Alignment of PBL and Assessment

Anette Kolmos¹ and Jette Egelund Holgaard²

Abstract - This paper addresses the alignment between assessment and Problem-Based and Project-Based Learning (PBL) at Aalborg University, Denmark. A situation in Denmark has occurred, giving the opportunity to compare two assessment systems used for PBL: 1) group-based assessment and 2) individual assessment - both with individual grading. In the spring of 2006, a longitudinal study was initiated to compare the two assessment systems using both qualitative and quantitative methods. This paper presents the preliminary results from the first collection of data, based on questionnaires send out to both staff, external examiners and students – in total more than 1100 respondents from the Faculty of Engineering, Science, and Medicine. The focus in this paper is especially on the general approach to the two assessment systems and the assessment of certain skills. The preliminary results, based on experiences from faculty staff, external examiners and students, show that the individual exam compared to a group exam does not test out the same level of knowledge and skills.

Index Terms – Assessment, Problem Based Learning, Project Based Learning.

Since 1974, Aalborg University has been running a curriculum based on problem-based and project-based learning (PBL), and the objectives embrace technical and contextual as well as process skills to reflect the complexity and inter-disciplinary collaboration of most engineering professions. As the students from the very beginning of their study are working in teams to complete their projects, it has been in a natural continuation of the PBL-philosophy and in correspondence with the theory of alignment to assess the project skills in a group-based examination [1].

Traditionally, project examinations at Aalborg University have therefore been group-based with individual grading. However, the institutional framework in Denmark has been changed considerably by the liberal government, and it was banned by law to assess students in a group setting, so therefore the examination should be carried out individually.

No matter what political reasons the government might have had for enacting this law, this created a situation where it was possible to compare two different assessment systems.

This paper presents selected results from the first part of a longitudinal study conducted at Aalborg University, comparing the two assessment systems as to whether these systems are able to capture knowledge and process skills, which are emphasised within a PBL philosophy.

Theoretically, we take our point of departure in the principle of alignment arguing that an alignment between assessment and PBL necessitates an assessment system which is able to address complex knowledge construction combining knowledge and process skills. To explain the institutional setting at Aalborg University further, we turn to describe the two assessments systems, which have been used in relation to problem based projects. After a short overview of the design and implementation of the empirical study, we turn to the analysis of results to address the alignment of PBL and the two specific assessment systems.

ASSESSMENT AND PBL

The principle of alignment among all elements in the educational process was presented for the first time in John Biggs’ work on curriculum development [2]. The principles of alignment entail the existence of consistency and logic among all the elements:

- The curriculum that we teach.
- The teaching methods that we use.
- The assessment procedures that we use and the methods of reporting results.
- The climate that we create in interaction with the students.
- The institutional climate, the rules and procedures we have to follow [2: 26].

This represents a holistic approach to analyse and develop the curriculum. It is not possible to change one element without rethinking the whole. This is in line with the Scandinavian relationship approach to curriculum development (in continental Europe called didactics). However, the Scandinavian approach also stresses the importance of the interaction between educational politics and curriculum development in order to analyse curriculum change [3]. Him and Hippe [3] operate with six factors, which make up the most important elements in teaching and learning analysis:

- Goals for learning.
- The student’s social, cultural, psychological, and physical prerequisites for learning
- Cultural, social and physical factors (including the prerequisites of the teacher)
- Content knowledge
- Learning process
- Assessment

¹ Anette Kolmos, Faculty of Engineering, Science, and Medicine, Aalborg University, Aalborg, Denmark, ak@plan.aau.dk
² Jette Egelund Holgaard, Faculty of Engineering, Science, and Medicine, Aalborg University, Aalborg, Denmark, jeh@plan.aau.dk


1st International Conference on Research in Engineering Education

June 22 – 24, 2007, Honolulu, HI

Page 1 of 9
Based on these elements Kolmos & de Graff [4] have developed a model, which emphasise the institutional level, see figure 1.

The point is that changing to PBL systems involves an alignment among all the elements in the teaching and learning process in order to establish an efficient system and avoid the creation of a hidden curriculum [5-8]. Graham Gibbs [5] emphasises that the assessment systems are the most powerful factor influencing students' learning process and that the assessment system can be used strategically in order to enhance students’ learning outcomes. Gibbs’ research shows that changing the criteria for assessment and assessment systems might have an effect on students’ learning.

However, PBL curricula objectives are often more complex in nature, so in order to align learning objectives and assessment methods, it might be necessary to look at both assessment criteria and assessment methods [9]. New criteria are needed in order to test students’ abilities to analyse and solve complex problems, e.g. criteria for applying theoretical knowledge, analysing problems and pointing out problem solving methods. Also new assessment methods might be needed in order to have time for testing the students’ knowledge, skills and knowledge management.

PBL is a relatively new area and recent research has contributed both to the development of new assessment systems and to the evaluation of implemented systems. A broad scope within this research has covered both the formative and the summative assessment related to a PBL learning environment. In the formative assessment, the results are used for feedback during learning, whereas summative assessments are used to grade students [2].

There is a tremendous body of research analysing, documenting, and developing formative assessment and peer assessment methods in order to establish new process oriented practices, as PBL has multiple objectives concerning content knowledge as well as process skills [4,10]. Formative assessment is very much in alignment with the principles of PBL, because formative assessment supports awareness of the learning process. However, it is argued that not only formative assessment but also summative assessment should be carefully constructed in order to support students’ learning and that focus should be on both learning and performance [11-12].

If assessment has to focus upon both learning and performance, it involves an oral assessment combined with types of written work like portfolios, projects, etc. [13]. There might be many reasons for using mixed assessment methods. One reason is to achieve more valid testing methods determined by both quantitative and qualitative paradigms. The written exams are typically based on the quantitative paradigm combined with subsequent control, whereas the oral examination is based more on the qualitative paradigm combined with dialogue [5-7, 13-14].

Research has documented that written exams do not always include a high validity. For example, a Danish research project compared students’ results at a written exam with a traditional focus on calculation with the results of a qualitative test focusing on students’ conceptual understanding. This comparison showed that only half of the students had been given a grade which correlated to their level of understanding. For around one quarter of the students the level of qualitative understanding was not living up to their grades, and for another quarter of the students it was the other way around [15].

Whereas there is much research on formative assessment, the body of research on summative assessment and PBL is not so impressive, especially not the research on group-based summative assessment. Most of the rather limited research on group-based assessment address the formative perspective. This perspective has the special aim to improve students’ cooperative skills by use of peer and self assessment [16-17].

As PBL has become more widespread at a global scale, the variation in assessment practices has increased. However, when looking at relatively well known institutional models, the traditional problem based model at Maastricht and McMaster [18] and the problem based and project based model at Aalborg University [1], there are significant differences in the use of group-based summative assessment models, see figure 2.

At Maastricht University, there is a widespread use of individual formative assessment – both individual and group-based, however the summative assessment is primary carried out as individual exams.

At Aalborg University, project exams are carried out by use of group-based assessment methods. This method has been chosen as it is considered suitable for testing the objectives of problem based, interdisciplinary, and cooperative learning which is a much more complex knowledge construction process. This process is knowledge-based as well as skill-based and requires resources of a team [9]. As it is skill-based, the learners should be able to utilise their knowledge and relate it to practice, and the assessment method should be able to capture this ability.
Dochy et al. [19] have made a review of literature from the 1990s evaluating the long-term effects of PBL. Their main conclusion is that the implementation of PBL has a considerable impact on the development of skills. The impact on the knowledge acquisition as provided in the traditional system is missing or not significant. However, PBL students do not acquire less knowledge compared to students educated the traditional way. Several studies come up with the same findings: that there is no significant improvement of the knowledge acquisition as provided by the traditional system but significant improvement of skills such as process skills, see for example the recent studies by Faland and Frenay [20] and Crosthwaite [21]. Furthermore, Schmidt and Moust [18] have done a review of existing literature and conclude that PBL seems to have an effect on long-term retention of knowledge such as remembering and understanding various concepts.

In a Danish context, research comparing graduates from Aalborg University and from other Danish institutions show that employers appreciate Aalborg candidates’ ability to cooperate, share knowledge, manage a project, apply knowledge and search for new knowledge [22]. However, if these objectives are to be obtained, the change of assessment system might create a risk for the future education.

**THE TWO ASSESSMENTS SYSTEMS AT AAU**

Traditionally, project exams at the faculty of Engineering, Science and Medicine at Aalborg University have been group-based with subsequent individual grading. Table 1 gives an overview of this process of group-based assessment with subsequent individual grading.

<table>
<thead>
<tr>
<th>Phase in assessment process</th>
<th>Activities for students</th>
<th>Activities for facilitator and external examiners</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral presentation. Approx. 1 hour</td>
<td>The students make the oral presentation by taking turn, max. 10 min. per student.</td>
<td>The facilitator welcomes everyone and the external examiner presents his professional profile.</td>
<td>The report and presentation constitutes the background for differentiating grades. First impression of the individual with written notes.</td>
</tr>
<tr>
<td>Feedback on presentation</td>
<td>The students answer by raising their fingers and reply when asked.</td>
<td>Clarifying questions to the content and comments on the form of the oral presentation.</td>
<td>Ongoing assessment</td>
</tr>
<tr>
<td>10-15 min. break</td>
<td></td>
<td></td>
<td>Ongoing voting an individual grading</td>
</tr>
<tr>
<td>Discussing the project Approx. 1½ hour</td>
<td>The students answer by raising their fingers to answer a question from facilitator/external examiner – or to follow up on another student’s comment.</td>
<td>Questions to problem definition, methodology and to the general strength and weaknesses of the choices made during the project. Passive students are asked directly.</td>
<td>Ongoing assessment</td>
</tr>
<tr>
<td>10-15 min. break</td>
<td></td>
<td></td>
<td>Ongoing voting and individual grading</td>
</tr>
<tr>
<td>Questions to specific individuals Approx. 1½ hour</td>
<td>The students by request respond individually. The other group members are passive.</td>
<td>Asks clarifying questions to the student, e.g. by asking them to clarify an issue by use of the blackboard.</td>
<td>Ongoing assessment</td>
</tr>
<tr>
<td>Voting 15 min.</td>
<td></td>
<td></td>
<td>Final voting and individual grading</td>
</tr>
<tr>
<td>The grades are announced and commented on 15 min.</td>
<td>Comments on the project as a whole and the overall level of the report and presentation. Comments to explain the reasoning behind each grade.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is an oral examination based on a written report. The students present their project and subsequently, facilitators and external examiners comment on their communication skills and assess whether the presentations serve to clarify or add something new to the written report. From that point on, the facilitators and external examiners ask questions to the team.
members related to their report, and the team members have the opportunity to answer the question or elaborate on a point made by one of their fellow students. Finally, after voting, the team members receive an individual grade. The report serves as a point of reference for differentiating the individual grades, which are based on the oral presentation and the performance during the questioning part. The facilitator is at all times in charge of the assessment and responsible for making sure that the questions asked are related to the material of the report or the oral presentation.

In 2006 the Ministry of Science, Technology and Innovation and the Ministry of Education, banned group-based exams. The consolidated act (LBEK 280 of 21/03/2006) came into force from the 1st of April 2006. Therefore, a new assessment system had to be developed. Table 2 gives an overview of the individual exam practiced at Aalborg University in the summer exam 2006.

### Table 2

<table>
<thead>
<tr>
<th>Phase in assessment process</th>
<th>Activities for students</th>
<th>Activities for facilitator and external examiners</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral presentation of the project</td>
<td>The students make the oral presentation by taking turn, max. 10 min. per student.</td>
<td>The facilitator welcomes all and the external examiner presents his professional profile.</td>
<td>This phase is considered as a part of the project and is not basis for grading.</td>
</tr>
<tr>
<td>Approx. 1 hour</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short discussion</td>
<td>Reflect on the comments given and eventually reply on the given comments.</td>
<td>Make comments on the presentation and give some overall comments on the project.</td>
<td>This is an activity seen as a part of the project and is not basis for grading.</td>
</tr>
<tr>
<td>Approx. 1/2 hour</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Break 10-15 min.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual exam of student 1-n</td>
<td>Each individual gets 2-3 predefined questions and answers.</td>
<td>Question the student to clarify or elaborate on the answers.</td>
<td>Immediately after the questioning, voting takes place and the students receive an individual grade.</td>
</tr>
<tr>
<td>35 min. each</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As shown in table 2, the group-based oral presentation has been retained, followed by some overall comments to the report addressed to the group as a whole, but this part is now considered as the last part of the project process and therefore is not a part of the exam. However, at the following individual questioning part, the facilitator and external examiner have the right to ask questions related to the presentation, as it is a part of the project.

In the individual part, the students come into the room one at a time and get 2-3 questions which they answer and elaborate on by request. After each individual exam voting takes place and the grade is given. The questioning, voting, and commenting to the grade are to be completed in approx. 30-35 minutes, leaving 20-25 min. for the questioning part.

### DESIGN AND IMPLEMENTATION OF THE EMPIRICAL STUDY

A thorough investigation was made at the summer exam 2006 to gather experience from students, facilitators and examiners, in order to compare the strength and weaknesses of individual versus group-based project exams in a problem based learning environment. In the following the methodology, considering the design of the study and the data collection and processing, is briefly outlined.

#### Design of the study

The overall design of the study is based on several data sources, see table 3. Data have been collected at all faculties at Aalborg University, but only at the engineering faculty has it been possible to compare the two different assessment systems at the summer exam 2006.

#### Table 3

<table>
<thead>
<tr>
<th>Overall Design of the Study</th>
<th>Data collection summer 2006</th>
<th>Data collection summer 2007 and winter 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaire</td>
<td>Group-based exam with individual grading</td>
<td>Individual exam only at engineering</td>
</tr>
<tr>
<td>Students</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Faculty facilitators</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>External examiners</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Video Observations</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Qualitative interviews</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

The questionnaire was designed to 1) document the group-based assessment and the potentials and problems connected with this assessment system, 2) report on the first experiences with the individual assessments system from students enrolled as bachelors, 3) get an impression of the attitudes and expectations towards the new assessment system which could be compared with the actual experiences when the system is fully implemented.

The questionnaire was designed to cover seven themes:
1. Background information about the respondents – gender, occupation, experience with the educational system and more specifically with the type of assessment system.
2. Expectations and experiences with the differentiation of grades within a group-based and individual assessment system.
3. Relationship between assessment systems and the impact on the learning process.
4. The opinions towards group-based exams versus individual exams.
5. The skills that are assessed within the different assessment systems, considering both the scientific-technological skills and the process skills.
6. The assessment process as a learning process focusing on whether the students enhance their skills during assessment.
7. The overall stand towards the two assessment systems, considering what the respondents would prefer based on an overall view and which argumentations they use to support their opinion.

The questionnaire did have both multiple choice and open questions. This paper reports on selected data focusing primarily on theme no. 5, 6 and 7.

It should be noted, that the results presented in this paper are preliminary, as more data has to be collected to address the experiences with the individual assessment system. Furthermore, we would like to stress that the questionnaire is designed to capture the respondents’ attitudes towards the two assessment systems and their perceptions of the knowledge and skills tested at a given exam. In other words, the investigation documents whether the two specific assessment systems in question, from the respondents’ point of view, creates a setting that enables assessment of knowledge and process based skills.

Data collection

The questionnaires were carried out using Survey Exact and send out at the Faculty of Engineering Science and Medicine. 1151 students and examiners participated, distributed on 794 students (14% international students), 202 internal examiners (staff) and 155 external examiners. The respondent rate is at an average 25%.

Data processing

The frequency analysis, which was accessible in Survey Exact, was exported to Excel in order to visualise the results. At this preliminary study a multivariate analysis has not yet been prepared. In the second data collection, summer 2007, more substantial material was collected on the individual assessment system and it will be feasible to make such analysis and point out significant differences between the two assessment systems. Besides processing the data to illustrate some preliminary tendencies, we have also used this first set of data to limit the amount of questions even further in round two of the data collection.

Results

All the participating examiners had practiced at the summer exam 2006. 96% of the internal faculty examiners (n=200) and 88% of the external examiners (n=154) had practiced within the group-based system, and respectively 42% and 40% had practiced within the individual system. Of the participating students only 15% had tried the individual exam (n=765).

71% of the participating students from the Faculty of Engineering, Science and Medicine were male and 29% female (n=786). Of the internal examiners 88% were men and 12% were women (n=202). For the external examiners the males were also the dominating part (91%, n= 154). 39% of the internal examiners and 26% of the external examiners had more than 10 years of experience as examiners at project exams at AAU.

Overall results

Figure 3 shows students’ overall preferences for assessment methods. The overall results show that 69% of the students (n=516) preferred a group-based exam instead of an individual exam. However, the students that have already experienced an individual exam are more positive towards the individual exam compared to the students that have participated in a group-based exam and not yet experienced the individual method.

Skills
In the study, a series of knowledge and process skills have been defined, see table 4. Some of these skills are corresponding to typical objectives in a PBL curriculum. All respondents have been asked the same question: To what extent do you feel that the following skills were evaluated at the exam? The skills tested in the study are outlined in table 4.

The assumption is here that it is difficult to assess the more complex knowledge based skills, (covering no. 4-9 in table 4) and the process skills (covering no. 10-12 in table 4) in an individual exam due to the relatively short time where it is possible for the facilitator and examiner to interact with each student and the students’ lack of possibility to interact with fellow students.

TABLE 4
SKILLS TESTED IN THE STUDY

| 1. Quickly answer the asked questions? |
| 2. Remember concepts and definitions? |
| 3. Relate different concepts to each other? |
| 4. Argue for methodological choices? |
| 5. Assume a theoretical overview? |
| 6. Apply analytical skills? |
| 7. Argue for and against a solution to a problem? |
| 8. Transfer knowledge gained in projects to other situations? |
| 9. Communicate knowledge through a presentation? |
| 10. Enter into dialogue and cooperation? |
| 11. Complement and expand on others answers? |
| 12. Participate in a group work? |

Clearly, the results show that for the first three levels of skills, there is no considerable difference between a group-based and an individual exam for any of the respondent groups, see figure 5. However, there is a minor tendency which show that the group-based exam has a higher ability to address whether the students are capable of relating different concepts to each other, at least if you ask the students (90% compared to 73%) and the staff (94% to 76%).

Thereby, the 3 lowest levels of academic skills such as to answer questions quickly, remember concepts and definitions, and relate concepts to each other, are by students, faculty and external examiners estimated to have almost the same significance in both an individual and a group exam. Nevertheless, it is levels, which are typical for a superficial rather than a deeper understanding.

For the next three levels of skills, considering if it is tested whether the student can argue for methodological choices, assume theoretical overview and apply analytical skills, there is a remarkable difference between the students’ and staff’s perception of a group based exam and an individual exam. In figure 6, the results from both students and staff show that such skills are better evaluated in a group-based exam than in an individual exam. However, there are only minor differences in the external examiners’ experiences of the two assessment systems in this regard.

Concerning skills as the ability to argue for and against a solution, transfer knowledge to one context to another, and to present one’s knowledge, figure 7 shows the students’, staff’s and external examiners’ estimations of, to what extent these skills will be evaluated at a group exam and an individual exam. In regard to these skills, there is an even clearer tendency that the group exam offers a better setting than an individual exam.
We are dealing with a significant difference in the estimation process skills as the ability to enter into dialogue and cooperation, elaborate on others answers, and participate in teamwork are perceived to be evaluated in the two assessment systems. The tendency for process skills is clearly the same. Results show that these are only tested to a more limited extent in an individual exam compared to a group exam where team play becomes a central element in the argumentation for and defence of the project.

Figure 8 shows to what extend the process skills, as the ability to enter into dialogue and cooperation, elaborate on others answers, and participate in teamwork are perceived to be evaluated in the two assessment systems. The tendency for process skills is clearly the same. Results show that these are only tested to a more limited extent in an individual exam compared to a group exam where team play becomes a central element in the argumentation for and defence of the project.

Figure 9 shows the students’ experience if their professional knowledge has been tested. The results clearly shows that 90% of the students who have been to a group-based exam do feel that they have really been asked questions to their professional knowledge, however only 37% of the students that have participated in an individual exam have the same feeling.

Figure 10 shows examiners’ and external examiners’ evaluations of, whether the possibilities for probing students’ deeper academic understanding has become better or worse. External examiners and examiners have different estimations. More than 65% of the external examiners and about 43% of the staff find that the individual exam gives possibilities which are significantly better, better or the same. On the other hand, about 15% of the external examiners and almost 50% of the examiners at the individual exam find that the individual exam situates worse conditions. This is a quite remarkable percentile of the examiners, which ought to give rise to further consideration for the system, in such a way that the academic profundity still can be tested.

Generally, the estimations made by the staff and students about which skills that are evaluated at a group exam and an individual exam, confirm the hypotheses for all types of skills. We are dealing with a significant difference in the estimation of the two types of exams. One of the reasons for this is that in a group-based exam the progression of questions going from the level of memorising and understanding to analysis and new synthesis of knowledge is moving on slowly during 3-4 hours. However, in an individual exam, it is so that staff and external examiners will have to start from scratch at the memorising level each time and thus the possibility to test deep knowledge is limited.

Testing of professional knowledge and response
In this paper, we have addressed the alignment between assessment and PBL at Aalborg University, Denmark, drawing on a longitudinal study initiated in the spring of 2006. The empirical investigation includes a questionnaire, which was handed out to more than 1100 students and scientific staff from the Faculty of Engineering, Science and Medicine, Aalborg University, and related external examiners. The questionnaire includes seven themes, whereas emphasis in this paper has been on: General experiences and the skills that are assessed within the different exam systems.

The result shows that a majority of respondents experience that it is difficult at an individual exam to test complex knowledge calling for overview, argumentations and reflections and especially process skills developed in teams. The individual exam fits better with a control approach to test disciplinary knowledge. The individual exam gives priority to surface knowledge by means of testing the students ability to answer quickly, memorise, or relate concepts to each other however, the deeper level as to argue for methods, theoretical overview, analytical skills, argue for and against solutions, and transfer knowledge from one context to another are not tested to the same level in a half hour individual exam compared to a 5 hours group-based exam. Furthermore, process skills as communication and dialogue with other peers and documentation of the students’ ability to share knowledge in project groups are no longer basis for grading. That is maybe why, half of the faculty staff does not think that it will be possible to test the academic understanding in an individual exam compared to their experiences from the group-based exam.

This paper only reports on the first preliminary results. At the summer exam 2007, all exams will be individual and there will be more solid data from the individual exam that can be compared with experiences from group based examinations, summer 2006. Furthermore, it will be possible to take qualitative data from interviews and video observations into consideration. The purpose is to address the following hypothesis even further:
- Choice of projects, where supervisors do not know the results in advance. Thus there will be a real risk with regard to acquisition of competences such as problem analysis, independence, and cross disciplinary complexity.
- Training of employability as there will not be emphasis on the ability to transfer knowledge to other areas.
- The students’ co-operation processes in the direction of more competition and a more individually based working environment.
- The project advisor will take on a more controlling role rather than being a facilitator.

However, at this point this preliminary analysis support the hypothesis that it is difficult to assess the more complex knowledge and skills in an individual exam due to the relatively short time where it is possible for the facilitator and examiner to interact with each student and students’ the lack of possibility to interact with fellow students. However, there is also a tendency that external examiners and students having experienced the individual exam seems much more positive towards the change to the individual assessment system – therefore it is also an interesting matter for further investigation to follow the opinions and perceptions of students and staff as they adjust to the system.

REFERENCES


