Teaming Assessment: Is There a Connection Between Process and Product?

Richard L. Upchurch and Judith E. Sims-Knight

Abstract - It is reasonable to suspect that team process influences the way students work, the quality of their learning and the excellence of their product. This study addresses the relations between team process variables on the one hand, and behaviors and outcomes, on the other. We measured teeming skill, project behavior and performance, and project product grades. We found that knowledge of team process predicts team behavior, but that knowledge alone does not predict performance on the project. Second, both effort and team skills, as assessed by peers, were related to performance. Third, team skills did not correlate with the students’ effort. This pattern of results suggests that instructors should address issues of teaming and of effort separately. It also suggests that peer ratings of teammates tap aspects of team behavior relevant to project performance, whereas declarative knowledge of team process does not.

Index Terms – Team Assessment, Software Engineering Education.

INTRODUCTION

The ability to work as an effective member of a development team is a primary goal of engineering education [4]. Industry surveys indicate

“... new graduates do not lack technical skills and scientific preparation. Rather, they lack people and process skills. New hires do not know how to communicate effectively, they have insufficient experience and preparation for working as part of a team, they lack the ability to manage their individual work efficiently and productively, and they do not understand or appreciate organizational structures or business practices.”

[5, p. 45]

Such reports resulted in teaming becoming one of ABET’s EC2000 learning outcomes [6] for engineering programs, including software engineering. To embed the learning outcomes into a continuous improvement loop, as required by EC2000, it is necessary to assess teaming itself. It is not sufficient to organize students into teams for project work, then assume that the skills needed for successful teams will be gained merely by team participation. The best way to assess teaming skills is still in question. One approach is to assess whether students have the basic knowledge of how successful teams operate, as addressed by the Team Knowledge Test (TKT). A second approach is to assess students’ skills as they operate in a team. This has been assessed successfully by peer ratings (e.g., [3][7][8]). It is likely that students whose team members view their teeming skills positively would also have high knowledge of effective teeming and that team skill should be positively related to project performance. It does not, however, follow that knowledge of good team process would also be related to project performance, because it reflects declarative knowledge only and students may be unable to translate declarative knowledge into action.

One of the major issues of teams has been that of social loafing, that is, the tendency for one or more team members to exert little effort, leaving the rest of the team to do their jobs. Thus, it might be that what makes teams effective is that they motivate their members to exert effort and to exert it consistently across the duration of product development. Thus, it is predicted that the peer team ratings will exert their influence on performance through their effects on effort. In this study students recorded their effort weekly, which allows us to explore the relation between teeming skills, effort, and performance.

THE STUDY

The software engineering course at UMass Dartmouth is a required course for majors in Computer Science and Computer Engineering and is taken during their senior year. The content of the course is typical of such courses in the range of topics [9][10]. The intent is to cover the software development lifecycle including practices typically associated with large-scale product development, e.g., configuration management and software quality assurance. The course has both lecture and laboratory components.

The course requires a semester-long team project. The lab for the course supports the project portion of the course, and provides a common time for project team members to meet to plan and organize, as well as meet with the instructor to review progress. Groups of students--in this particular semester we had ten five-person teams, one six-person team, and one four-person team--work together to specify and build a product during the semester. Each project had a customer with a real need. Student teams worked with the customer to determine the requirements. The instructor recruited project managers from the enrolled students. This process involved researching those enrolled, querying faculty to determine a
pool of students with the skills needed to fill this specific role. The project leaders met with the instructor collectively to
dissect their responsibilities and expectations. Once all
agreed to participate, the instructor met with individuals to
determine the project. Having completed the preliminaries,
each project manager received a problem statement typically
described as a statement of work.

Students in the course were asked to prepare a résumé.
The résumés were stored in the course’s electronic portfolio
[11] detailing their experiences and technical qualifications for
a position on a software team. The project managers reviewed
the résumés of those enrolled in the course and prepared a list
of those they wanted for their team. The team selection was a
round-robin process. During the selection activity project
managers were asked to specify why particular people were
chosen. Thus, the project manager had to defend a choice based on project criteria and the contents of the résumé.

Each person on a team was assigned a process role by the
project manager for the duration of the project. The process roles [12] (development manager, planning manager, support
manager, and quality/process manager) were distributed
among team members through negotiations between the team
members and the project manager as were the development
activities, e.g., design and coding. The course required
students to participate in both the process, by fulfilling the
duties of a particular process role, and development, by
engaging in one or more design and implementation activities.

As part of the course requirements, each student was
required to submit to their portfolio a weekly effort log which
detailed the time, in hours, spent on project-related activities.
The effort log was divided into two major sections: 1) process
and 2) product. The process section collected effort data
regarding actions in process related activities, such as their
primary role on the team, whereas the product section
provided a record of their participation in more technical
aspects of product development. The planning manager
collected the weekly logs and maintained the team’s effort
summary in the team’s portfolio.

**Method**

**Participants**

The 60 students in the software engineering course
participated in the study as part of the course requirements.
The participants included thirty-five computer science majors
and twenty-five computer engineering majors, all seniors in
their respective majors. The students were organized into 12
teams as per the process detailed earlier in this paper.

**Measures**

The Teamwork Knowledge Test (TKT) [1][2] was developed
locally and was loosely based on the format used by Stevens
and Campion [13] for their Teamwork KSA Test. The current
assessment, however, was designed for use with an
undergraduate college population rather than an industrial or
corporate population, as was the original. Its 21 items are
designed to sample students' understanding of the four

domains--team process, decision making, communication, and
conflict resolution--covered on the team training website,
which the students were directed to read early in the semester.
The questions were multiple choice questions with one right
answer. The test was web-enabled which allowed students to
take the test via a web browser at their convenience. The test
was administered near the middle of the course, after the
students had reviewed the training material.

Team peer ratings were collected using the scale
developed by Farh, Cannella and Bedeian [3] and given to
students in the developmental context recommended by the
authors, that is, the students were told that the purpose of the
evaluation was to provide feedback to students and would not
be used to grade. The responses can be divided into three
subscales: task performance, group maintenance, and
individual orientation.

Project grades were a combination of process and product
activities. The total possible project score was 400 points,
divided into 200 points for process and 200 for product. The
process score for each student was based on the artifacts
produced as part of the individual’s process role. For instance,
the individual responsible for quality assurance was required
to produce a quality assurance plan, conduct reviews, insure
standards were met, etc. When reviewing the student’s
portfolio the instructor was mindful of the particular process
roles and reviewed items that supported the assigned role. The
product points were based on deliverables the students
submitted to their portfolios indicating what contributions they
made toward the technical development of the product. Since
projects moved from inception to delivery, students would
provide evidence of effort made from requirements elicitation
through implementation and testing. Again it was the student’s
responsibility to provide evidence in each category. The
portfolio was structured in a manner to facilitate the placement
of deliverables in categories consistent with those in which the
material was sought. Students were responsible for organizing
the material in their portfolios so as to highlight the
contributions for both process and product.

The team effort logs were retrieved from each student’s
portfolio and reviewed. The instructor notified both the
planning manager and individual students if effort reporting
was not complete. This tactic resulted in effort data for all
students. The analyses described below are on the weekly
effort data for each student covering a ten-week period
beginning the second week of the course. The first week was
eliminated due to semester/course organizational issues and
preparing students to complete the requisite effort forms.

**Results**

**Team Skills**

The TKT was correlated with the three Farh et al. [3]
subscales (r(58) = .32, p = .02 for task performance, r(58) =
.29, p = .02 for group maintenance, and r(58) = .32, p = .03 for
individual orientation. TKT was not, however, correlated with
project performance, r(58) = .06 for project process scores and
r(58) = .09 for project product scores. This is consistent with
the viewpoint that knowledge of teaming process may facilitate team behavior, but that declarative knowledge alone is not sufficient to influence performance.

The three subscales of the peer ratings were very highly interrelated (see Table I). To decide whether sufficient discrimination existed among the subscales to warrant keeping them, we examined the correlations of the three subscales with process and product scores on the project. All of the correlations were between .30 and .40, except for the correlation of group maintenance with product scores, which was .08, p = .54. Thus, the bulk of the findings suggest that students form a global impression of their teammates’ teaming skill and we summed across the three scales to create one peer team rating.

Effort Data

Figure I provides a distribution of the total effort reported by the students (N=60). The maximum total effort reported was 171.5 hours while the minimum was 39.95 hours for the ten-week period. The mean for total effort was 81.2 hours with a standard deviation of 29.1 hours. Figure II shows the distribution of total effort by team. The mean team effort was 406.2 hours with a standard deviation of 110.6 hours.

Figures IV, V and VI provide the distribution of effort by week by team, with four teams plotted in each figure. The process model predicts that teams who give steady, consistent attention to process will produce better performance than teams whose effort fluctuates. To measure both total amount of effort and consistency, we calculated the first two moments of the effort data. The overall amount of effort was measured by the sum of each participant’s ten effort reports.

TABLE I
INTERCORRELATIONS AMONG THE TEAM PEER RATING SUBSCALES

<table>
<thead>
<tr>
<th></th>
<th>TP</th>
<th>GM</th>
<th>IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task</td>
<td>1</td>
<td>.65*</td>
<td>.70*</td>
</tr>
<tr>
<td>Performance (TP)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>1</td>
<td>.73*</td>
<td></td>
</tr>
<tr>
<td>Maintenance (GM)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Orientation (IO)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Correlation is significant at the .01 level (2-tailed).
Steadiness—or lack thereof—was measured by the standard deviation of each individual’s ten effort reports (see Figure VII). These two variables exhibit substantial correlation with each other, r(58) = .51, p = .001. Yet the 74% unshared variance is substantial enough to test the hypotheses.

We would expect in a process with defined roles to see variability in individual student effort over the ten-week period as specific roles have differing obligations at particular periods during development. Yet, differing levels of individual activity should result in a more consistent weekly team effort if the team process is enacted and viable.

From analysis of the effort data, several things are evident: 1) there is a great deal of variability among individuals, 2) there is a great deal of variability among teams, 3) some teams appear much more consistent over the course of the semester than others (compare team 1 and team 5), and 4) the modal team and individual pattern is a spike at the end; only three teams showed relative consistency of effort over the ten-week period.

**Predicting Performance**

Process performance and product performance were analyzed separately. These two scores correlated, r(58) = .37, p = .004, which, although sizeable, means that 86% of the variance is unshared. Thus, it cannot be assumed that the effort and teaming variables will behave in the same way for both performance variables.

When total effort and consistency of effort were considered together, they predicted performance, both product and process (see Appendix). Thus, the more effort students
exerted and the less variable they were in effort across the 10 weeks, the better their scores on both process and product.

If team effectiveness influences performance by affecting effort, team ratings should be related to effort. As indicated in Table III, they were not. Team ratings did correlate with both process and product performance, however, which suggests that teaming effectiveness affects performance by some mechanism other than effort.

**DISCUSSION**

This study demonstrates that peer ratings of students’ team skills correlate with both process and product performance. We show that while students’ knowledge of team skills is related to their team behaviors, as assessed by their team members, knowledge of team process is not sufficient to help improve performance. This should discourage instructors and researchers from believing that declarative knowledge is a sufficient measure of actual skills.

We also show that team skills do not have their effect on performance by affecting effort. The amount and consistency of effort does predict both process and product performance, but it is a separate effect. That students’ effort predicted project performance independent of teaming suggests that instructors should address it directly. Furthermore, instructors should attend to consistency of effort as well as total effort. A good way of doing this is to incorporate it as an assessment in a continuous improvement loop, but that will only work if both total amount and consistency of effort are measured. Consistency of effort may well be the key to improving students discipline in development activities. Such discipline may serve to prevent the end-of-term effort spikes typically in student projects.

It is notable that all the findings were similar for both process performance and product performance. This may be because students tend to meet course demands consistently, that is, if students know they will be graded on both process and product good students will make sure that they do a good job in both domains and weak students will fail to attend to both. It might also reflect that the instructor assigned points to both.

This study provides evidence that the team peer rating scale, administered under developmental instructions, was a valid method of assessing team skills. First, it correlated with TKT performance, indicating that it is measuring something about teaming. Second, it correlated with performance on the team project, which suggests that it taps aspects of teaming important to project performance.

Finally, this study demonstrated the use of a variety of assessment strategies in informing the instructional process. We demonstrated how different assessment strategies tapped different aspects of the learning process. Effort data and team skill assessment identified different components of the instructional context from which the continuous improvement paradigm can be applied. Most importantly, if process is deemed critical in the instructional context, explicit attention should be devoted to both defining the meaning of process and assessing student performance in that area at both the team and individual level. As suggested in this study, effort data provide an enormous opportunity for faculty to view the software process used by students. Our suggestions are consistent with others [14]-[17] in recommending that software engineering educators find assessment criteria and methods that allow them to ascertain the level and role process plays in their projects courses. If product is the singular assessment criteria in a software engineering course, there will be no way to address process deficiencies and improve them.

**REFERENCES**


APPENDIX

This was determined by entering effort total and effort variability into a simultaneous multiple regression. Together they predicted both process and product scores, $R^2 = .10$, $F(2,57) = 3.08, p = .05$ and $R^2 = .11$, $F(2, 57) = 3.62, p = .03$, respectively. Individually, both effort measures predicted process and product although the $sr$ for variability on process just missed significance ($sr = .29, t = 2.32, p = .02$ for total effort on process, $sr = .30, t = 2.44, p = .02$ for total on product, $sr = -.24, t = -1.95, p = .06$ for variability of effort on process, and $sr = -.28, t = -2.22, p = .03$ for variability of effort on product).

These results must be interpreted carefully, because the zero-order correlations of effort total and effort SD with the performance scores were not significant. The pattern suggests that it would be undesirable to use either total effort or variability of effort to predict performance without including the other.