IDENTIFICATION AND CHARACTERISATION OF EXEMPLARY SCIENCE ASSESSMENT PRACTICES

Rekha B Koul

The research described in this paper was aimed at identifying exemplary assessment practices in secondary science classes. In the first stage, following a review of the literature, a six-scale instrument of 48 items was trialled with a sample of 470 students from grades eight, nine and ten in 20 science classrooms in three Western Australian schools. Based on internal consistency reliability data and exploratory factor analysis, refinement decisions resulted in a five-scale instrument that was named the Student Perceptions of Assessment Questionnaire (SPAQ). In the second stage, the SPAQ was used with an attitude scale, and a self-efficacy scale. This survey was administered to a larger sample of 960 students from 40 science classes from the same grades as in the first stage. Statistical analyses confirmed the validity and reliability of the SPAQ. Based on the results of this survey exemplary teachers were identified. In the third and last stage interviews with teachers and students were conducted. Classes of these exemplary teachers were also observed. These exemplary teachers were found to be thorough in their teaching, giving students enough time to prepare for the assessment, giving students freedom to choose from a variety of assessments and were flexible in teaching and assessment. They also demonstrated in-depth understanding of the science topics they were teaching.

KEYWORDS : Science Assessment, SPAQ, Exemplary Science Teachers

INTRODUCTION

A constructivist view of learning supports the use of clear goal statements and success criteria, targeted feedback and student self-assessment (Sadler, 1989). However, little contemporary evidence exists to support the view that students are genuinely involved in decision-making about their assessment tasks (Dorman, Waldrip & Fisher, 2008) as forms of assessment and specific assessment tasks employed in schools are usually decided by teachers and administrators. Furthermore, even though reports like The Status and Quality of Teaching and Learning in Australia (Goodrum, Hackling, & Rennie, 2001) have asserted that assessment is a key component of the teaching and learning process, teachers tend to utilize a very narrow range of assessment strategies on which to base feedback to parents and students. In practice, there is little evidence that teachers actually use diagnostic or formative assessment strategies to inform planning and teaching (Randnor, 1996).

There are conflicting views about the role and nature of assessment practices in education. Harlen (1998) advocates that teachers should use both oral and written questions in assessing student's learning. While as, experts (Dorr-Bremme & Herman, 1986; Stiggins, 1994) encourage alternative assessment strategies, such as teacher observation, personal communication, and student performances, demonstrations, and portfolios, for greater usefulness of evaluating students and informing classroom instruction. Tobin (1998) asserted that assessment could be used to provide opportunities for students to show what they know. Reynolds, Doran, Allers and Agruso (1995) argued that for effective learning to occur, congruence must exist between instruction, assessment and outcomes. This paper represents a context-specific investigation of this congruence.

An effective assessment process should involve a two-way communication system between teachers and their students (Black & William, 1998). Historically, teachers have used testing instruments to transmit to the student and their parents what is really important for the student to know and do. While this reporting tends to be in the form of a grade, the form and design of the assessment can send subtle messages on what is important. There has been a substantial amount of research into types of assessment but very little research into students' perceptions of assessment (Black & William, 1998; Crooks, 1998; Plake, 1993, Popham, 1997).

Use of Student Perceptual Data: Until the late 1960s a very strong tradition of trained observers coding teacher and student behaviours dominated classroom research. Indeed, it was a key recommendation of Dunkin and Biddle (1974) that instruments for research on teaching processes, where possible, should deal with the objective characteristics of classroom events. Clearly, this approach to research which often involved trained observers coding teacher and student behaviour was consistent with the behaviourism approach of the 1960s. The study of classroom psychosocial environments in the late 1960s broke this tradition and used student perceptual data. Since then, the strong trend in classroom environment research has been towards this high-inference approach with data collected from the teachers and students. Walberg (1976) supported this methodological approach where student learning involves student perceptions acting as mediators in the learning process. Walberg (1976) also advocated the use of student perception to assess environments because students seemed quite able to perceive and weigh stimuli and to render predictively valid judgments of the social environments of their classes.

Classroom Learning Environment: The notion that a learning environment exists which mediates aspects of educational development began as early as 1936 when Lewin (1936) recognised that the environment and the personality

of the individual were powerful determinants of behaviour and introduced the formula, B=f(P,E). Since Lewin's time, international research efforts involving the conceptualisation, assessment, and investigation of perceptions of aspects of the classroom environment have firmly established classroom environments as a thriving field of study (Fraser, 1994, 1998, Fraser & Walberg, 1991). For example, recent classroom environment research has focused on constructivist classroom environments (Taylor, Fraser, & Fisher, 1997), cross-national constructivist classroom environments (Aldridge, Fraser & Huang, 1999), science laboratory classroom environments (McRobbie & Fraser, 1993), computer laboratory classroom environment (Newby & Fisher, 1997) computer-assisted instruction classrooms (Stolarchuk & Fisher, 1999) and classroom environment and teachers' cultural back grounds (Koul & Fisher, 2006).

A great deal of classroom learning environment research has been carried out over the past 40 years and evidence from these studies reveals that classroom learning environment dimensions are good indicators of teaching and learning processes and have predictive power on a number of learning outcomes pointing towards the possibility of improving students' outcomes through changing classroom environments (Fraser, 1994, 1998; Fraser & Walberg, 1991;Wubbels & Levy, 1993).

The present interpretive study involved a multi-method approach in exploration of factors associated with students' perceptions of assessment.

Attitude towards Science Classrooms: The impact of students' attitudes towards their science assessments is regarded as an important goal in the present study. Attitudes towards science, has been defined as "a learned disposition to evaluate in certain ways objects, people, actions, situations or propositions involved in learning science" (Gardner, 1975 p. 2). This learned disposition refers to the way students regard science, such as interesting, boring, dull or exciting. Positive student attitudes are then measured by the degree of motivation and interest reported by the students. Klopfer (1971, 1976) went further and developed a structure for evaluating attitudes related to science education. He included four categories in his structure i.e. events in the natural world, activities, science and inquiry. Klopfer's (1976) second category, relating to students attitudes towards their science assessments, was a focus of the present study.

Academic Efficacy: Over the past two decades the broad psychological concept of self-efficacy has been a subject of interest (Bandura, 1997; Schunk, 1995). Within this field, one particular strong area of interest is that of academic efficacy, which refers to personal judgments of one's capabilities to organize and execute courses of action to attain designated types of educational

performances (Zimmerman, 1995). Research studies have provided consistent, convincing evidence that academic efficacy is positively related to academic motivation (e.g.,Schunk & Hanson, 1985), persistence (Lyman, Pretice-Dunn, Wilson & Bonfilio, 1984), memory performance (Berry, 1987), and academic performance (Schunk, 1989).

Gender and Year Level: It is well documented in reviews of literature that women are under-represented in science and technology courses and careers (Greenfield, 1996; Kahle & Meece, 1994) and that boys outperformed girls in science (especially physical science) (Bellar & Gafini,1996; Kahle & Meece, 1994; Murphy,1996). Among the sources that may cause these differences are individual, cognitive, attitudinal, socio-cultural, home and family, and educational variables (Farenga & Joyce,1997; Kahle & Meece, 1994). In the classroom context, boys and girls may not have equal opportunities in science activities, and this could cause gender differences in science achievement (Fraser, Tobin, & Khale, 1992; Harding, 1996; Warrington & Younger, 1996). Because educational variables are one of the important sources for accounting for gender differences in students' achievement in science, and for participation in science activities, the perspective of gender differences needs to be understood.Previous studies have reported gender-related differences in students' perceptions of the learning environment (Fraser, Fisher, & McRobbie, 1996; Koul & Fisher, 2006). Therefore in keeping with these lines of research, gender-related differences in students' perceptions of their assessment were explored in this study.

Year level as well as gender differences in students' perceptions, other learning environment research studies in science classrooms have indicated differences between perceptions of students in different years of school (Kim, Fisher & Fraser, 2000). In this study, differences between the perceptions of students in different years of lower secondary were examined for trends.

OBJECTIVES OF THE STUDY

The overall aim of the study was to investigate relationships among students' perceptions of their assessment tasks, classroom learning environments, academic efficacy and attitude to science in years eight, nine and ten.

The objectives of this study were:

1. to provide further validation data on the instrument for assessing students perceptions of assessment tasks;
2. to investigate differences between students' perceptions in terms of year levels and gender;
3. to investigate associations between students' perceptions of their assessment tasks and their attitude to science and academic efficacy

outcomes;

4. to identify exemplary teachers on the basis of students perceptions of their assessment tasks; and

5. to describe the form and design of assessment tasks used by exemplary science teachers.

INSTRUMENTS AND PROCEDURE USED

The study was carried out in phases over a period of three years using a multi-method research approach:

1. In the first phase a pre-existing and validated questionnaire, Perceptions of Assessment Tasks (PAT) a six-scale instrument of 55 item developed by Schaffner, Burry, Cho, Boney and Hamilton (2000) was administered to 470 students from grades eight, nine and ten in 20 science classrooms in three Western Australian schools. Close-ended interviews were conducted with randomly selected 40 students to look at student perceptions of their assessment tasks.

2. In second phase based on internal consistency reliability data and exploratory factor analysis, refinement decisions of PATT resulted in a five-scale instrument that was named the Student Perceptions of Assessment Questionnaire (SPAQ). This study was part of a larger study carried out in three states of Australia. The SPAQ was used with an attitude scale, and a self-efficacy scale. This survey was administered to a larger sample of 960 students from 41 science classes from the same grades as in the first stage.

3. In the final stage of the study five teachers identified on the basis of students showing most positive perceptions on the scales of SPAQ were interviewed and their teaching observed. Informal interviews were also conducted with students from the classes identified.

Students' Perceptions of Assessment Questionnaire (SPAQ): Students' perceptions of assessment were assessed with the 30-item SPAQ. These items are assigned to internally consistent scales namely Congruence & Planned Activity, Authenticity, Student Consultation, Transparency and Diversity. Table 1 shows the scales, descriptions and sample items from the SPAQ. Validation statistics performed on the data collected are presented in the results section. Responses in the SPAQ were recorded on a four point Likert type response format for each item (e.g., Almost Never, Sometimes, Often, and Almost Always).

Two outcome scales namely Attitude to Science and Academic Efficacy were also employed in present study. A review of literature revealed a large pool of science-related attitude scales. Of particular interest to this study is the Test of Science Related Attitudes (TOSRA) developed by Fraser (1978) to

measure students, attitudes towards their science classes. Fraser based the subscales of this instrument on Klopfer's (1976) taxonomy of the affective domain related to science education. Attitude to Science was assessed on an 8-item scale adopted from the Test of Science-Related Attitudes (TOSRA: Fraser, 1981). Responses were recorded on a four-point format ranging from 1 (Disagree) to 4 (Agree).

Table 1

Description and Example of Items for Each Scale of Students Perceptions of Assessment Questionnaire (SPAQ), Attitude Scale and Academic Efficacy.

Scale	Description	Item
Congruence and Planned Activity	Extent to which assessment tasks align with the goals, objectives and activities of the learning program.	My assignments/tests are about what I have done in class.
Authenticity	The extent to which assessment tasks feature real life situations those are relevant to the learner.	I find science assessment tasks are relevant to what I do outside of school.
Student Consultation	The extent to which students are consulted and informed about the forms of assessment tasks being employed.	I have a say in how I will be assessed in science.
Transparency	The extent to which the purposes and forms of assessment tasks are well-defined and clear to the learner.	I am clear about what my teacher wants in my assessment tasks.
Diversity	The extent to which all students have an equal chance at completing assessment tasks.	I have as much chance as any other student at completing assessment tasks.
Attitude to Science	The extent to which students are interested in, enjoy and look forward to lessons in that subject.	I enjoy the activities we do in science.
Academic Efficacy	Students' judgments of their capabilities to organize and execute courses of action to attain designated types of educational performances.	Even if science is hard, I can learn it.

Perceived Academic Efficacy refers to students' judgments of their ability to master academic tasks that they are given in their classrooms. A 6-item scale using items developed by Midgley and Urdan (1995) was used to assess perceived academic competence at science class work. Items were modified to elicit a response on academic efficacy in science. All items in the academic efficacy scale had a four-point response format with anchors of 1 (Disagree) and 4 (Agree).

RESULTS

Development and Validation of SPAQ

A principal components factor analysis followed by varimax rotation confirmed a refined structure of the SPAQ instrument comprising of 30 items in

5 scales and 14 items in two outcome scales. All the 44 items had a loading of at least 0.40 on their a priori scales (see Table 2). The percentage of the total variance extracted with each factor is also recorded at the bottom of Table 2. The percentage of variance varies from 3.55% to 26.03% for different scales, with the total variance accounted for being around 50%.

The validity and reliability information of the instrument developed in this study are presented in Table 3. To determine by the degree to which item in the same scale measure the same aspects of students' perceptions of assessment tasks, attitude to science and academic self-efficacy, a measure of internal consistency, the Cronbach alpha reliability coefficient was used. For the scales of SPAQ the highest alpha reliability of 0.83 for the scale of Authenticity, and the lowest of 0.63 for the scale of Diversity was recorded. The scale of student attitudes to science has alpha reliability score of 0.85 and scale of Academic Efficacy of 0.90. Since the reliability values for all the scales of SPAQ were consistently above 0.63 the instrument developed is therefore reliable for use .

High mean scores ranging from 2.16 for the scale of Student Consultation to 3.17 for the scale of Congruence with Planned Learning on a four point Likert type scale confirm that students generally have a positive perception of their assessment tasks. Scale of Student Consultation having the lowest score confirms that students generally do not have a say in their assessment tasks.

Overall culture of each class is different and the ability of SPAQ to differentiate between the classes in the study was considered important. The instruments ability to differentiate in this way was measured using one-way analysis of variance (ANOVA). The eta^2 statistics was calculated to provide an estimate of the strength of the association between class membership and the dependent variables as shown in Table 3. The eta^2 statistic for the SPAQ indicates that the amount of variance in scores accounted for by class membership ranged from 0.12 to 0.28 and was statistically significant ($p<0.001$) for all scales. It appears that the instrument is able to differentiate clearly between the perceptions of students in different classrooms.

Associations between SPAQ and Attitude towards Science and Academic Efficacy

One of the aims of the study was to investigate associations between students, perceptions of assessment tasks and their attitude to science classes. These associations were explored using simple and multiple correlation analyses. The results of the analyses are shown in Table 4. For all the scales of the SPAQ associations are positive and statistically significant.

It was found that the scales of Congruence and Planned Activity, Authenticity, Transparency and Diversity were positively and significantly associated whereas, scale of Student Consultation was negatively and significantly associated with attitude to science. The multiple correlation (R)

between the set of SPAQ scales and attitude to science class was 0.55. The R^2 value which indicates the proportion of variance in attitude to science class that can be attributed to students' perceptions of their assessment tasks given by the teachers was 30%. To determine which SPAQ scales contributed most to this association, the standardized regression coefficient (β) was examined for each scale. It was found that the scales of Congruence and Planned Activity, Authenticity, Transparency and Diversity were positively and significantly associated whereas, scale of Student Consultation was negatively and significantly associated with attitude to science.

Table 2

Factor Loadings for the SPAQ Questionnaire.

Item No	Congruence and Planned Activity	Authenticity	Student Consultation	Transparency	Diversity	Attitude to Science	Academic Efficacy
1	0.46						
2	0.52						
3	0.51						
4	0.68						
5	0.69						
6	0.46						
7		0.43					
8		0.60					
9		0.69					
10		0.64					
11		0.72					
12		0.52					
13			0.47				
14			0.49				
15			0.66				
16			0.59				
17			0.44				
18			0.74				
19				0.66			
20				0.63			
21				0.68			
22				0.69			
23				0.64			
24				0.57			
25					0.41		
26					0.48		
27					0.54		
28					0.44		
29					0.46		
30					0.52		
31						0.77	
32						0.47	
33						0.66	
34						0.71	
35						0.44	
36						0.78	
37						0.60	
38							0.42
39							0.61
40							0.73
41							0.69
42							0.77
43							0.76
44							0.71
% Variance	26.03	8.12	7.05	4.13	3.55	3.43	2.46
Eigen value	11.45	3.57	3.10	1.81	1.56	1.50	1.08

Table 3

Scale Mean, Standard Deviation, Internal Consistency (Cronbach Alpha Reliability) and Ability to Differentiate between Classrooms (ANOVA Results) for the SPAQ, Attitude to Science and Academic Efficacy.

Scale	Mean	S D	Alpha Reliability	ANOVA (eta^2)
Congruence and Planned Activity	3.17	0.51	0.76	0.13*
Authenticity	2.24	0.64	0.83	0.17*
Student Consultation	2.16	0.58	0.71	0.19*
Transparency	3.06	0.63	0.83	0.14*
Diversity	2.56	0.54	0.63	0.12*
Attitude to Science	2.56	0.80	0.85	0.28*
Academic Efficacy	2.96	0.76	0.9	0.13*

n= 960 students in 40 classes *$p<0.001$

Table 4

Associations between Scales of SPAQ and Attitude to Science in terms of Simple Correlations (R), Multiple Correlations and Standardized Regression Coefficient (β)

Scale	Attitude to Science Class		Academic Efficacy	
	r	β	r	β
Congruence and Planned Activity	0.32**	0.1***	0.33**	0.09***
Authenticity	0.45**	35***	0.36**	0.22***
Student Consultation	0.25**	-0.17***	0.21**	-0.16***
Transparency	0.39**	0.18***	0.46**	0.28***
Diversity	0.39**	0.18***	0.43**	0.25***
Multiple Correlations	R = 0.55* R² = 0.3		R = 0.56* R² = 0.32	

*$p<0.0001$, **$p<0.01$, ***$p<0.05$ n=960

GENDER DIFFERENCES

Differences between the students' perceptions of the scales of the SPAQ and the gender of the students were analysed. The gender differences in students' perceptions of classroom learning environment were examined by splitting the total number into female (388) and male (572) students involved in the study.

Table 5

Item Mean and Standard Deviation for Gender Differences in Students' Perceptions on the scales of SPAQ

Scale	Item Mean		Item SD		Difference	
	Female	Male	Female	Male	t	Effect Size
Congruence with Planned Activity	3.21	3.14	0.51	0.51	0.05	0.13
Authenticity	2.18	2.28	0.67	0.61	5.28*	0.15
Student Consultation	2.12	2.19	0.58	0.58	0.06	0.12
Transparency	3.07	3.06	0.62	0.65	0.52	0.01
Diversity	2.54	2.58	0.54	0.54	0.26	0.07
Student Attitudes	2.52	2.58	0.81	0.78	0.01	0.07
Academic Efficacy	2.88	3.01	0.77	0.74	0.48	0.17

$p < 0.05$, females (n = 388); males (n = 572)

To examine the gender differences in students' perceptions of the classes, the within-class gender subgroup mean was chosen as the unit of analysis as this aims to eliminate the effect of class differences due to males and females being unevenly distributed in the sample. In the data analysis, male and female students' mean scores for each class were computed, and the significance of gender differences were analysed using an independent t-test. Table 5 shows the scale item means, male and female differences, standard deviations, t-values and Cohen's d effect size. The purpose of this analysis was to establish whether there are significant differences in perceptions of students according to their gender.

As can be seen in Table 5, out of five scales of the SPAQ and two Attitude scales, the gender differences in the perceptions of males and females were found to be statistically significantly different only on the scale of Authenticity.

YEAR LEVEL DIFFERENCES

One of the aims of the study was to investigate the differences in the perceptions of the scales of SPAQ and the two sides of attitude and efficacy in students from different year levels. This was explored by splitting the students in their year groups (yr 8 = 347, yr 9 = 328, yr 10 = 285).

The results of the analyses are shown in Table 6. In the data analysis, mean scores for each of the three-year groups were computed. Table V shows the scale item means and F values of the scales of the TROFLEI with the perceptions of students from the three year groups in study. The purpose of this analysis is to establish whether there are significant differences in the perceptions of students according to their year groups.

As can be seen in Table 6, the differences in the perceptions of students on

the scales of SPAQ and Attitude, five out of seven scales are statistically significant confirming that year level does impact significantly on students' perception of their assessment.Tukey's post hoc test (p<0.05) revealed that for the Congruence with Planned Activity scale the Year Eight students were dominant and had statistically significant higher means while the Year Ten students had the highest means for the scale of Diversity.

Table 6

Item Mean, Item Standard Deviation and Ability to Differentiate Between Levels (ANOVA Results) for Year Level Differences in Students' Perceptions Measured by the SPAQ

Scale	Mean			SD			Difference
	Year 8	Year 9	Year 10	Year 8	Year 9	Year 10	F
Congruence with Planned Activity	3.24	3.14	3.12	0.51	0.49	0.51	5.64**
Authenticity	2.32	2.24	2.15	0.66	0.65	0.59	5.13**
Student Consultation	2.34	2.15	1.95	0.56	0.59	0.51	37.06***
Transparency	3.10	3.09	2.99	0.64	0.64	0.62	2.48
Diversity	2.62	2.58	2.47	0.54	0.55	0.53	6.21**
Student Attitudes	2.65	2.44	2.60	0.81	0.85	0.71	6.38**
Academic Efficacy	2.99	2.96	2.91	0.74	0.76	0.78	0.83

Sample Size = 34 7(Year 8), 328 (Year 9) and 285 (Year 10).
$**p<0.01$, $***p<0.05$

INTERVIEWS

Based on the findings of the quantitative data five exemplary teachers (three male and two female) were identified from the total sample of 40 and their teaching observed and informal interviews conducted.These five teachers represented Private, Public and Rural schools in Western Australia. These selected teachers had been rated by their students' more than one standard deviation above the mean for at least three of the five scales. This process has been described previously by Waldrip, Fisher and Dorman (2009).

Furthermore, four students from the classes of each of the five selected teachers were also interviewed. The students' interviews were structured and conducted in three phases on the same day. The interview phases occurred before, during and after an activity in the classroom. Similar questions regarding the activity were asked to assess students' initial perceptions about the task, during the task and when the task was completed.

The students were asked few general questions followed by questions relating to each of the five scales of SPAQ questionnaire. This approach enabled the researcher to draw on a variety of paradigms to inform their

interpretation in a bid to explain the positive student perception of assessment tasks. The interview schedule along with stages and scales is represented in Table 7.

Table 7
Student Interview Schedule

Scale	Beginning	During	After
General	What do you think of the task? Why do you feel that way? How do you feel you will achieve in this task? Why do you feel that way?	What is challenging about the task? Why do you think that is? What are you learning during this task? Have your feelings towards the task changed? What has changed? What would you like to do differently?	What have you learned during this task? Did you like learning in this way? Why do you think that? Was it challenging? Why do you think that is? Was there anything different about the task? What would you like to have done differently?
Congruence	Is the task related to what you have been learning in class?	Is the task related to what you have been learning in class?	Was the task related to what you have been learning in class?
Authenticity	Is the task useful or helpful to you? Is it meaningful to you?	Is the task useful or helpful to you? Is it meaningful to you?	Was the task useful or helpful to you? Was it meaningful to you?
Student Consultation	Do you have some say in this task?	Are you having some say in this task?	Did you have some say in this task?
Transparency	Are you clear about what you need to do in this task?	Are you still clear about what you need to do in this task?	Were you clear about what you needed to do in this task?
Diversity	Are there different ways you can complete this task? What will happen if you are confused?	Are there different ways you can complete this task? What will happen if you are confused?	Were there different ways to complete this task? What happen if you are confused?

LEARNING AND ASSESSMENT

The results, which emerged from, the interviews with teachers and students are presented in this section. Interviews and observations reflected that the exemplary teachers were engaging constructivist ways of teaching underpinning formulations of formative assessment (Sadler, 1989). As supported by the quantitative results, students of these teachers had very positive perceptions of the assessment practices employed by their teachers and it was observed that social interactions within these classes were generally very strong. Assessment practices employed by these teachers not only look at what students know, but also at developing student identities as capable and competent learners. These teachers take into consideration what, why, and how students are learning as well as showing a shift in their views of

assessment in science by keeping themselves informed on the changing nature of the outcomes of the science education. Some of the comments supporting these claims are:

Teacher: *I formulate assessments very early in the year keeping science intended outcomes in view. My assessments are designed to let me know what students know, not what they do not know. Thus, assessment becomes a part of learning.*

Student 1: *Beginning of each term he gave us details of all the assessment and what is expected of us. This approach gives us clear guidelines for learning. After the assessment is evaluated, often, he runs a session on our misconceptions.*

Student 5: *We knew in the beginning of the term that we are required to make an information poster or pamphlet or flyer regarding the infectious diseases. I kept on collecting the related information and stuff you know… It was easy to compile all the information close to the date of submission*

Student 8: *Since we know what is required and even expected form us…we learn accordingly.*

Student 11: *If I decide that I want good marks, we have to work for it. I cannot say what the marks should be but If I have worked according to teachers' guidelines I am sure that my work will get a high grade.*

CURRICULUM AND ASSESSMENT

The teachers when interviewed commented on the way they considered assessment and curriculum to be related and interact in complex ways. They believed that a well perceived curriculum that incorporates assessment also narrows the gap between intended and implemented curriculum resulting in an achieved curriculum. Exemplary teachers also researched and used the available relevant assessment resources.

Their comments in this regard were:

Teacher: *I do not separate assessment from the curriculum. Both are different but lead to same object-student/teacher learning. It is complex but once understood can be practiced successfully. These days there are lot of ideas and materials available.*

Student 4: *She tells us what will be asked to do. So we prepare accordingly. She also gives us an evaluation criterion for each assessment.*

Student 9: *You know this was different. I exactly knew what is required in this project. It turned out to be the biggest project I had ever done.*

Student 14: *The last work sheet he gave us was confusing to start with. Lot of application mixing topics in machines, light and heat. I thought about it and did well.*

CLASSROOM AND ASSESSMENT

The exemplary teachers believed that there is a need to recognise the roles and responsibilities of both teachers and students. This view resonates with Sadler's (1989) view that formative assessment is based on the principle that students need to become consumers as well as the objects of assessment activities. This socio-cultural view of learning enhances positive classroom interactions. Assessments also reflect a power relationship in classroom. The teacher questions and students respond. However, in an exemplary teacher's class, teacher provides enough resources for students to respond to the questions and create knowledge. These resources could be books, the World Wide Web, peers or other resource persons.

Teacher: *I provide many resources to students so that they can research and find answers to the investigations we do. It is interesting to see how many resources students find on their own and enter classroom with different world views.*

Students 3: *Teacher directs us to the reading material. We also do a lot of web surfing. I find many useful links on You Tube.*

Student 7: *Last night when I was chatting with my friends on Facebook [internet interaction site] we looked at viruses, bacteria, protozoa, worms and fungi. That was cool. We all learnt a lot about the lesson we are doing in class.*

Student 10: *First, I thought we are not going to learn much in this year's science unit. It seemed he was boring. I had not done much research. Now that we have started researching and we find the importance of substances like the mining in up north. We get the crude material and useful things come out of that.*

TEACHERS AND ASSESSMENT

Although these selected teachers had emancipatory views about assessment and stood apart generally from their counter-parts, they were feeling concerned about the external influences on them. They felt answerable to various stake holders namely students, parents, administrators and the community at large. To establish their accountability their students had to perform well in national and international science tests. They could use these test results as evidence of efficiency for their performance. The teachers also believe that knowledge and expertise of various assessment activities is mandatory for all science teachers who need to have an in-depth understanding of the topic being taught and students' existing knowledge. The exemplary teachers recommend that this can be achieved through planning of the course content which should include teaching, learning,

assessment and curriculum and their interrelationship.

Teacher: *I feel responsible for student learning. I am answerable for the student learning and on top of that we have science Olympiads, national testing and international testing. It is complex.*

Student 2: *He knows his stuff well and also how to teach. For example last topic on renewable energy he talked about many ways, how energy can be renewed and also conserved. It was great. I enjoyed the lesson and writing the project. With the result I got good grade.*

Student 16: *She is thorough with the content of all the lessons.*

Student 15: *Later this year we will be writing the international science test and she wants us to do well in that. This is a science extension class and many of us also participate in science Olympiads.*

STUDENTS AND ASSESSMENT

The final and last section of this study identified the students as active and intentional participants in classroom assessment practices. Cowie (2005) highlights the multiple consequences of classroom assessment for students as: importance of trust and respect, the influence of their goals and learning motivations, and equity issues. Our study also found parallels with each of these factors. Continued teacher support and positive classroom learning environment contribute towards what students consider important to learn. Mutual trust and respect among teachers and students is central to student learning. Students should believe that assessments are designed to help them and they view assessment as a joint teacher-pupil responsibility. The students also pointed out the importance of using a variety of assessments.

Teacher: *I have to be very careful about what I speak in classroom. I try to look at students' positive points and build on that. I tend to add plurality in the assessments we (students and I) design. This gives all students from different cultural backgrounds and ability levels to demonstrate their learning. It also keeps them interested in science.*

Student 6: *What I love about our teacher is the respect and belief she has for us. She designs assessments, which she is confident that we have learnt and can do well. Last assignment when she thought that I could improve upon it, she talked privately and respectfully to me. I am learning, and that is her job.*

Student 18: *During the question/answer session every student has equal chance of being asked for a response. He will only ask those students who have raised their hands. In the class (while teaching) he never shows individual preference.*

Student 12: *We are free to do our assignments the way we want. We don't get a choice on the things where teacher has already planned an activity and if we change it that could affect our learning.*

CONCLUSION

In this study a new instrument, the Students' Perception of Assessment Questionnaire (SPAQ) was further validated. For the five scale instrument the Cronbach alpha reliability scores ranged from 0.63 to 0.83, thus making these scales acceptable for use in future. Also, the SPAQ's ability to distinguish between classes was also established, which was an important contribution of the study. Additionally, scales of attitude to subject and academic efficacy were further validated.

It was found that student perceptual data can be used to identify exemplary teacher and SPAQ was a valid instrument to use for this. The exemplary teachers were identified as those who scored more than one standard deviation above the mean for at least three of the five scales of SPAQ.

Qualitative data added a new rich layer of understanding to already existing knowledge gained through quantitative data. While developing the SPAQ different dimensions of assessment were identified namely, Congruence with planned learning, Authenticity, Student Consultation, Transparency and Diversity were identified. Observations and interview data identified the same dimensions existing within different sections of assessment process. The identified sections namely, learning, curriculum, classroom and assessment, teacher, and student are integral part of assessments. The qualitative data identified the importance and role of these sections in learning and assessment.

Assessment for learning has emerged as central theme in this study. Identified exemplary teachers were found to be very thorough in their teaching, giving students enough time to prepare for an assessment, allowing students freedom to choose from a variety of assessments and were flexible in teaching and assessment. They also demonstrated an in-depth understanding of science topics they were teaching.

This study demonstrates that scales of learning environment can be used in complex studies where many interrelated variables are assessed. By identifying good science teachers and describing what they do in their classrooms, we have an opportunity to use this information in professional development of other interested teachers. This is one of the ways to bring about desired changes in the educational system.

REFERENCES

Aldridge, J.M., Fraser,B.J., & Huang, T. (1999). Investigating classroom environments in Taiwan and Australia with multiple research methods. *Journal of Educational Research*, *93*,48-57.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.

Bellar, M., & Gafni, N. (1996). The 1991 international assessment of educational progress in mathematics and sciences: The gender differences perspective. *Journal of Educational Psycology, 88*, 365-377.

Berry, J. M. (1987). A self efficacy model of memory performance. Paper presented at the annual meeting of the American Educational Research Association, New York.

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, *5* (1), 7-74.

Cowie, B. (2005). Student commentary on classroom assessment in science: A sociocultural interpretation. *International Journal of Science Education, 27*(2), 199-214.

Cronbach, D. J. (1951). Coefficient alpha and internal structure of tests. *Psychometrika, 16*(3), 297-334.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research 58*, 438-481.

Cronbach, D. J. (1951). Coefficient alpha and internal structure of tests. *Psychometrika, 16*(3), 297-334.

De Vellis, R. F. (1991). Scale development: Theory and application. Newbury park: Sage Publications.

Dorr-Bremme, D., & Herman, J. (1986). Assessing student achievement: A profile of classroom practices. Los Angeles, CA.

Dunkin, M. J., & Biddle, B. J. (1974). The study of Teaching: New Yok: Holt, Rinehart & Winston.

Farenga, S.J., & Joyce, B. A. (1997). What children bring to the classroom: Learning science from experience. *School Science and Mathematics, 97*, 248-252.

Fraser, B. J. (1981). Test of Science-Related Attitudes: Handbook. Melbourne, Australia: Australian Council for Educational Research.

Fraser, B. (1994). Research on classroom and school climate. In D. Gabel (Ed.), Handbook of research on science teaching and learning (pp. 493-541). New York: Macmillan.

Fraser, B. J. (1998). Science learning environments: Assessments, Effects and determinants. In B. J. Fraser & K. G. Tobin (Eds.), International handbook of science education (pp. 527-564). Dordrecht, The Netherlands: Kulwer.

Fraser, B.J., Tobin, K., & Kahle, J.B. (1992). Learning science with understanding: In search of holy grail. *Research in Science and Technological Education, 10*, 65-81.

Fraser, B.J., Fisher,D.L., & McRobbie, C.J. (1996). Development, validation and use of personal and classforms of a new classroon environment instrument. Paper presented at the annual meeting of the American Educational Research association, April in New York, USA.

Fraser, B.J., & Wallberg, H.J. (Eds.)(1991). Educational environments: Evaluation, antecedents and consequences. Oxford; Pergamon Press

Gardner, P.L. (1996). The dimensionality of attitude scales: A widely misunderstood idea. *International Journal of Science Education, 18,* 913-919.

Goodrum, D., Hackling, M., & Rennie, L. (2001). The status and quality of teaching and learning in Australian schools. Canberra.

Greenfield, T. A.(1996). Gender, ethnicity, science achievement and attitudes. *Journal of Research in Science teaching, 33*, 901-933.

Harlen, W. (1998). Teaching for understanding in pre-secondary science. In B. Fraser & K. Tobin (Eds.), International handbook of science education (pp. 183-198). Dordrech, The Netherlands: Kulwer.

Harding, J. (1996). Girls' achievement in science and technology-Implicatins for pedagogy. In P. F. Murphy & C. V. Gipps (Eds.), Equity in classrooms (pp. 111-123). London: The Falmer Press.

Henderson, D., Fisher, D. L., & Fraser, B. J. (1995). Gender differences in biology students perceptions of actual and preferred learning environments. Paper presented at the annual meeting of the National Association for Research in Science Teaching, April in San Francisco, CA.

Kahle, J. B., & Meece, J. (1994). Research on gender issues in the classroom. In D. Gabel (Ed.), Handbook of research on science teaching and learning (pp.542-557). New York.

Kim, H., Fisher, D., & Fraser, B. J. (2000). Classroom environment and teacher interpersonal behaviour in secondary classes in Korea. Evaluation and Research in Education, 14, 3-22ork: Macmillan.

Klopfer, L.E. (1976). A structure for the affective domain in relation to science education. *Science Education, 60*, 299-312

Koul, R., & Fisher, D. (2006). An investigation into teacher interaction, perceptions of learning environment cultural differences and science achievement in India. In D. Fisher, D. Zandvliet, I. Gaynor & R. Koul (Eds.), Sustainable communities and sustainable environments: Envisioning a role for science, mathematics and technology education (pp. 323-335). Perth: Curtin University of Technology.

Lewin, K. (1936). Principals of topological psychology. New York: McGraw.

Lyman, R. D., Prentice-Dunn, S., Wilson, D. R., & Bonfilio, S. A. (1984). The effect of sucess or failure on self-efficacy and task persistence of conduct-disordered children. *Psycology in the Schools, 21,* 516-519.

McRobbie, C. J., & Fraser, B. J. (1993). Associations between student outcomes and psychosocial mathematics laboratory environments. *Journal of Educational Research, 87,* 78-85.

Midgley, C., & Urdan, T. (1995). Predictors of middle school students' use of self handicapping strategies. *Journal of Early Adolescence, 15,* 389-411.

Murphy, P. (1996). The IEA assessment of science achievement. *Assessment in Education,3,* 213-232.

Newby, M., & Fisher, D.L. (1997). An instrument for assessing the learning environment of a computer laboratory. *Journal of Educational Computing research, 16,* 179-190.

Plake, B. S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher, 6,* 21-27.

Popham, W. J. (1997). Consequential validity: Right concern-wrong concept. *Educational Measurement: Issues and Practice, 16*(2), 9-13.

Radnor, H. (1996). Evaluation of key stage3 assessment in 1995 and 1996 (research report). University of Exceter.

Reynolds, D. S., Doran, R. L., Allers, R. H., & Agruso, S. A. (1995). Altenative assessment in science: A teacher's guide. Buffalo. NY: University of Buffalo.

Sadler, D. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18,* 119-144.

Schaffuer, M., Burry-Stock, A., Cho, G., Boney, T., & Hamilton, G. (2000). What do kids think when their teachers grade? Paper presented at the annual meeting of the American Educational Research Association, April in New Orleans, LA.

Schunk, D. H. (1989). Self-efficacy and cognitive skill learning. In C. Ames & R. Ames (Eds.), Research on motivation in education 3, 13-44). San Diego, CA: Academic.

Schunk, D. H. (1995). Self-efficacy and education and instruction. In J. E. Maddux (Ed.), Self-efficacy, adaptation, and adjustment: Theory, research, and application (pp. 281-303). New York: Plenum.

Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psycology, 77,* 312-322.

Stiggins, R. (1994). Student-centered classroom assessment: Macmillan

College Publishing Co.

Stolarchuk, E., & Fisher, D. (1999). The effect of using laptop computers on attitude and achievement in science. *Computers in New Zealand Schools, 11*, 40-43.

Taylor, P.C., Fraser, B.J., & Fisher, D.L. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research, 27*, 293-302.

Tobin, K. (1998). Issues and trends in the teaching of science. In B.J. Fraser and K. Tobin (Eds). International handbook of science education (pp. 129-151). Dordrecht, The Netherlands: Kluwer

Walberg, H. J. (1976). Psychology of learning environments: Behavioral, structural, or perceptual. *Review of Research in Education, 4*, 142-178.

Waldrip, B. G., Fisher, D., & Dorman, J. P. (2009). Identifying exemplary science teachers through their student perceptions of assessment process. *Research in Science and Technological Education, 27*(1), 117-129.

Warrington, M., & Younger, M. (1996). Goals, expectations and motivation: Gender differences in achievement at GSE. Curriculum,17,80-93.

Wubbles, T., & Levy, J. (1993). Do you know what you look like/: Interpersonal relationships in Education. London: Falmer.

Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), Self-efficacy in changing societies (pp. 202-231). Cambridge, UK: Cambridge University Press.