INTERNATIONAL CONFERENCE ON ENGINEERING AND PRODUCT DESIGN EDUCATION
2 & 3 SEPTEMBER 2010, NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY, TRONDHEIM, NORWAY

SYSTEMATIC GRADING PROCEDURE BASED ON SUBJECTIVE VALUES

Anders BERGLUND and Phillip TRETTEN
Division of Industrial Design, Luleå University of Technology

ABSTRACT
In Design education there can be a challenge in grading students when the task is based upon the student's ability to learn new knowledge and apply it. Thus the grading criteria needs to be and is most often subjectively focused. For example, one requirement is that, "the object should have realistic lighting". Since there is no way to, with absolute values, measure whether an image has realistic lighting, the teacher's subjective values are the base for assessment and the student's score. An observed result is that these types of assessments can vary between teachers, thus, increasing the chance for varied and possibly improper scores. Much has to do with the level of experience and knowledge a particular teacher has and the individual differences to which parameters make a good picture or animation. It is also true that two different teachers can find separate items in student's work that are praiseworthy. This paper proposes that it is possible to systematize the evaluation process with a weighting method. A Systematic Grading Procedure (SGP) can be used in design projects to separate and weigh design criteria against each other without losing the overall picture of the work. This method will be tested to see if it can help teachers in grading students more accurately.

Keywords: Assessing 3D-art, subjective values, systematic grading procedure

1 INTRODUCTION

The use of grades is a common measure to show how well a student has performed in a course(s). Schools and perspective employers also use grades as guidelines in judging an individual's capabilities. How well or how poorly a student performs in school can have long-term effects on future schooling and job opportunities, therefore it is very important that students are correctly judged in their courses. Research has shown that an unsatisfactory grading scheme can be the principal source for unreliable grades. [1] The final grade given is often a numerical value based upon how the student performed in relation to the level of work required and the other student's performances.
Students are often graded on a 5-point scale, in which, students course grades can be compared to each other and then ranked. This ranking, on a larger scale, is useful for determining which students are to be chosen for highly competitive programs and/or acceptance into the university system. From the overall administrative perspective this is a simple process, in which the students final grades are averaged and weighted based upon the difficultly of the courses, thereafter, the students with the highest average grades are given a higher priority in choosing high school and/or university programs.
In Sweden, two models of grading are used. In the first model are the grades based upon a letter system, failing (IG), passing (G), above average (VG). In the second model are the grades based on a number and letter system, U – failure, 3 - average, 4 – above average, and 5 – exceptional. These grades can be based on assessment of exam tests, lab work, reports, and assignments. In grading it is not uncommon that the student's performance is based on a scoring system which knowledge is measured as correct or incorrect, right or wrong. In addition to that are reports, written work, labs, etc. grades based upon a complete or incomplete grade. When the correct number of assignments is complete with passing grades then a final grade is given. To receive a better than average grade is the student's performance rated on the amount of correct answers and the amount of correct information reported in the assignments. Most often are the students graded upon objective measures and on occasion on more subjective measures. There are courses of artistic nature in which the student is not graded by a pass/fail standard or a clear objectively accessible grading scale. It is common that everyone who completes the task gets a passing grade although the scoring that determines the

"average" work from the "above average" has much to do with the teacher's knowledge and experience. That knowledge and experience is in many cases an intangible quality that cannot be specifically measured. In some instances what the teacher "feels" is better than another is very difficult to argue in point. Just as it is in the case of a 3D-generated image, it is difficult to determine what proportions are correct, how correct the imaging, and lighting are.

For the teachers help in these cases is the Biggs & Tangs Structure of Observed Learning Outcomes (SOLO) taxonomy useful [2]. This shows that by defining different levels of student achievement, the examiner's assessment of students work can be simplified and be made more tangible. The Biggs SOLO has five levels: Pre-structural, there information is acquired but not organized in understandable ways, *unistructural*, in which simple and obvious connections are made but not significantly grasped, *multistructural*, in which connections are made but meta-connections are missed, *relational*, in which the student is able to appreciate the significance of the parts in relation to the whole, and finally, *extended abstract* level, in which the student can make connections to other subject areas [2]. This is a good guideline in helping to grade the student's comprehension of knowledge.

To further help in the grading process CF Leung used the *Biggs SOLO taxonomy* to assess design and technology students. He argues that the design problems students face are complex and that there are any completely fail-safe methods for assessing students. [3] Elisabeth Ahlstrand stated that the student should know in advance what is examined, which aids are permitted, who examines, and how and when the examination is carried out, so high reliability can be ensured in subjective assessments. [4] This was stated to both help the students and teachers in the process so that both could use the same frame of reference when relating to the level of work being judged. This is also supported by Richard James, Craig McInnis and Marcia Devlin in their work "In Assessing Learning" in Australian Universities which states that well-designed and well-planned assessments are a strategic tool to clarify what learning is rewarded and even guide students into an effective approach to study [5]. Ponn, Kreimeyer and Lindemann conclude that a methodical approach to assessment is extremely useful documentation when feedback is given to students [6]. The template they used has been found easily applicable, making the rather subjective assessment more objective and comparable. These methods are a good starting point and do need refining for the teaching of 3D-generated images where both technical and artistic elements are graded.

In the courses for 3D-generated images there are no detailed presentation of how the final grade is set, thus, the student does not know how the image elements and criteria scored. It is also not uncommon that the students disagree with the examiner's assessment, which leads to small conflicts that arise when the examiner must defend the grading process. Even though, the examiner has the best and most relevant arguments some students do continue to disagree. At such times, a document presenting the arguments in a more detailed way could be useful. A method to systematize and standardize the grading of creative work is needed. The goal of this work is to facilitate a method that could help the examiner in grading the student's work, at the same time, helping the students understand what criteria are used and how the process is conducted.

Figure 1. Remote control modelled and visualized by a student

The aim of this paper is to show how a systematic grading procedure was tested for assessment of student's works in 3D-art. We also want to investigate whether it is possible to systematize grading based on subjective assessment and compare the grade distribution in previous courses with the outcome of a systematic grading procedure (SGP).

2 METHOD

This method was tested in a Computer Aided Industrial Design course in January 2010. The course teaches the students how to use the design tool to produce images with as high degree of photorealism as possible. The students were given six 3D assignments, where the first three are tutorial based and of a fundamental nature. These aim to provide students with sufficient knowledge to complete the remaining assignments. In the final three assignments the students were given the tasks to modelling and visualize: a *still life*, a *remote control*, and a *computer projector*. The final assignment has a different theme each time the course is offered. Beside the themes of the assignments the students have full creative freedom to design and compose the 3D models. The purpose of these final three assignments is; firstly, that the students should use and improve upon the knowledge gained in the three first assignments and, secondly, that the 3D rendered images are the basis for examination and thus graded accordingly.

The assignments where given fundamental assessment criteria for modelling, lighting and visualization. Each assignments objectives differ to some extent, meaning that the assignments where given slightly different assessment criteria, e.g., modelling is valued higher in the *remote control* assignment than in the other assignments, in the *still life* assignment is the composition of the image given greater weight, and in the *computer projector* assignment the student must produce a poster that presents the product in a selling way. Criteria can be given different weights, depending on the purpose of the specific assignment. The primary purpose of the *remote control* is that students should learn to use more advanced modelling tools and the secondary purpose is to produce a 3D image. Therefore, the modelling criterion is weighted 50% while the criteria for the materials and light are 25% each. The criterion used to grade the student's performance had both absolute measures and subjective measures. The absolute measures can be, i.e., "*Remote control* must have at least 9 keys" or " *Computer projector* must have split lines". Subjective criteria can be "Materials should give a realistic impression" or "The *Still life* should have a harmonic composition". An example of the grading procedure used for the *remote control* is shown below (Table 1).

The student's work was assessed in these two ways. If the absolute measures were incomplete a final grade, U, was given. With the subjective measures were the scores U, 3-5 given for each specific criteria. The Swedish grading system with four levels gives the examiner a high level of freedom, thus, the ratings had to be refined so grades could be given with greater precision in this method. So instead of just using integers decimal numbers were also used with increments of 0.5. The criteria scores were then multiplied with the criteria weight, which generated the specific grade for each assignment. Finally the assignments where also weighted so the *still life* was 25%, the *remote* 25% and the *computer projector* 50% of the final grade in the course. Thus, the student received the grade U, 3, 4 or 5, depending on the outcome of SGP-scheme (Figure 2 and table 1).

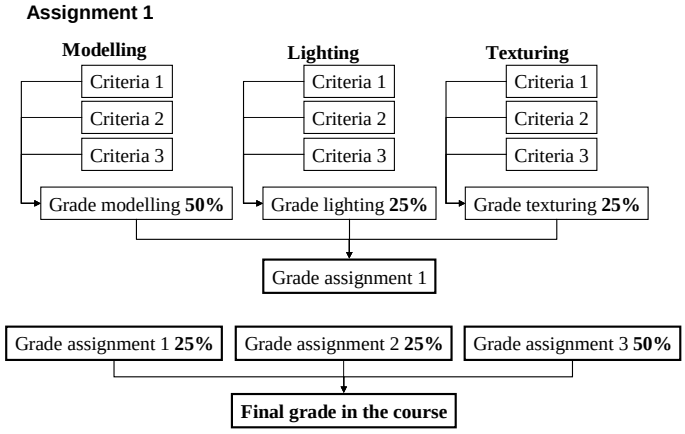

Figure 2. The flow of the SGP scheme

The students were presented the grading method before the final three tasks. Students who handed in the images for examination before the deadline received a grade determined by the SGP-method. Due to time constraints were only images handed in before the deadline included in this paper. The student's work process neither how quickly the student produced the images are considered. The SGP scheme model was developed in an analyzed in Microsoft Excel.

Table 1. Grading scheme used to grade the image in Figure 2

Remote control

Absolute criteria	At least 9 keys	Yes	
	Right size of image	Yes	
Modelling	The casting (grade 3-5)	5	
	The keys (grade 3-5)	4,5	
	Total	**4,75**	**50 %**
Lighting	Realism (grade 3-5)	5	
	Shadows (grade 3-5)	4,5	
	Total	**4,75**	**25 %**
Material	Materials (grade 3-5)	5	
	Texurering (grade 3-5)	5	
	Total	**5**	**25 %**
	Assignment grade	**4,8**	**100 %**

3 RESULTS & DISCUSSION

Overall, did 12 students participate in the study, of which their grades resulted in three 3's (25%), eight 4's (67%) and one 5 (9%). The ratings were reviewed after the use of SGP method and after consideration the examiners corrected the grade for a student from the grade 4 to the grade 5. All other final grades were considered properly assessed. Although the few number of students tested, the use of SGP scheme did not improve the distribution of grades but it give the examiners support for judging the grades and none of the students asked for more detailed information on their grades. This could be interpreted that the students experienced that the grades were satisfactory and the grading scale helped them to feel that they had greater control of their work.

The previous years results showed that the final grade, 3, has become the result of almost inadequate student work, while the grade 4 has become a sort of average score. The scores of the last four years were; 3's (23%), 4's (55%), and 5's (22%) (Table 2). This course is offered three times per year and the number of registered students each year has varied from 69-126. The variation of students through the years is due to curriculum restructuring which resulted in the course having twice as many students as usual during 2006-2007 and fewer during 2008-2009. The small amount of students being tested by the SGP method is due to the fact that only one course group was tested.

Table 2. The comparison between grades 2005 – 2009

Grade	2005-06		2006-07		2007-08		2008-09	
3	10	**12%**	29	23%	15	18%	28	**41%**
4	53	63%	65	52%	54	66%	28	40%
5	21	25%	32	25%	13	**16%**	13	19%
Total	84		126		82		69	

There is a risk that student's follow a SGP scheme rigorously and, thus, lose creativity and the overall composition of the work. The SGP scheme should not be presented to the student in the initial stage of the course, but is merely a tool for the teacher. Criteria should instead be presented to the students in a more aggregated form, which feels less structured. Much effort must be placed on developing appropriate assessment criteria. With incomplete or inaccurate criteria could the SGP scheme become so inefficient that it does not work as an assessment method. It is important that each criterion are assessed systematically so no weight is added on criteria already weighted in the SGP scheme, otherwise there is a risk that the rating is inaccurate due to double-weight. The range that gives the grade 3 and 5 is smaller (0.49) than for grade 4 (0.99) (Table 3).

Table 3. Grades range of rounding

3	4		5
3-3,49..	3,5-3,9..	4-4,49..	4,5-5

To overcome the skewed distribution of scores the scale can be extended 0.5 on both ends so the range of 2.5 – 3.499 .. gives the grade 3 and 3.5 – 4.499 .. gives the grade 4 and 4.5-5.5 gives grade 5. These extended intervals provide a more equitable distribution of grades (Table 4).

Table 4. The extended range of rounding

3	4	5
2,5-3,49..	3,5-4,49..	4,5-5,5

Criteria scores can be set with decimals which gives a more detailed rating. The final rounding occurs when the grades for each assignment are summarized. To rate the individual parts of an image or animation might not give the correct grade every time. Therefore the accuracy of the final grade must be considered so that the overall impression and how the image is composed are not missed. SGP does not provide an absolute measure of student's skills, but should probably be seen as a guide for grading. The main purpose of the SGP method is to grade students, but it also facilitates feedback to students,

as the SGP method provides a good basis for discussion. All students who received feedback on their work acknowledged the validity of the grade and credibility of the SGP method. This method gave them a good understanding of how the grades were set and they responded positively to the specific feedback given for each grade. The SGP method can also be used to quantify the student's results producing statistics that can be compared in a longitudinal study.

4 CONCLUSIONS

To draw accurate conclusions after only testing the SGP model once is difficult. However, the SGP model needs to be modified to work in the assessment of 3D art. A modified SGP model will be used in the course the next time it is given to test the extended grading range. Furthermore, the SGP model will also be tested in the course Basic Animation in spring 2010. It is important to spend effort on creating as relevant and useful assessment criteria as possible. Otherwise, the method will fail with incorrect grades as a result. The first impression of an image or animation is usually higher at the beginning of the assessment rather than later. When the components of the 3D art is broken up and assessed individually the overall impression tends to decrease. To capture that first impression of an image or animation a wow factor with 1.0, 1.1 or 1.2 will be applied to the SGP scheme. The final grade of each assignment will then be multiplied with the wow factor set at the very beginning of the assignment assessment. As Josef Ponn el al [6] says a standardized method for the systematic grading facilitates better results for new teachers. The teacher must, however, have much experience with 3D art so the overall impression of the student work won't be missed.

REFERENCES

[1] Meadows M. and Billington L. A Review of the Literature on Marking Reliability, 2005 (National Assessment Agency, UK)

[2] Biggs J. and Tang C. Teaching for Quality Learning at University (3rd edn), 2007 (SRHE and Open University Press Buckingham)

[3] Leung C. F. Assessment for Learning: Using Solo Taxonomy to Measure Design Performance of Design & Technology Students, International Journal of Technology and Design Education, 2, 2000, pp 153

[4] Ahlstrand E., Examination and assessment / teacher training, Quality Conference 10-11 October 2007 Umeå University.

[5] James R., McInnis C. and Devlin M., *Assessing Learning in Australian Universities*, Centre for the Study of Higher Education, www.cshe.unimelb.edu.au/assessinglearning. retrieved Mars 3 2010.

[6] Ponn J., Kreimeyer M. and Lindemann U., Methodical Evaluation of Single and Group Projects, *International Conference on Engineering and Product Design Education*, September 13-14, Northumbria University, Newcastle upon Tyne, United Kingdom, pp. 190.