What Makes Students Succeed in Performing Assessment Tasks? Evidences from a Physical Science Class[1]

Eddie G. Fetalvero[2]

ABSTRACT

Successful learning is brought about by an interaction of a lot of factors. An important factor is what the learners already possess before a new learning experience begins. This paper presents evidences from a physical science class about the significant learner-related factors that could explain why students are successful or unsuccessful in performing assessment tasks. Self-assessment rubrics of teachers are included as tools in checking the students' concepts and understanding while assessment and learning are going on simultaneously. It also addresses some important issues in selecting assessment tasks; designing cognitive structures to ensure that learning can occur while students are being assessed; and developing, validating and testing the reliability of rubrics.

Key words: *Alternative Assessment, Performance Assessment, Rubrics Assessment, Scientific Attitude, Self-Concept and Learning Styles*

---

[1] Paper presented during the 2010 International Conference in Science and Mathematics Education held at the University of the Philippines National Institute for Science and Mathematics Education (UP NISMED) on October 28-30, 2010.

[2] MA in Science Education; Director for Research, RSU.

BACKGROUND OF THE STUDY

Assessment of student performance is emerging as a crucial ingredient in the recipe for on-going improvement of school science. As programmatic change is occurring, there is a need to align student assessment practices with curricular aims, instructional practices, and performance standards. In the words of Iris Carl (1993), "What we teach must be valued; what we test is what must be taught."

Before considering alternative approaches to assessing student performance, it is important to consider the various functions that assessment serves. Various reasons for assessing student performance have been described in both specific and general terms, with distinctions being made between assessment for reporting purposes and for purposes of diagnosis and program evaluation.

On assessment in the service of instruction, its purpose is to help students, teachers and parents monitor learning (Haury, 1993). Assessment plays an important part in the learning process, having both formative and summative aspects. Formative assessment involves the use of assessment as a diagnostic tool so teachers may appropriately cater to the individual needs of their students and so students can determine their areas of strengths and weaknesses, celebrating their strengths and giving greater attention to improving their weaknesses. Summative assessment is used to report progress, for certification, for accountability or for monitoring (Caygil & Eley, 2001).

Performance Assessment

Performance assessment is one of the newest forms of testing that requires students to perform a task rather than select an answer from a ready-made list. Rudner and Boston as cited by Wangsatorntanakhun (n.d) defined performance assessment as a continuum of assessment formats which allows teachers to observe student behavior ranging from simple responses to demonstrations to work collected over time.

As opposed to most traditional forms of testing, performance assessment does not have clear-cut or wrong answers. Rather, there are degrees to which a student is successful or unsuccessful. This can be accomplished by creating rubrics.

Airasian, (1991) Popham, (1995) and Stiggins, 1994) were cited by Brualdi (2000) on their definition of rubric as a rating system by which teachers can determine at what level of proficiency a student is able to perform a task or display knowledge of a concept. With rubrics, the different levels of proficiency for each criterion can be defined. Scoring rubrics are descriptive scoring schemes that are developed by teachers or other evaluators to guide the analysis of the products or processes of students' efforts (Brookhart, 1999). Scoring rubrics are typically employed when a judgment of quality is required and maybe used to evaluate broad range of subjects.

In the article by Heidi Goodrich Andrade published in American Leadership in 1999, he defined rubric as a scoring tool that lists the criteria for a piece of work or what counts. It also articulates gradations of quality for each criterion, from excellent to poor. The term defies a dictionary definition, but it seems to have established itself.

Types of Rubrics

There are many types of scoring rubrics. An analytic scoring rubric allows for the separate evaluation of each independent criterion scored on a different descriptive scale. But when there is an overlap between the criteria set, a holistic rubric is preferable. In this type of scoring rubric, the criteria are considered in combination on a single descriptive scale which supports broader judgments concerning the quality of the process or the product (Brookhart, 1999). Scoring rubric may either be general or task specific. General scoring rubrics are designed to evaluate broader category of tasks while task specific scoring rubrics are designed to evaluate students' performance on a single-assessment event. However, scoring rubrics may contain both general and task-specific components (Moskal, 2000). Wiggins (1993) as cited by Brualdi (1993) has cautioned that not all hands-on activities can be used as performance-based assessments. Stix (1997) as cited by Brualdi (1993) further stressed that in constructing rubrics, the varying levels of proficiency must be properly communicated. This may be done by using impartial words instead of numerical or letter grades.

Why Rubrics Appeal to Teachers and Students

Rubrics appeal to teachers and students for many reasons. First, they are powerful tools for both teaching and assessment. Rubrics can improve student performance, as well as monitor it. By making teachers' expectations clear and by showing students how to meet these expectations, the result is often marked improvements in the quality of student work and in learning (Marcus, 1995 as cited by Andrade, 1997).

A second reason that rubrics are useful is that they help students become more thoughtful judges of the quality of their own and others' work.

When rubrics are used to guide self and peer-assessment, students become increasingly able to spot and solve problems in their own and one another's work. Repeated practice with peer assessment and specially self-assessment, increases students' sense of responsibility for their own work and cuts down on the number of "Am I done yet?" questions (Andrade, 1997).

Third, rubrics reduce the amount of time teachers spend evaluating student work. Teachers tend to find that by the time a piece has been self and peer assessed according to a rubric, they have little left to say about it. When they do have something to say, they can often simply circle an item in the rubric, rather than struggling to explain the flaw or strength they have noticed and figuring out what to suggest in terms of improvements. Rubrics provide students with more informative feedback about their strengths and areas in need of improvement (Andrade, 1997).

Fourth, teachers appreciate rubrics because their accordion nature allows it to accommodate heterogeneous classes. Finally, rubrics are easy to use and explain (Andrade, 1997).

Rubrics are becoming increasingly popular among educators moving towards more authentic, performance-based assessments. Andrade suggested some steps in rubric design process among students to boost the learning leverage of rubrics: look at models, list criteria, articulate gradations of quality, practice on models, use self and peer assessment, revise and use teacher assessment using the same rubric students used in assessing their work.

In 2006, a comprehensive semester long investigation on the antecedents of performance in rubrics assessment among physical science students was conducted at Romblon State College Main Campus in Odiongan, Romblon.

OBJECTIVES

This study was conducted to determine the factors that explain performance in assessment tasks among physical science students at Romblon State University. Specifically, the study sought to:

- Develop assessment tasks in physical science, determine the validity and reliability of rubrics, and use rubrics in determining student performance;
- Determine the factors that explain the performance in assessment tasks; and

- Propose a novel approach in the construction and use of rubrics

This study is anchored on Jerome Bruner's Constructivist Theory (1966). In Bruner's theoretical framework, learning is viewed as an active process in which learners construct new ideas or concepts based upon their current or past knowledge. The learner selects and transforms information, constructs hypotheses, and makes decisions, relying on a cognitive structure to do so. Cognitive structure provides meaning and organization to experiences and allows the individual to "go beyond the information given".

In this study, students performed tasks based on agreed criteria as to what levels of performance were expected of them. Most tasks were open-ended. Students' skills and intelligences were demonstrated in the way they wanted to construct their own understanding of the scientific concepts in physical science. These levels of understanding were reflected in their outputs.

As far as instruction is concerned, a constructivist instructor tries and encourages students to discover principles by themselves. The instructor and students engage in an active dialogue. The task of the instructor is to translate information to be learned into a format appropriate to the learner's current state of understanding.

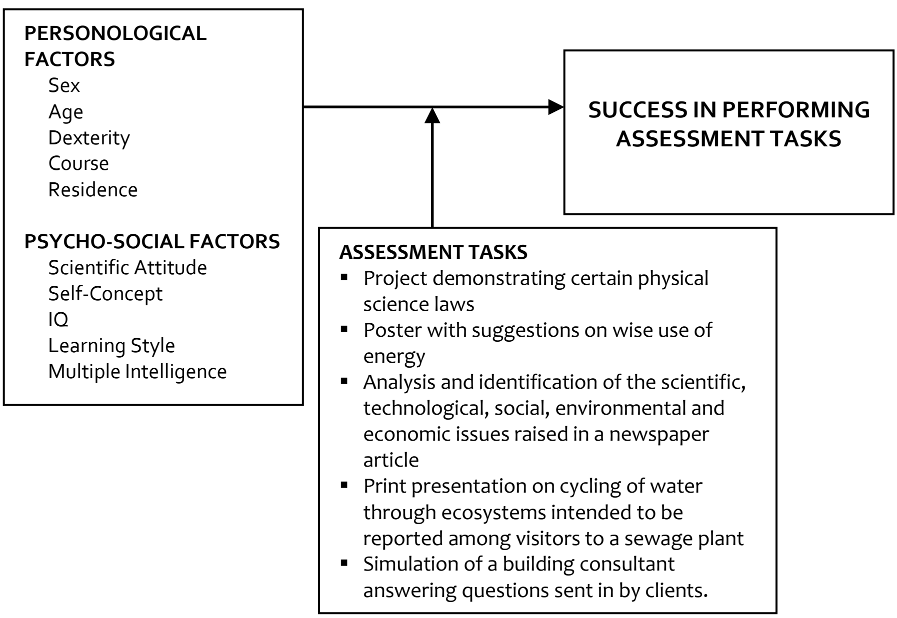


Figure 1. Conceptual Framework

One of the offshoots of the aforementioned theory was the birth of performance assessment. Performance assessment is an umbrella term that embraces both alternative assessment and authentic assessment. Alternative assessment was coined to distinguish it from what it was not: traditional paper-and-pencil testing - assessments which are meaningful in an academic context while "authentic" assessment are those which have meaning and value in the context of the real world (Rudner & Boston as cited by Wangsatorntanakhun). From the outset, one thing must be made clear, assessment encompasses more than testing, and much more than standardized testing.

RESEARCH INSTRUMENTS

The variables included in this study were all student variables which are categorized into three: personological factors which include sex, age, dexterity, course and residence; psycho-social factors which include the respondents' scientific attitude, self-concept, IQ, learning style and multiple intelligence; and aptitude factors which include their science and over-all ratings in the college admission test.

Inventory instruments and tests were used to measure some of the antecedents. The following validated instruments were adopted with permission from authors in measuring some of the variables included in this study.

Salmorin's Scientific Attitude Scale. This 47-item Likert-scale instrument developed by Dr. Lolita M. Salmorin, contains positive and negative statements where students were asked to indicate their reaction indicative of their scientific attitude like curiosity, questioning attitude, believing in cause-and-effect relationship, open-mindedness, respect for evidence, honesty, humility, patience and determination, resourcefulness and creativity, and intellectual responsibility. Answer choices were strongly agree, agree, disagree and strongly disagree with their corresponding numerical values of 4, 3, 2 and 1 respectively.

Self-Concept Test. This 36-item dichotomous test developed by Tan in 1991, measures the student's self-concept which includes the student's self-image, self-confidence and self-esteem.

Manila Self-Administering Test of Mental Ability. This 80-item multiple choice test is used in determining the student's intelligence quotient (IQ). The instrument measures the student's verbal, abstract, numerical and logical abilities.

Learning Style Inventory. This is adapted from Barsch Learning Style Inventory by Jeffrey Barsch, Ed D. and Sensory Modality Checklist by Nancy A. Haynie, which are both based on Dunn & Dunn Physiological Element of Learning Style Model specially on Perceptual Modalities. This determined the learning styles of the respondents as to visual, auditory, tactile and varied or multiple.

Multiple Intelligences Inventory. This instrument was developed by Walter McKenzie and was used in determining the dominant intelligence of the respondents based on Howard Gardner's Theory of Multiple Intelligences which includes: naturalist, musical, logical, existential, interpersonal, kinesthetic, verbal, intrapersonal, and visual.

DEVELOPMENT OF ASSESSMENT TASKS

Task Identification. The assessment tasks were identified by carefully studying the list of topics in the Physical Science course syllabus. A specific task was matched with a particular topic in the syllabus. The task list was presented to five Physical Science instructors at Romblon State University in a brainstorming session. After incorporating their comments, suggestions, and recommendations, the task list was again shown them for the identification of each of their top 10 preferred tasks with consideration to the science skills that students would develop in the process. Their responses were tallied and the tasks with the most number of responses were considered for selection. Table 1 shows the assessment task per representative topic.

Table 1. The Assessment Tasks

Assessment Task	Representative Topic	Category
1. Project demonstrating certain physical science laws	Introduction to Physical Science (Scientific Method)	Displayed Medium
2. Poster with suggestions on wise use of energy	Physics (Energy and Its Forms)	Displayed Medium
3. Analysis and identification of the scientific, technological, social, environmental and economic issues about an oil spill raised in a newspaper article	Chemistry (Physical and Chemical Change and Basic Chemistry of Petroleum)	Written Essay

Assessment Task	Representative Topic	Category
4. Print presentation on cycling of water through ecosystems intended to be reported among visitors to a sewage plant	Meteorology (The Water Cycle and Sewage Treatment)	Displayed Medium
5. Simulation of a building consultant answering questions asked by clients	Geology (Classification and Properties of Rocks)	Written Essay

Development of the Assessment Tasks and Rubrics. The tips on task and rubric preparation from various literatures read were considered in framing the mechanics of each of the five tasks. Each task has six (6) major components: nature of the assessment task, objectives, background learning, science skills, tasks which include the lead-up activities and the assessment task itself, and the rubrics for student and teacher's use.

VALIDITY OF RUBRICS

The mechanics of the assessment tasks and rubrics were validated by four recognized experts in the field of science education: a high school principal, a department head, the Vice President for Academic Affairs and the Dean of RSU graduate studies. They validated the assessment tasks and rubrics using the criteria suggested by Religioso (2002). The experts unanimously agreed that the criteria below be the basis for the development of student and teacher rubrics.

1. It relates to the outcome being measured.
2. It covers important dimensions of student performance.
3. Expected quality reflects current conceptions of excellence in the field.
4. The indicators of student performance (scale points) are well defined.
5. There is a basis for assigning scores in each scale point.
6. It can be used consistently by different scorers.
7. It can be understood by the students.
8. It can be applied to a variety of tasks.
9. It is fair and free from bias.
10. It is useful, feasible, manageable and practical.

RELIABILITY OF THE TEACHER RUBRICS

The reliability of the rubrics defined in this study is the inter-rater's reliability. This is the degree of agreement in scoring between two or more raters who used the same rubrics in assessing the task accomplished by the students. The twenty-five samples of students' work from the researcher's science class were rated by the researcher himself and another physical science teacher using the teacher rubric.

Table 2 . Inter-Rater's Reliability of the Teacher Rubrics (n = 25)

Teacher Rubric	r value	Interpretation
1. Project demonstrating certain physical science laws	0.95	Excellent
2. Poster with suggestions on wise use of energy	0.91	Excellent
3. Analysis and identification of the scientific, technological, social and environmental issues about an oil spill raised in a newspaper article	0.85	Good
4. Print presentation on cycling of water through ecosystems intended to be reported among visitors to a sewage plant	0.92	Excellent
5. Simulation of a building consultant answering questions asked by clients.	0.90	Good

The scores given by the two teachers to the 25 work samples were analyzed using the *Pearson's r* correlation. Results of reliability testing are shown in Table 2.

In interpreting values for reliability analysis, a rule of the thumb that applies to most situations indicates that an r value of >.9 is an excellent instrument, >.8 is a good instrument, >.7 is an acceptable instrument, >.6 is a questionable instrument, >.5 is a poor instrument, and <.5 is an unacceptable instrument (George & Mallery, 2000).

ADMINISTRATION OF THE ASSESSMENT TASKS

Five assessment tasks which are representative of each of the major components of Physical Science were carried out by the respondents. They were provided with the mechanics of each task as well as the lead-up activities. The student and teacher rubrics were discussed among them. The ratings of the students are valued in this study since self-assessment offers reliable and valid strategy in assessing students (Butcher and Stefani as cited by Kilic, 2003). The student rubric served as an immediate feedback generator of performance and comprised 40 percent of their rating in each task. On the other hand, the score given by the teacher using the teacher rubric comprised the remaining 60 percent. The following sections described how the tasks were administered to the respondents.

Assessment Task 1. The first assessment task required the respondents to prepare a project that demonstrated certain physical science laws. As a lead-up activity, the class was grouped into five and given a problem to be solved using the scientific method. The solutions that they had come up with will be the presented to the class after a given time. Other lead-up activities are discussions on the list of discoveries and significant events in science, science superheroes from the prehistoric date up to present, Filipino scientists and their contributions in science and technology After carrying out the lead-up activities, the assessment task was then introduced. Their outputs were rated using the student and teacher rubrics which contain the following parameters: science concept and understanding, aesthetic appeal and creativity.

Assessment Task 2. The second assessment task required the respondents to prepare a poster with suggestions on wise use of energy at home. Two activities were done before the assessment task was carried out. The first activity enabled the respondents to identify the different forms of energy depicted in a certain picture. A game called 'energy cards' was also played. This game strengthened the concept of the respondents in identifying the form of energy used by the appliance that was drawn on a card. The second activity was conducted to determine whether the different forms of energy used wisely around the home of the respondents. They were given a form to be filled up. This activity required them to observe, describe, record and look for patterns of use of the appliances they have at home. The data gathered from this activity were used in performing the assessment task. Their posters were rated based on the rubrics containing the following parameters on expected quality: science understanding, making sense of information, applying understanding and communicating understanding.

Assessment Task 3. The third assessment task required the respondents to analyze and identify the scientific, technological, social, environmental and economic issues about a major oil spill raised in a newspaper article. Before this task was administered, the class was divided into groups. They were given a copy of the newspaper report about a major oil spill. This was discussed in the group level and various issues were identified from the report. Issues that needed further clarification and further investigation were marked. There were series of activities performed: investigating the physical and chemical properties of oil; simulating an oil spill; and investigating possible methods of cleaning it up. After doing these activities, the assessment task was introduced to them. Their experiences in the activities performed served as inputs so that the issues that needed further clarification would be answered. Their written works were rated based on the following parameters on expected quality: identification of issues, science knowledge and understanding, application of understanding, and communication of information and understanding

Assessment Task 4. The fourth assessment task called for the respondents to prepare a print presentation about water cycle intended to be reported among visitors to a sewage plant. Before the task was carried out, a grouped activity was performed investigating the change of state in water in terms of the behavior of water particles associated with energy input and energy loss. Their understanding of the change of state was applied to a natural phenomena - the water cycle. A diagram was studied and the factors that could likely interrupt or affect the water cycle were enumerated and discussed. After performing those activities, the assessment task was explained to them. Their print presentations were rated according to the subsequent parameters on expected quality: science concept and understanding, applying understanding, and communicating information. Indicators of performance were expert, developing and beginner.

Assessment Task 5. The fifth assessment task required the respondents to simulate a building consultant answering questions asked by clients. Prior to the task administration, activities were carried out by the respondents. Through experiments, they investigated some physical properties of different kinds of rocks. They also simulated the action of chemical weathering on different types of rocks. The knowledge they learned from these activities served as inputs in preparing the assessment task. Four questions were asked by the clients and the respondents gave their advices based on the lead-up activities. Their works were rated based on the following parameters on expected quality: science concept and understanding, applying understanding and communicating information.

The indicators of performance for the five (5) assessment tasks are the following: beginner, developer and expert.

RESPONDENTS OF THE STUDY

The assessment tasks were administered to 88 students of Romblon State University taking up physical science subjects chosen through incidental sampling, a purposive process of selecting samples based on their availability and researcher's control to ascertain their participation until the completion of the study.

RESULTS AND DISCUSSION

Performance on Assessment Tasks

The table below presents the performance of the respondents in doing the assessment tasks based on the student and teacher rubrics used. Result of the first assessment task appears in Table 3.

Table 3. Student Performance in the Five Assessment Tasks

Performance Indicators	Self Rating		Teacher Rating		Joint Rating	
	F	%	F	%	F	%
TASK 1						
Expert	41	46.6	36	40.9	42	47.7
Developing	47	53.4	35	39.8	42	47.7
Beginner	0	0	17	19.3	4	4.5
Total	88	100	88	100	88	100
TASK 2						
Expert	48	54.5	17	19.3	19	21.6
Developing	39	44.3	61	69.3	62	70.5
Beginner	1	1.1	10	11.4	7	8.0
Total	88	100	88	100	88	100
TASK 3						
Expert	15	17	14	15.9	13	14.8
Developing	67	76.1	37	42	56	63.6
Beginner	6	6.8	37	42	19	21.6
Total	88	100	88	100	88	100
TASK 4						
Expert	19	21.6	20	22.7	20	22.7
Developing	66	75	49	55.7	53	60.2
Beginner	3	3.4	19	21.6	15	17
Total	88	100	88	100	88	100
TASK 5						
Expert	22	25.0	20	22.7	18	20.5
Developing	62	70.5	54	61.4	60	68.2
Beginner	4	4.5	14	15.9	10	11.4
Total	88	100	88	100	88	100
OVER-ALL PERFORMANCE						
Expert	19	21.6	7	8.0	22	25.0
Developing	69	78.4	77	87.5	62	70.5
Beginner	0	0	4	4.5	4	4.5
Total	88	100	88	100	88	100

Overall Performance. As reflected in Table 3, 62 or 70.5 percent of the respondents were developers, 22 or 25 percent were experts and 4 or 4.5 percent of the respondents were beginners. Difficulty of the respondents in expressing themselves in the English language and the heterogeneity of the class influenced to some degree the students' achievement in the five assessment tasks as shown by their outputs.

Factors Explaining Student Performance

A factor analysis was employed to explain which student-related factors accounted for the demonstrated performance the students in the five assessment tasks given them. Table 3a shows the researcher's description/label of each factor.

Table 3. Factor Analysis of the Antecedent Variables

	COMPONENT					
	1	2	3	4	5	6
FACTOR 1						
Intrapersonal Intelligence	**.757**	.125	.028	.049	.112	-.021
Logical Intelligence	**.747**	-.166	-.028	.002	-.169	-.217
Visual Intelligence	**.740**	-.015	.107	.257	-.054	.222
Verbal Intelligence	**.736**	.001	-.198	.094	.006	-.025
Interpersonal Intelligence	**.729**	.052	.249	.004	.067	.006
Existential Intelligence	**.703**	-.152	.083	-.047	.112	.233
Naturalistic Intelligence	**.700**	-.148	-.017	.040	.235	-.030
Musical Intelligence	**.675**	-.102	.078	-.049	-.101	-.214
Kinesthetic Intelligence	**.624**	-.154	-.086	.284	.108	.348
FACTOR 2						
ROSCAT Rating	-.029	**.885**	-.046	-.078	-.124	-.024
ROSCAT Science Score	-.078	**.799**	-.048	.010	-.072	.091
Mental Ability	-.083	**.678**	.195	.212	.260	-.008
Course	-.174	**.635**	.181	-.195	.185	-.091
FACTOR 3						
Sex	-.011	-.058	**-.819**	-.120	.209	-.099
Self Concept	.120	.360	**.664**	.105	.150	-.278
Age	.120	-.082	**.538**	.052	-.028	.409
FACTOR 4						
Auditory Learning Style	.073	.046	.246	**.735**	.043	-.129
Tactile Learning Style	.020	-.234	.145	**.692**	.103	.035
Visual Learning Style	.193	.232	-.333	**.642**	-.065	.177
FACTOR 5						
Residence	-.065	.059	-.005	.010	**-.780**	-.089
Scientific Attitude	.135	.314	-.223	.183	**.632**	-.176

FACTOR 6						
Handedness	-.041	.033	.047	-.012	-.002	**.790**

Extraction Method: Principal Component Analysis
Rotation Method: Varimax with Kaiser Normalizati

Table 3a. Factor Label / Description

FACTOR #	Description
Factor 1	Multiple Intelligence
Factor 2	Intellectual Competence
Factor 3	Idiosyncratic Factor
Factor 4	Learning Style
Factor 5	None
Factor 6	None

Regression Analysis

To determine the factor that reflects best student performance in the five assessment tasks, the stepwise regression analysis was run and produced four regression models. The fourth model with its characteristics shown in Tables 4, 5 and 6 was considered.

Table 4 shows the four predictor variables that meet the entry requirement to be included in the regression equation (scientific attitude, self-concept, RSUCAT Rating, and tactile learning style). The B values are the non-standardized regression coefficients of the independent variables and the constant for the regression equation that measures predicted values of student performance. B may be thought of as a weighted constant that describes the magnitude of influence of a particular predictor variable on the criterion variable. A positive value for B indicates an increase in the value of the criterion variable, a negative value for B a decrease. The regression equation for a student performance in the five assessment tasks in physical science is shown below.

$$\text{Performance in Rubrics Assessment}_{\text{(predicted)}} = -0.36 + 0.443(\text{scientific attitude}) + 0.024\ (\text{self concept}) + 0.014\ (\text{tactile learning style}) + 0.010\ (\text{RSUCAT rating})$$

This could be interpreted that in every unit increase in the scientific attitudes of respondents, a corresponding 0.443 increase occurs proportionately in their performance, provided that the values of the other three variables like self-concept, RSUCAT rating and tactile learning style remain constant. The same interpretation is true for the other three variables.

Table 4. Regression Analysis of Student Performance in Five Assessment Tasks

	B	Sig.
(Constant)	-.136	
Scientific Attitude	.443	.001
Self-Concept	.024	.030
RSUCAT Rating	.010	.003
Tactile Learning Style	.014	.022

Tables 5 and 6 present the regression model summary and the significance of the regression model.

Table 5. Regression Model Summary

R	R Square	Adjusted R Square
0.572	0.327	0.295

Table 6. Significance of the Regression Model

	Sum of Squares	Df	Mean Square	F	Sig.
Regression	2.517	4	.629	10.096	.000
Residual	5.174	83	.062		
Total	7.692	87			

As shown in Table 5, multiple correlation (R) shows a moderate relationship between the four predictor variables taken as one and the criterion variable which is the students' performance in rubrics assessment (R = .572). The R-Square value (0.327) indicates that about 32.7 percent of the variance in student performance is explained by the four predictor variables. The R-square or the multiple coefficient of determination is the proportion of variance in students' performance that is explained by the combined influence of the four predictor variables that entered the regression equation, namely, scientific attitude, self-concept, RSUCAT Rating, and Tactile Learning Style (Table 4). The adjusted R square is for population estimate purposes only. This regression model was significant at the 0.05 level as shown by the one-way analysis of variance in Table 6.

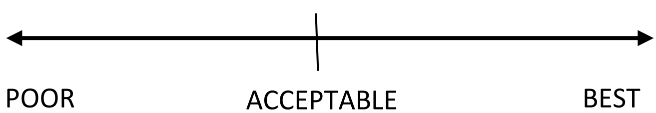
RUBRICS: THE RSU EXPERIENCE

This study was one of the initial efforts done to investigate performance assessment in Romblon State University. After this investigation, rubrics were popularized in the campus and issues like objectivity, scoring and reliability were given solutions. In RSU, general rubrics like an oral report for example, found place in other disciplines like psychology, biological sciences, education, etc. With the number of teachers resorting to rubrics, a modification was proposed to resolve issues surrounding the use of this form of assessment. Rubrics, in the Romblon State University experience, resolved some of the following issues:

Student involvement in the preparation of rubrics. After a task is determined, teachers should ask the students: "What do you think is an excellent *oral report*? How would you describe a very good *oral report*." Students would surely give varied answers but the role of the teacher is to group which of the characteristics measure the same thing. The characteristic should be given a common name and be placed into the first column (expected quality). Then the teachers should guide the students in determining which of these qualities matters most. The class together with the teacher should decide how many points should be apportioned in each criterion (expected quality). This will be the basis in making a scale. It should be taken into consideration that the narrower the range of the scale, the better judgment could be made on the work.

Ease in making indicators of performance. As opposed to most rubrics where each gradation of quality is described, RSU rubrics just describes the best performance. For example, how would you say that an *oral report* is delivered excellently? This is done so that students will aim for the best and would try to satisfy what is expected of them in a certain criterion.

Establishment of a 'continuum' scale where performance could be graded. Unlike other rubrics where indicators of student performance are fixed to beginner, developer or expert, the RSU rubrics provide a 'continuum scale'. The midpoint of the scale is the 'acceptable' performance, leaning towards its right is the best performance, and on its left is the poor performance. A scorer, upon inspecting students' work will determine from the scale the level of student performance for a particular task and plot this against the point scale allotted for a particular criterion where half of the value means an 'acceptable performance'. As much as possible, raters should refrain from giving decimal scores. The use of whole numbers is recommended.



Sample Scales: 0 – 3; 0-4, 0-5

Figure 2. The Continuum Scale

Ease in converting the score into percent and grade points. Since RSU's grading system is commonly in percent form and then converted only to a grade point, it is recommended that the total score from the rubrics be transmuted into a percent grade using the equation:

$$\text{Grade} = \frac{\text{Score (100-Base Grade)}}{\text{Highest Possible Score}} + \text{Base Grade}$$

In this case, the common problem of how a letter/descriptive grade are converted into numerical measure is given a creative solution.

CONCLUSIONS AND RECOMMENDATIONS

Since the respondents of the study were non-probability samples, findings cannot be inferred to the population. However, evidences from this study support the following big concepts in performance assessment:

a. Not all hands-on activities can be used for rubrics assessments. Extra care particularly in the procedural aspect is needed in selecting assessment tasks.
b. Assessment tasks should be structured in such a way that students will have cognitive guides on how to go about performing a task. Lead-up activities related to the task are also recommended to be done first by the students before the actual assessment.
c. Experts' opinions and students' inputs should be validated and tested for reliability before use.
d. The use of performance assessment gives students the chance to know beforehand what is expected of them and to improve on areas which would meet the specified criteria. Though rubrics mirror real-life performance assessment, issues on language facility still hamper students from performing at a higher level.
e. Scientific attitude is the independent variable that explains to a greater degree student performance in rubrics assessment in physical science. Other significant predictors are self-concept, tactility and rating in the College Admission Test. In the rally for science literacy and culture, the best way to start is with students' scientific attitude which should be greatly encouraged by science teachers.
f. Rubrics can be modified and tailored according to the culture of the school. Issues around objectivity, standardization and transmutation can be settled by providing manuals on how to use teacher-made rubrics and receiving feedbacks from users.

REFERENCES

Andrade, H.G. (1997). Understanding Rubrics. *Educational Leadership*, *54(4)*

Brookhart, S.M. (1999). *The art and science of classroom assessment: The missing part of pedagogy.* ASHE-ERIC Higher Education Report (Vol. 27, No. 1). Washington, DC: The George Washington University, Graduate School of Education and Human Development.

Brualdi, A. (2000). Implementing performance assessment in the classroom. *Practical Assessment, Research & Evaluation, 6 (2).* Retrieved March 18, 2003 from http://ericae.net/pare/getvn.asp?v=6&n=2

George, D., Mallery P. (2000). *SPSS for windows step by step: A simple guide and reference 9.0 update.* Massachusetts: Allyn & Bacon Pearson Education Company

Kilic, G. B. (2003). *Application of self and peer assessment in an elementary science teaching methods course.* Retrieved October 20, 2004 from http://www.upd.edu.ph/~ismed/

Moskal, B.M. (2000). Scoring rubrics: What, when and how? *ERIC CSMEE Digest.* Retrieved March 20, 2003 from http://ericae.net/pare/getvn.asp?v=7&n=3

Religioso, T.F. (2002). Performance assessment: A St. Scholastica's Experience. *Sangguni, 13 (1&2)*, 256-268

Salmorin, L. M. (1987) *Some Factors Asscociated with Academic Achievement in Physics of 4th year Mapa High School Students.* Unpublished Doctoral Dissertation. CEU Manila

Wangsatorntanakhun, J.A. (____). *Designing performance assessments: Challenges for the three-story intellect.* Retrieved March 15, 2003 from http://www.geocities.com/Athens/Parthenon/8658

________________. (____). *Performance assessment.* Retrieved March 15, 2003 from http://www.exit109.com/~learn/perfases.htm