

## An Evaluation of Criteria-Based Assessment and Grading in Architecture Design

<sup>1</sup>Nangkula Utaberta, <sup>2</sup>Badiossadat Hassanpour, <sup>3</sup>Mohd. Arsyad Bahar

Department of Architecture, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia (UKM), Malaysia

**Abstract:** The main objective of this study is to evaluate the basis of an assessment in the architectural design studio and try to provide a new assessment method that is more precise and objective. Determining the criteria-based assessment methods and grading is very important in architectural design. This aims to avoid a holistic assessment, subjective, introverted and less explicit. The criteria-based assessment and grading refers to the process of forming a decision about the quality and level of student achievement or performance in a transparent, truthful and fair. So it will not cause disputes and dissatisfaction that have an impact on the mental and spirit of the students themselves. Students deserve to know which of their works and under what kind of criteria will be assessed. Coupled with the views and assessment of a work of a judge with other judges would have a different perception, so as to avoid subjective judgments, it is necessary to create a standard model that regulates how the appropriate assessment criteria and helps the creativity of the students in the architecture studio. With the criterion-based grading in architecture and design, students will know what the criteria will be assessed and trying to reach assessment targets that must be fulfilled. Then the jurors will also be easier to determine the appropriate value based on the assessment criteria that have been previously defined.

**Keywords:** Architecture education, criteria-based assessment, grading

### INTRODUCTION

Nowadays Assessment is known as a part of education process and not as a distinct part. It has an important impact on learning process. Assessment of students' achievement is an important factor in encouraging students to adopt deep learning strategies (Utaberta *et al.*, 2011a). In order to improve education and student learning, assessment must be appropriately designed and implemented. It has long been observed that the evaluation in the architectural design studio UKM as in other departments, are using a standard value of the letters A, B, C and so on, where the letter A is the highest value and E is the lowest. But before getting into that value, there is not enough discussion on the assessment framework (Hassanpour *et al.*, 2011). Assessment in architecture design studio is more holistic and subjective (Utaberta *et al.*, 2011b). The current assessment standards include the availability of materials such as assignments, attendance, activity and creativity. Though each of these criteria has a sub basic valuation should be done objectively and regularly.

Most of the students felt aggrieved and think they have fulfilled all these requirements while still getting

bad grades. Then there are some students who subjectively have given good value and completeness of the task when they are almost same. Student frustration due to lack of transparency of the assessment procedure and previous communications.

In an educational institution, a very important factor is the determination of assessment standards. Assessment standards should be accepted and applied properly (Hassanpour *et al.*, 2010). With a system of assessment, curriculum and education will be achieved on target. It can also be used as a benchmark for the development of learning systems at each institution, majoring in architecture especially. Because the architecture majors are not the same as other majors are mostly based on an assessment benchmarks that have been established definitively.

Architecture has the correlation between art and engineering. Art based on the taste, whereas if the technique is based on mind or idea. Both are different things and sometimes there is a contradiction that it needs a proper learning system to incorporate both of them. The authors feel the need to write this evaluation as the starting point to increasing the quality of architectural education (Utaberta *et al.*, 2011c).

Design process in architectural studios is based on some small-small well defined projects during the semester and on final project at the end which is ill defined and in larger scale. Students should finalize their project before deadline and present it in submission day with proper documentation. In this day they have a chance to see other student's project and get the comments from peers and experts and finally they will get mark. Experiences show that students are worry about their grades insofar as they won't attend in discussions if they think their comments will affect grades and with small negative comments or finding fault in their project they get disappointed and loose other statements and suggestions coming after. Most of the student's complain is about the unfairness and inequitable of grades. This may rout in unawareness of the way they evaluate and graded (Utaberta *et al.*, 2011a).

On the other hand analysis shows that there is no common understanding of what grading process is in architecture and what occurs in faculties are just instructors experience from what their own professors did. This has inhabited high-quality discourse, research and development of grading system in architecture education. First of all, we have to investigate about the past and current implemented grading systems in architecture faculties to find the characteristics and attributes of idealistic grading systems Since different definitions of some terms related to the discussion are used differently in different countries and even within a single country, in different education sectors, finding an appropriate terminology to use in analysis of assessment and grading is essential. For instance, 'assessment' in some contexts in the USA refers to the evaluation of a wide range of characteristics and processes relating to higher education institutions, including entry levels, attrition rates, student services, physical learning environments and student achievements. In the UK, assessment can mean what students submit by way of project reports, written papers and the like as distinct from what they produce under examination conditions. Similarly, a 'grade' may refer to the classification of the level of a student's performance in an entire degree, the summary of achievement in a single degree component or the quality of a single piece of work a student submits in response to a specified task.

The main objective of this study is to initiate a discourse on how to evaluate and construct a basis of an assessment of the architectural design studio and propose better assessment method which is more precise and objective.

## **MATERIALS AND METHODS**

Evaluation of the assessment has been performed to the students of 2<sup>nd</sup> year architectural design studio,

Department of Architecture, the National University of Malaysia (UKM). Evaluations performed during the sessions of criticism and assessment collection presentation at the end of the task design.

Evaluation conducted in the studio by the jury of internal and external to the individual students who presented their work. The jury assessed based on the guidance and evaluation sheets which had previously been given by the lecturer and coordinator of the studio.

### **The importance of criteria-based assessment and grading models in architecture design studio:**

Assessment in this study refers to the process of forming a judgment about the quality and extent of student's achievement or performance. Such judgments are mostly based on information obtained by requiring students to attempt specified tasks and submit their work to instructors or tutors for an appraisal of its quality. Scoring and marking are used interchangeably in this study to refer to the processes of representing student achievements by numbers or symbols. Scoring includes assigning a number to reflect the quality of a student's response to an examination item. In most cases, scoring and marking apply to items and tasks rather than to overall achievement in a whole course (Sadler, 2005).

Grading refers to the evaluation of student achievement on a larger scale, either for a single major piece of work or for an entire course. Scores or marks often serve as the raw material for grade determinations, especially when they are aggregated and the result converted into a different symbolic representation of overall achievement (Sadler, 2005). Grading symbols may be letters (A, B, C, D, etc.) descriptive terms (such as Distinction, Honors, Credit, Pass, etc.), or numerals (such as 7, 6, ..., 1). Numerals are usually deemed to represent measurements and this provides a straightforward rout to the calculation of Grade Point Averages (GPAs). The other symbols need a table of numerical equivalents.

Students deserve to know which of their works and under what type of criteria will be assessed. This will enable students to shape their work appropriately during the design process and specifying the bases for grading help to provide a rationale for grading judgments after they have been made and the results given back to the students.

In all studio based educating systems such as architecture studios, we can find different grading models, which their principles may deduced from either the policy document or from accepted practice. One of these systems is comparative method. In this appraisal model the student's projects will compare with each other. In fact jurors or the related tutors that are going to give marks in submission day, judge the quality of

projects holistically then they rank the projects. Grades follow in descending form best project to worth one. This method is unfair. Students deserve to be graded on the basis of the quality of their work alone, uncontaminated by reference to how other students in the studio perform on the same or equivalent tasks and without regard to each student's previous level of performance.

In comparative system, the holistically attitude to the projects judgment leads to neglect Student's Creativity and abilities in some contexts. Students can't be aware of their weak and strong points and by this way and they can't do any effort to increase their marks and just lucky students who are skillful in graphic design are able to impact jurors for better grades. On the other hand making pair-wise comparisons just among small set of students submissions is possible. It will be very difficult in large amount of projects and students. Albeit this method is not objective based and we can know it as a subjective method, this method is still use by instructors all around the world.

In recent years, universities have made explicit overtures towards criteria-based grading and reporting. Under these models, grades are required to how well students achieve the juror's expectations. These expectations can be explain in different form. We name these expectations as course objectives. The objectives are assumed to provide the basis for the criteria, but exactly what the criteria are is in essence left undefined (Sadler, 2005). These objectives should be known by instructors, students and especially external jurors. Because invited jurors have their certain tendency and assumed objectives that would be the base of their grading. This incoherency may lead to variant in given marks by different instructors and students dissatisfaction.

One of the implemented methods under this way is grading system base on marking forms. These grading criteria sheets (Montgomery, 2002) typically do not map in any simple way into course objectives. They are scoring rubrics which shows some tasks and their marks portion. These tasks outline some of the knowledge and skills students ideally should be able to exhibit by the end of the course. For instance, 3D model and executive details, boards, oral presentation as tasks and 5 mark for each of them. The given mark is based on the quality of presented documentation. This holistic method cannot explain about the expected details in each task and will leave the doors open to enter the personal opinions and subjective decisions in evaluation. An underlying difficulty is that the quality of performance in a course, judged holistically on the basis of the quality of work submitted, may not be determinable well with the attainment of course objectives.

It is obvious that in all grading models transforming students work to marks, grades or scores is very difficult because whenever the projects encoded with symbols the connection between course objectives and projects has broken and after that just the grade exists and can show the student's success amount. This has lead to do many efforts to define and implement some norms and criteria in appraisal methods. Despite the broad desirability of criteria-based grading in educating systems to implement these methods and ways, there are different conceptions of what it means in theory and practice. This study is based on a review of the most common grading policies and will try to hybrid the criteria based models to introduce a new appraisal method in evaluating architectural projects in universities.

**Definition of criteria-based assessment and grading models:**

Since criteria are attributes or rules that are useful as levers for making judgments, it is useful to have a general definition of what criterion is. There are many meanings for criterion (plural criteria) but many of them have overlap. Here is a working dictionary style definition, verbatim from (Sadler, 2002) which is appropriate to this discussion and broadly consistent with ordinary usage (Sadler, 2002).

**Criterion (n):** A distinguishing property or characteristic of anything, by which its quality can be judged or estimated, or by which a decision or classification may be made. (Etymology: from Greek criterion: a means for judging.)

**Criteria based grading models:** Grading models may be designed to apply to whole course or alternatively on specific assessment tasks and some can be appropriate for both. For all grading models explained below, the interpretation of criteria is same with the general definition given above and all of them make a clear connection between the achievement of course objectives and given grades, without reference to other students achievements.

**Verbal grade description:** In this model, grades are based on student's achievement to the course objectives. In this form, the given grades are base on interpretations which clarify the attainment amount of course objectives Table 1. This kind of grading method is based on holistically attitude in evaluations.

**Objective achievements:** In this form the course objectives will be portioned into major and minor and the achievement of each can be determined by yes or No and the achievements of each objective will be

Table 1: Form of verbal grade description

Grades	Interpretation
A	Clear attainment of all course objectives, showing complete and comprehensive understanding of the course content, with development or relevant skills and intellectual initiative to an extremely high level.
B	Substantial attainment of most course objectives, showing a high level of understanding of the course content, with development of relevant analytical and interpretive skills to a high level.
C	Sound attainment of some major course objectives, with understanding of most of the basic course content and development of relevant skills to a satisfactory level.
D	Some attainment of a range of course objectives, showing a basic understanding of course content with development of relevant skills.

Table 2: Form of objective achievements

Grades	A	B	C	D	E
Major objectives achieved	All	All	Most	Some	Few or none
Major objectives achieved	All	Most	Some	Some	Few or none

computed (Sadler, 2005) Table 2. Both of these two objective base models make clear connections between the attachments of course objectives and the grades awarded but students can't easily ant close connection between the course objectives and assessment items and they are not in strong position to judge how much they reached to the objectives.

Therefore these types of models have little prospective value for students. Also there are no indications of whether given grades are for attainment in objectives of a special task or for whole objectives and it will be assessed by its own or in combination to other objectives.

Most educational outcomes and attainments amount cannot be assessed as dichotomous states like yes or no or zero and one, because learning is a continuous process that in contrast with discrete scales it can just be divided into segments satisfactory and dissatisfactory (Sadler, 2005).

**Qualitative criteria:** Teachers specify the qualitative properties as criteria to be closer to teaching and learning and assessment grading. In this method teachers are obliged to make a judgment about the quality of student responses to each assessment task and objectives.

In this model the grades are given in simple verbal scale for each task such as poor, acceptable, good and excellent. But since in reality student's works are not perfect and there are different descriptions for these verbal scales and some teachers believe that Excellent and A is just for god and no one deserve grade A, the distribution of grades and marks can't be appropriate.

In this model scores in different assessment tasks are added together and finally the 100 point scale may divided into segments according to the number of grades.

## RESULTS AND DISCUSSION

**Proposed criteria based model in architecture assessments:** All aforementioned methods have weak and strong points. For instance, first model has tried to avoid dispersion of interpretations for grades between different assessors which can affect the given marks. But there is no room for expected objectives and their definitions in design process and final projects. So doors of subjective judgment will be still open.

Second model is based on dividing the expected objectives into major and minor and the evaluation is completely related to the student's achievements to these objectives but as mentioned before it is not possible to judge about the attainments and achievements in continuum process just by yes or no.

In third form by introducing tasks as criteria for grading and verbal definitions for students achievements amount has improved two previous models but objectives and importance amount of them are still unclear for students and external assessors. So we have to hybrid these methods to reach the improved model.

What makes the definition of different projects (their scale, title, objectives) during architecture education is transmitting new knowledge and experience based on learned related topics, issues and projects in continues process of learning. So the aim of each project is unique to it and has different layers.

In all submission days, students prepare needed documentation such as sheets included plans, evaluations, sections, perspectives etc. and 3D models which may determine by instructors or leave arbitrary. But these are not just the things that are going to be assessed by jurors. Primary goals that were the basis of problem solving process are the most important part of assessment. So the criteria to be used in assessment and grading are linked directly to the way objectives expressed (Biggs, 1999).

Since this approach has some conceptual parallels with the behavioral objectives movement, according to Mager (1962), a behavioral objective is not properly formulated unless it includes a statement of intent, descriptions of the final behavior desired, the conditions under which this behavior is to be demonstrated and the minimum acceptable level of performance that signifies attainment of that objective.

Table 3: Form (c)

	Fail ----- Little or no evidence	Poor ----- Beginning	Average ----- Developing	Good ----- Accomplish	Excellent ----- Exemplary	Grade
Graphic presentation						10%
• Composition	---	---	---	---	---	
• Focus and explanation	---	---	---	---	---	
• How clear is the information	---	---	---	---	---	---
• .....	---	---	---	---	---	
• .....	---	---	---	---	---	
• .....	---	---	---	---	---	
Critical explanation						40%
• Process and idea development	---	---	---	---	---	
• Detail explanation	---	---	---	---	---	---
• .....	---	---	---	---	---	
• .....	---	---	---	---	---	
• .....	---	---	---	---	---	
• .....	---	---	---	---	---	
Logical development						30%
• .....	---	---	---	---	---	
• .....	---	---	---	---	---	
• .....	---	---	---	---	---	---
• .....	---	---	---	---	---	
Proposal and recommendation						20%
• .....	---	---	---	---	---	
• .....	---	---	---	---	---	
• .....	---	---	---	---	---	---
• .....	---	---	---	---	---	
• .....	---	---	---	---	---	
Final grade						----- 100

Defined architecture assignments, Depends on their type, scale and duration, have different objectives and expectations to assess the student’s submissions and different tasks are required. These tasks are based on some practical necessity and some personal standards aligned with course objectives. These tasks will create policies for assessors to intend to take into account in judgment. Eyeballing different evaluation sheets in variety of studios for different projects bring us to this result that the rubric of the tasks is as follow:

- Critical explanation
- Logical development
- Proposal and recommendation
- Oral and graphic presentation

The potential number of tasks relevant to the projects is large but these are enough to be illustrated and discussed in this study. For each rubric and task some criteria will be defined. Segregating evaluation extent to more tasks will increase student’s opportunities to show their capabilities and sufficiency and gain more chance to get better marks. But in contrast the more objectives are expressed for each

task, the more they will operate isolated and will recede from the overall configuration that constitutes a unit of what the students are suppose to do. In addition it will restrict assessors between these defined borders and will confine their authority and experiences in cognition and analyzing students hidden intends in their designing. This is completely in opposition with the main target of inviting external jurors which is benefit from diversity of expert ideas and critical attitudes. So characteristic of objectives are more effective that their numbers in defining flexible evaluation borders.

Since not all criteria types are same, there is no necessity for the number of criteria to be same in different tasks. In fact these are subtitles for what is expected from students to do and they elaborate the borders of course objectives for assessors. For instance in Table 3 we can see tasks with some of their criteria

Table 4: Form (d)

Grades	Major objectives achieved	Major objectives achieved
A	All	All
B	All	Most
C	Most	Some
D	Some	Some
E	Few or none	Few or none

which have defined by related instructor base on course objectives and implemented strategies in studio. Each of criteria is included in marking grid.

On the other hand according to main focus of education process in certain period, different priorities with different attention portion will be dedicated to each objective. This kind of precedence will import to assessment criteria and evaluation sheets (Montgomery, 2002). Therefore each task would have dedicated percentages to show the major and minor objectives and grade amount. Table 4 illustrates this type of grading model.

Since students perform in continuous path, the result of their performance just can be revealed in continuum that can be divided between satisfactory and dissatisfactory. Student's locus this vector derives from quality of their work in response to defined criteria in each task. So it is needed to define some qualitative levels to apply as a norm to the assessment. Descriptions should have the best overall fit with the characteristics of the submitted projects. The assessor does not need to make separate decisions on a number of discrete criteria, as is usual list form. Such as little or no evidence, beginning, developing, accomplish, exemplary.

However these descriptions are very helpful and effective in appraisal system but finally the qualitative assessment should be able to be transmitted into grades and marks. So we need to coordinate this model to one of the common grading system. As we mentioned before, using grading systems such as (1-100) or (A, B,..) are not appropriate ways to import to criteria based assessment model because after transmitting students work to numerical grades the connection between course objectives and grades will be completely broken. Since marks and grades do not in themselves have absolute meaning in the sense that a single isolated result can stand alone as an achievement measurement or indicator that has a universal interpretation.

Assessment and grading do not take place in a vacuum. Quality of student's work together with interpretations of such judgments can be known as comprehensive model in judgments. So alternatively, a simple verbal scale could be used for each criterion such as Fail, Poor, Average, Good and Excellent but in this type verbal grade description applies to given assessment task, with a separate description for each grade level (as mentioned before). So each list of criteria can be elaborated into a marking grid. Finally components of grades will be weighted before being added together to reflect their relative importance in the assessment program.

There are several forms to show the final grades. The simplest is a numerical rating scale for each criterion, in which case the ratings could be added to arrive at an overall mark or grade for the work. Using numerical ranges gives the impression of precision and the system is easy to make operational.

Introduced model contains most of the strong points of other criteria based models and none criteria base models. These strong points are revealed in Fig. 1. This method does not depend on ranking or sorting student's projects. It means there is no explicit reference to other student's performance. But final grades are assigned by determining where each student stands in relation to others.

Also since this model is completely base on course objectives and instructor's expectations and strategies in conducting the project, it makes opportunities for instructors to discuss and criticize their implemented methods in teaching and defining assignment and their objectives. This may lead to improvement in education level.

Although judgments can be made either analytically (that is, built up progressively using criteria) or holistically (without using explicit criteria), or even comparatively, it is practically impossible to explain a particular judgment, once it has

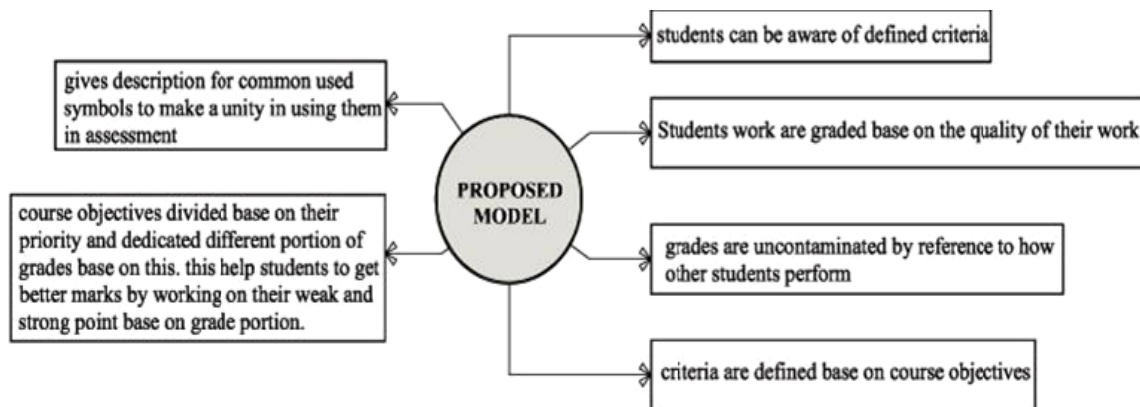


Fig. 1: Proposed model

been made, without referring to criteria. So it is needed to investigate about all evaluation and assessment methods and find used criteria and hybrid their potentials to current methods and upgrade the existing models.

## CONCLUSION

Evaluation and grading system in art and architecture and especially in their studio-based courses are more difficult than other majors and field. Since their teaching and learning process are different and more complicated than theory courses, it is admissible. But there is common thought that believes there is no criterion and norm in their grading and assessing system, in the other word the grading system is holistically and subjective. This statement also is not incoherent. There is no special criteria and norm among jurors and instructors in evaluating and grading student's project and if they have it is not known and explained to students. Students themselves are inducted directly into the processes of making academic judgments so as to help them make more sense of and assume greater control over, their own learning and therefore become more self-monitoring.

In recent years, more and more universities have made explicit overtures towards criteria-based grading to make assessment less mysterious and more open and more explicit. But whenever there is no discussion and contribution, there is no way to improve and development in this model and many institutions may employ identical or related models without necessarily calling them criteria-based. A further framework can be self-referenced assessment and grading, in which the reference point for judging the achievement of a given student is that student's previous performance level or levels. What counts then is the amount of improvement each student makes.

## REFERENCES

Biggs, J., 1999. *Teaching for Quality Learning at University: What the Student Does*. UK, SRHE and Open University Press, Buckingham.

- Hassanpour, B., N. Utaberta, M.M. Tahir, N.A.G. Abdullah, N. Spalie and A.I. Che-Ani, 2010. Developing sustainable approach in architectural education: Lesson from critique session experience of architecture studio program. 9th WSEAS International Conference on Education and Educational Technology (EDU '10), Iwate Prefectural University, Japan, October 4-6, pp: 325-330.
- Hassanpour, B., N. Utaberta, A. Zaharim and N.A.G. Abdullah, 2011. Student's preception of the evolution system in architecture studios a case study of 2nd year studio in architecture. *J. ICES World Acad. Sci. Eng. Technol.*, 77: 885-891.
- Mager, R.F., 1962. *Preparing Instructional Objectives*. CA Fearon, Belmont.
- Montgomery, K., 2002. Authentic Tasks and Rubrics: Going Beyond Traditional Assessments in College Teaching. *College Teaching*, pp: 34-39.
- Sadler, D.R., 2002. Ah!... so that's 'Quality. In: P. Schwartz and G. Webb (Eds.) *Assessment, Case Studies, Experience and Practice from Higher Education*. Kogan Page, London.
- Sadler, D.R., 2005. Interpretations of criteria-based assessment and grading in higher education. *Assessment Educ. Hig. Educ.*, 30: 175-193.
- Utaberta, N., B. Hassanpour and S. Reza, 2011a. Investigation in effective assessment models in architecture design studios national university Malaysia as case study. *Aust. J. Basic Appl. Sci.*, 5(9): 571-577.
- Utaberta, N. and B. Hassanpour, 2011b. Authentic Assessment as a Sustainable Approach In Critique Session of Architecture Design Studio: A Case Study of 2nd Year Architecture Student of Universiti Kebangsaan Malaysia. *Proceedings Environmental Talk: Toward A Better Green Living*. Mercu Buana University Jakarta Indonesia, March 9-13, pp: 46-57.
- Utaberta, N., B. Hassanpour and A. Zaharim, 2011c. A comparative study on traditional design process with computer based second year design studio in National University Malaysia (UKM) as case study. *Proceedings of the 10th WSEAS International Conference on Education and Educational Technology (EDU '11)*, Penang, Malaysia, October 3-5, pp: 224-227.