IMPROVED SCIENCE ASSESSMENTS USING STUDENT PERCEPTIONS

Rekha B. Koul and Darrell L. Fisher

Curtin University, Perth, Australia

R.Koul@curtin.edu.au, D.Fisher@curtin.edu.au

An instrument to assess secondary students' perceptions of assessment was developed in a three-stage study. In the first stage, following a review of literature, a six-scale instrument was trialed with a sample of 470 students from grades eight, nine and ten in 20 science classrooms in three 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 five scales of the 'What is Happening in this Class' (WIHIC) questionnaire, an attitude scale, and a self-efficacy scale. This survey was administered to a sample of nearly 1,000 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 survey results exemplary teachers were identified. In the third and last stage interviews with teachers and students were conducted. The classes of these exemplary teachers were also observed.

Keywords: Student perceptions, Evaluation, Learning environment

INTRODUCTION

Despite the growth in emancipatory conceptualizations of classrooms that embrace a constructivist epistemology, little contemporary evidence exists to support the view that students are genuinely involved in decision-making about their assessment tasks. That is, forms of assessment and specific assessment tasks employed in schools are overwhelmingly 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 (Radnor, 1996). Teachers feel that they need to 'sacrifice learning with understanding for the goal of drilling students in the things for which they will be held accountable' (Hobden, 1998).

Historically, teachers have received substantial levels of advice on assessment practices. Harlen (1998) advises teachers that both oral and written questions should be used in assessing student's learning. The inclusion of alternative assessment strategies, such as teacher observation, personal communication, and student performances, demonstrations, and portfolios, have been offered by experts as having greater usefulness for evaluating students and informing classroom instruction (Dorr-Bremme & Herman, 1986; Stiggins, 1994). Tobin (1998) asserted that assessment can be used to provide opportunities for students to show what they know. Reynolds, Doran, Allers & Agruso (1995) argued that for effective learning to occur, congruence must exist between instruction, assessment and outcomes. A sociocultural view of learning directs attention towards classroom interaction as a locus for formative assessment. This paper represents context-specific investigation of this congruence.

An effective assessment process should involve a two-way communication system between teachers and their students. Student opportunities to actively participate in assessment for learning interactions are inextricably entangled with the discourse of power that is operational in a particular classroom. 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 & Wiliam, 1998; Crooks, 1988; Plake, 1993; Popham, 1997).

For students to generate knowledge as part of social practices they must be given the authority for and the resources with which to build knowledge. The idea of authorative and accountable positioning with conceptual agency suggests being entitled and expected to move about the environment freely, with access to resources throughout the environment and with the authority to use, adapt and combine those resources in unconventional ways.

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. The study of classroom psychosocial environments broke this tradition and used student perceptual data in the late 1960s. 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) advocates 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: Literature reviews (Fraser, 1994, 1998) show that science education researchers have led the world in the field of classroom environment research over the last three decades, and that this field has contributed much to understanding and improving science education. For example, classroom assessments provide a means of monitoring, evaluating and assessing science teaching and curriculum. A key to improving student achievement and attitudes is to create learning environments that emphasize those characteristics that have been found to be linked empirically with student outcomes.

Academic Efficacy: Over the past two decades the broad psychological concept of self-efficacy has been the 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).

Methodology

Aim

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 from grade eight, nine and ten.

Objectives of this study were:

— to validate the instrument for accessing students perceptions of assessment tasks.

— to investigate differences between students' perceptions in terms of year levels and gender; and

— to investigate associations between students' perceptions of their assessment tasks and their attitude to science and academic efficacy outcomes.

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

— 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:

In first phase *Perceptions of Assessment Tasks (PAT)* a six-scale instrument of 48 items from a 55 item questionnaire developed by Schaffner, Burry, Cho, Boney and Hamilton (2000) was administered on 470 students from grades eight, nine and ten in 20 science classrooms in three Western Australian schools. Close ended interviews were conducted with 40 students to look at student perceptions of their assessment tasks.

In the second phase, 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). The SPAQ was used with an attitude scale, and a self-efficacy scale. This survey was administered to a larger sample of nearly 1,000 students from 41 science classes from the same grades as in the first stage. 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.* Validation statistics performed on the data collected are presented in results section. Responses in 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 *Enjoyment to Science* and *Academic Efficacy* were also employed in present study. Attitude to Science was assessed on a 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).

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 & 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 scale of academic efficacy scale had four-point response format with anchors of 1 (Disagree) and 4 (Agree).

Results

Validation of SPAQ

To determine the degree to which items 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 (Cronbach, 1951) 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 Attitude to Science* and scale of *Academic Efficacy* showed alpha reliability score of 0.85 and 0.9 respectively. Since all the reliabilities for the scales of SPAQ were consistently above 0.63 the instrument developed is reliable for use (De Vellis, 1991).

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 very positive perception of their assessment tasks. *Scale of Student Consultation* having the lowest scores confirms that students generally do not have a say in their assessment tasks.

The ability of SPAQ to differentiate between the classes is 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. 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 o.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. Significant correlations ($p<0.01$) were found between the scales of the SPAQ and scales of Student Attitude and Academic Efficacy. For example, Congruence with Planned Learning was positively related to and was positively associated with all the other scales in analyses.

Student attitudes

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. For all the scales of SPAQ associations are positive and statistically significant.

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 teacher was 30%. To determine which SPAQ scales contributed most to this association, the standardized regression coefficient (â) 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.

Qualitative data

Based on the findings of the quantitative data five exemplary teachers were identified out of 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. Denzin and Lincolin (1994) bricolage method influenced me while interpreting the information, which was collected using a variety of research methods. These selected teachers had been rated by their students with significantly higher means. In this study these teachers were called exemplary teachers and these teachers scored more than one standard deviation above the mean for at least three of the five scales.

Further, four students from the classes of each of the selected five teachers were also interviewed. Students' interviews were structured and were conducted in three phases on the same day. Interview stages were before, during and after a selected activity in the classroom. Similar questions regarding a selected activity were asked to assess students' initial perceptions about the task, during the task and when the task was completed. Randomly selected students were asked few general questions followed by a question relating to each of the five scales of SPAQ questionnaire. This approach enabled researcher to draw on a variety of paradigms to inform their interpretation in explaining the positive student perception of assessment tasks.

In the next section the gist of interviews with teachers and students are represented. The sections are derived and conclusions drawn. The main sections, which emerged from the qualitative data, are:

Learning and assessment

Interviews and observations reflected that the teachers were engaging a constructivist way of teaching underpinning formulations of formative assessment (Sadler, 1989). As supported by the quantitative data students of these teachers had a very positive perception of the assessment practices

employed by their teachers and it was found that social interactions within these classes were generally very strong. Assessment practices employed by teachers try not only to 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 student are learning. These exemplary teachers showed a shift in their views of assessment in science, by keeping themselves informed on the nature of the outcomes of the science education.

Curriculum and assessment

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

Classroom and assessment

Exemplary teachers believed that there is a need to recognise the roles and responsibilities of both teachers as well as students. This view resonates with Sadler's (1989) view, formative assessment is based on the principle that students need to become consumers as well as the object assessment activity. This socio-cultural view of learning enhances positive classroom interactions. Assessment also reflects a power relationship in the classroom. Teacher questions and students respond. However, in an exemplary teacher's class, teacher should provide enough resources for students to respond to the questions and create knowledge. These resources could be books, the internet, peers or other resource persons.

Teachers and assessment

Although these selected teachers had emancipatory views about assessment and stood apart generally from their counterparts, but they were feeling concerned about the external influences on them. Teachers felt answerable to various stake holders namely students, parents, administrators and community at large. For establishing their accountability their students had to perform well in national and international science tests. Teachers would use these test results as evidence of efficiency for their performance. Teachers also believed that knowledge and expertise of various assessment activities is mandatory for all science teachers. A teacher needs to have an in-depth understanding of the topic being taught and students existing knowledge. They recommend that this can be achieved through planning of the course content which should include teaching, learning, assessment and curriculum and their interrelationship.

Students and assessment

The final and last section of this study is identifying the students as active and intentional participants in the classroom assessment practices. Cowie (2005) highlights the multiple consequences of classroom assessment for students; 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 contributes 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. Students also tend to view assessment as a joint teacher-pupil responsibility. Students also pointed out on variety of assessment tasks.

Conclusion

In this study a new instrument, the Students' Perception of Assessment Questionnaire (SPAQ) was validated. A five scale instrument with Cronbach Ü reliability score ranging between 0.63 to 0.83, makes these scales acceptable for use in future. Also, SPAQ's ability to distinguish between classes was established, which was an important attribute 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 teachers and SPAQ was a valid instrument to do so. The exemplary were those teachers 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 SPAQ different dimensions of assessment namely Congruence with planned learning, Authenticity, Student consultation, Transparency and Diversity were identified. Observations and interview data identified same dimensions existing within different sections of the assessment process. The identified sections namely, Learning, Curriculum, Classroom, Teacher, and student are integral part of assessment. Qualitative data identified the importance and role of these sections in learning and assessment.

Assessment for learning has emerged as a 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, freedom to choose from a variety of assessments and flexibility 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 way to bring about desired changes in the educational system.

Notes

Due to page limit, tables showing the statistical results and details of interviews could not be included in this paper. These

details can be provided on request made to the author at R.Koul@curtin.edu.au.

References

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

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.

Denzin, N.K., & Lincolin, Y.S. (1994). Introduction: Entering the field of qualitative research. In N.K. Denzin & Y. S. Lincolin (Eds.), *Handbook of qualitative research,* 1-18. Thousand Oaks, CA: sage.

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: University of California.

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

Fraser, B. (1994). Research on classroom and school climate. In D. Gabel (Ed.), *Handbook of research on science teaching and learning*, 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,* 527-564. Dordrecht, The Netherlands: Kluwer Academic Publishers.

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

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

Hobden, P. (1998). The role of routine problems in science teaching. In B. Fraser & K. Tobin (Eds.), *International handbook of science education,* 219-232. Dordrecht, The Netherlands: Kluwer Academic Publishers.

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

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

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 stage 3 assessment in 1995 and 1996* (Research Report). University of Exceter, UK.

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

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

Schaffner, M., Burry-Stock, J.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: New Orleans

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*, 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 Psychology,* 77, 312-322.

Stiggins, R. (1994). *Student-centered classroom assessment.* Macmillan College Publishing Co.

Tobin, K. (1998). Qualititative perceptions of learning environments on the world wide web. In B.J. Fraser & K.G. Tobin (Eds.), *International handbook of science education,* 139-162. United Kingdom: Kluwer Academic Publishers.

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

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