# An Assessment Model for Large Project Courses

Maria Vasilevskaya
Linköping University
Sweden
maria.vasilevskaya@liu.se

David Broman
UC Berkeley, USA and
Linköping University, Sweden
davbr@berkeley.edu

Kristian Sandahl
Linköping University
Sweden
kristian.sandahl@liu.se

## ABSTRACT

Larger project courses, such as capstone projects, are essential in a modern computing curriculum. Assessing such projects is, however, extremely challenging. There are various aspects and trade-offs of assessments that can affect the quality of a project course. Individual assessments can give fair grading of individuals, but may loose focus of the project as a group activity. Extensive teacher involvement is necessary for objective assessment, but may affect the way students are working. Continuous feedback to students can enhance learning, but may be hard to combine with fair assessment. Most previous work is focusing on some specific assessment aspect, whereas we in this paper present an assessment model that consists of a collection of assessment activities, each covering different aspects. We have applied, developed, and improved these activities during a six-year period and evaluated their usefulness by performing a questionnaire-based survey.

## Categories and Subject Descriptors

K.3.2 [**Computers and Education**]: Computer and Information Science Education—*computer science education, self-assessment*

## Keywords

Project Courses; Assessment; Software Engineering

## 1. INTRODUCTION

Project-based courses are essential in a successful computing curriculum; capstone projects [16, 17] or large scale software engineering projects [2, 3, 12] make students learn how to apply technical skills in a simulated real-world scenario. Yet teaching such project courses is difficult. In particular, assessing student performance in terms of learning outcomes is a challenging, non-trivial task.

Project courses are typically organized by dividing students into groups, where each group develops their own software product. Assessing project results solely on a group's final product does not, however, recognize the difference between individual students' contributions [10]. Certain individuals may contribute significantly, whereas others fly under the radar, adding little to the project. Strictly individual assessment and grading may, on the other hand, result in too much individual focus, thus sacrificing the common group goal of producing a high quality product. As a consequence, good assessment techniques must balance between *individual* and *group assessment* aspects.

Teachers can only assess project parts that are observable. Students are usually working with little teacher involvement, making it very hard to assess what is really happening in a project. Getting teachers extensively involved in all project activities may not be possible within resource constraints. Teacher involvement can also change students' behavior: a group meeting with or without an observing teacher can be drastically different. An alternative is that the students perform the assessment themselves, using peer or self assessments [4, 18]. Clearly, assessment techniques need to involve both *teachers* and *students* and it is essential that such assessment techniques are both fair and accountable [5].

The primary purpose of assessment activities is usually to assess student performance, that is, to generate *feedout* for grading purposes. Such assessments are said to be *summative*. Alternatively, assessments may also be *formative*, meaning that the assessment activity gives *feedback* to students for improved learning [9]. Both summative and formative assessments are essential, but can be hard to combine.

Although tempting, it is very hard to find a single assessment activity that balances between individual and group assessment, involves both teachers and students, and is summative as well as formative. Previous work is focusing on specific aspects, such as self and peer assessments [4, 7, 18], formative vs. summative assessment [9], or grading of written projects [15]. Instead, we propose that a project course should include a collection of assessment activities, each covering certain aspects of the assessment landscape. In particular, we make the following contributions:

- We present an *assessment model* consisting of nine different activities. The key aspect of the model is that the assessment landscape is covered by categorizing the activities in three dimensions: (1) group/individual assessment, (2) teacher/student involvement, and (3) formative/summative assessment. We describe the essential parts of each activity and give rationales for the categorization (Section 2).
- The suggested assessment model is the result of applying and refining these activities in a larger software engineering project during a six-year period. We evaluate the model and its activities by performing a questionnaire-based survey with closed questions (Sections 3 and 4).


*SIGCSE'14,* March 5–8, 2014, Atlanta, GA, USA

http://dx.doi.org/10.1145/2538862.2538947.

## 2. ASSESSMENT MODEL

In this section, we first give an overview of the proposed assessment model and explain how it may be used. This is followed by a brief explanation of each assessment activity within the model.

### 2.1 Overview of the Model

Figure 1 depicts the proposed assessment model. Each activity is categorized within three dimensions: formative/summative assessment (x-axis), teacher/student involvement (y-axis), and group/individual assessment (represented with G, I, and G/I labels). For instance, *individual student interviews* is a summative assessment activity where both teachers and students are involved, but where the interviewees (the students) are in focus. This activity is only assessing an individual student (labeled I), whereas other activities may be group-based, meaning that the whole group is assessed as one entity (e.g., the *grading criteria* activity).

This assessment model may be seen as a *template* when designing assessment activities for a project bases course. For instance, when a teacher is designing a new capstone course, he/she can pick assessment activities that are relevant in the specific context[1]. One benefit of this model is that it emphasizes different aspects of the activities. For instance, if an activity should be used for grading purposes, an individual student report (partially summative) is more suitable than student coaching (only formative). It is important to select a wide range of activities, such that different aspects of assessments are covered. It is also essential that the assessment activities are aligned with the intended learning outcomes of the course [1]. For example, if oral communication skills are part of the learning outcome, oral feedback at student meetings may be more important than individual student reports.

The proposed model for assessing large student projects is general and may be applicable outside the field of software engineering. However, we find it particularly useful for handling the dynamics of this kind of projects that are characterized by changing requirements, flexible management needs, and multiplicity of roles.

We have applied and evaluated this model on a larger software engineering project course (see Section 3.1 and previous publication [3] for details of the course approach). The assessment activities have evolved and been refined over a six-year period. In particular, we have used student feedback in the form of course evaluations, emails, and informal interviews to continuously improve the activities[2].

We contend that this model is general and may be applied to other larger project-based courses. In the rest of this section, we describe the essence of each assessment activity of this model and motivates its categorization.

### 2.2 Grading Criteria

*Description:* The grading criteria activity is basically a list of requirements that a student group must fulfill to get a certain grade. This list should be available to all students before the course starts. Example of grading criteria are software engineering practices that shall be adapted (e.g., continuous integration and Scrum [13]) and artifacts that shall be produced (e.g., test plan and architecture). The grading criteria are formulated by teachers, but the student group decides on a target grade. Thus, students need to estimate and balance the workload and available resources (the number of members in a group, their skills, and available time).

*Categorization*: When grading students at the very end of a course, teachers should follow up the grading criteria to check if groups are meeting the requirements for a certain grade. Consequently, this activity is a tool for *summative* assessment where *teachers* are involved in the activity. On the other hand, grading criteria can also be used by students along the course as a checklist of things to do. Therefore, this activity also contributes to *formative* assessment that is characterized by *student* involvement. Finally, this activity is a *group* assessment; the requirements apply for the project group as a whole.

### 2.3 Oral Feedback at Student Meetings

*Description:* Oral feedback is a simple and efficient assessment activity. To implement it, teachers need to physically attend student meetings, such as status, project, or review meetings. At these meetings, a teacher monitors group progress and contributions of individuals. Teachers should give oral feedback to the whole group or directly to individuals.

*Categorization:* The main purpose of oral teacher feedback is to guide students towards learning goals. Based on this feedback, students have the opportunity to take appropriate actions. Thus, such oral feedback is mostly a tool for *formative* assessment. However, the information about how students address and identify problems may also be useful when grading students. Therefore, this activity is also partially *summative*. Finally, oral feedback contributes to both *group and individual* assessment; feedback may be given to the whole group or to separate individuals.

### 2.4 Student Coaching

*Description:* Student coaching accounts for all forms of peer assessment activities conducted by students within project groups. In particular, students are encouraged to assist each other by giving internal feedback. For example, students can give short tutorials, conduct hands-on sessions within their groups, or review produced artifacts.

*Categorization:* Peer assessment for grading is extensively explored for small group projects [6, 11]. Similarly to peer assessment, student coaching includes evaluation and monitoring by peers. However, student coaching included in our assessment model is a tool for pure *formative* assessment. Thus, the goal of student coaching is to give/receive feedback during the course rather than to grade work of others. Student coaching is conducted without teacher intervention and is, therefore, categorized as a *student* only activity. Both groups and individuals can be targets of student coaching. Therefore, we categorize it as a tool for both *group and individual* assessment.

### 2.5 On-Demand Artifact Feedback

*Description:* On-demand artifact feedback means that students may request, at any time during the course, feedback on their artifacts. Depending on the size of the artifact and timing in the course, the teacher may give oral feedback, written feedback, or both. The purpose of this activity is to give students early feedback, so that the quality of these artifacts can be improved during the course.

*Categorization:* As mentioned above, students are not obliged to request feedback on their artifacts. The final grade is neither affected by requesting feedback nor by the quality of intermediate versions of the artifact. Thus, this activity is a *formative* assessment. This activity requires both *teacher* and *student* involvement: it is the student who requests feedback, whereas the teacher gives

[1]Note that the collection of assessment activities that we present here is neither exhaustive nor complete. There are without doubt many more suitable assessment activities that may be categorized within this model.

[2]Some of the assessment activities were not present the first years that the course was given (e.g., individual student interviews and individual student reports), whereas some other activities have been gradually improved over the years.

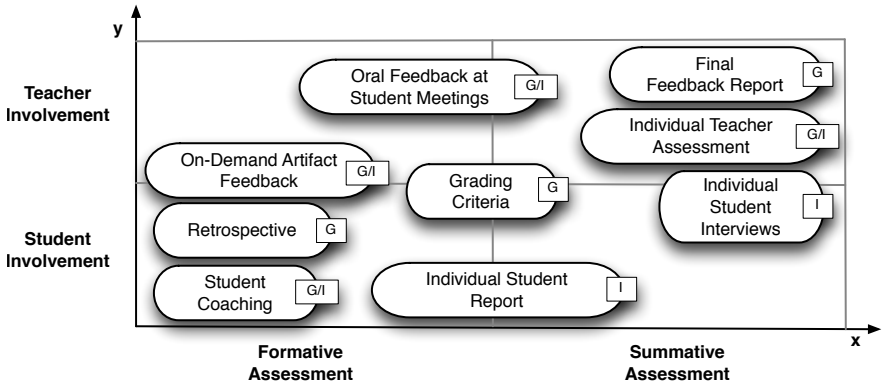



**Figure 1: An assessment model of the relationship between formative/summative assessment and teacher/student involvement. Each oval defines an assessment activity. The different activities are marked as group-based (G), individual (I), or a combination (G/I).**

the actual feedback. Reviews can be requested by both individual students and their groups. Therefore, this procedure contributes to both *group* and *individual* assessment.

## 2.6 Individual Student Report

*Description:* In this activity, each student prepares an individual report where he/she reflects on his/her own performance during the course. This report consists of the two parts: (a) what the student claims he/she has contributed to the group, and (b) what are the most important things that he/she learned.

*Categorization:* An individual report is meant to be used by students as a self-evaluation tool. As a consequence, students have the opportunity to improve their own learning. Hence, a student performs *formative* self-assessment through an individual report. Additionally, this document is used by teachers to analyze individual student contributions, which is also used for student grading. Therefore, individual student reports are also used for *summative* assessment. Naturally, only *students* are involved in this activity because each student writes the report individually.

## 2.7 Individual Teacher Assessment

*Description:* When a team of teachers is involved in the assessment of project groups, each teacher will have their own understanding of the different groups' performance and results. It is, therefore, important to take into account individual teachers' opinions when deciding on final grades. Therefore, we introduce an individual teacher assessment activity into the model, where each teacher first independently assesses groups and individual students. These individual assessments, including individual grading, are then discussed by all teachers together, so that a final grade can be agreed upon[3].

*Categorization:* Individual teacher assessment is carried out at the end of the course. Therefore, it contributes only to *summative* assessment where only *teachers* are involved. As previously mentioned, both *groups and individuals* are assessed in this activity.

## 2.8 Final Feedback Report

*Description:* Written teacher feedback is an efficient way of communicating student performance [15]. In our assessment model, this technique is represented as a final feedback report that is handed out at the last meeting. This report consists of two main components: a grading table and closing comments. The grading table consists of different grading areas, where each area is given a specific grade and a weight factor. This table is then used to calculate the final group grade. The closing comments include detailed feedback for each grading area. In particular, closing comments specify who performed the assessments (which teachers), what information was used (what documents and observations), positive and negative observations, pointers for further improvement, and rationale of the given grade with respect to other grades.

*Categorization:* The final feedback report is categorized as a tool for *summative assessment* because the received feedback (closing comments) cannot be used by students for improvements within the same course. Only teachers are involved in the assessment and the report concerns only the whole group, not individuals.

## 2.9 Individual Student Interviews

*Description:* It is always challenging to evaluate individual students when they work in large groups. Direct evidence of student work is often invisible for teachers. Therefore, we introduce an additional assessment activity where individuals are the main focus. At this event, each student is briefly interviewed at an informal meeting (10-15 minutes) with respect to their role in the group. Typically, only a subset of all students is interviewed. The selection can be based on the individual reports or information received during student meetings. The main purpose is to judge if individual students have contributed more or less than the average student in the group.

*Categorization:* All interviews are conducted at the end of a course. Therefore, this activity is only meant to contribute to *summative assessment*. Both *teachers* and *students* are involved in the interviews as interviewers and interviewees, respectively. The main goal of this activity is to gain better understanding of individual student contributions. Therefore, this activity contributes to assessment of *individuals*.

## 2.10 Retrospective

*Description:* If the project is organized into iterations (for instance by using an agile approach such as Scrum), an important assessment activity is to perform one or more retrospectives. In this activity, students discuss how they have been working during the last iteration. The main questions that should be discussed are: (1) what was good in the iteration? (2) what did not went well?, and (3) how can we improve? Retrospectives can also be used even if the project course is not based on iterations (for example by having the retrospective in the middle of the course).

[3]This activity has similarities to the Delphi method, which was originally designed for expert group judgments.

*Categorization:* A retrospective is meant to be *formative*. It is important for the students to acknowledge what is wrong early in the project so that they can improve. Retrospectives are pure student activities, although teachers may help the students to organize their first meeting. The activity may be seen as a self-assessment of a group. In particular, retrospectives are meant to improve the process (the way of working) for the whole group, not assessing specific individuals.

# 3. DESCRIPTION OF THE STUDY

In this section, we first describe the course that is studied, followed by methods used for collection and analysis of data for the survey.

## 3.1 Context

The assessment model presented in the previous section is the result of a six-year period of experimentation with a software engineering project course (see Broman *et al.* [3] for a detailed description of the course approach). The course is given as a master's level course. The students come from different curricula and countries. In 2012, when this study was performed, 88 students participated in the course and about 15% of the students were female.

The major goal of the course is for students to obtain fundamental understanding of a larger software engineering project and experience its challenges. Students should also learn organizational, process, and communication perspectives. Initially, students are divided by lottery into 3-4 groups of 20-30 students. One reason for having such large groups is that most functions and roles in a real company shall be present. All groups get the same task and the same customer representative. In 2012 the customer was an external product development company. The course runs during an entire semester and comprises two phases:

1. A prestudy phase where the students organize the group, assign roles, start education, interview the customer, and create a preliminary version of the architecture. The prestudy ends with a *tollgate* meeting where the continuation of the project is formally determined.

2. A series of 2-week iterations for implementation, prototype evaluation, testing, documentation, and configuration management. The project ends with a demonstration and experience seminar with all groups.

The students are free to organize themselves. However, we recommend them to organize into smaller, cross-functional teams after the prestudy and use a SCRUM-based process framework.

In total, four teachers/teaching assistants are involved in the course. During the course, each group meets one teacher for approximately one hour at least once a week. Depending on their needs, some individuals and small groups also arrange additional short meetings with the teachers (e.g., to ask questions about project planning and requirements documents). The summative assessment (grading and writing final feedback reports) is done in 2-3 consequent days where all teachers are involved. Performing all these activities may seem costly with a lot of communication overhead, but when all teachers know their role in the course our experience is that the course can be executed quite efficiently.

The course is, according to the official student evaluation, very popular and students are satisfied with their learning.

## 3.2 Data Collection

The evaluation is based on a questionnaire-based survey with closed questions. The students answered each statement by selecting one of the six options: strongly disagree, disagree, neutral, agree, strongly agree, and N/A. The blanks with questions were handed out during the first 15 minutes of the last mandatory meeting, just before the final grades were announced and the final feedback reports were given to the students. The answers from each group were collected in a closed envelope by an assistant not teaching the project and the group identities were randomly coded. No identification of the individual student was possible. The survey was conducted in December 2012. In total, 72 questionnaires were collected that corresponds to approximately 82% response rate.

## 3.3 Analysis Method

To process the obtained data, we used descriptive methods of SPSS (Statistical Package for the Social Sciences). The answers of students were analyzed as nominal scale measurements. There were seven independent variables (year of the course, age, sex, industrial experience, software development experience, curriculum, and result from theory exam) and 32 dependent variables (the questions). Among dependent variables, 21 were used for further analysis while the rest of variables were related to the general course evaluation. Out of these variables, 11 are detailed in this paper.

# 4. RESULTS AND DISCUSSIONS

In this section, we discuss the results of the study. We present the results in terms of the previously introduced dimensions.

## 4.1 Individual and Group Assessment

Although students work together in group-based courses, each student should be given an individual grade at the end of a course. As previously described, the assessment activities in our model (Figure 1) contribute to both group and individual assessment. A natural question then arises whether individual or group performance should be emphasized when deciding on the final grade.

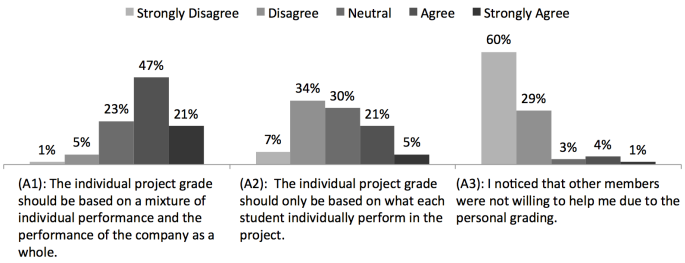
According to the results of statement A1 in Figure 2(a), a strong majority believes that a mixture of individual and group-based performance gives fair grading. In particular, 69% of the students agrees or strongly agrees to statement A1, whereas only 6% of the students disagree or strongly disagree. Only 26% of students think that individual project grades should be based on their individual performance (A2), whereas 41% disagree. These two results show that most students do believe in mixed individual and group based assessments, although a minority still thinks that the assessment should be individual.

Collaboration among students is one of the primary learning goals in many group-based courses. Too strong focus on individual grading may, however, force students to focus solely on their individual performance, thus prevent fruitful collaboration. Fortunately, the results of this study do not point in this direction; 89% of the students confirm that such prevention of collaboration did not exist (statement A3).

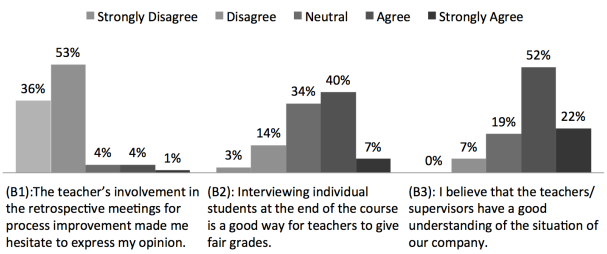
## 4.2 Teacher and Student Involvement

To enable fair and valuable student feedback, teachers need to be present at certain occasions where they can observe the students' work. In the project course of this study, teachers are involved in student planning sessions and retrospective meetings, where students plan their work and discuss emerging problems.

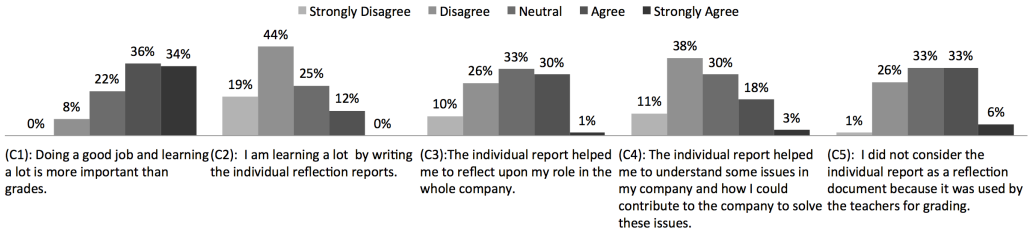
A disadvantage of such an approach is that students may hesitate to express their opinions when a teacher is present. Consequently, this could prevent students from improving their way of working and learning. Statement B1 in Figure 2(b) shows, however, that this threat is quite mild; 91% of the students state that the teacher involvement does not restrain them from honest discussions.




(a) Group and Individual Assessment




(b) Teacher and Students Involvement




(c) Formative and Summative Assessment

**Figure 2: Results of the quantitative study. The diagrams are categorized according to the three dimensions of the assessment model. Missing and N/A answers are omitted for the sake of clarity.**

When only teachers are involved in the assessment, it is extremely difficult to get a good understanding about individual students' contributions. This may lead to unfair summative assessment. To tackle this deficiency, our assessment model includes several activities with both student and teacher involvement (see Figure 1). In particular, we want to draw extra attention to the assessment activity of individual student interviews (see Section 2). The results of statement B2 clearly shows that individual interviews are appreciated; 47% of the students agree that this activity is a good way to assess and give fair grades (34% were neutral).

Finally, the result of statement B3 shows that 74% of the students believe that the teachers have a good understanding of the situations within the group. Although the students' belief of teachers' understanding does not express the teachers' actual understanding, we conclude that the teacher involvement in the course is vital and that the involvement suggested in this approach does not affect the team work negatively.

## 4.3 Formative and Summative Assessment

One possible consequence of too much focus on summative assessments is that students focus on obtaining good grades rather than effective learning. Our survey shows, however, that 70% of the students think that doing a good job and learning are more important for them than getting good grades (statement C1 in Figure 2(c)).

The individual student report (Section 2.6) is both a summative and formative assessment activity. Our experience (as teachers) confirms that an individual report is a useful tool for teachers to get better understanding of individual student performance. As a tool for formative assessment, an individual report is meant to be an instrument for self-reflection, helping students to obtain insights of the group and their individual role. Such self-reflection may also help students improve their way of working. The results of our survey show, however, that this goal is not completely achieved. Only 12% of the students confirm that writing an individual report helps them to learn (statement C2 in Figure 2(c)), whereas 63% of the students disagree with this statement.

Statements C3 and C4 show that an individual report is more successfully used by students to reflect upon their own way of working, compared to the whole group. Still, a large percentage (36% and 49% for statement C3 and C4, respectively) do not believe that individual reports are useful as a formative self-assessment activity.

Before we performed our survey, our hypothesis was that the use of an individual report for summative assessment may prevent students from honest reflection. This hypothesis is partly confirmed. In particular, 37% of the students did not consider the individual report as a reflection document because it was used for grading (statement C5 in Figure 2(c)), whereas 27% stated the opposite. However, to completely prove this hypothesis, further studies need to be performed. Moreover, additional education of students on self-reflection (e.g., as lectures or seminars) could also be used to better prepare students for self-assessment.

# 5. RELATED WORK

There is much research reported on capstone projects that are relevant to large project courses. In this section we compare our results.

## 5.1 Individual and Group Assessment

A student survey by Farell *et al.* [5] shows similar results as our study about student perception: 54% prefer a mix of individual and group marks and 30% want their mark only based on individual achievements. Teachers' means of getting individual results are written contributions statements (as suggested by Clark *et al.* [4]) and peer assessment [7]. A comprehensive introduction to several practically useful techniques is described by Wilkins and Lawhead [18] and Lejk *et al.* [11].

## 5.2 Teacher and Student Involvement

From Figure 1 we note that the right-bottom corner is empty. This is where peer assessment would go, an assessment method that is

frequently researched. The reports point out both that students have confidence in peer assessment [5], and that the accuracy is acceptable [7, 10]. There seems to be no significant difference between students of different gender and nationality [7]. One of the challenges is to create smooth processes and tools that motivate the students to participate [4, 5]. Computer based tools such as SPARC have been well received by a majority of students [19] and tools combining assessment with assessor's confidence have scaled up to group sizes of 20 students [14].

In our approach, students are involved in a dialog with teachers and giving each other formative assessments. So far, we have not experimented with peer assessment, but we view it as interesting future work to evaluate and extend our model with peer assessment.

## 5.3 Formative and Summative Assessment

Providing formative assessment is challenging; the response on writing individual reflection reports shown in Figure 2(b) indicates an unexpectedly low degree of learning. Many students' submitted reports because it was required information for grading. Similar results have been reported by Hernandez [8] in a study of undergraduates in Hispanic studies. Both student and teacher perspectives were considered. When asked to rank the most frequent purpose of performing assessments students responded "grading students" in 24% of the cases, whereas teachers answered "grading student" in 13% of the cases. The study also investigated what students do with the assessment. Even though teachers gave feed-back by marking weak points and suggesting improvements, many students just took part of the summative information of the assessment. To make the students learn more from assessments, the assessments needs "feed-forward", that is, information about of what to do with the feed-back. This is probably most applicable to our course, especially since we expect students to make a self-assessment; we must teach them what to do with the results of their self-assessment. The inherent complexity of formative assessment is well described in the theoretical work of Yorke [20] who provides a multi-faceted tool for reasoning about course improvement.

# 6. CONCLUSIONS

In this paper, we present an assessment model for assessing large software engineering project courses. The model consists of nine assessment activities that have been evaluated by a questionnaire-based survey. We contend that this model can be useful when analyzing assessment activities in various software engineering courses.

## Acknowledgement

The authors would like to thank all the students who participated in the software engineering course. This research was funded by the Department of Computer and Information Science, Linköping University, Sweden. The second author was funded by the Swedish Research Council #623-2011-955, with support from the Center for Hybrid and Embedded Software Systems (CHESS) at UC Berkeley.

# 7. REFERENCES

[1] J. Biggs. Enhancing teaching through constructive alignment. *Higher Education*, 1996.

[2] D. Broman. Should Software Engineering Projects be the Backbone or the Tail of Computing Curricula? In *Proceedings of the 23th IEEE Conference on Software Engineering Education and Training*, Pittsburgh, USA, 2010.

[3] D. Broman, K. Sandahl, and M. Abu Baker. The company approach to software engineering project courses. *IEEE Transactions on Education*, 55(4):445–452, 2012.

[4] N. Clark, P. Davies, and R. Skeers. Self and peer assessment in software engineering projects. In *Proceedings of the 7th Australasian conference on Computing education - Volume 42*, ACE '05, pages 91–100, Darlinghurst, Australia, Australia, 2005. Australian Computer Society, Inc.

[5] V. Farrell, G. Ravalli, G. Farrell, P. Kindler, and D. Hall. Capstone project: fair, just and accountable assessment. In *Proceedings of the 17th ACM annual conference on Innovation and technology in computer science education*, ITiCSE '12, New York, NY, USA, 2012. ACM.

[6] D. Hemer. Peer assessment of group-based software engineering projects. In *Australian Conference on Software Engineering*, 2008.

[7] N. Herbert. Quantitative peer assessment: can students be objective? In *Proceedings of the ninth Australasian conference on Computing education*, ACE '07, pages 63–71, Darlinghurst, Australia, Australia, 2007. Australian Computer Society, Inc.

[8] R. Hernández. Does continuous assessment in higher education support student learning? *Higher Education*, 64(4):489–502, 2012.

[9] P. T. Knight. Summative Assessment in Higher Education: Practices in disarray. *Studies in Higher Education*, 27(3):275–286, 2002.

[10] N. LeJeune. Assessment of individuals on CS group projects. *J. Comput. Sci. Coll.*, 22(1):231–237, Oct. 2006.

[11] M. Lejk, M. Wyvill, and S. Farrow. A survey of methods of deriving individual grades from group assessments. *Assessment & Evaluation in Higher Education*, 21(3):267–280, 1996.

[12] F. Meawad. The Virtual Agile Enterprise: Making the Most of a Software Engineering Course. In *Proceedings of the 24th IEEE Conference on Software Engineering Education and Training*, pages 324–332, 2011.

[13] K. Schwaber and M. Beedle. *Agile Software Development with Scrum*. Prentice Hall, 2001.

[14] L. Slivovsky, L. Jamieson, and W. Oakes. Evaluating Multidisciplinary Design Teams. In *Proceedings of the 2003 American Societey for Engineering Education Annual Conference & Exposition*, 2003.

[15] L. J. Smith. Grading written projects: what approaches do students find most helpful? *The Journal of Education for Business*, 83(6):325–330, 2008.

[16] R. H. Todd, S. P. Magleby, C. D. Sorensen, B. R. Swan, and D. K. Anthony. A survey of capstone engineering courses in north america. 84:165–174, 1995.

[17] D. A. Umphress, D. Hendrix, and J. H. Cross. Software process in the classroom: the capstone project experience. 19(5):78–85, 2002.

[18] D. E. Wilkins and P. B. Lawhead. Evaluating individuals in team projects. In *Proceedings of the thirty-first SIGCSE technical symposium on Computer science education*, SIGCSE, pages 172–175, New York, NY, USA, 2000. ACM.

[19] K. Willey and A. Gardner. Improving self- and peer assessment processes with technology. *Campus-Wide Information Systems*, 26(5):379–399, 2009.

[20] M. Yorke. Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher education*, 45(4):477–501, 2003.