the relative lack of Team Worker ... and effective Chairpersons ..., computing teams and managers generally lack the required interpersonal skills," thus making the teams less effective than they could be.

Unfortunately, there are some deficiencies with this investigation; the analysis is completely qualitative and is not quantitatively substantiated. Therefore, Thomsett presents an interesting study with some initial findings that needs further investigation. He presents general information that appears very useful but has only intuitive supporting arguments and no empirical evidence. Quantitative data needs to be examined in order to support his ideas.

3. The controlled experiments

This paper is concerned with a single experiment that is part of a much larger research investigation. The fundamental issue that the overall project addresses is how to improve software development teams using knowledge of role theory and personality characteristics. This includes analyzing extant teams and forming new teams. This first phase of the project consists of a controlled experiment in a software engineering laboratory class in which participants are arranged on teams to accept or reject a specific hypothesis that tests the importance of a selected Belbin role on teams.

The importance of this research lies in the cause-and-effect relationships that can be used to improve extant teams or to form good teams. The effects are observed when some teams are specifically designed to be better than other teams in terms of team effectiveness. The

primary aspect of effectiveness that is of interest for these experiments is performance. For a quantitative analysis, some objective measure of performance needs to be used to provide a "measure of success." In this research, that measure is the time to *correctly complete* a programming problem, viz. the time-to-completion of *correct* solutions. The mean completion time of the teams anticipated to perform well are compared to the mean completion time of the teams anticipated *not* to perform well. A statistical comparison of these means determines whether there is a significant difference between them. One could certainly argue that time-to-completion is a small measure of success, but this measure was chosen because it is objective, quantitative, and easy to collect. A solution was not accepted until it was correct, so no measure of correctness was possible, which would have the deficiency of being subjective. Similarly, for rating the "quality" of the code, such a measure would be completely subjective. Thus, this part of the research demonstrates that teams that contain the specified roles perform better, i.e. finish sooner on the average, than teams that do not possess those roles.

The specific purpose of these experiments is to produce empirical results to support some of the concepts central to the research. Several hypotheses were proposed to focus the research on issues that could be directly tested. The hypothesis that was ultimately selected investigates the importance of a team containing one leader on a team. Specifically, the null and alternate hypothesis for a statistical test are:

H_0: Teams containing one and only one "leader" perform equivalently to teams with no leader or multiple leaders

H_1: Teams containing one and only one "leader" *do not* perform equivalently to teams with no leader or multiple leaders.

As mentioned in Section 2.3, the participants in this experiment are senior college students majoring in computer science, having previously had courses requiring major development projects. They typically include the top of the class and are the highly motivated students. Most of them have had internships or coop positions in industry or have done part time programming for companies during the school year as well as summers. Also, the course has a distinct reputation for being an extremely time-consuming course and for being very impressive to potential employers.

3.1. Establishing teams and groups

To understand how the experiments were conducted, the terms *team* and *group* must be identified as distinct. Each *team* was composed of three members who indicate certain role characteristics under observation on the Belbin Self-Perception Inventory. Each *team* was placed in a *group* that was being used to test the hypotheses in question. Therefore, each team was established to be a "Leader" team or a "No Leader" team in terms of how it was expected to perform in the experiment. In other words, some teams were established with a single, identifiable leader and were in the successful, "Leader" group; other teams were established with no leaders or all leaders and were part of the unsuccessful, "No Leader" group.

Using a table like Table 3 but including all of the teams, the participants were formed into teams and groups. Approximately half of the teams were placed in a group expected to be successful, i.e. quicker to complete a programming problem; the other teams were formed with the anticipation that they would perform worse than the successful teams in terms of the mean *completion time* of the groups.

In order to ensure that the role under scrutiny was the reason the groups performed differently, the other roles were made as equivalent as possible across the teams, and the team members' self-rated experience was used as a blocking factor.

As shown in Table 3, the table used to accomplish this team formation consisted of the following information. Each row under a team, e.g. Team X, contains the information for one team member. The first column designates the team names. The second column shows members' numbers. The subsequent *columns* represent the members' data for the Belbin roles: Chairman (CH), Shaper (SH), Plant (PL), Resource Investigator (RI), Monitor-Evaluator (ME), Company Worker (CW), Team Worker (TW), and Completer-Finisher (CF). The final column shows the members' self-rated experience on a one to five (1–5) scale, low to high, which is used to make the teams equivalent, mitigating the chances that the successful teams simply contained better programmers.

The data in each cell in the table represents that members' score on the Belbin Self-Perception Inventory, which indicates the Belbin roles that the members have a

Table 3
Example of data for a single team

Team	Member	CH	SH	PL	RI	ME	CW	TW	CF	Exp
Team X										
	1	c1	–	–	–	–	W2	–	F9	5
	2	–	s8	p2	–	m0	–	–	–	4
	3	–	s0	p7	–	m5	w0	–	–	4

propensity toward. Cells with a dash (–) in them indicate that the score was not significantly high enough to take into account when forming the teams. One should note that if an individual takes the test and distributes the points as evenly as possible, each cell would have the average of 8.75; therefore, a score of 9 or less is insignificant, and a 10 is weak. Each cell contains a letter and a digit. The letter represents the role, which is also shown by the column headers abbreviations, and the digit indicates the score on the Belbin test. Scores in the range 10–19 are indicated with a lower case letter for the role, e.g. c3 and p2 indicate a Chairman score of 13 and a Plant score of 12; scores of 20 and above are indicated with a capital letter, e.g. W2 and F9 indicate a Company Worker score of 22 and a Completer-Finisher score of 29, which are very high scores. This scheme is used because it removes a lot of extraneous information, making comprehensible the enormous amount of data that is needed to form the teams evenly (except for the leadership role under study).

To determine each member's potential roles, several factors need to be taken into consideration:

1. An individual fills more than one role by having significant scores for multiple roles, typically two or three roles.
2. Although an individual might have a number that appears to make him or her fill that role, he or she may not fill that role because he or she has a stronger tendency to fill other roles. For example, in Table 3, member 1 probably does not fill the Chairman role.
3. Other fellow members can keep an individual from filling a role in the case that the fellow member is stronger in the role. This is particularly true if the member has other roles that are stronger. For example, member 3 in Table 3 would not fill the Shaper role because of his or her other stronger roles as well as member 2 being such a strong Shaper.
4. Because the scores are relative, not absolute, sometimes a score of ten (10) or eleven (11) is significant, sometimes not. If a person has "high" scores of 12 in two roles and 10 in another role, then the 10 would be considered significant, because the individual has a tendency to assign a few points to all of the statements in the test. Member 1 (Table 3) is a good example of when a low score of 11 is not significant.

Part of the basic assumption of this research is that some roles can conflict within a team, such that a team is "better" with only one of that role type on the team. This appears to be true for certain roles, such as Shaper, where conflicts can occur. Other roles appear not to demonstrate this effect, such as Plants, where the team improves with more members who are innovative.

It is very important to note that the results of this test indicate individuals who have a natural propensity to these roles. This is particularly important for the Shaper leadership role that is the focus of this paper. Teams were established with no, one, or multiple leaders based on the results of Belbin tests taken by the team members. The teams *were not* formed based on previous leadership experience or any other factors, and no member was designated as the leader for each team. The leadership role on each team had to be assumed (or not assumed in the case of unsuccessful teams) by one of the members. The participants had no knowledge of the Belbin roles nor of any particular leadership qualities that members might possess unless they discussed it amongst themselves.

3.2. Conducting the experiments

The teams and groups were formed prior to the commencement of the experiments (see Table 4), and the participants met in a laboratory where there were nine computers used in the experiment. The twenty-four (24) student participants had been formed into eight teams, and each team acquired one computer apiece. One computer was used by the experimenter to test and accept problem solutions. Each participant was given a copy of the problem to be solved that day. The problems were intended to be solved in a little over an hour, although some teams ended up requiring up to two hours. In order to make everything as balanced as possible among the teams, each team could only work on one workstation and were told that they could use whatever editors, C or C++ compilers, or utilities they wished. The systems being used were running Free BSD 1.1.6. All of the facilities on the machines were equivalent. Teams e-mailed their solutions to the experimenter as soon as they felt that they had a correct solution, and the experimenter tested the solutions with his own acceptance data, which was not shown to the participants. A completion time for each team was only recorded once a submitted solution tested *correctly* with the acceptance data. Although the problem statements included some test data, the teams were informed that they must create test data themselves in order to have their solution pass all of the acceptance tests.

The experiment consisted of four problems that were solved in four separate lab sessions. During the four sessions, the teams were given one problem at a time to complete during that lab session. The means of all four problems for each group was calculated and compared in the following analysis.

The two Belbin roles Chairman and Shaper are both types of leaders; they *can* be complementary roles, where the Chairman observes the team, knows the strengths and weaknesses of each member (and hence the team), and knows how to take maximal advantage of the team's human resources. The Shaper plays more of an active part in the team, pushing the team members by questioning opinions and decisions, making members take extra effort in their work because a Shaper is very

Table 4
Teams used in the leadership experiment

Team	Member	CH	SH	PL	RI	ME	CW	TW	CF	Exp
Team D										
	1	c3	–	–	–	–	W2	–	f9	5
	2	–	s4	p2	–	m0	–	–	–	4
	3	–	s3	p7	–	m3	w0	–	–	4
Team F										
	1	–	s1	–	–	m4	–	–	–	5
	2	c0	–	p0	–	–	w0	t4	–	5
	3	–	–	–	–	M0	W5	T0	–	2
Team H										
	1	–	s1	P0	r2	–	–	–	–	4
	2	–	–	P0	–	–	w1	–	–	4
	3	–	–	–	–	m1	w1	t8	–	4
Team G										
	1	c0	–	–	–	m4	w1	–	–	4
	2	–	–	p0	–	–	w4	t3	f0	3
	3	–	–	–	–	–	w9	t5	–	6
Team E										
	1	–	s0	p3	–	–	w1	t6	–	2
	2	–	–	–	–	–	w4	–	f1	5
	3	–	–	p3	r0	m1	w4	–	–	5
Team C										
	1	–	s2	–	–	–	w2	–	–	3
	2	c0	s4	–	–	–	w1	–	f2	6
	3	c4	–	–	–	–	w5	t4	f2	5
Team B										
	1	–	s2	–	r2	–	w0	–	–	5
	2	–	s0	p3	–	–	w4	–	–	5
	3	c0	s2	–	r2	m2	–	–	–	3
Team A										
	1	–	s0	P2	–	–	–	–	f5	5
	2	–	s3	–	–	m6	w7	–	f0	6
	3	–	s9	–	–	m3	w3	–	–	2

motivating, in a confrontational way. Chairmen tend to be introverted and thoughtful; Shapers tend to be extroverted and motivating, focusing members to extend their capabilities. "The 'Chairman' of (three-person) team(s) would be something of a misnomer" (Belbin, 1981) and does not have a significant impact on the team because coordinating the resources of two subordinates is not significant. Therefore, this experiment focuses on the Shaper types of leaders. Further cause for not considering Chairman leaders is the virtual dearth of Chairman types in the data in the study.

There are two items that should be noted regarding how the experiments were conducted: the use of one workstation per team and the focus on the Shaper type of leader. Although forcing each team to use only one workstation may *appear* to favor single-leader teams, this was not the intention of the experimenters. The intent was to force individuals to work as a team, instead of simply investigating groups of individuals; if the members were allowed to separate and work independently, then the experiment would simply measure which member on each team could solve the problem first. This could be considered a limitation of the experiment. As stated above, due to a scarcity of Chairman types in the participant population, this experiment focuses on the Shaper type of leader. This is definitely a limitation of this experiment; the Chairman role also requires investigation, as do all of the other Belbin roles.

3.3. Quantitative analysis

The analysis of the controlled experiments focuses on the first part of the definition of the effectiveness of a team, i.e. performance. The experiment demonstrates that teams with a single, recognized leader perform better than those with no leader. The "Leader" teams were established with a single, recognized leader on each team with the other factors roughly equivalent; the "No Leader" teams consisted of teams that either had *no* leader type present *or* had no *clear leader*, i.e. 2 or all 3 were leader types, which caused conflicts over who was in charge and which "direction" the team should pursue. The basic hypothesis is that the mean of the completion times for the teams in the "Leader" group is significantly lower than for the "No Leader" group. In other words

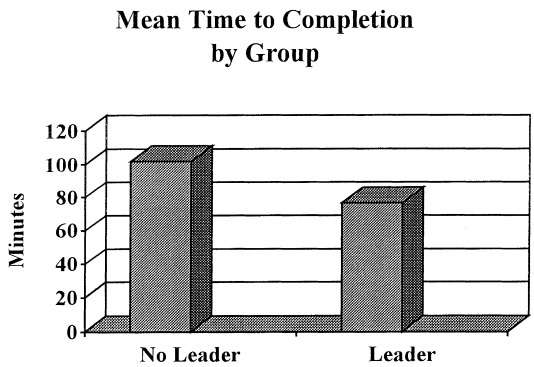

Fig. 1. Means with Team G in No-Leader Group.

on the average, leaderless teams should take longer to solve the problems.

After transforming the data with a log function to form a linear model for statistical analysis, the data showed a significant difference between the groups. In other words, as shown in Fig. 1, teams with leaders solved the programming problems in 75.27% of the time needed by the leaderless teams. Specifically, an ANOVA using Proc GLM [1] in SAS (Littell et al., 1991) showed that the two groups are statistically different ($p = 0.0068$); the means are 101.81 and 76.63 min for the two groups, as shown in Fig. 1. Therefore, the null hypothesis is rejected and the alternate hypothesis is accepted; the groups are not equivalent. Because the groups are not equivalent and the mean completion time of the "Leader" group is better, this is interpreted as teams with a single, recognized leader perform better than the teams with no clear leader.

The "Leader" group consists of Teams D, F, and H (Table 4) and contain a single leader. These teams form the "successful" group. Teams A, B, and C contain multiple leader roles, who tend to introduce conflict into the teams composition, reducing the effectiveness of the team. Teams E and G have no obvious leader who can be identified by Belbin's roles, although the c0 member of Team G might be interpreted as a leader. The information in Fig. 1 was generated using Teams A, B, C, E, and G as the "No Leader" teams. These include both the no-leader teams as well as the multiple leader teams. As it turns out, the individual on team G with a Chairman score of 10, i.e. a c0, acted as the leader of the team. This can be explained, post hoc, by the following.

1. The individual had a weak measure of the leadership role.
2. The other two roles that the individual should have filled were not sufficiently stronger than the Chairman role.
3. The Chairman role needed to be filled.

[1] GLM means General Linear Model, which is a standard statistical procedure.

Typically when examining the roles in order to set up the teams, only the top two roles were used for each individual. Fig. 2 shows the results of an analysis of the data putting Team G in the "Leader" group, which had a single leader, although it was not obvious from the Belbin data. This analysis has even better results, as shown by a p value of 0.0034 in an ANOVA, with means of 100.90 and 69.75 min.

Some subjective observations by the experimenter were noted both during and immediately after each lab session. Some notes about the teams that were predicted to perform well include: "The team obviously works well together". "One person (is) at the helm, with the other two supporting technically, non-passively; the leader listens fairly well but is not the bright one." "The leader is bright and doesn't listen well but listens well enough to get the good ideas from the other members; he tends to become almost too focused sometimes." By the fourth (and last) problem for this experiment, Team F said they "had it down" in terms of coordinating together. Some comments from and about the "unsuccessful" teams include: "...very separated group," "Two dominant individuals with the third being very passive; one member takes the leadership role but only to argue with the other dominant member over everything, although the leader did make attempts to listen to the quite one." By the fourth (and last) problem Team E had "completely given up, apparently feeling that their team had been 'set up' not to work well together."

4. Conclusions and continuing work

Obviously, the conclusion of this study is the alternate hypothesis stated above: Teams containing one and only one "leader" perform better than teams with no leader or multiple leaders, as indicated by the Shaper role. Intuitively, this appears obvious, however, controlled experiments are necessary to support the idea. As this research continues, more definitive statements will be verified. This research provides guidance to managers

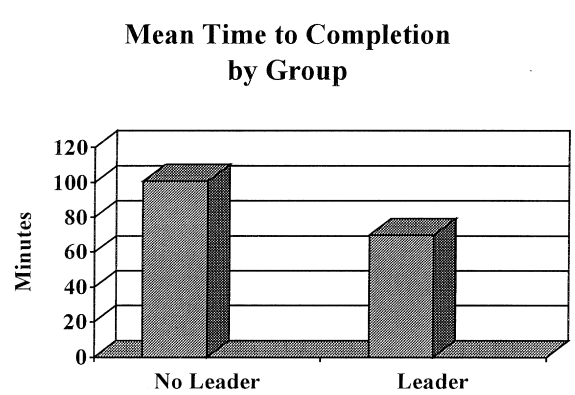

Fig. 2. Means with Team G in the Leader Group.

in forming successful teams as well as in evaluating extant teams to identify deficiencies, with the intent of rectifying such deficiencies. It should be noted that this was a limited controlled study. It certainly implies useful information, but we did not study extended-time performance, such as on a typical project that take several months or years of development. Results may be different for long term projects simply because human interactions over an extended period tend to be different than for short term projects. For example over time, people act differently in terms of what they are willing to do, go along with, and put up with.

Although some managers *might* use personality information directly in that they want to hire "computer science" types, this research shows that team improvement can be better accomplished by combining personality information with the concepts of roles and the need to have diverse teams. Team roles need to be the focus of this information, and the roles need to be filled with the appropriate individual. The above analysis demonstrates the utility of this approach to forming or evaluating software development teams. Further, if teams are properly formed using roles, then the team members are happier and more likely to remain with an employer, thus increasing the viability of the team. All in all, fulfilling the goals of this research will produce a model that can be used to improve both the performance and the viability of teams and, therefore, team effectiveness.

Appendix A. Description of the Belbin roles

A *Chairman* is an individual who controls the team in a typical head-of-the-table manner. This role guides the team towards what he or she perceives as the teams objectives in the best manner he or she can determine. A Chairman is calm and listens to other team members well, having a natural ability to get the most and best out of everyone's potential. Additionally, he or she is very aware of the strengths and weaknesses of the team as a whole. The Chairman is one type of team leader.

The second role, *Shaper*, is also a leader type but has a completely different personality and managerial style than a Chairman. A Shaper is a slave driver, questioning members to find the best approaches to problems. This role leads the team by stimulating the members to "challenge inertia, ineffectiveness, and complacency". Shapers tend to be nervous, extroverted, competitive, and argumentative, just to name a few of their stronger characteristics.

The question arises of who is leader of the team if individuals of both roles are present in a team. The answer is that either they cooperate in their leadership, trading off at designated appropriate times, or they produce a great deal of conflict in the team and can make the team less successful. Ideally, both roles are present on a team, and they complement each other. The Chairman of the team tries to keep progress moving in a steady, conservative fashion, and the Shaper likes to rile things up to make sure that all possibilities are being considered. For each of the eight roles, a similar relationship exists – there is a complement of some sort for each role.

The third role is the *Plant*, who is the innovator of the team; he or she advances new approaches and ideas with special attention to major issues. A Plant is typically introverted, unorthodox, imaginative, and intelligent but "inclined to disregard practical details or protocols". The Plants are the brainchildren who must be nurtured and occasionally drawn back into the real world because they tend to have their heads in the clouds. They are considered one of the intellectual types in a team.

The complementary role to the Plant is the *Monitor-Evaluator*, the other intellectual. The Monitor-Evaluator is the analyzer of the team; the member who evaluates all of the alternatives for all decisions, methods, and approaches so that the team is positioned as competitively as possible. This role can be under-appreciated because members that fill this role tend to be unemotional, dry, over-critical, hard-headed, and unmotivational to others; this is terribly unfortunate because the role is necessary for the team to succeed. The Monitor-Evaluator plays counterbalance to the Plant: Plants come up with pie-eyed ideas and the Monitor-Evaluator is the only type of role that can debate successfully with a Plant to make correct choices.

The fifth role is *Resource Investigator*, who is considered one of the negotiator types of roles. In one aspect, Resource Investigators are very similar to Plants, although not considered complements. Both of these roles are seen as creative members who bring innovation to the team. The difference is that Resource Investigators get their innovations from sources external to the team, primarily due to their extroverted nature. Other attributes associated with this role are enthusiasm, curiosity, and communication skills. These members also have a tremendous "capacity for contacting people and exploring anything new;" unfortunately, they tend to lose interest in a problem or situation once the novelty of it wears off. Their strength lies in that they know how to get what they want out of a person.

The complementary negotiator role is the *Team Worker*, who makes sure that the team works together towards their goals. Whereas the Resource Investigator negotiates outside of the team to get what the team needs, the Team Worker facilitates or negotiates within the team. For example, no matter how brilliant a Plant may be or how successful a team leader may be, team members need to get along in order for the team to be effective or successful. Further, members may get

irritated by the slow, thoughtful decision-making process of the Monitor-Evaluator; members may not let the introverted Plant express herself or himself. Therefore, some member role is necessary to make sure that everyone gets along and is provided the opportunity to accomplish their functions. Team Workers tend to be very socially oriented, mild, and sensitive; they have a strong ability to respond, communicate, and deal with people and situations, the ultimate facilitator.

The Seventh role that Belbin discusses is the *Company Worker*, who is the meat-and-potatoes member of the team. Company Workers' primary function is "turning concepts and plans into practical working procedures"; they focus on "carrying out agreed (upon) plans systematically and efficiently". Company Workers tend to be conservative, dutiful, predictable, good organizers, hard-working, and self-disciplined. On the negative side, they are inflexible and resistant to unproven ideas, views, and approaches; they support the status quo a bit extremely. Although Company Workers tend to exhibit some negative aspects, one should bear in mind that negative aspects are part of what defines the individual who can fulfill the role successfully.

The final role, *Completer-Finisher*, is the complement of the Company Worker in a manager-worker relation. Completer-Finishers focus on a project delivering on-time and within cost estimates; the Company Worker carries out work to accomplish this. Primarily, team members in this role focus on avoiding "mistakes of both omission and commission, actively searching for aspects of work (that) need a more than usual degree of attention". They stress a sense of urgency to the team in order to meet deadlines and goals. Team members who fill this role are painstaking, orderly, conscientious, and anxious; they demonstrate a strong sense of follow-through and perfectionism.

References

Belbin, R.M., 1981. Management Teams. Wiley, New York.

Belbin, M., 1993. Team Roles at Work. Butterworth-Heinemann, Oxford.

Biddle, B.J., 1979. Role Theory: Expectations, Identities, and Behaviors. Academic Press, New York.

Boehm, B.W., 1981. Software Engineering Economics. Prentice-Hall, Englewood Cliffs, NJ.

Bostrom, R.P., Kaiser, K.M., 1981. Personality differences within systems project teams: Implications for designing solving centers. In: Proceedings of the Eighteenth Annual Computer Personnel Research Conference, 3–5 June. ACM (Assoc. Comput. Machinery), New York, pp. 248–285.

Cherlin, M., 1981. The Toll of Turnover. Datamation (April), 208–212.

Curtis, B., Soloway, E.M., Brooks, R.E., Black, J.B., Ehrlich, K., Rudy Ramsey, H., 1986. Software psychology: The need for an interdisciplinary program. Proc. IEEE (August), 1092–1106.

DeMarco, T., Lister, T., 1987. Peopleware: Productive Projects and Teams. Dorset, New York.

Henry, S.M., 1983. A project oriented course on software engineering. In: Proceedings of the 1983 ACM SIGCSE Symposium on Computer Science Education, February, pp. 57–61.

Holt, R.W., Boehm-Davis, D.A., Schultz, A.C., 1987. Mental representations of programs for student and professional programmers. In: Olson, G.M., Sheppard, S., Soloway, E. (Eds.), Empirical Studies of Programmers: Second Workshop, Washington, D.C., 7–8 December, Ablex, Norwood, NJ, pp. 33–46.

Jeffery, D.R., 1987. The relationship between team size, experience, and attitudes and software development productivity. In: Proceedings of COMPSAC87, Tokyo, Japan, 7–9 October. IEEE Computer Soc. Press, Silver Spring, MD, pp. 2–8.

Jeffrey, H.J., Putman, A.O., 1994. Relationship definition and management: Tools for requirements analysis. Journal of Systems and Software (March), 277–294.

Knowles, E.S., 1982. From individuals to group members: A dialectic for the social sciences. In: Ickes, W., Knowles, E.S. (Eds.), Personality, Roles and Social Behavior. Springer, New York.

Littell, R.C., Freund, R.J., Spector, P.C., 1991. SAS System for Linear Models, 3rd edition. SAS Institute, Cary, North Carolina.

Mills, H.D., 1983. Software Productivity. Little, Brown and Company, Boston.

Myers, I.B., 1987. Introduction to Type. Consulting Psychologists Press, Palo Alto, CA.

Nichols, J.A., 1982. Forward. In: Proceedings of Human Factors in Computer Systems, Gaithersburg, Maryland, 15–17 March. ACM (Assoc. Comput. Machinery Press), New York, p. iii.

Sarbin, T.R., 1954. Role Theory. In: Gardner Lindzey (Ed.), Handbook of Social Psychology. Addison-Wesley, Reading, MA.

Scott, T.J., Cross II, J.H., 1995. Team selection methods for student programming teams. In: Software Engineering Education: Proceedings of Eighth SEI CSEE Conference, New Orleans, Louisiana, 29 March–1 April. Springer, Berlin, pp. 295–303.

Shneiderman, B., 1980. Software Psychology. Winthrop Publishers, Cambridge, MA.

Sunstrom, E., De Meuse, K.P., Futrell, D., 1990. Work Teams. American Psychologist (February), 120–133.

Thomsett, R., 1990. Effective Project Teams. American Programmer (July/August), 25–35.

Trower, J.K., Straub Jr., D.W., 1991. Improving the performance of technologists and users on interdisciplinary teams: An analysis of information systems project teams. Computer Personnel (November), 62–72.

von Mayrhauser, A., 1984. Task analysis and the selection of software development team structures. In: Proceedings of COMPSAC84, Chicago, IL, 7–9 November. IEEE Computer Soc. Press, Silver Spring, MD, pp. 120–126.

Walz, D.B., Elam, J.J., Krasner, H., Curtis, B., 1987. A methodology for studying software design teams: An investigation of conflict behaviors in the requirements definition phase. In: Olson, G.M., Sheppard, S., Soloway, E. (Eds.), Empirical Studies of Programmers: Second Workshop, Washington, D.C., 7–8 December. Ablex, Norwood, NJ, pp. 83–99.

Weinberg, G.M., 1971. The Psychology of Computer Programming. Van Nostrand Reinhold, New York.

Zahniser, R.A., 1990. Building software in groups. American Programmer (July/August), 50–56.