

EDITORIAL ADVISORY GROUP

Dr. B. Brereton	Mr. E. Brandon
Prof. G. Rohlehr	Dr. D. Broomes
Ms. S. Craig-James	Dr. D. Clarke
Mr. M. Morrissey	Dr. J. Glasgow
Dr. V. Pollard	Dr. E. Newton

EDITORIAL COMMITTEE

Janet Fullerton-Rawlins
Jeanette Morris
June George
Ian Robertson

EDITOR: Lynda Quamina-Aiyejina

SUBSCRIPTION INFORMATION: Caribbean Curriculum is published once a year. **Price per issue:** TT \$20.00 in Trinidad and Tobago. US \$5.00 in the Caribbean, and US \$10.00 elsewhere. *Prices include postage and handling.*

© School of Education, UWI, St. Augustine, 2008

ISSN 1017-5636

CARIBBEAN CURRICULUM

Volume 15

2008

CONTENTS

Addressing Social Participation as a Major Goal in Social Studies: A Case Study of a Fifth-Form Group Pursuing the CXC/CSEC Social Studies Programme in a Barbadian School	1
<i>Anthony D. Griffith and Sonia St. Hill</i>	
Using Role-Play to Develop Science Concepts	25
<i>Rawatee Maharaj-Sharma</i>	
Extra-Lessons: A Comparison Between “Different Sides of the Track” in Trinidad and Tobago	45
<i>Samuel Lochan and Dorian Barrow</i>	
Can Standards-Referenced, Large-Scale Assessment Data Lead to Improvement in the Education System? Judging the Utility of Student Performance Standards in the Primary School National Assessments of Educational Achievement	71
<i>Jerome De Lisle</i>	
What Leadership Styles Have to Do With Teacher Job Satisfaction	115
<i>Jasmattie Yamraj with Bertram Hudson B. Ross</i>	
The Potentials of Embodied and Ethnomathematical Perspectives for Caribbean Mathematics Education	133
<i>Steven Khan</i>	
Notes on Contributors	
Information for Contributors	

**ADDRESSING SOCIAL PARTICIPATION
AS A MAJOR GOAL IN SOCIAL STUDIES:
A Case Study of a Fifth-Form Group
Pursuing the CXC/CSEC Social Studies Programme
in a Barbadian School**

Anthony D. Griffith and Sonia St. Hill

This paper, in the form of an exploratory case study of a select group of fifth-form students, employs triangulation techniques to (a) examine the extent to which the Caribbean Examinations Council's Secondary Education Certificate (CSEC) programme addresses social participation, and (b) explore the students' responses to opportunities to engage in active involvement in the social issues and affairs of their community. The findings revealed that, at the school involved, social participation is one of the goals least emphasized during the teaching of social studies. There is also a general heavy emphasis on traditional modes of instruction such as reading and notetaking. After the use of more participatory activities introduced during the unit, there were some noticeable behavioural and attitudinal changes among the participating students. The students themselves also expressed a preference for more active learning tasks and for activities that directly involve them in examining social issues in their own local community. The findings of the study could have implications for addressing what is perceived as varying levels of alienation and social disaffection among Barbadian and Caribbean youth.

Introduction

The consensus that emerges from the literature is that citizenship education is the primary, overarching goal of Social Studies, and that social participation is one of the major goals of the subject (Parker, 2005; Wraga, 1993), and the key to testing how effective social studies learning has been. Social participation is defined as active involvement by students in investigating social issues, performing certain social tasks and activities, and in demonstrating certain pro-social capabilities and dispositions both in and out of school (Griffith, 1995, p. 55).

In Social Studies, therefore, it is important that students are given the opportunity to actively participate in classroom learning activities that encourage students to think critically as they read newspapers and magazines, watch television, or monitor political or policy debates in their community. Students would thus be engaged in reflecting thoughtfully on what they are learning, in asking questions, sharing opinions, and participating in public discourse about important social issues and concerns. Social participation also engages students in the process of confronting ethical and value-based dilemmas, and encourages them to think critically and creatively, and to make personal and civic decisions based on information derived from a variety of sources.

In order to achieve this vision, what takes place in the classroom must go beyond the teaching of content only. Education is not simply about preparing students to pass examinations; it is about preparing them to be active participants in their society. The nation depends on a well-informed and civic-minded citizenry to sustain its democratic traditions, especially now as it adjusts to its own heterogeneous society and its shifting roles in an increasingly interdependent and changing world. (National Council for the Social Studies [NCSS], 1993, p. 213). Students thus need to be provided with opportunities to actively participate in social roles and practices—practices that promote growth and development, and enhance civic mindedness (Barber, 1989; Conrad, 1991; Newmann, 1989; Thornton, 2004). Social Studies, therefore, should be viewed more in terms of its meaning as an active verb—that is, something one does—rather than as a noun, or as content to be studied.

Social participation also serves the purpose of preparing students to function effectively as citizens. It is a vehicle for engaging students, at the classroom level, in social criticism and authentic deliberation on issues of public import. According to Janzen (1995, p. 135), students do not really understand the notion of citizenship unless they become actively involved in the social and political affairs of the local community, since democratic theory is tested and understood by students only as they become active participants in civic endeavours.

This form of authentic instruction (Newmann & Wehlage, 1993), where students are engaged in significant and meaningful learning activities,

and which directly links the school and the community, is an effective strategy for increasing their competence to think critically and to function as citizens in a democratic environment. These activities help to produce participating citizens who play their role as voters, taxpayers, and members and leaders of social groups within their community (Griffith, 2006; Morris, Morrissey, & King, 1996).

Exemplary social studies programmes, therefore, seek to develop social and civic participation skills that engage students in working effectively in groups to address problems by examining alternative strategies, making decisions, and taking action on them. They also involve students in negotiation and compromise, as well as active participation in civic affairs. Participation in informed public discussion of policy issues is direct preparation for active citizenship—especially when it culminates in decisions and actions that have real consequences (NCSS, 1993, p. 215)—that students themselves can experience and observe. This, according to the NCSS (p. 213), is the essence of social understanding and civic efficacy. However, if the classroom is to become meaningful to students, it must function as a laboratory for students to engage in social and authentic activities (Boyle-Baise, 2003). Therefore, the purpose of social studies—and more particularly of social participation—becomes critical if students are to understand and contribute to the society in which they live.

Context of the Issue

Given the nature and purpose of Social Studies, it is relevant here to offer a word on the social and institutional context of the problem under examination. Social Studies is taught in all primary schools in Barbados and throughout the Caribbean, as well as in most secondary schools. However, there is evidence that students in our schools are not displaying the level or type of participation and involvement associated with exemplary Social Studies.

Recently, a number of disturbing behaviours have been observed among some of our school-age children. In its White Paper on Education, the Barbados Ministry of Education, Youth Affairs and Culture [MOE], (1995) has expressed concerns about evidence of a general decline in youth behaviour in terms of values, positive attitudes, and appreciation

for existing social services and amenities. Social commentators and media reports have also highlighted a sense of apathy and disenchantment among Barbadian and Caribbean youth, with both political and social institutions. Reports in the newspaper and on radio indicate that there is a lack of respect for law, authority, property, and people in general. Students are also accused of displaying a lack of self-esteem, motivation, and a general sense of belonging to family, community, and nation (MOE, 1995, p. 18). Recently, there has also been a public outcry about the level of violence in our schools, especially violence against teachers (Rising violence, 2005, p. 6A).

This outcry is a continuation of what the literature attributes to the failure of the educational system to address the actual and felt needs of children and society, instead of placing the emphasis on transmitting fixed bodies of knowledge that allow students to pass an examination. This growing level of violence in schools may thus relate to the high level of frustration experienced by children as a result of what they perceive as a dysfunctional curriculum and of their failure to succeed in the school system (Rising violence, 2005, p. 6A).

An approach to the teaching of Social Studies that is sensitive to, and informed by, the social participation paradigm outlined earlier, it is argued, can perhaps help to address some of these concerns.

Further, with respect to the secondary schools in Barbados, there seems to be a lack of coherent policy or practice in terms of the place of Social Studies in the school curriculum. Though Social Studies is taught in most secondary schools, the Form levels at which it is offered vary widely between schools (Pennegan, 2006). In most of the schools (52%), the subject is taught from Forms 1 to 5; but in some (21%), it is taught in Forms 4 and 5 only; and in others (9%), in Form 1 only. In other schools (9%), Social Studies is taught only from Forms 1 to 3; and it is not taught at all in 9% of the schools (St Hill, 2005).

There is also a perception, both among the public and among some officials in education, that Social Studies is intended for the weaker students, and that anyone can teach it (Griffith, 1995). One thus finds the subject, on occasion, timetabled for one or two periods per week, and often during the last period of the day. Such practices exist in spite of the

view recently expressed in the Barbados policy document entitled *Curriculum 2000: Rationale and Guidelines for Curriculum Reform in Barbados* (MOE, 2000) to the effect that:

The revised curricula at the primary and secondary levels seek to ensure that all students receive quality education which will enable them to be productive citizens capable of creative and critical thinking, and of applying the problem-solving skills relevant to the complex challenges of our modern society. (p. 3)

The evidence also suggests that some secondary school principals have a very positive perception of the value of Social Studies and its contribution to student development—particularly with respect to developing their critical thinking skills and to understanding their culture and heritage (Pennegan, 2006). These principals are, however, often constrained in fully offering the subject on the curriculum, due partly to the competing demands of, and for, other more traditional subjects such as geography and history, or newer subjects such as computer studies and business studies.

The Caribbean Examination Council (CXC), which came into being in 1973, was established to examine and certificate students at the secondary level in Caribbean schools. Social Studies was first examined in 1981, and currently has the third highest number of candidate entries of all subjects (CXC, 2005). The most recent Social Studies syllabus (CXC, 2002) rationalizes the inclusion of the subject in the secondary school curriculum on the premise that adequate provision should be made for enabling students to gain the knowledge and skills that will prepare them to be effective citizens. The syllabus further specifically mentions effective social participation, which it notes as integral to the learning and teaching of Social Studies, and should “be informed by a sense of commitment to the development of the community” (CXC, 2002, p. 1).

Unfortunately, however, the teaching of Social Studies seems to be heavily content-oriented and teacher-centred, rather than activity-oriented and learner-centred. One notes, for example, the frequency with which objectives in the CXC and local syllabi are prefaced with lower-level cognitive terms such as: define, state, describe, identify, and list.

This contrasts with the relative paucity of higher-level cognitive terms such as, for example: account for, justify, evaluate, assess, propose a plan for .., present a case for (or against) ..., evaluate your experience as a member of ... [a named community organization].

Brophy and Alleman (1993) have emphasized the use of activities as a critical means of complementing what is taught in schools, through participation in the community and the use of its resources and environment. Instead of engaging students in these types of activities, however, the local Social Studies syllabus, as taught, seems to focus more on the acquisition of factual information rather than on engaging students in practices that reflect social participation and enable them to develop meaningful and fulfilling ties within their communities (Griffith, 1999). Such an approach may be described as being grounded more on “passivities” rather than on “activities.” The emphasis is therefore more on the passive acquisition of knowledge, and on lower-level forms of knowledge intake such as reading, listening to lectures, and notetaking.

A possible consequence of this is that students are not adequately aware of what citizenship is about. However, if our students are expected to become strong leaders and decision makers of the future, then their classroom experiences should provide them with opportunities to think critically, make decisions, solve problems without violence, and actively participate in and outside of the classroom. This paper raises the hitherto unexplored instructional approach of teaching Social Studies to specifically address the goal of active social participation and social involvement. It thus examines the extent to which the CXC/CSEC Social Studies programme can, in fact, promote/provide students with opportunities for such social participation.

Purpose of the Study

In spite of the many efforts by the Ministry of Education in Barbados to upgrade the social studies syllabi at both primary and secondary level, as well as the training of social studies teachers, many teachers continue to experience great difficulties in conceptualizing and teaching social studies as it ought to be taught. Though social participation is often explicitly stated in the local Social Studies syllabus, this is apparently just as often not reflected in most Social Studies classes. This paper will

analyse the response of a selected group of fifth form students, at a secondary school in Barbados, to the infusion or use of social participation activities in their Social Studies class.

This paper will attempt to examine (a) the types of activities that these students undertake in the CXC/CSEC Social Studies programme which promote social participation; (b) the views of the Social Studies teachers at the school on promoting social participation activities in the CSEC Social Studies programme; (c) the views of the students in the group on incorporating social participation activities in the CSEC Social Studies programme; and (d) whether the students exhibit any greater interest or sense of involvement in social participation activities **after** these have been introduced in this particular CSEC Social Studies class.

Methodology

This study utilizes a case study design that involves the observation and study of an individual or bounded unit (Gay & Airasian, 2003)—in this case, a select class of students who were preparing for the CXC examination in Social Studies. As such, the study is also essentially exploratory in nature.

The Participants in the Case Study

The student body of the Barbadian secondary school in question totalled 965 students (410 males and 555 females). The teaching staff of 64 teachers is 55% female and 45% male. At this institution, all the students in Forms 1 and 2 are required to take Social Studies, whereas the choice is optional in Forms 3 to 5.

The selected group was identified on the basis of the following criteria (Burns, 2000, p. 465). The unit consisted of one group of Form 5 students who were following the CXC/CSEC Social Studies syllabus. Based on their academic history, this group of students was classified as “under-performers” who displayed a lack of interest in school. They were also often absent from school and were frequently cited for disciplinary problems. This class was made up of 10 male students between the ages of 15 and 17, all of whom had taken social studies prior to Form 4. Another criterion used in selecting this class was the fact that

the Social Studies teacher was, at the time, pursuing a postgraduate degree in Social Studies education at The University of the West Indies. Eighty percent (80%) of the students in the class indicated that they had last taken social studies in Form 2, whereas 20% stated that they had last taken the subject in Form 3.

Five teachers participated in the study. They all teach Social Studies throughout the school and are familiar with the CSEC Social Studies syllabus. All of the teachers hold graduate degrees from approved universities and four of the five are trained teachers.

As a case study, and given both the nature and the size of the unit, it is recognized that there are clearly certain inherent constraints on the extent to which any findings or conclusions can be generalized to the school or the country as a whole.

Instruments

Five instruments were used to collect data for this study: (1) a Pre-Test and Post-Test, (2) a Teaching Unit, (3) a Student Questionnaire, (4) a Teacher Questionnaire, and (5) a Classroom Observation Schedule. The student questionnaire was administered both before and after the unit was taught.

The test was based on the unit topic, “The Family,” and was designed to evaluate not merely the students’ knowledge about the topic but, even more importantly, their attitudes, interest, and response with respect to the use of participatory activities in Social Studies lessons. The test also required them to articulate their feelings on an issue related to the family.

The Teaching Unit

The unit, of six weeks duration, was based on the CSEC Social Studies Syllabus 2002, and focused on the use and promotion of social participation activities among students, through engaging them in both out-of-class and out-of-school learning experiences. These experiences involved the use of a variety of “participating” activities such as debates, social investigations, projects, community activities, and class presentations; as well as a variety of media such as local music,

newspapers, magazines, and news reports. For the purpose of the study, the teaching of the unit thus moved away from the more content-oriented, teacher-centred approach, and adopted a more activity-oriented, learner-centred paradigm.

The Findings

The major findings of the study suggested that:

1. Social participation, as a major goal of Social Studies, was hardly addressed, much less emphasized, by the teachers who took part in the study.
2. The selected group of students seemed to prefer to be more actively involved in their Social Studies classes, and to have a positive attitude towards activities such as fieldtrips, project work and class presentations.
3. After these students were exposed to classroom activities that required more active participation, they showed greater willingness to become involved in issues relating to their local group and community.

Discussion of Findings

1. Teaching for Social Participation in the Social Studies Class

One of the findings of this research clearly indicates that, even though the teachers and administrator are aware that social participation is one of the major goals of social studies, there seemed to be no correlation between this perception and the actual incorporation of this goal into the teaching of the CSEC social studies syllabus. The social studies classroom in the case study initially exhibited rather limited evidence of participatory activities. Students were mostly engaged in passive forms of knowledge intake, such as notetaking, question-and-answer, and reading. Table 1 indicates that these activities were rated, by both teachers at the school (4.0, 3.8, and 3.4 respectively) and by students in the study (5.0, 4.1, and 4.0 respectively), as those most frequently engaged in. These passivities offered little or no active participation and were also only marginally connected to the world beyond the school.

Lessons that have little or no value beyond the classroom are geared only for success within school, since the students' work would serve only to certify their level of compliance with the norms of formal schooling (Newmann & Wehlage, 1993, p. 10).

Table 1. Frequency and Ranking by Teachers and Students of Instructional Techniques Used by Teachers (Pre-Unit)

Instructional Techniques	By Teachers		By Students	
	X	Rank	X	Rank
Field Trip	2.4	8	1.2	9
Lectures	4.2	1	1.7	4
Notetaking	4.0	2	5.0	1
Discussion	3.8	3	4.1	2
Reading	3.4	4	4.0	3
Sociodrama	1.8	10	1.3	7
Debates	2.0	9	1.3	7
Class Presentation	3.2	5	1.5	6
Group Work	3.2	5	1.6	5
Projects	3.0	7	1.1	10

X = Mean Frequency

Thus, the evidence collected did not reveal any high levels of either active or reflective participation by students. Active participation requires students to be actively involved and participating in social organizations such as groups, in investigating social issues, and participating in community service work (Griffith, 1995, p. 56). This was, however, not reflected in the pre-unit data collected from both students and teachers. In fact, as indicated above, the type of participation observed was more passive in nature; and Table 1 reveals rather low ratings given by both teachers and students to such real "classroom activities" as group work, field trips, projects, and debates. The more traditional teaching strategies, such as "chalk and talk," were more commonly in evidence, and these did not provide opportunities for students to communicate with each other or to become actively involved in their community.

Further, few examples of reflective participation or civic mindedness were observed from the pre-unit data. There was, in fact, very little evidence of students being placed in settings and roles that require them to think critically and participate in public processes which influence and guide decision making on public issues. According to Griffith (1995, p. 58), this type of participation is an effective and meaningful strategy for not only increasing the competence of students to think critically, and to function and participate as citizens, but also for equipping them with adequate knowledge of their own social, legal, political, and economic subsystems.

The data from the teacher's and students' questionnaire, therefore, suggest that social participation was definitely one of the goals least emphasized during the teaching of Social Studies at the school. This conclusion is reflected in the literature, which posits that social studies instruction in the Caribbean has traditionally focused almost exclusively on the first three goals of knowledge, skills, and attitudes, to the exclusion of the higher goals of social participation and citizenship education (Griffith, 1995, p. 52). Griffith goes on to state that, even with respect to these latter goals, these are mainly addressed in an academic/intellectual sense and almost exclusively within the artificial confines of the classroom.

Howe and Marshall (1999, p. 15) further note that many of the social studies syllabi in the Caribbean focus on "intake activities," or low-level, "passive type" activities such as reading and notetaking. The bigger goals of social participation and citizenship education, even though mentioned in the national statements on education, have been studiously neglected, in any explicit form, in social studies curricula and instructional practices (Griffith, 1995). Activities to promote social participation and citizenship education, therefore, are clearly not deliberately articulated in the syllabus.

In relation to the CSEC Social Studies syllabus (2002), most (some 60%) of the teachers at the school stated that there were no participatory activities written into the objectives of the syllabus, even though there are topics that provide opportunities for such activities. The data from the teachers' pre-unit questionnaire showed that most (60%) of the teachers were also cognizant of the opportunities available through the CSEC

syllabus for introducing some participatory activities in the classroom. Some examples given of these were: (a) visiting homes for the elderly and children, (b) organizing a family day for problem students and parents, and (c) organizing a clean-up campaign within the school and the community. However, in spite of these examples, the teachers appear to have done little to encourage or introduce such activities in the classroom.

The teachers were, in fact, portrayed as the traditional instructors who dominated the classroom, keeping students quiet and passive. Various explanations were given for this. Some teachers observed that participatory activities would take up too much valuable teaching time and would therefore prevent them from completing the requirements of the CSEC syllabus. The teachers also felt that the syllabus was too broad for students to complete if they were to be engaged in such activities. Another issue mentioned was the fear of students being injured while taking part in participatory activities, and the consequent issue of liability. Teachers noted, for example, that on field trips the legal requirement is for one teacher to supervise each group of 20 students; and, with the generally large size of classes this was often not possible. Field work, however, can be readily planned and executed on the school compound, or in the immediate vicinity of the school; and the debate on “depth” versus “breadth” of coverage (Olsen, 1995) offers teachers some guidance in maintaining a balance in their approach. Opportunities for such active involvement in the local community are critical in the development of social understanding and civic efficacy, including such critical learnings as values, and social and interpersonal skills.

2. Student Response to More “Active” Participation

The second finding suggests that these students prefer, and were requesting, more participatory activities in their social studies lessons. Table 2(a), for example, indicates that prior to the unit being taught, reading was highly rated (4.0) by students as a preferred learning activity, and was followed at some distance by group work (1.6), and dramatizations and debates (1.3). After the unit was introduced, however, field trips (4.8), group work (4.5), and dramatizations (3.8) emerged as the preferred learning activities among the students—and all with markedly increased ratings. Clearly, the students were asking for

classroom activities that place them directly in the community and directly address social issues. The students therefore seemed to be asking educators to, in fact, fulfil the initial purpose for the introduction of social studies to the Caribbean classroom.

Table 2. Students' Preference for the Use of (a) Activities and (b) Media in Their Social Studies Classroom (Pre- and Post-Unit Data)

(a) Activities	Pre-Unit		Post-Unit	
	X	Rank	X	Rank
Field Trips	1.2	5	4.8	1
Reading	4.0	1	3.6	4
Dramatizations	1.3	3	3.8	3
Project Work	1.1	6	3.5	5
Debates	1.3	3	3.1	6
Group Work	1.6	2	4.5	2
(b) Media				
Newspapers	1.8	2	3.5	3
Magazines	1.6	3	4.4	1
Local Music	1.3	5	4.3	2
Textbooks	4.7	1	3.5	3
News Broadcasts	1.6	3	2.9	5

While the teaching of social studies is conceptually linked to life beyond the classroom, it did not seem to be effectively taught in order to provide students with the practical application and experience necessary for fulfilling the stated rationale of the CSEC Social Studies syllabus. The pre-unit data, from both the teachers and students, support this conclusion. While the teachers maintained that the topics on the syllabus offered opportunities for active social participation, the data show that there was little evidence of such opportunities being provided in practice.

In order for social studies to deliver its goals effectively, it is not only the curriculum that matters, but also the teaching and learning strategies employed. This was noted by Brophy and Alleman (1993), who state that:

A curriculum is not an end in itself but a means, a tool for accomplishing educational goals....the ways that content is represented and explicated to students, the questions that will be

asked, the types of teacher-student and student-student discourse that will occur, the activities and assignments, and the methods that will be used to assess progress and grade performance – will be included because they are believed to be necessary for moving students towards accomplishment of the major goals. (p. 27)

Even more importantly, at the heart of social participation is the meaningful contribution that students can make to their society through the resolution of social problems, decision making, and problem solving. These cannot be effectively achieved if students are kept in the sterile environment of the classroom, learning and regurgitating facts, as needed to pass an examination.

The role of the teacher in all of this is critical. Griffith (1999, p. 13) notes that a number of theorists have identified teacher attitude as a critical variable in both the teaching and the learning of social studies. The teacher's love and enthusiasm for the subject and his/her effort to make it interesting, alive, and exciting can be transferred to students. Students, it would seem, are somewhat more sensitive to the value and benefits of "activities" and active participation in the Social Studies classroom than teachers may realize.

At the moment, however, it appears that it is only in respect of the school-based assignment, which is compulsory for passing the CXC examinations, that teachers encourage students to engage in any meaningful participatory activities. This project is intended to be a research paper done by students on a designated area of the syllabus. Through this project, students would not only be exposed to skills in collecting and processing data, but would have the opportunity to use the higher-order skills of data analysis, synthesis, and problem solving. The NCSS (1993, p. 213) argues that teaching social studies is powerful when it is meaningful, active, challenging, value-based, and integrated. However, the strategies and techniques used by the Social Studies teachers at the school were largely traditional, with very little emphasis being placed on meaningful activities such as community involvement and out-of-class field work.

This finding further suggests that there was a noticeable absence of opportunities for students to develop their communication skills and

public speaking abilities. None of the students in the group had, prior to the unit, participated in public speaking or in activities that allowed them to develop their communication skills. The teaching and learning of social studies, however, require that students be exposed to co-curricular activities such as student government, peer leadership programmes, the school newspaper, and the debating society (Wraga, 1993). These activities promote student-to-student interaction as well as opportunities for engaging students in public speaking and in other forms of communicating their own views and ideas. Ferguson (1991) found, for example, that a positive correlation exists between involvement in school activities such as governance and other co-curricular activities, on the one hand, and later involvement in civic and community affairs. These activities can also provide valuable sources of information about the world beyond the school.

The heavy reliance on traditional ways of teaching also demonstrates some deficiency in the skills exhibited by these teachers. The teachers in this population, for example, did not utilize a variety of media in their classroom on a regular basis, but rather relied heavily on the textbook and accompanying notes as major teaching techniques. This deficiency was highlighted in the teachers' and students' responses to items on the questionnaire about the use of various media in the classroom. According to Table 2(b), prior to the unit being taught, students indicated a clear preference for the use of textbooks (4.7) as, by far, the major form of teaching media, with lowest ratings given to local music and magazines—perhaps reflecting their own classroom experiences. After the unit, however, local magazines (4.4), and music and newspapers (3.5) were the most highly rated forms of media. Though the textbook remained fairly highly rated, it was the only media to have received a lower frequency rating than it did prior to the unit.

The suggestion, therefore, seems to be that after students have been introduced to the active and meaningful use of a variety of local instructional media, other than the textbook, they seemed to better appreciate the relevance and value of these media to the study and understanding of local issues and concerns, and to making Social Studies more alive, interesting, and authentic.

Activities that involve the use of a variety of resource materials and local media help to motivate students and provide opportunities for them to interact with curricular content and accomplish curricular goals (Brophy & Alleman, 1993). The use of various media and resources, for example, the overhead projector and local newspapers as utilized during the unit, allows students to interact with each other and use higher-order thinking skills, as opposed to the more passive traditional approach of relying on textbooks and worksheets. Dependence on the traditional approach to teaching places the students in a passive role and makes social studies content remote. Brophy and Alleman also argue that hands-on inquiry and discovery approaches to social studies encourage children to be active learners. The NCSS (1993, p. 218) states that powerful social studies requires the use of a variety of instructional materials such as photographs, maps, illustrations, and videos. Through using these resources and media, students learn by doing, and both their attention to the subject matter and their retention of the material being studied increases.

3. Students' Attitude to Social Participation After the Unit Was Taught

The post-unit questionnaire and the observation schedule were designed to observe which, if any, indicators of social participation occurred during the teaching unit, and whether there were any changes in students' active participation and interest in the Social Studies lessons. The third finding, according to both instruments, shows that the students' responses and attitudes to the unit exhibited active enthusiasm and interest. The students participated willingly in all activities and indicated that they particularly liked the (to them) "new" activities such as role play and community field work, or out-of-class activities that were introduced in their classes.

Whereas prior to the unit, a number of passive, non-participatory activities and behaviours were evident, the activities in the unit involved students in very little "seat work," based on reading and written assignments. Rather, the unit promoted student involvement in both the school and the wider society, through activities that fostered creativity, active participation, and problem solving. Students were exposed primarily to behaviours such as communicating with each other,

investigating social issues, public speaking, and policy formulation—both within the school and in relation to the local community. Instead of merely compiling notes on social issues, students were involved in actually investigating and analysing these issues through projects and field activities. The focus therefore shifted from learning facts to be regurgitated for tests—as evident at the beginning of the unit—to questioning these facts (Evans, Newmann, & Saxe, 1996), and analysing and applying information from newspaper articles and other relevant sources to current social issues. In addition to the above activities, as the unit progressed, students were also exposed to different instructional techniques, such as group work and sociodrama, which stimulated their interest and strengthened their desire to participate. These teaching techniques were complimented by the use of media such as the computer and the Internet.

After the unit, the students exhibited more interest in local social issues and were more willing to discuss these in class. They also displayed a growing level of creativity in the form of the poems, songs, group work, and debating skills displayed during class presentations. Students further indicated that they (a) socialized more with their schoolmates outside of the classroom than they did prior to the unit (100% after the unit as opposed to 70% before); (b) had greater opportunities to engage in public speaking (80% vs nil) and to develop their communication skills (90% vs 30%); and (c) engaged in more Social Studies lessons that involved activities in the community (80% vs 10%). In addition, the students in the group performed significantly better ($p < 0.01$) on the post-test (mean score 14.3) than they did on the pre-test (mean score 10.6); and this was especially so in the cases where their written responses to items that required extended answers were more detailed, better developed, and more fluently written. The community activity undertaken by the class, where students engaged in setting up a family life booth to distribute information on family planning, was also indicative of their interest in active social participation at both the school and community level. Through such activities, it has been noted, students develop new understandings through the process of active construction of knowledge (NCSS, 1993; Sunal & Haas, 2007).

During the initial period of the unit, the researcher had difficulty in eliciting student involvement or interest in the political affairs of their

country. In most cases, students indicated that they were disenchanted with the governing of the country and therefore were not interested in taking part in the process of choosing parliamentary representatives. By the end of the unit, however, students were exhibiting a greater interest in political affairs and a slightly greater willingness to vote, if given the opportunity (40% as opposed to 30% before the unit).

The literature reveals, in fact, that traditionally taught civics and government classes are sterile. These classes are associated with diagrams of “How a bill becomes a law,” and with details on the three branches of government. In these classrooms, democratic values are discussed in the abstract, conflict is absent, and participation invisible. As a result, rather than promote active political involvement, the traditional civics curriculum portrays government as a complex entity quite removed from students’ lives (Avery, Sullivan, Smith, & Sandell, 1996, p. 22). When participation is discussed, it is usually linked to the conventional forms of participation such as voting and letter writing. Students therefore tend to view their political role as that of a passive bystander. This assessment was also noted by Griffith (1995, p. 10), who stated that these participation activities tend to focus on things such as the election of class prefects, some involvement in school clubs, and a few cases of student societies. These activities and groups, however, tend to be under the direct supervision of teachers, and therefore offer students very little in terms of autonomy or involvement in decision making or the formulation of policy. Thus, students are rarely given the opportunity to engage in authentic participatory activities that promote the acquisition and construction of knowledge, or in the process of governing. Guttman (1987), for example, argues that deliberation about public issues is essential to democratic education. Secondly, students also need to become familiar with a repertoire of strategies for meaningful participation in the democratic process (Evans et al., 1993).

All social studies classes should therefore include community service activities that would engage students in making meaningful choices in relation to issues that affect their community. Such activities have been also proposed by Griffith (1995), who notes that these and other civic educational activities must be more specifically and deliberately written into curriculum practice. Thus, provision for such social action, together with linkages to relevant social organizations, should be tied to the

curriculum in such a way as to meaningfully engage the students in direct involvement in, and reflection on, the functions, activities, and processes of the various organizations and institutions in their society.

Conclusion

This study sought to address the notion that, in the teaching and learning of Social Studies, social participatory activities are necessary to engage students in meaningful learning activities, which would not only increase the competence of students to think critically, but would also enable them to function and participate as citizens in a democratic environment. At the same time, such activities would equip students with an in-depth knowledge of their own social, legal, political, and economic subsystems.

Evidence from this research suggests that the teaching unit, which employed a variety of activity-based and authentic learning exercises, appeared to be related to a greater willingness by students to participate and become involved in the affairs of their community. After the unit, the students in the group also seemed to have improved their communication skills and appeared more willing to vote and to be involved in the governance of their country.

Given that these improvements seem to have occurred within a small group of academically underachieving students at one school, this certainly raises intriguing questions about the potential outcomes of such an approach with a larger, more representative group of students.

In a recent newspaper article (Comments reported, 2005), Rosina Wiltshire, Resident Representative of the United Nations Development Programme, in responding to points raised about the need to put curricula in place to respond to the social problems within the Caribbean, stated that:

the educational process has to determine whether the goal is to pass exams or shape the human being for life,... from the time we determined that the whole purpose of education was simply to pass exams, we lost our footing. (p. 3)

According to Newmann and Wehlage (1993, p. 8), reflective participation is equated with authentic instruction and with the NCSS position statement (1993, p. 213), that “powerful social studies teaching... enables students to... develop social understanding and civic efficacy.” This type of participation is also both constructivist and transformative.

The growing problems among our youth suggest that more should be done with them during their formative years. If one is to address the general decline in the behaviour of our youth, as alluded to earlier, all stakeholders in education must address the relevance, need, and worthwhileness of a curriculum infused with social participatory activities. Already in Barbados, under Curriculum 2000, Social Studies has been identified as one of the core subjects required by all students at the secondary level. The CSEC Social Studies syllabus has the ability to address this issue, if the higher goals of social participation and citizenship education are seen as an integral and central purpose in the teaching and learning of Social Studies.

It is acknowledged that the nature and size of the selected group for this case study clearly limit any generalizations on the teaching of Social Studies and on the implications of introducing more social participation activities in the Social Studies classroom in Caribbean schools. The study, nevertheless, offers a new and intriguing line of research that illuminates the importance of teaching Social Studies more, or even essentially, for social investigation and for active social involvement and participation, rather than for mere knowledge and factual information.

Further research, using a larger, more representative sample of students and Social Studies teachers, would certainly provide another layer of understanding with respect to these initial, tentative conclusions.

References

- Avery, P. G., Sullivan, J. L., Smith, E. S., & Sandell, S. (1996). Issues-centered approaches to teaching civics and government. In R. W. Evans & D. W. Saxe (Eds.), *Handbook on teaching social issues* (NCSS Bulletin No. 93, pp. 199–211). Washington, DC: National Council for Social Studies.
- Barbados. Ministry of Education, Youth Affairs and Culture (1995). *White paper on education: "Each one matters — QUALITY education for all."* Bridgetown, Barbados: Author.
- Barbados. Ministry of Education, Youth Affairs and Culture (2000). *Curriculum 2000: Rationale and guidelines for curriculum reform in Barbados.* Bridgetown, Barbados: Author.
- Barber, B. R. (1989). Public talk and civic action: Education for participation in a strong democracy. *Social Education*, 53(6), 355–356, 370.
- Boyle-Baise, M. (2003). Doing democracy in social studies methods. *Theory and Research in Social Education*, 31(1), 51–71.
- Brophy, J., & Alleman, J. (1993). Elementary social studies should be driven by major social education goals. *Social Education*, 57(1), 27–32.
- Burns, R. B. (2000). *Introduction to research methods.* Thousand Oaks, CA: Sage.
- Caribbean Examinations Council. (2002). *Secondary Education Certificate Examinations: Social studies syllabus, May/June.* St. Michael, Barbados: Author.
- Caribbean Examinations Council. (2005). *Report on the administration of the CSEC and CAPE 2005 examinations: The performance of candidates and related matters* (Statistical Bulletin for the Meeting of Council, 2005). Kingston, Jamaica: Author.
- Comments reported by the Resident Representative for the UNDP. (2005, July 24). *Barbados Advocate.* p. 3.
- Conrad, D. (1991). School-community participation for social studies. In J. Shaver (Ed.), *Handbook of research on social studies teaching and learning* (pp. 540–546). New York: Macmillan.
- Evans, R. W., Newmann, F. M., & Saxe, D. W. (1996). Defining issues-centered education. In R. W. Evans & D. W. Saxe (Eds.), *Handbook on teaching social issues* (NCSS Bulletin No. 93, pp. 2–4). Washington, DC: National Council for Social Studies.
- Ferguson, P. (1991). Impacts on social and political participation. In J. Shaver (Ed.), *Handbook of research on social studies teaching and learning* (pp. 385–399). New York: Macmillan.
- Gay, L. R., & Airasian, P. (2003). *Educational research: Competencies for analysis and applications* (7th ed.). Columbus, OH: Merrill Prentice Hall.

- Griffith, A. D. (1995). Social participation as a major goal in social studies: Towards authentic instruction and learning in Caribbean schools. *Caribbean Curriculum*, 5(2), 47–66.
- Griffith, A. D. (1999). *Social studies in Caribbean schools: Some challenges for instruction and assessment* (EFA in the Caribbean: Assessment 2000. Monograph Series; No. 2). Kingston, Jamaica: UNESCO.
- Griffith, A. D., & Barth, J. (2006). *Social studies curriculum and methods for the Caribbean*. Mona, Jamaica: University of the West Indies Press.
- Guttman, A. (1987). *Democratic education*. Princeton, NJ: Princeton University Press.
- Howe, G., & Marshall, D. (1999). *Citizenship education, democracy and global shifts: Re-thinking Caribbean social studies* (EFA in the Caribbean: Assessment 2000. Monograph Series; No. 7). Kingston, Jamaica: UNESCO.
- Janzen, R. (1995). The social studies conceptual dilemma: Six contemporary approaches. *The Social Studies*, 86(3), 134–140.
- Morris, P., Morrissey, M., & King, R. (1996). The development of social studies in the Commonwealth Caribbean. In D. Craig (Ed.), *Education in the West Indies: Development and perspectives, 1948–1988* (pp. 232–248), Mona, Jamaica: Institute of Social and Economic Research, UWI.
- National Council for the Social Studies. (1993). A vision of powerful teaching and learning in the social studies: Building social understanding and civic efficacy. *Social Education*, 57(5), 213–223.
- Newmann, F. M. (1989). Reflective civic participation. *Social Education*, 53(6), 357–360.
- Newmann, F. M., & Wehlage, G. (1993). Five standards of authentic instruction. *Educational Leadership*, 50(7), 8–12.
- Olsen, D. (1995). “Less” can be “more” in the promotion of thinking. *Social Education*, 59(3), 130–134.
- Parker, W. C. (2005). *Social studies in elementary education* (12th ed.). Upper Saddle River, NJ: Prentice Hall.
- Pennegan, M. (2006). *An investigation into the attitudes and perceptions of administrators to social studies education in secondary schools in Barbados*. Unpublished master’s thesis, The University of the West Indies, Cave Hill, Barbados.
- Rising violence against teachers*. (2005, January 16). *Sunday Sun*, p. 6A.
- St. Hill, S. (2006). *Investigation into the extent to which the Caribbean Examinations Council’s social studies syllabus provides students with opportunities to achieve the goal of social participation*. Unpublished master’s thesis, The University of the West Indies, Cave Hill, Barbados.
- Sunal, C., & Haas, M. (2007). *Social studies for elementary and middle grades: A constructivist approach*. Boston, MA: Allyn & Bacon.
- Thornton, S. (2004). *Teaching social studies that matters: Curriculum for active learning*. New York: Teachers College Press.

Wraga, W. G. (1993). The interdisciplinary imperative for citizenship education.
Theory and Research in Social Education, 21(3), 201–231.

USING ROLE-PLAY TO DEVELOP SCIENCE CONCEPTS

Rawatee Maharaj-Sharma

There is an urgent need in science education to explore new and innovative ways to teach science to children at all levels of the education system in ways that will pique their curiosity and hold their interest. This article reports on the use of role-play to develop children's ideas and understandings about energy concepts and energy conservation issues/practices. The findings show that children retain many abstract science concepts when taught using this approach and, further, that this method of instruction allows them to make linkages and distinctions among several energy concepts and to appreciate the importance of energy conservation.

Introduction

Role-play is an exciting and creative way to develop children's science ideas by allowing them to physically participate in their science learning, thereby making the experience more memorable.

This article explores how children's science concepts and ideas can be developed using role-play. Recent developments in science teaching and learning have suggested that while practical hands-on activities are useful and effective ways to foster conceptual development in science, other activities that allow for psychomotor engagement as well as expression of emotions and feelings can be equally rewarding and stimulating for science students (Banister & Ryan, 2001). The focus in this project is on learning about energy concepts, including energy conservation issues and practices, in a Standard 4 primary school classroom. The approach adopted will be described, the findings as well as the benefits of the approach will be presented, and some of the challenges encountered in implementing the approach will be identified.

The aims of this project are to investigate the ability and potential that role-play has to:

1. develop students' conceptual understandings about energy and energy-related concepts;
2. explore students' views/feelings about energy conservation practices.

Literature Review

Best practice in science education, as reported by Haury and Rillero (1994), is often perceived as being based on practical activities to facilitate hands-on learning, in order to allow children to build knowledge and develop scientific understandings. The role of the teacher in such instances becomes that of facilitator for such scientific experimentations. Goldring and Osborne (1994) and Millar (2002) have suggested that while there are tremendous benefits linked to the use of activities of this type, practical activities might not always be the best way to introduce a new idea or to facilitate the grasp of conceptual understanding.

Craven and Penick (2001) have suggested that role-play in the classroom has the potential to capture children's attention and to stimulate their imagination, thereby making it a useful strategy for science teaching. Graber et al. (2001) have further suggested that role-play is a good way to liven up science teaching and develop the students' interest in science. The versatility of role-play as a teaching/learning strategy makes it appealing at all levels of the educational system and for a variety of abstract science topics. Francis and Byrne (1999) reported that using role-playing exercises in the teaching of physics and astronomy, even at the undergraduate level, "deepens student understanding and dramatically increases the level of classroom interaction" (p. 206).

McSharry and Jones (2000) provided a theoretical basis for the use of role-play in science teaching and also presented some ideas that can help science teachers develop role-play scenes. Bonnet (2000) used some of the ideas suggested by McSharry and Jones to develop role-play scenes that were used to analyse the impact of the strategy on the learning of children aged 8–10 years. He found that not only did the strategy actively engage students in science learning but it also facilitated the emergence of values such as tolerance, responsibility, and autonomy among students.

In work on using role-play in the classroom, involving three teachers and a university academic, with mixed-ability classes from three high schools in New South Wales, Ladrousse (1989), reported that role-play had tremendous potential to allow students to demonstrate their

understanding, explore their views, and develop deeper understanding of phenomena. Building on the work done by Ladrousse, Aubusson, Fogwill, Barr, & Perkovic (1997) found that while role-play provided tremendous affective gains for students, its use to portray ideas, promote discussion, and facilitate scientific analogical reasoning among students could not be ignored.

Dallman-Jones (1994) posited that there are two different kinds of role-play that can be used in the classroom. Type I involves having children act as if they are components of a physical or biological system, for example, atoms of a molecule or organelles in a cell. Type II involves an ethical issue, and students act as humans in a situation where a decision must be made, for example, making a decision about whether or not a teenager should have an abortion. In this type of role-play, different children are given brief descriptions of the role they will assume and may be asked some stimulating questions during the role-play exercise to reveal insights about their views and feelings about the topic of the lesson. Adopting Type II role play in a study involving primary school science students, Dallman-Jones further reported that the intricate relationships among the role-players in a variety of situations allow children to show their understandings about the particular character they are impersonating, and also to recognize the relationship they have with the other role-players.

In this work, Type II role-play will be used as children explore their understandings about energy and energy-related concepts as well as their views and feelings about energy conservation practices.

Methodology – The Approach

The investigation was carried out in the science classroom of a Standard 4 primary school, located in central suburban Trinidad. The class consisted of 29 children aged between 9 and 11 (males and females) of mixed abilities. An overview of previous formative and summative assessments of the performance of students in tests done the previous term revealed that there was a range of performance among the students, with some scoring very high marks (85% and above) and some scoring very low marks (20–35%); but with 60% of the class scoring in the 40–80% range.

To maintain confidentiality, a number was used to refer to each child. The topic “Energy — Concepts and Conservation” was chosen because this was the topic the teacher was teaching in the normal course of the term at the time the project was undertaken.

The entire project consisted of five segments as shown in Table 1.

Table I. The Five Segments of the Project

Segment	Timing	Activity
1	Week 1, Day 1	Gauging children’s prior knowledge/ understandings about energy, energy consumption, and transformation processes
2	Week 1, Day 3	Introducing the role-play about energy consumption and transformations
3	Week 2, Day 2 & Day 4	Children act out the role-play
4	Week 2, Day 5	Finding out what new knowledge and understandings children gained about energy, energy consumption, and transformation processes
5	Week 6, Day 2	Interview with students

Segment 1: Gauging Prior Knowledge/Understandings and Feelings

In this segment, the children wrote responses to five open-ended questions:

1. What is energy?
2. Where does energy come from?
3. What is energy used for?
4. What happens to energy when you “use” it?
5. How do you feel about energy that is wasted?

Each question was read aloud to the class and the children had about five minutes to individually respond, in writing, to each of the questions. Children with spelling, writing, and other literacy challenges had their responses tape-recorded.

Segment 2: Introducing the Role-Play

The information gained from Segment 1 was used to create a role-play [by the researcher in close collaboration with the class teacher] called “The life of Eddy Energy,” which incorporated a number of scientific ideas the children had learned earlier in the term, as well as ideas and understandings emerging from their written/verbal responses to the questions in Segment 1. Scene 1 formed part of the introduction in which the play started by addressing the first question in a scene in which Eddy Energy is a new student and is asked to introduce himself to his classmates. Some of the other students also introduce themselves as various devices that need energy in different forms, for example, Anthony Airplane, Susan Skater, and Terry TV.

[Throughout the enactment that followed, the researcher sat in the classroom, mostly as an observer, taking notes and interjecting occasionally—when requested by the teacher either to assist children with their costumes or to repeat a response made by a student—but it was primarily the class teacher who facilitated the enactment by guiding students through their roles, asking questions, probing responses, and writing students’ responses on charts posted on the walls of the classroom.]

Segment 3: Children Act Out the Scenes

In the following four scenes, which constituted Segment 3, Eddy interacted with his new friends in a variety of situations all tailored to introduce and expose students to concepts, ideas, and relationships aimed at developing the understandings embedded in the last four questions in Segment 1. Each scene addressed one of the questions. In Scene 2, Eddy is asked by his friends to explain more about himself, and the conversation continues along with the acting between Eddy and his new friends leading to an exploration and an eventual uncovering of the sources of energy. In the following scene, Scene 3, Eddy is asked to

share his talents and to show his friends some of the things he can do. In this scene, Eddy interacts extensively with his friends, showing them that while each of them has a particular skill or talent, there comes a time when they would each need help to continue to do their respective jobs. The scene unfolds as Eddy shows himself off, indicating his numerous abilities and that he can help each of them do their jobs either faster or for a longer period. In the fourth scene, Eddy and friends are having a wonderful time, with his friends repeatedly calling upon him for assistance to carry out their respective tasks, until, in the latter part of the scene, Eddy begins to feel exhausted. He is unable to help them with the same rigour as he did at the start of the scene; he becomes hot and sweaty and eventually he has to stop because he cannot continue to go on. His friends, too, no longer having that continuous help, eventually have to stop performing their respective tasks. They too become tired, hot, and sweaty.

Each scene of the role-play was complemented with large charts on the wall, on which the teacher wrote the responses of students to structured questions she posed to the students at selected points during and at the end of each scene enacted. The questions posed to the students for each scene revolved around the main idea in the scene, which emerged directly from the questions in Segment 1. For example, throughout Scene 3, the teacher asked several questions encouraging students to think about how is it possible for Eddy to help each of them even though they were each doing different things. The idea was to get students to realize that while Eddy is one individual [one thing], he can function in different roles or forms, so that the idea of energy being transformed from one form to another, depending on the demands of the situation, was brought into focus in this scene. At the end of Scene 4, the teacher asked the students to think about why Eddy becomes hot. She used their responses to lead into a discussion about energy being transformed into heat instead of being “lost.”

In the last scene, Eddy is faced with overuse and abuse arising from his interaction with his friends. ”An expert” [*The teacher acted as the expert in this scene. She assesses the situation and then explains to the students that in order for each of them to achieve their respective task, energy was transformed from one form to another and that some of the energy is converted into heat, which explains the “hot” they each felt*] stepped in

to suggest what Eddy and his friends should do to carefully control the use and to stop the abuse. Once again, the teacher asked prompting questions and the children's responses were written on the charts. *[Though not an intended outcome of this project, eight students indicated that Eddy and his friends would need "food" (or a "drink") to "get back energy," suggesting that they understood the concept of energy sources.]* A total of five scenes were enacted.

Segment 4: What Ideas Did Children Develop?

In this segment, the children revisited the responses they had given in Segment 1, and compared their original ideas with those presented on the charts and those gained from the enactment experience. They discussed these, noting whether their ideas had changed, broadened, or remained the same. After the discussions, the children were asked to respond to the questions again. This was useful as both teacher and children were able to see the extent to which children's understanding of the concepts had developed or changed. It also allowed the teacher to detect any misconceptions that children might still have had and if there was any child/children whose ideas had remained unchanged after the intervention.

Segment 5: Gauging Children's Views and Feelings With Interviews

This segment consisted of short semi-structured interviews with the students. These were conducted one month after the role-play session, and the aim was to discover what knowledge and understandings, as well as the levels of awareness about the topic, had been retained by the students. The interviews were also used to probe the impact that the role-play strategy had on students. Each child was questioned using the following list of interview questions;

1. What science topic did we do using role-play?
2. What were three things you enjoyed in that science topic and why?
3. What is energy?
4. Where does energy come from?
5. What is energy used for?

6. What happens to energy when you use it?
7. Why do you think it is important to use energy wisely [the term *wisely* was explained to the children to mean conservatively]?
8. Which character in the role-play impressed you the most and why?
9. If you had to choose a character to be, which one would you choose and why?

The interviews lasted about 15 minutes, as sufficient time was given for all children to collect their thoughts and to respond completely. Some children had to be prompted to get the responses focused on what the question was asking. The anticipated outcome in this segment was two-fold; firstly, to gauge students' conceptual understandings about energy and its usage one month after the role-play strategy, essentially through the first six questions asked in this segment; and secondly, by way of the final three questions, to gauge students' feelings/views about energy conservation and energy conservation practices. Conceptual understandings revealed orally during the interviews were compared to those obtained from students' written responses in Segment 1, in order to determine whether students' original ideas/understandings had broadened, changed, or remained the same.

The interviews were all tape-recorded and subsequently transcribed and coded for analysis.

Outcomes — The Findings

Data were obtained from all 29 children, as there were no absences during the course of the project. Children had a variety of ideas before the role-play but, generally, did not display great concern for the controlled consumption of energy or any great awareness of energy conservation practices. When questioned after the role-play, using the questions asked in Segment 5 [which emerged from Segment 1], it was found that many children had either gained new ideas/understandings leading to a change in their original ideas, or that their original idea/s had broadened to reflect "more scientifically correct" understandings and explanations. For four of the children, there were no changes in their original ideas after the role-play. Of these four children, two of them knew quite a lot about forms and uses of energy and were aware of the

need for conservation, so that their ideas at the end of the project were very much what they came into the project with. These two students, however, indicated how enjoyable and exciting the role-play experience was for them [when interviewed in Segment 5].

The two other students knew very little about energy in the scientific context, but they had what might be described as a religious/cultural perception that energy was used to “do things” and “to live,” and that energy was “given to us by God.” During the course of the project, they participated in the role-play strategy and they volunteered responses on occasion, thus giving the impression that they were developing “new scientific understandings” of energy and its related concepts. Despite this perceived new learning, which seemed to have run parallel with their “cultural views,” their responses to the questions at the end of Segment 5 of the project indicated that it was their cultural understandings/beliefs that were the preferred responses [e.g., “energy comes from God”] given to the questions.

The following acronyms were used to label the impact [if any] that role-play had on students’ ideas:

GNIAP — Gained New Idea After Play

IBAP — Idea Broadened After Play

IRSAP — Idea Remained the Same After Play

Table 2 is a qualitative table that shows some of the students’ responses before (Segment 1) and after (Segment 5) the role-play intervention, in support of the impacts—GNIAP, IBAP, and IRSAP.

Table 3 further summarizes the outcomes obtained in respect of the change in ideas expressed by each student for questions 3, 4, 5, and 6 (Segment 1 compared with Segment 5) after the role-play intervention.

Table 2. Some Students' Responses in Support of Each Impact

Impact	Segment 1	Segment 5
GNIAP	"energy disappears when used"	"energy is converted to different forms to do work ... in the end it changes to heat energy"
IBAP	"energy is fuel"	"energy is the ability to do work"
IRSAP	"energy comes from God" "energy is needed to do all work"	"God provides energy for us to live" "energy is needed to do work"

Table 3. Students' Understandings After the Role-Play

Student	Q 3	Q 4	Q 5	Q 6
S1	GNIAP	IBAP	IBAP	IBAP
S2	GNIAP	IBAP	GNIAP	GNIAP
S3	GNIAP	IBAP	IBAP	IBAP
S4	IBAP	IBAP	IBAP	IBAP
S5	IBAP	IBAP	IBAP	IBAP
S6	IRSAP	IRSAP	IRSAP	IRSAP
S7	GNIAP	IBAP	IBAP	IBAP
S8	IBAP	IBAP	IBAP	IBAP
S9	GNIAP	GNIAP	GNIAP	GNIAP
S10	GNIAP	IBAP	IBAP	GNIAP
S11	IBAP	IRSAP	IRSAP	IRSAP
S12	IBAP	IBAP	IBAP	IBAP
S13	IBAP	IBAP	IBAP	IBAP
S14	GNIAP	IBAP	GNIAP	GNIAP
S15	GNIAP	IBAP	GNIAP	GNIAP
S16	IBAP	GNIAP	IRSAP	IRSAP
S17	IBAP	IBAP	IBAP	IBAP
S18	GNIAP	IBAP	GNIAP	GNIAP
S19	IRSAP	IBAP	IBAP	IBAP
S20	IBAP	GNIAP	IBAP	GNIAP
S21	GNIAP	GNIAP	GNIAP	GNIAP
S22	GNIAP	IBAP	IBAP	GNIAP
S23	GNIAP	GNIAP	GMIAP	GNIAP
S24	GNIAP	GNIAP	GNIAP	GNIAP
S25	GNIAP	IBAP	IBAP	IBAP
S26	IBAP	GNIAP	GNIAP	GNIAP
S27	GNIAP	GNIAP	GNIAP	GNIAP
S28	IRSAP	IBAP	IBAP	IBAP
S29	GNIAP	IRSAP	IRSAP	IRSAP

Table 4 is a translation of Table 3, which depicts the respective percentages of the class in response to each of the four questions asked.

Table 4. Students' Understandings After the Role-Play (in percentages)

Ideas	% GNIAP	% IBAP	% IRSAP
What is energy?	55%	35%	10%
Where does energy come from?	28%	62%	10%
What is energy used for?	34%	52%	14%
What happens to energy when it is used?	45%	41%	14%

What is Energy?

Responses like “it make[s] me run” and “it make[s] things move,” given before the role-play, suggested that many students believed that energy was a concept associated with motion.

In looking at the responses given by the students after the role-play, it was clear that the role-play exercise increased the understandings of 90% of the students, either by giving them new insight (55%) or by broadening the understanding they already had (35%). At the end of Segment 5, many responded to this question by saying things like “you must have energy to do work,” “energy is stored in high things,” and “electricity is a kind of energy.”

Where Does Energy Come From?

Twenty-eight percent of the students were uncertain as to the various sources of energy and their understandings developed only after the role-play exercise, while 62% of them had some ideas such as energy coming from “food” and from “the sun,” which were broadened, so that the

responses given in Segment 5 of the project included many of the traditional fossil fuel sources like coal, gasoline, and oil, as well as “food” and “the sun,” as the following interview segment illustrates:

Interviewer: “Can you tell me where energy comes from?”

Student: “Energy comes from the sun, the food we eat and the gas in the car...”

Interviewer: “Can you think of or remember any other sources of energy?”

Student: “In the play we learn that energy come from fossil fuels which is like crude oil...”

What is Energy Used For?

Before the role-play exercise, just over half of the students surveyed (52%) had some ideas [as gleaned from their actual responses] about what energy was used for. However, their ideas were broadened after the exercise. On the other hand, 34% of the students had very little idea [no responses given before the role-play] about specific tasks or activities that energy was needed for. However, they evidently gained new knowledge/understandings in this regard after experiencing learning through role-play, as they were able to identify actual situations in their everyday life where energy was used to get a task or activity completed. The following interview segment illustrates this point:

Interviewer: “Can you give me some examples of what energy is used for?”

Students: “Energy is used to do work around the house ... and to move things...”

Interviewer: “What do you use energy for?”

Students: “...to ride my bike ...to play cricket ... and to dance...”

What Happens to Energy When It Is Used?

Responses to this question given in Segment 1 of the project indicated that 45% of the students surveyed had misconceptions about what happens to energy when it is used. Some of these misconceptions were conveyed in responses such as “energy disappears” and that when energy

is used it is “lost.” An almost equal number of students (41%) had some idea—which was captured in responses such as “it goes into the thing”—about what happens to energy when it is used. At the end of Segment 5, those with misconceptions had most of these “corrected” through the acquisition of new ideas, and those with some initial ideas had these broadened, as could be gleaned from responses such as “...energy is change to other forms...” and “...most times it turns to heat in the end....”

How Do You Feel About Energy That is Wasted?

With respect to their views about energy conservation practices, which was the focus of the last three questions in Segment 5 of the project, the data obtained from the taped interviews conducted one month after the project indicated that most students realized that energy conservation practices were important considerations when using energy. Students’ responses were transcribed and grouped into several categories based on the messages communicated by the individual responses. The categories were reviewed several times to ensure that the category label reflected the ideas being conveyed in the truest sense. Similar categories were then grouped together and an overarching theme assigned to each group.

Three prominent themes emerged from the data collected in response to the question “Why do you think it is important use energy wisely?” They were: “securing energy for future generations,” “it is not good to waste,” and “energy will be depleted very soon if wasted.” These themes suggested that students understood the importance of using energy conservatively. Responses obtained from two students could not be coded into any of these categories and the intention was to analyse these on their own merit. Just fewer than half of the class had the view that energy wastage is not a good thing and, therefore, that it was important to use energy conservatively. The numbers of students responding in each theme as well as actual verbatim responses given by some of the students are shown in Table 5.

Table 5. Emerging Themes of Students' Views on the Need to Conserve Energy

Emerging Theme	No. Students Responding	Verbatim Responses
Securing energy for future generations	8	"...[we] need to save for the future..." "...save for later on..." "...keep some for later..." "...other people ...later...will need energy..."
It is not good to waste	13	"...do not waste it..." "Wasting energyis not good..." "... do not waste ...use energy careful[ly]"
Energy will be depleted very soon if wasted	6	"...if we waste it will done..." "...it will finish ...one day..." "...what we go do when it gone through..." "One day it will end" "...it will end soon..."
None of the above	2	

As suggested from the verbatim responses, it was clear that most of the students emerged from the project with the understanding that energy was a valuable commodity and that it was worth conserving energy for future generations. In their explanations, students indicated that "careful use" meant things like turning off the fan (or AC unit) and opening the windows on a hot day, or hanging clothes outside to dry instead of using an electric clothes dryer on hot days. Their ideas of not wasting were indicated in responses such as not wasting food, not leaving the lights on if you are not in a room, and not leaving the refrigerator door open for long periods. Some students indicated that turning the television [and radio] off were ways of "saving" energy. Generally, their responses suggested that conserving energy meant using less in all instances. The

idea of alternate sources of energy, for example, solar or wind instead of electrical, was not explicitly expressed by any of the students, but some seemed to have an implicit understanding that was captured in responses such as hanging clothes outside to dry instead of using an electric dryer.

Benefits

As reported by Aubusson et. al. (1997), perhaps the most notable benefit of using role-play to teach energy-related concepts in this project was that students' alternative ideas were clarified, either through a broadening of prior understandings or by the development of new, "scientifically correct" understandings. The interaction during the role-play also allowed for students to make meaningful linkages among several related concepts such as sources of energy, forms of energy, uses of energy, and transformations of energy. With respect to transformations of energy, in particular, students developed clearer understandings about the transformation process in the first instance, but more importantly about the nature of the different forms of energy at each step of the transformation process. The use of common everyday situations in the scenes allowed students to see the relevance of the learning and the ways in which their understandings could be applied to familiar situations in their life. This is similar to the argument put forward by Kofoed (2006) that role-play not only develops students' subject competencies, but that it helps to make them more "competent consumers" of life.

Though it was not intentional in the conceptualization of the project, the interactions and the questions led students to an appreciation of the law of conservation of energy, so many of them came to the realization that energy could not be created or destroyed, but that it is simply converted from one form to another. Linked to this learning, the project revealed that students became aware of the need to conserve energy, either simply to prevent wastage or to preserve supplies for future generations. Students' sense of personal responsibility, in terms of their indication to adopt various energy conservation practices, seems to be a significant benefit emerging from this project.

Challenges

Among the challenges faced during the course of this project was gauging students' prior knowledge. The responses obtained from Segment 1 were helpful to the extent that the specific understanding/s conveyed by them were, in fact, a true reflection of the students' actual understanding at the time. Getting students to physically participate in the "acting" was a challenge, but only for the first two scenes. By the third scene students had become more enthused and much less shy. There were a few instances in Scenes 3 and 4 when the acting was not in exact synchronization [in terms of the scientific concepts and understanding conveyed] with the intended roles of the actors, and it became necessary for the teacher to intervene on a few occasions to get students to focus on their roles. Furthermore, at times students became so involved in the drama that it was necessary for the teacher to explicitly make the link between the drama and the learning. The charts on the wall for each scene were very helpful in this regard. One of the greatest challenges encountered in this project was getting students to articulate their feelings [about energy conservation] in Segment 5. It was clear that this was not a kind of thinking that they were familiar with—it seemed to be an issue to which not much consideration was given. Most of the students thought at length before answering, and the tone of their voices as well as their facial expressions suggested that they were responding with some degree of uncertainty. Of course, the diverse levels of literacy among the students also presented a challenge, but this was minor in comparison to some of the others and was well aided by tape-recording the interviews. Inclusive of pre-planning and debriefing, the project ran throughout the term and so there was no additional time to probe into ambiguous aspects of the findings; for example, to delve deeper into the understandings of the two students who seemed to have a cultural/religious perception or understanding/view about energy.

Conclusion

This project revealed that role-play is an effective way to engage students in science learning at the primary school level, but it must be carefully planned and diligently implemented to achieve meaningful learning outcomes. It is truly a strategy that puts the students at the centre of the learning process and transforms the teacher into a facilitator or

guide. In the present context, teachers do not naturally see themselves in these “new” roles, and so any shift towards this method of teaching will prove to be challenging for teachers in the first instance. The implication, therefore, is that teachers will have to be trained and/or retrained in ways of implementing strategies of this type in their classrooms. At the personal level, there is the further suggestion that teachers would need to be creative and innovative when choosing this teaching/learning strategy to effect learning in their classrooms.

This particular project revealed that the group of students surveyed had a range of ideas coming into the project [before the role-play intervention], and that it was important for the researcher to note these and to use some of these in preparing the scenes to take students from the known to the unknown. The strategy allowed many students to either broaden their old ideas or to develop new ideas. Many of the students shared their new understandings in a variety of ways [as gauged by the interviews conducted one month later], most of which were in line with conventional science ideas.

The major findings at the end of this project may be summarized as follows:

1. 90% of the students had a scientifically sound understanding of what is energy.
2. 90% of the students could identify at least three different sources of energy.
3. 86% of the students could suggest at least five different situations of energy being used.
4. 86% of the students were able to satisfactorily explain what happens to energy when it is used, suggesting in all cases that it is transformed from one form to another.
5. 93% of the students expressed an appreciation for the need to conserve energy, most of them suggesting that wastefulness is not a good habit and others suggesting that it is important to conserve energy for future generations.

Discussion

It is important to remember that this project was conducted in one school with one group of students and, therefore, aspects of the findings cannot

be generalized across all primary schools or even at different levels in the same school. It is unfortunate, however, that the role-play sessions were not video-taped or audio-taped. This was an oversight on the part of the researcher, the shortcomings of which were only recognized late in the data analysis phase of the work. The overarching aim was to gauge the impact of role-play on science learning for a chosen topic, and towards this end the project yielded a positive outcome. In respect of the affective benefits, the findings here are similar to those reported by Bonnet (2000), in that the strategy facilitated a considerable degree of social interaction and bonding among the students of the class. Furthermore, it effectively captured students' attention and stimulated their imagination, as was reflected in the intense levels of dramatization observed in each of the scenes enacted, similar to the experiences reported by Dallman-Jones (1994).

This project offers possibilities for further work, for example, the impact of the strategy on students' learning at other levels of the education system such as infants, and possibly at secondary school as well. It will be interesting to see what the findings at various levels of the education system will reveal. Furthermore, the versatility of the strategy makes it adaptable to a range of science topics, and it will be useful to identify topics in which the strategy might be most effective as a teaching/learning tool.

An emerging question, however, is how do we encourage or motivate science teachers to use role-play strategies in their science teaching? While this project did not probe into the teachers' views and feelings about the strategy [apart from an articulation of the benefits observed and the challenges encountered], it might be useful to explore these in another aspect of this study, in order to understand the issues that interplay in the teacher's domain. This might be an excellent topic for further research.

References

- Aubusson, P., Fogwill, S., Barr, R., & Perkovic, L. (1997). What happens when students do simulation-role-play in science. *Research in Science Education*, 27(4), 565–579.
- Banister, F., & Ryan, C. (2001). Developing science concepts through storytelling. *School Science Review*, 83(302), 75–83.

- Bonnet, C. (2000). The relevance of role playing in environmental education. *Proceedings of the International Union of Biological Sciences Commission for Biological Education (IUBS-CBE): International Symposium on Biological Education, IUFM Versailles, Centre de Cergy, France, 15-18 May, 2000*. Available at: <http://intl.concord.org/cbe/pdf/bonnet.pdf>
- Craven, J. A., & Penick, J. E. (2001). Preparing new teachers to teach science: The role of the science teacher educator. *Electronic Journal of Science Education, 6*(1). Retrieved from <http://unr.edu/homepage/crowther/ejse/ejsev6n1.html>
- Dallman-Jones, A. S. (1994). *The expert educator*. Fond Du Lac, WI: Three Blue Herons Publishing.
- Francis, P. J., & Byrne, A. P. (1999). Use of role-play exercises in teaching undergraduate astronomy and physics. *Publications of the Astronomical Society of Australia, 206*–211.
- Goldring, H., & Osborne, J. (1994). Students' difficulties with energy and related concepts. *Physics Education, 29*(1), 26–32.
- Graber, W., Nentwig, P., Becker, H-J., Sumfleth, E., Pitton, A., Wollweber, K., & Jorde, D. (2001). Scientific literacy: From theory to practice. In H. Behrendt, H. Dahncke, R. Duit, W. Graber, M. Komorek, A. Kross, & P. Reiska (Eds.), *Research in science education – Past, present and future* (pp. 61–70). Dordrecht: Kluwer Academic.
- Haury, D. L., & Rillero, P. (1994). *Perspectives of hands-on science teaching*. Columbus, OH: ERIC Clearinghouse for Science, Mathematics and Environmental Education.
- Kofoed, M. H. (2006). The Hiroshima and Nagasaki bombs: Role-play and students' interest in physics. *Physics Education, 41*(6), 502–507.
- Ladrousse, G. P. (1989). *Role play*. Oxford: Oxford University Press.
- McSharry, G., & Jones, S. (2000). Role-play in science teaching and learning. *School Science Review, 82*(298), 73–82
- Millar, R. (2002). Thinking about practical work. In S. Amos & R. Boohan (Eds.), *Aspects of teaching secondary science: Perspectives on practice* (pp. 53–59). London: *RoutledgeFalmer*.

EXTRA-LESSONS: A COMPARISON BETWEEN “DIFFERENT SIDES OF THE TRACK” IN TRINIDAD AND TOBAGO

Samuel Lochan and Dorian Barrow¹

This paper attempts to compare the participation in extra-lessons by the students of a senior comprehensive school and a traditional seven-year girls' school in Trinidad and Tobago. Survey data were gathered from 25 students at each level from Forms 4, 5, and 6 in each school. Data are presented on the socio-economic status of different students, participation rates by students of the two schools, subjects chosen by students for extra-lessons, costs per subject, reasons for seeking extra-lessons, and the effects of extra-lessons. This study therefore presents some data on private lessons not previously available. It also draws some comparisons between the two schools that are critical for understanding school performance in Trinidad and Tobago.

Introduction

The extra-lessons industry in Trinidad and Tobago has a long heritage. Both at the primary and secondary level, there has always been the practice of students seeking the assistance of private tutors outside of the school's normal timetable. This practice has no doubt been driven on both the demand and supply side. Our colonial heritage of an examinations-driven school system, both at the primary and the secondary level, feeds the demand for extra-lessons. In the past, comparatively low salary levels of teachers fed the supply side.

Periodically, there are protests, in the form of letters to the Editors of the daily newspapers from irate parents who have problems with the practice of extra-lessons. Ministry of Education officials, from time to time, also make pronouncements indicating disapproval of extra-lessons. This

¹ The authors wish to acknowledge the help of Ms. Eathra Stephen in preparing this report.

practice, however, has become a permanent part of the informal schooling process in Trinidad and Tobago.

No doubt, this practice exists all over the world. Countries like the United Kingdom (UK), the United States (US), China, Japan, and Singapore, as well as Vietnam and Mauritius, all have this practice embedded in their education system (Bray 2006). However, the form it takes and the effect it has on the education system varies from country to country.

To date, only one researcher had conducted empirical research on the extra-lessons industry in Trinidad and Tobago (Brunton, 2000, 2002). Brunton's research notwithstanding, there has been no systematic study of this phenomenon by schools, levels of the school system, and levels within the school system. That breach is now being filled by a baseline study, which is being done in two phases. The first phase involves the secondary school level, while the second phase will involve the primary level. In the first phase, a survey of 30 secondary schools is being undertaken (from a total of 134 secondary schools in Trinidad and Tobago).

This paper is an attempt to capture some of the emerging patterns as this study progresses at the secondary level. It sets out to compare the extra-lessons phenomenon in two secondary schools in Trinidad and Tobago from which data have been collected. One school is representative of the old sector secondary schools that were built before Independence and which are run mainly by religious boards. This school is a 7-year, traditional, single sex, girls' school. The other is a new sector, comprehensive, 7-year, co-educational school built after Independence and run by the State.

Definition of Extra-Lessons

We use the term *extra-lessons* to refer to all teaching/learning activities outside of the normal school timetable that attempt to cover the formal school curriculum at a cost to the student or parent. Sometimes, it is also referred to as *private lessons*. In the literature, it is referred to as *private tutoring*. Free assistance given by the school to children is not considered extra-lessons, nor is payment for extra-curricular activities such as music

and dance when they are located outside of the school's timetable. This is the definition adhered to in the international study of private tutoring done by Bray (2006).

The Secondary School Context

The historical evolution of the secondary school system in Trinidad and Tobago has produced an unhealthy division into "prestige" and "non-prestige schools." The simplest basis for division seems to be historical time—those secondary schools built before 1970 are considered more prestigious than those built after 1970. Given the way selection and placement of students are done at the end of primary schooling, the end result is that the prestigious stream absorbs the children of the affluent in the society and the non-prestigious stream absorbs the children of the poor.

The Jules report (1994) found a clear, positive relationship between students' socio-economic background and distribution across school types. The study confirmed that students attending prestige schools were mainly from families where parents had overall higher occupational levels and greater educational qualifications. Students attending junior secondary and senior comprehensive schools, on the other hand, came from families where parents were employed in manual jobs and possessed lower levels of education.

The two schools selected, therefore, are representative of the two streams in a system stratified into a dominant sector and a weaker sector—the dominant sector producing the national scholarship winners and the future professionals of the society, and the weaker sector associated with generally low academic standards.

Background to the Study

Trinidad and Tobago has been implementing a new wave of education reform since the early 1990s. First, there was "Quality in Basic Education," which was based on improving education at the primary level. This involved school expansion as well as curriculum change. Then there is the still ongoing Secondary Education Modernization Project (SEMP), which involves curriculum reform, administrative

reform at all levels of the system, and the building of additional secondary schools.

In addition to these measures aimed at making universal primary and secondary education a reality, the Government of Trinidad and Tobago has also instituted a range of supporting measures aimed at making it possible for the poor in the society to access these educational opportunities. This is all part of the Government's commitment to the universal goal of "education for all." These measures include a programme of school feeding, free school transport for the needy, and a textbook rental programme. In other words, the Government of Trinidad and Tobago has expended significant resources to expand educational opportunity and to make such opportunities more accessible to the poor.

The existence of a private lessons industry raises doubts about the efficiency and equity goals of this reform agenda. To the extent that a large extra-lessons industry still exists despite the reform initiatives, it may mean that the formal school system does not adequately prepare children to attain the goals of schooling, or that the poor are still at a disadvantage in the system since only those who can afford can participate. Additionally, it raises questions about the basis for the well-entrenched and accepted division of schools into prestige and non-prestige, which is part of the educational landscape of Trinidad and Tobago.

Efficiency Issues

To the extent that there is a significant market for extra-lessons, it raises questions about the adequacy of the delivery system for preparing students in the schools. If there is insufficient time, resources, or competent teaching to complete the programme of work set for schools and the demand for extra-lessons is an attempt to fill this gap, then this could be viewed as an indictment on the ability of the State to provide schooling. Therefore, the motives for accessing extra-lessons are important.

Equity Issues

Since the access to extra-lessons involves a significant expense, it means that despite the interventions of the State, education is still not as free as is generally assumed. The provision of free meals, books, or transport, therefore, does not equalize educational opportunity since only those who can afford to pay can access extra-lessons. Knowing the extent and costs of extra-lessons, therefore, will help us to better understand the true private costs of schooling in Trinidad and Tobago.

Accountability Issues

In addition, the existence of an extra-lessons industry makes it difficult to evaluate and compare school performance. Given ready access and commitment to extra-lessons on the part of students and parents, the achievements of a school, if measured by success in examinations, is not a reliable indicator of either the instructional capability of the school or the intellectual ability of the students. In schools where parents seek out extra-lessons for their children, the schools' efforts are complemented by efforts external to the school. High pass rates in examinations, therefore, may not be reliable indicators of sound instructional practices.

Purpose of the Study

This paper seeks to explore the extra-lessons phenomenon in two schools in Trinidad and Tobago. These two schools are significant in terms of their standing within the educational system of Trinidad and Tobago—one representative of the older, traditional, more successful schools and the other representative of the new sector schools, built more recently and perceived as less successful in achieving high levels of success in final examinations.

Research Questions

1. When the participation in extra-lessons by students of both schools is compared, can it be said that education is efficient and equitable?
2. Are measures of comparison based on performance in examinations reliable for comparing the performance of both schools?

Specific sub-questions include the following:

1. What is the extent of participation in extra-lessons by secondary school students?
2. What is the cost of extra-lessons?
3. What are the forms that extra-lessons may take?
4. What are the motivations for seeking extra-lessons by students?
5. Who are the decision makers in the process?
6. What are the proposed benefits of extra-lessons?

Literature Review

Some significant pieces of literature have been encountered on the issue of extra-lessons. These include the very recent publication in February 2008 by Dang and Rogers; Hallak and Poisson (2007); Mark Bray (2006); Silova, Budiene, and Bray (2006); Heyneman (2004); Brunton (2000, 2002); Buchmann (1999); and Stevenson and Baker (1992). The publication by Hallak and Poisson is an international comparative study with a series of articles on all aspects of corruption in education. The article by Heyneman is a survey of the same sets of issues covered by Hallak and Poisson. Stevenson and Baker concentrate on the private lessons experience in Japan, while Mark Bray does a world survey of “supplemental private tutoring.” The articles by Brunton on the extra-lessons phenomena in Trinidad and Tobago are the only local articles on the extra-lessons industry in Trinidad and Tobago. Buchmann describes the situation in Kenya, a post-colonial society, which closely resembles the situation in Trinidad and Tobago.

The overarching conceptual framework for the entire publication by Hallak and Poisson is the concern with corruption in educational provision in its widest sense, and its effect on the efficiency of resources in educational planning and provisioning at all levels of the education system. The authors see the elimination of corruption at all levels as a necessary precondition for the achievement of the goals of “education for all” set in Jomtien, Thailand in 1990 and re-affirmed in the Dakar Framework for Action in 2000.

Hallak and Poisson (2007) summarize the findings from an international project that was carried out with the participation and support of many

institutions and individuals from about 25 countries worldwide. The partners included ministries of education, universities, researchers, and members of civil society organizations. The guiding idea behind this research effort is that there exists corruption in education systems across the world and such corruption affects all levels of operation of education systems: teacher management and behaviour, school construction improvement and repairs, textbooks and teaching materials, school feeding and nutrition programmes, exams accreditation and credentials, and private tutoring.

In the examination of the private tutoring phenomenon, data are provided from countries all over the world, giving evidence of the universality and frequency of this practice. The main conclusion of the writers is that private tutoring is not bad in itself; it can be a healthy practice but this depends on the conditions and circumstances under which it takes place. It can complement the mainstream system but it can also corrupt it.

The impact of private tutoring could mean that teachers do not cover all topics during normal class time. This would result in children seeing the classroom work as unnecessary, and could possibly lead to a narrowing of the goals of teaching and learning. In such a case, the effects on the mainstream is bad. However, to the extent that it assists students with special needs or complements the work of the schools, then it is positive.

Heyneman (2004) is but a precursor of the study by Hallak and Poisson, and does no detailed investigation of the extra-lessons phenomenon except to classify it as a form of corruption under the rubric of professional misconduct.

The study by Stevenson and Baker (1992) is relevant to the Trinidad and Tobago situation. The researchers establish three preconditions for the existence of the shadow education system that exists in Japan: 1) the allocation of school places by formal examinations, 2) the use of centralized systems for doing so, and 3) the clear connection between future occupational and social status and access to certain schools. In Japan, the competitive pressure exists most strongly in the transition from high school to university. Selection to a prestigious university almost guarantees students access to the good life. The extra-lessons industry is an institutional response to this social arrangement. This is a

vast industry and the participation rate by students is extremely high at the secondary level and immediately after secondary school when students may enrol in special schools for preparation to sit university entrance examinations.

The cultural/institutional context of Japan very much resembles the situation of Trinidad and Tobago, except that in our case the quest for extra-lessons begins at the primary level since the connection to occupational and social status begins with the secondary school that students can access through competitive examinations at the end of primary schooling.

Bray (2006) provides data to show the universality of the extra-lessons phenomenon throughout both the developed and developing countries of the world. Bray's position is that it is a shadow system which requires understanding by policy makers. It is referred to as a shadow since it owes its existence to the mainstream and responds to needs of the mainstream. Bray explores the levels of participation, forms of supply, motives for participation, and effects of private tutoring. Bray's findings are inconclusive on the instructional benefits of private tutoring, but he concludes that this medium is a form of social reproduction as wealthier families are able to afford the opportunities that this facility makes available.

The report by Silova and Bray (2006) was the result of a study of the nature, causes, and consequences of the private tutoring market in some former socialist countries of Eastern Europe, including, Azerbaijan, Bosnia, Herzegovina, Croatia, Georgia, Lithuania, Mongolia, Poland, Slovakia, and Ukraine. As these former socialist countries attempt to transform themselves into market economies, the growth of the private tutoring market has been significant. This is a likely challenge to the development of free and open societies and, therefore, it is a phenomenon requiring in-depth study. This study was a first attempt to document systematically the general characteristics of private tutoring: its scale, cost, geographic spread, and subject matter, as well as the factors underlying the demand for private tutoring and its educational and social impact.

In the countries surveyed, it was found that private tutoring had a number of negative consequences. It exacerbated social inequities, distorted curricula and teacher performance, fostered corruption, skewed the university admissions process, and deprived the state of revenues. On the positive side, it helped students to compete in the education market place, facilitated private investment in human capital, and provided additional options for out-of-school youth.

A survey of first-year university students done in all these countries revealed a high participation rate in private tutoring in late secondary school. The participation rates ranged from a low of 56% for Croatia to a high of 93% for Azerbaijan. Students complained that they were sometimes pressured to take private tutoring. University professors in charge of entrance examinations abused their positions by setting up private tutoring services. By far the strongest driver of the private tutoring business was the need to perform well in qualifying examinations for university. This report recommended actions to policy makers for the control of private tutoring, which included: raising public awareness, regulation of private tutoring, development of a code of ethics, and the licensing of private tutors.

The recent study by Dang and Rogers (2008) confirms the universal and growing nature of the phenomenon of private tutoring. Firstly, while the greatest incidence of private tutoring exists in the East Asian countries, it is now an important phenomenon in many countries of different sizes, levels of development, political institutions, and geographical locations. Secondly, it exists at all levels of the education system, from primary to upper secondary. Thirdly, both in absolute terms and relative to the formal education system the private tutoring industry seems to be growing rapidly.

From a survey of 23 countries, which included Bangladesh, Cambodia, Cyprus, Kenya, Egypt, Japan, the US and UK, it was determined that there was a common core of factors which were the drivers of the private tutoring industry at the level of households. These were: household income, level of education of parents, and urban location. Other factors may affect the demand by households as a result of cultural differences between countries, but the common core holds across all countries.

At a macro level, the key drivers of the private tutoring industry were found to be: firstly, the transition to a market economy in countries where it did not exist earlier; secondly, the existence of tight linkages between schooling and work; thirdly, the existence of a deficient public education system; and fourthly, cultural values, such as in the case of East Asia.

Reviewing several econometric studies that use grades to ascertain whether students benefit from private tutoring, the authors found the results inconclusive. But they did find remedial-type support programmes that yielded positive returns. At the end of the study, the authors made some very qualified conclusions. They felt that under certain conditions, private tutoring can complement the formal school system. In addition, private tutoring increases the welfare of households and society overall, but only if a free market prevails. The private tutoring industry may, however, put poor households at some disadvantage.

In terms of state policy, the authors advocate a policy of market cleansing rather than attempts to impose a complete ban. This would mean making the market competitive and free by ensuring that there is freedom of information, freedom to choose, and a public awareness of standards.

In 1997, using a quantitative approach, Brunton sampled, by questionnaire, 500 secondary school students throughout Trinidad. The sample was derived from eight secondary schools representing urban/rural, semi-urban, north/south/central, and traditional and non-traditional type schools. The findings proved his two hypotheses that: (a) participation in extra-lessons in Trinidad increases as secondary school students face greater examination pressure, and (b) participation in extra-lessons reflects patterns of educational and social stratification in Trinidad and Tobago. The findings also showed that students in the 15–18 age group were three times more likely to be enrolled in private lessons than students below 15. So the search for private lessons was strongest at Forms 5 and 6.

Brunton's study confirmed his social reproduction hypothesis. From his sample, 56% of students in prestige schools sought extra-lessons while

the figure for comprehensive schools was 44%. There was a high correlation as well between father's occupation and participation in extra-lessons. The higher the status of the father's occupation the greater was participation in extra-lessons. The connection between school type, income levels, and participation in extra-lessons was clearly established by the Brunton study. From his national sample, 65% of those with fathers in the professional category of work participated in extra-lessons while 29% of those with fathers in the unskilled area participated in extra-lessons.

While Brunton's study used a national sample, this paper seeks to examine the phenomenon of extra-lessons more comprehensively at the level of two specific schools. While the social reproduction issue is of great significance for this research paper, more details about the time spent, the subjects chosen, and the motivations for and the effects of extra-lessons are also needed to explore the phenomenon from an educational perspective.

Buchmann (1999) attributes the existence of a large extra-lessons industry in Kenya to three factors. Firstly, after Independence, schooling became a cultural criterion for being modern for most Kenyans and this led to an insatiable demand for schooling overall. Secondly, the existence of a two-tier secondary system—one perceived as better than the other—meant that there was competition for the few schools perceived as good. Competitive examinations became the device for selection for school placement. Buchmann describes an extra-lessons culture that is costly to parents, time consuming for children, and incapable of control by the State.

Methodology

The Two Schools

The two schools chosen for this comparison are significant because they typify two separate categories of secondary schools in Trinidad and Tobago. One school, a traditional, single-sex, girls', 7-year, denominational secondary school, in existence for a relatively long period of time—established 12th January, 1951; the other, a co-educational, new sector, government secondary school, a senior

comprehensive school, was one of the secondary schools built during the school expansion programme undertaken during the 1968–1983 plan period.

In the educational landscape of Trinidad and Tobago, the former would be referred to as a prestige school while the latter would not be so designated. Senior comprehensive schools, unlike prestige schools, are usually portrayed in the media as schools with greater student indiscipline and poor student performance in examinations. Generally, the Form 1 intake into a traditional prestige school, as the one in this sample, would have to attain marks in the range above the 90th percentile in the selection examinations at the end of primary schooling. The Form 1 intake into senior comprehensive schools, however, falls within a much lower range of examination scores in the selection examinations at the end of primary school, and may even include students who scored below the 30th percentile in the selections examinations.

Given a system of secondary school placement based on parental choice and examination scores, the historical fact has been that parents prefer the placement of their children at the older traditional schools. It would not be untrue to say that in the public mind there exists a hierarchy of schools and the new sector schools are at the bottom of the heap.

Target Group

In each of these two schools, students from Forms 4, 5, and 6 were surveyed. These levels were chosen because these are the upper levels of the secondary school system when the concern for success in examinations is great and the participation in extra-lessons is highest. Twenty five students were surveyed from each level. In each school, therefore, 75 students were targeted. Figure 1 shows the actual response rate of the survey in both schools by Form level.

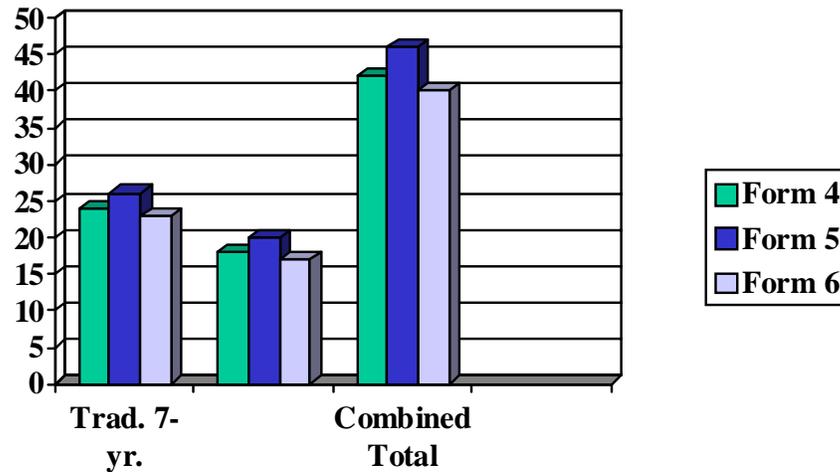


Figure 1. Sample size 1: Number of responses by school/form level.

Instrument for Data Collection

A questionnaire was designed with 21 questions covering the different research objectives of the study. In addition, socio-biographical data were collected through the survey instrument, which contained questions on the subjects that students sought extra-lessons in, the costs of extra-lessons, the motivations for taking extra-lessons, and the benefits of extra-lessons. The rate of return of questionnaire by school was 100%, whereas the rate of response to questionnaire by student was 85.3%.

Administration of the Instrument

The approach to data collection was based on winning the support of the principals. The two researchers visited the schools, and after explaining the purpose of the study to the principals left the questionnaires with them to be administered at the school's convenience. The researchers were overwhelmed by the positive response and level of interest by the administrators in this study. In these two schools, the support of the administrators was very positive.

Findings

1. Socio-Economic Comparisons

Figure 2 shows a summary of the distribution of parent occupations by categories: professionals (i.e., parents who are doctors, lawyers, accountants, upper level managers, etc.); skilled professionals (programmers, technicians, supervisors, middle-level managers, etc.); manual workers (labourers, farm workers, store-clerks, domestic workers, etc.); and housewives. Figure 2 shows the summary of the occupations of the parents of students attending both schools. Note the skewing to the left towards the professional occupations end for the parents of students attending the traditional all girls' school, suggesting that these students largely come from more materially affluent households than the students attending the senior comprehensive school.

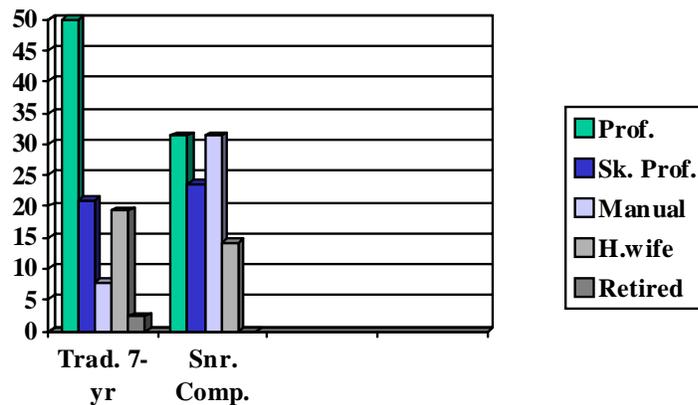


Figure 2. Parents' occupation.

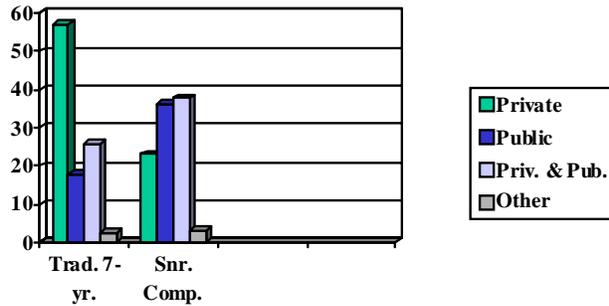


Figure 3. Students' travel habits.

Figure 3 depicts the travel-to-school habits of the student population of the two schools. Note that the majority of students attending the traditional school (56.8%) are driven to school via a privately owned vehicle, whereas of the students attending the senior comprehensive school, only 23% travel to school in this way. This suggests that the home resources available to these two sets of students are different, with the students attending the traditional girls' school having access to greater material resources.

2. Comparison in Participation Rates in Extra-Lessons

Overall, 73.4% of the students sampled from these two schools take extra-lessons in one or more subjects. When the level of participation in extra-lessons is compared by school, the data suggest that, overall, 89% of the students in the traditional 7-year school take extra-lessons in one or more subjects, whereas only 52.7% of students attending the senior comprehensive school attend extra-lessons, a mean difference (ΔM^+) of 36.3% of the traditional 7-year school students over the students attending the senior comprehensive school.

Figure 4 illustrates the difference in participation rates by both schools according to form level. It shows that 100% of the students of the traditional 7-year school in Forms 4 and 5 attended extra-lessons classes in one or more subjects, with this level of participation decreasing to 65.2% among the students in Form 6. This pattern of participation is

reversed among students attending the senior comprehensive school, where in Form 4, the level of participation in extra-lessons is 33.3%; in Form 5, it increases to 45%, and at Form 6, the level increases to 82.4%.

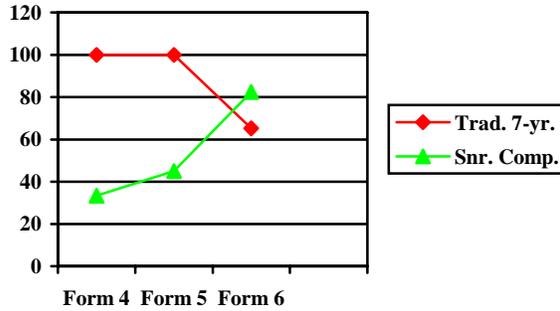


Figure 4. Students taking one or more subjects in extra-lessons by school/form level.

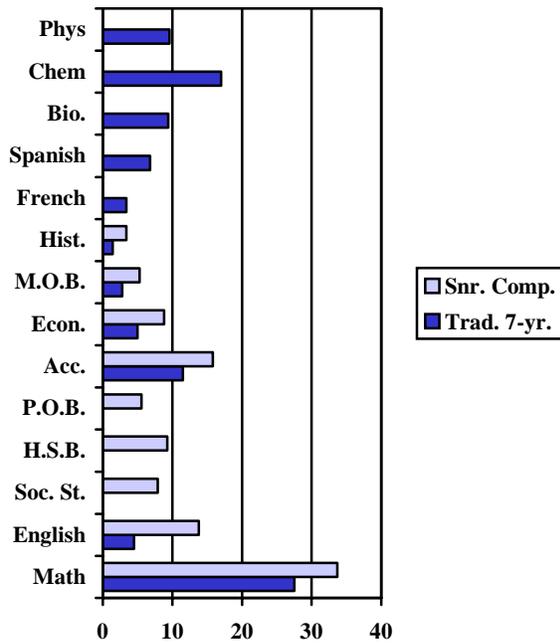


Figure 5. Participation rates by schools/by subjects.

The students in this sample attended extra-lessons in 14 separate subjects, which can be grouped into four broad categories, namely, mathematics, science, business, and “others,” which includes English, social studies, French, Spanish, and history. As Figure 5 shows, mathematics is the most popular option (30.6%), followed by business subjects (27.5%), and science (22.7%). The other subjects make up the remaining 19.2%, of which English is taken up by 9.2% of the students.

The students attending this traditional 7-year secondary school not only participated in extra-lessons at a greater rate than those students attending the 7-year senior comprehensive school, but they also enrolled in a much wider range of subjects. Students from the traditional 7-year school took lessons in 1–7 subjects, with the majority doing lessons in three 3 subjects. Whereas, students attending the senior comprehensive school took lessons in 1–3 subjects, with the majority (55.6%) doing lessons in three subjects.

3. Comparative Expenditures on Extra-Lessons

Table 1. Cost Per Subject Per Month by School

Subjects	Senior Comprehensive	Traditional 7-Year
Accounts	\$155.00	\$159.00
Economics	\$160.00	\$136.67
H.S. B. /Biology	\$204.00	\$130.11
Mathematics	\$172.00	\$125.00
History	\$275.00	\$125.00
M.O.B.	\$200.00	\$125.00
English	\$196.67	\$121.67
Actual mean = \$194.67 vs. \$131.79; Overall mean = \$163.23		
Overall actual range: \$275 (history); \$121.67 (English)/month/subject		

The price paid for extra-lessons by students reveals a range from a high of \$275 for history to a low of \$121.67 for English per month. The average overall cost per subject per month is \$163.23. The students attending the senior comprehensive school in the sample took lessons in a narrower range of subjects than students from the traditional school. Students from the senior comprehensive school participated in a range of

9 different subjects, whereas the students from the traditional 7-year school in the sample took extra-lessons from a range of 11 subjects.

There were seven subject fields—accounts, economics, biology (human and social biology), mathematics, history, management of business, and English—common to both sets of students. The cost to access extra-lessons in these common subject fields was different for the two sets of students, with the students from the senior comprehensive school paying a higher average cost (\$194.67 per subject per month) than the students attending the traditional 7-year school, whose average cost was \$131.79 per subject per month.

In summary, the overall unit cost per subject that these students pay for extra-lessons ranges from \$275 per subject per month to \$122 per subject per month. The average cost is about \$163 per subject per month. Since students take, on average, lessons in three subjects, they spend, on average, about \$500 per month on extra-lessons services.

4. Purpose for Extra-lessons: Enrichment and Remediation

The following are the five reasons given by students from both schools for taking extra-lessons, ranked in order of priority:

1. Enhancement and remediation
2. Mitigating against the teaching/learning deficit
3. Examination drills
4. Extending the school week
5. Lessons as a quality assurance policy

Enrichment and remediation. A detailed analysis of the students' (N = 128) responses to this issue revealed that students from these two schools do lessons in the subjects they are doing well in, as well as those that they are having difficulty with. This suggests that extra-lessons is an avenue that students use for both academic enrichment and for remediation. When students purchase extra-lessons services in the subject areas in which they are doing well in school, it is to secure for themselves the advantage that they, seek especially in getting good

grades when competing in external examinations. When they purchase extra-lessons services for remediation purposes, it is largely an attempt to make up for the deficit in the teaching/learning experiences they are having with the subject in school.

Mitigating against the teaching/learning deficit. These students appear to experience this teaching/learning deficit in a variety of subjects and for a variety of reasons, including poor teaching; the student's personal history and expectations; and the school context, including its culture and ethos. At the core of the "poor teaching" factor is the perception that students have of some subject area teachers as being unable to explain their subject properly. This was the most frequent reason cited by the students for difficulties they experienced in a subject area.

Consequently, one of the major reasons why these students went to extra-lessons classes was to work with another teacher, whom they perceived as being able "...to provide a better explanation of the subject matter"; one who would be better able to "...relate the subject matter better to the student"; and a teacher who would "...push and encourage students ...who would answer questions when asked ...would correct home work ...[and] ...who would be sensitive and not make you feel stupid." This teacher would also "...cover the whole syllabus" but not at such "a fast pace" where a deep understanding of the subject matter was sacrificed.

Examination drills. One of the ways these students seemed to evaluate the effectiveness of their teachers was by the extent to which the teacher provided them with opportunities to do past examination questions. Teachers who provided them with lots of opportunities to attempt past examination questions were considered favourably. Those teachers who didn't, became suspect.

This becomes even more significant when one considers some of the other fundamental difficulties students have with many subjects. Whether it is with factual knowledge, conceptual knowledge, or procedural knowledge, many of our students' time is spent on recall, that is, on remembering the bits and pieces of facts, concepts, or procedures associated with a subject. Invariably, not enough time is allocated to understanding, applying, and analysing these facts, concepts, and

principles. Hence, these students noted that one of the reasons they sought remediation through extra-lessons was partly because the lessons teacher provided much more opportunities for them to “apply what they have learned, especially to solve past examination questions.”

Extending the school week. A sample of the students said that “...class size” was a factor that impacted on how they learned the content of a subject. Several students from the sample said that “...too many students in the class made it difficult for them to focus” or “...to concentrate” in class. This, of course, contributed to them “...not understanding the teacher or the subject” and so helped to propel them on to extra-lesson classes.

The teaching time allocated to subjects was also another factor that contributed to the deficit in the quality of the educational experience they had while at school. This was especially so when students thought that not enough teaching time was allocated to a subject. Therefore, students went to lessons, in part, to make up for this time deficit, so that extra-lessons also served the additional function of extending the school week for students who accessed the services.

On average, through extra-lessons, the students’ school week is extended by six hours of contact time with teachers. This is, for all intents and purposes, one full school day. These extra-lessons students, therefore, enjoy the advantage of a six-day school week. These students are willing to make this additional sacrifice, in part, because they want to do well, even in those subjects that they perceive themselves to be weak in.

‘Lessons’ as a quality assurance policy. Even if students were getting “excellent grades” in certain subjects in school, a solid majority of them (66.7%) would still do extra-lessons in some or all these subjects. These subjects commonly included mathematics, the natural and physical sciences, and certain core business courses such as principles of accounting. Clearly, here, these students are using extra-lessons as an assurance strategy for success in those subjects in both the internal and external examinations. “These days,” one student said, “...in order to pass exams you have to have the initiative to go the extra-mile, and to pay the extra cost, for a quality education.”

5. Who Decides? The Role of Parents

Because parents are the ones who ultimately decide in some cases (in 36.7% of the cases) on whether a student will purchase lessons services, parents' ambivalence about the teacher's capacity to assist their children in meeting the contested academic goals through scheduled in-school teaching alone is one of the many forces driving the extra-lessons explosion in Trinidad and Tobago.

Extra-lesson services have to be purchased. Students are charged on average about \$163 per month per subject, and, on average, students take lessons in three subjects, at a total cost of approximately \$500 per month. It is invariably the parent who has to come up with this resource. Hence, how many subjects a student accesses is in part determined by the capacity and the willingness of the parent to pay. Once the parent has the capacity and is willing to pay, then other factors come into play in the decision process, including the students' perceived need for the service. For example, one of the observations made in comparing these two schools is that the role of the parent in the overall decision-making process regarding extra-lessons varies for the two schools. The data show that students who attend the senior comprehensive school rely less on parental inputs (by a margin of 9.2%) than those students attending the traditional 7-year secondary schools.

6. Does Extra-Lessons Matter?

The correlation between the responses these two sets of students gave to this question was both high and positive ($r = +0.89$). Almost all the students (97%) sampled from these two schools, regardless of school type, confirmed the claim that extra-lessons did help to improve their performance in the subject. The limited times when lessons didn't work it was because the students did not apply themselves, or were distracted by friends, or that they tried their best "... and still could not understand." Hence, the few cases when extra-lessons did not work it was primarily because the students were unable to make it work.

Students from these two schools said that lessons worked for them "...primarily because the lessons teachers knew their work well." Furthermore, lessons provided more exam practice, and the lessons

teacher "...provided good notes." Students also indicated that they got the type of individual attention they needed and that "...repeating the school work in the subject" from the perspective of another teacher at these lessons sessions also "...helped."

Finally, students had a mixed response to how attending lessons affected their social lives outside of school. In a ratio of 8:1 [88.9%], students in the sample attending the senior comprehensive school said that lessons curtailed their time from extra-curricular activities like sports, recreation, and so on. Whereas, in the traditional 7-year school, the majority of the students (62.5%) made the opposite claim, saying that doing extra-lessons did not affect "their social life" outside of school.

Discussion and Conclusion

While the data presented here are rich, it begs more exploration about the range of subjects, differences in the type of participation and rate of participation between both schools, the role of parents in decision making, and so on. This discussion here, however, will confine itself to the main research questions.

The Equity Factor

Findings from socio-economic data confirm Jules' (1994) claim that students from comprehensive schools are generally materially less well off than students from traditional schools (see Figures 2 and 3).

Figures for participation in extra-lessons show that 89% of the sample from the traditional school took lessons compared to 52.7% of the sample from the comprehensive school (this is much higher than 56% for the former and 44% for the latter in Brunton (2002)).

It is immediately obvious that the two schools have a significant dependence on extra-lessons, but that the reliance on it is greater in the traditional school. In the traditional school, findings show that at Forms 4 and 5, the participation rates are at 100% and dip in Form 6. In the comprehensive school, the participation rate at Form 4 is 33.3%, 45% at Form 5, and 82.4% at Form 6. It is clear, therefore, that more students from the traditional school begin accessing extra-lessons much earlier.

The comparative decline in participation at the sixth form level in the traditional school could be due to the fact that students feel more confident to manage on their own while their counterparts at the comprehensive schools do not.

While there are variations in cost for different subjects and the range of subjects for which extra-lessons are sought (see Table 1), on average, students spend about \$500 per month on extra-lessons regardless of which school is considered. Most students in both schools attempt three subjects at extra-lessons.

The participation rates and costs, as well as the number of subjects for which extra-lessons are sought, suggest a high private cost of schooling on the part of parents. Parents with children in the traditional school in this study, which has all the advantages associated with high prestige, spend more money on extra-lessons than parents of children from the comprehensive school. That the participation in extra-lessons is significant in the “poorer” school means that parents also accept the need for extra-lessons and will pay when they can afford.

The perception of the need for extra-lessons and its costs suggest that education in Trinidad and Tobago is not free and those who can afford are at an advantage in the system. Children in the stronger school in this study accessed more extra-lessons.

The Efficiency Factor

The participation rates in private lessons and the range of subjects pursued already mean that students access significant assistance over and above what is given in the school system.

The data collected on motivations for extra-lessons speak more pertinently to this issue. There are two separate forces driving the process. One is the concern for good grades as a result of competition for scholarships or for job opportunities in a narrow labour market. Therefore, even if the system was efficient at delivering instruction, this factor would still hold. This is dependent on parental anxiety and a whole lot of cultural factors. Students agree that they would still seek extra-lessons in subjects in which their performance was good in order to

ensure high exam grades. Students from both schools agreed on this issue.

But there is a second factor driving the process, which has to do with student perception that teaching is inadequate and teachers are seen as not up to the task. They cite many bad practices from their teachers. In addition, they feel that instructional time is sometimes not sufficient for completion of their courses. In both schools, therefore, there is evidence to suggest that the system is not delivering quality instruction. Hence, students from both schools agreed that extra-lessons impacted positively on their school performance.

Accountability

In all education systems, success in formal examinations is the most popular yardstick by which schools are evaluated. It is an accepted fact in Trinidad and Tobago that the traditional schools achieve higher pass rates in examinations than senior comprehensive schools. And so, in the public mind, senior comprehensive schools are generally perceived as less successful than traditional schools.

The far greater reliance on extra-lessons by the traditional school compared to the comprehensive school in this study, if it is a typical pattern in the system, suggests the need to re-evaluate the way schools are assessed. Not only does the traditional school enjoy the advantages of stronger school culture, students with stronger academic skills, and more established alumni, they are also able to access greater support from the extra-lessons market. Therefore, comparisons of final examination marks cannot be relied upon to reflect the instructional capability of different schools.

It was stated at the beginning of this article that the Government of Trinidad and Tobago has been expending large sums of money to improve access to education in keeping with the drive for “Education For All.” It is clear from these findings that the existence of an extra-lessons industry contradicts the goals of democratization of education—both on grounds of efficiency and equity.

Due to the strong competitive nature of the examinations system and shortcomings in the delivery of instruction, candidates and their parents spend significant amounts of money in order to achieve success in examinations. Consequently, success becomes dependent on the ability to pay.

Policy makers must, therefore, pay more attention to the extra-lessons market in considering education reform.

References

- Bray, M. (2006). Private supplementary tutoring: Comparative perspectives on patterns and implications. *Compare*, 36(4), 515–530.
- Brunton, R. (2000). Extra-lessons in Trinidad and Tobago: Qualification inflation and equality of opportunity. *Caribbean Curriculum*, 7(2), 1–17.
- Brunton, R. (2002). The effects of extra-lessons on the qualification inflation and equality of educational opportunity in Trinidad and Tobago. In N. Mustapha & R. Brunton (Eds.), *Issues in education in Trinidad and Tobago* (pp 162–186). School of Continuing Studies, UWI, St Augustine.
- Buchmann, C. (1999). The state and schooling in Kenya: Historical developments and current challenges. *Africa Today*, 46(1), 95–117.
- Dang, H., & Rogers, F., H. (2008) *How to interpret the growing phenomenon of private tutoring: Human capital deepening, inequality, or waste of resources?* (Policy Research Working Paper; no. WPS 4530). Washington, DC: World Bank.
- Hallak, J., & Poisson, M. (2007). How to avoid the corruption of mainstream education. In *Corrupt schools, corrupt universities. What can be done?* (pp. 257–271). Paris: International Institute for Economic Planning.
- Heyneman, S. P. (2004). Education and corruption. *International Journal of Educational Development*, 24, 637–648.
- Jules, V. (1994). *A study of the secondary school population in Trinidad and Tobago: Placement patterns and practices: A research report*. St. Augustine, Trinidad: Centre for Ethnic Studies, UWI.
- Silova, I., Budiene, V., & Bray, M. (Eds.). (2006). *Education in a hidden marketplace: Monitoring of private tutoring: Overview and country reports*. New York: Open Society Institute.
- Stevenson, D. L., & Baker, D. P. (1992). Shadow education and allocation in formal schooling: Transition to University in Japan. *American Journal of Sociology*, 97(6), 1639–1657.

**CAN STANDARDS-REFERENCED, LARGE-SCALE
ASSESSMENT DATA LEAD TO IMPROVEMENT IN THE
EDUCATION SYSTEM?**

**Judging the Utility of Student Performance Standards in the
Primary School National Assessments of
Educational Achievement**

Jerome De Lisle

This paper documents the development of performance standards for the Trinidad and Tobago primary school national assessments of educational achievement. Performance standards are written expectations of student achievement operationalized in defensible cutscores. A major argument in this paper is that these standards are necessary to evaluate quality in the education system because they directly address the question of “*how good is good enough?*” Standards-referenced measurement systems are the basis of both compensatory and accountability systems. Standard-setting procedures in the 2005 and 2006 national assessments of educational achievement are described, followed by an evaluation of procedural validity using quantitative and qualitative data gathered from panellists. The findings indicated that while panellists appear confident about procedures and outcomes, cognitive complexity and organizational inefficiency could prove to be critical constraints. While the introduction of student performance standards appears useful; by itself, it cannot lead to significant education improvement unless there is also a coherent policy for effective data use within a national evaluation system. In developing such a policy, consideration must be given to choosing between (a) low versus high examination stakes, (b) compensatory versus accountability policies, and (c) school-based versus centralized management of test processes. At the very least, stakeholders must understand the purpose of national assessments of educational achievement.

In the enthusiasm for national testing, there is sometimes a fallacy that measurement by itself will induce positive change in education. Even when advocates recognize that testing must be linked to further action, just what that action should be is often unclear. (Chapman & Snyder, 2000, p. 458)

Examining the Context of Large-Scale Assessments

Eleven Plus Testing as Legacy and Tyranny

History has it that the first public examinations were the rigorous civil service examinations of ancient China (Miyazaki, 1976; Wilbrink, 1997). Just as with high-stakes examinations today, the results were released publicly with either favourable or grave consequences for both individuals and their families. By the 18th century, competitive public examinations were also introduced into a number of schooling systems. Today, while public examinations are common in many schooling systems, the way in which measurement tensions from different assessment purposes are balanced largely depends upon tradition and policy priorities (Broadfoot & Black, 2004). Sadly, in many Caribbean education systems, high-stakes public examinations are the dominant large-scale assessments (Payne & Barker, 1986). In particular, selection at Eleven Plus is, as a corrupting legacy, reducing the emphasis on classroom assessments and diminishing the value of newer instructionally supportive large-scale assessments. Perhaps, this distortion is most evident in the organization and mission of the primary school, with its manic focus on ensuring that students are prepared for Eleven Plus success, at all costs. Indeed, in order to guarantee success, early segregation of students is required to facilitate homogenous classrooms and appropriately paced instruction (Hanushek & Wobmann, 2006). Under these conditions, there is little early diagnosis and intervention. While these arrangements ensure that some high-ability students excel, teaching and organizing to a high-stakes test has negative consequences for most (Evans, 2001).

In the past, without large-scale assessments designed specifically for monitoring learning outcomes, Caribbean education systems lacked the data necessary to judge quality (World Bank, 1993). Therefore,

instituting national assessments of educational achievement became an important pillar in the education reform strategy promoted by international funding bodies (Sebatane, 2000). Ironically, at the same time, in developing countries, proposals for measurement-driven education reform using high-stakes testing were being promoted by some (Eisemon, 1990; Heyneman, 1987). For example, Chapman and Snyder (2000) had argued for the “positive impact” of high-stakes public examinations, citing the work of London (1997) in Trinidad and Tobago. However, London’s claim that the Eleven Plus essay component encouraged the teaching of writing in the classroom was based purely on anecdotal data. Kellaghan and Greaney (2004) were more realistic in assessing the use and consequences of high-stakes public examinations. While alluding to the possible benefits of measurement-driven instruction, they readily admitted to the negative consequences. Perhaps, with hindsight, it was extremely naïve to believe that a single examination, high stakes or not, could have a sustainable, positive, and system-wide impact. Such a conclusion underestimates the complexities and unpredictability of the washback phenomena (Broadfoot, 2002; Wall, 2005).

Can a High-Stakes Examination Tell Us About Achievement Standards?

National assessments of education achievement are standardized, large-scale measures designed primarily to describe levels of achievement in parts of or the whole education system (Greaney & Kellaghan, 2007). National assessments originated in the decisions made at the World Conference on Education for All (WCEFA) held in Jomtien, Thailand in March 1990 (Kellaghan & Greaney, 2001). However, by 2000, assessments designed to monitor achievement standards were still relatively rare in the region. Indeed, in 1992, only Jamaica and St. Lucia had developed adequate monitoring systems (World Bank, 1993). In the absence of such measures, some policy makers have resorted to using data from high-stakes public examinations to evaluate achievement standards (Trinidad and Tobago Chamber of Industry and Commerce, 2006; Trinidad and Tobago. Ministry of Education [MOE], 2005). While this approach might appear useful, the implication of such a strategy has not always been adequately gauged. Certainly, the quality of information derived from placement examinations like the Eleven Plus is limited by

content validity, cognitive complexity, reliability, teaching to the test, and student opportunities to learn.

Even if using data from a public examination was appropriate in evaluation, the information from the Eleven Plus is entirely norm-referenced and so cannot answer the question, “*how good is good enough?*” (Brandon, 2005). This question is an important one and relates to the adequacy of student performances. An answer always requires either criterion- or standards-referenced data. Unfortunately, without formal criterion-referenced cutscores, some technocrats have chosen to invent capricious cutscores, such as 30% or 50% of the total score, without reference to the difficulty of the test (MOE, 2005). Indeed, the usefulness of the 30% cutscore on the Eleven Plus as an indicator of “achievement standards” is limited because the working of the placement system depends primarily on ranked scores and not mastery of content. There is no assurance that students below the 30% cutscore are performing poorly; rather they are simply doing worse than the others. Of course, as highlighted earlier, the core problem is misuse of data from high-stakes public examinations for evaluation. National assessment policy, therefore, should distinguish between public examinations for selection and certification, national assessments for monitoring system quality, international assessments for international benchmarking, and classroom assessments for learning and feedback (Braun & Kanjee, 2006; Fiske, 2000).

The increasing role of national assessment systems within developing countries is an example of globalization and international transfer of educational reform strategies from industrialized societies (Benveniste, 2002; World Bank, 1993). Unfortunately, in the Caribbean, this transfer has usually been implicit, contained in the policy documents of the funding agencies. Less often, Caribbean Ministries of Education are active borrowers of assessments. It is rare that local policy makers seek out appropriate strategies and modify protocols to ensure contextual relevance and effectiveness (Sebatane, 2000). This is worrying because implementation fidelity for assessments will be affected by the values a society holds towards tests and test use (Oakland, 1995). Indeed, Greaney and Kellaghan (1996) have noted the many constraints for implementing national assessments within developing countries, including lack of measurement capacity. Ultimately, then, building

institutional capacity is one of the keys to successfully implementing a viable national assessment system. Currently, in Trinidad and Tobago, the Division of Educational Research and Evaluation (DERE), in the Ministry of Education (MOE), holds the primary responsibility for national assessments of educational achievement (T&T. DERE, 2006). Many of the recent improvements in system functioning have been due to better staffing, resourcing, and training coupled with leadership at technical and political levels. However, new issues are emerging to threaten further progress, including lack of stakeholder understanding, the absence of policies to guide evidence-based decision making, and inappropriate or limited data use.

Coming to Grips With the Purpose of National Assessments of Educational Achievement

Despite the implementation difficulties, it seems likely that data from national assessments of educational achievement are essential to improving education in Latin America and the Caribbean. The role of a national assessment system is magnified, perhaps, by the pervasive problem of low quality and high inequality in the region. Notable, too, is the apparent complexity of achievement patterns across geographical regions and social groups (Winkler, 2000). Indeed, the first step towards resolving inequity is the provision of disaggregated data needed to identify the nature and size of achievement gaps (Lewis, 2005). From the standpoint of the major funding agencies, the need for such an efficient national assessment system in Trinidad and Tobago was never in doubt. Therefore, building capacity became an early goal of the reform policies of international funding agencies (World Bank, 1993). World Bank reports also implied that the lack of national monitoring and evaluation systems was to blame for the pervasive inattention to equity and efficiency issues (World Bank, 1995). These concerns fuelled the reforms in the DERE in the late 1990s.

The need for quality data on the performance of the education system should have been obvious to local technocrats and academics on examining findings from the first international assessment in which Trinidad and Tobago participated. The assessment of reading literacy conducted by the International Association for the Evaluation of Educational Achievement (IEA) in 1990 and 1991 showed that Trinidad

and Tobago performed poorly (In the Grade 3 assessment T&T was ranked 25th among 27 other countries, which were mostly members of the OECD.) (Elley, 1992). The data clearly pointed to low quality and pervasive inequality across schools and regions (World Bank, 1995). Sadly, 16 years later, as evidenced in the country's performance in the 2006 Progress in International Reading Literacy Study (PIRLS), equity and quality are still critical issues (Mullis, Martin, Kennedy, & Foy, 2007). Therefore, a functional national monitoring and evaluation system is needed to (a) identify achievement gaps, (b) determine factors associated with these achievement gaps, (c) provide information on improvements resulting from education reform initiatives, and (d) facilitate the development of policies to support underperforming institutions and regions. A measurement system that supports these functions requires quality instruments, appropriate sampling strategies, performance standards, efficient analysis of data, and appropriate policies for data use (Greaney & Kellaghan, 1996).

Prior to 2004, a rudimentary system of "national testing" existed in the form of a centrally designed test distributed to all schools, with teachers expected to score student papers and provide feedback to students. Divisional facilitators were sometimes asked to assist in this process. However, this system could not meet all the objectives of a "real" national assessment system because little information on achievement standards was provided. For example, prior to 2004, only 10% of scripts were centrally scored (Manning, 2004). The system lacked integrity and purpose, and the function of the test remained nebulous. Consequently, in 2004, the MOE chose to install a more efficient, centrally administered national assessment programme. Instruments were designed to measure both achievement and non-achievement variables. The achievement tests in language and mathematics were administered to the entire population of Standard 1 (6- to 7-year-olds) and Standard 3 (9- to 10-year-olds), whereas the measures of teacher and student attitudes were administered to a sample of schools. Compared to the earlier system, there was a greater degree of centralization coupled with radical improvement in the core aspects of test development, such as item writing, scoring, setting performance standards, and data analysis.

Installing a “New” System for Monitoring Achievement Outcomes

Clarifying the Purpose and Intent of the Trinidad and Tobago National Assessments of Educational Achievement

The 2004 national assessments of educational achievement included pen and pencil tests covering four content strands in mathematics and six in language arts, using a combination of discrete and extended constructed-response formats. Centralized test development and scoring facilities were developed and coordinated by the DERE in conjunction with the Division of Curriculum Development (DCD). Curriculum facilitators played a major role at all levels of test development. Initially, there was some ambiguity about the role of the test and the stakes to be attached. As late as 2006, it was reported that the national achievement tests were to be a part of the continuous assessment process. The implication was that scores from the tests might contribute to high-stakes placement decisions at Eleven Plus (Hazel backs down, 2006). Such a strategy would have significantly increased test stakes, with consequent threats to defensibility and integrity. To be fair, however, this ambiguity of purpose was also present prior to 2004, especially within implementation proposals for the Continuous Assessment Programme (CAP). CAP’s vague assessment policy implied that classroom, school-based, and national assessments were similar in purpose and function. It may be that the implication of using scores from school-based assessments for high-stakes placement decisions had not been adequately considered (Jones, 2001).

The design choices of the 2004 national achievement tests led to the following decisions: (a) centralization of test development, (b) census annual administration, and (c) low-stakes implementation. There were advantages and disadvantages to each decision. On the one hand, the advantages of greater centralization were that the quality of the examination would be significantly improved and national benchmarks could be established. Additionally, with census administration, school performance information would be available for all institutions, facilitating the development of a comprehensive national monitoring system. On the other hand, the main disadvantage was that teachers would be faced with an externally mandated large-scale test that they might perceive as unnecessary. As in the United Kingdom (UK), teacher

resistance would then become a significant obstacle (Grant, 2006; Mayo, 2005). The perception that national assessments of educational achievement are an external imposition might also restrict ownership and limit data use. Additionally, with a low-stakes examination, some schools might simply opt out or choose to ignore the results, even with mandatory administration (Heubert & Hauser, 1999). Indeed, the tendency of schools to opt out of government-led reforms had occurred in the past, especially in the government-assisted sector (Braithwaite, 1981).

When a Norm-Referenced System is not Good Enough

In some aspects, the norm-referenced reporting system developed for the 2004 national tests was a significant advance over that used in the Eleven Plus. For example, the use of normal curve equivalents (NCEs) like the standard score ensured interpretable data that could be easily processed. The NCE standard score has a mean of 50 and a standard deviation of 21.06. The NCE scale of 1–99 coincides with a percentile rank scale at 1, 50, and 99. The advantage of the NCE scores is that the intervals between scores are equal and scores can be meaningfully combined and averaged. A national mean score of 50 serves as a benchmark or reference point for comparing the performance of regions and institutions. However, a central argument in this paper is that standards-referenced measurement is still required for sound evaluation. In terms of cost-effectiveness, the question then becomes, *How do we justify the additional expertise, training, cost, and time spent in devising and implementing a new system of performance standards?* Any judgement of cost-effectiveness must consider the inability of norm-referenced data to adequately answer the critical question of “*how good is good enough?*” (Brandon, 2005). Norm-referenced data can only provide information on the relative performances of schools and districts, but not on criterion-referenced standards.

Answering the question of “*how good is good enough?*” requires a benchmark for student performance that is tied to the curriculum. A standards-referenced measurement system is able to provide such an answer because it reports information anchored in substantive statements about the levels of student performance (Linn, 2005). Standard-referenced systems make use of cutscores that are absolute but not

capricious. A capricious cutscore, like the 30% threshold used in the Eleven Plus, lacks meaning because it is simply a point on a score scale, without reference to any particular educational criterion. While professional judgements are central to standards-referenced system, these judgements are valid and defensible to the extent that qualified, trained judges are used, appropriate criteria are constructed, and explicit procedures are closely followed during the exercise. Therefore, establishing the defensibility and validity of standards-referenced cutscores becomes the focus of the next part of this paper.

Developing Defensible Student Performance Standards

The fundamental element of any standards-referenced measurement system is the student performance standard (Hansche, 1998). Kane (2001) defined such a student performance standard “as a level of performance described in terms of what examinees at a particular level know and can do” (p. 55). Therefore, performance standards might be regarded as informed expectations of student proficiency, based on the objectives of the content standards and test items. These standards allow a qualitative description of different levels of performance, which might be considered the “normative” aspect of the standard. Additionally, there is also a substantive aspect of a standard, which refers to the written descriptive statements of the expected performance (Haertel & Lorie, 2004). Standards-referenced systems are similar to criterion-referencing systems because they are based upon student skills and knowledge rather than comparative performance. However, standards-referenced systems differ from criterion-referenced systems because the focus is on *expectations of proficiency* rather than on mastery of specific content. Whereas traditional criterion-referenced systems make use of quantity and category indices, such as the percentages correct or the numbers failing and passing on a particular task or class domain, standards-referenced systems provide qualitative descriptions of student performance (Cizek & Bunch, 2007; Nitko, 1980).

Standard setting is the test procedure used to develop a system of student performance standards. In this process, the cutscore for each proficiency level is obtained. These cutscores are at the heart of the process, providing an operational definition for each proficiency level, with a single cutscore at the lower boundary. By definition, individuals who

have attained marks at or above the cutscore would have met the specified standard. Therefore, a system for performance standards should include (a) definitions of each level of proficiency, (b) standard-setting procedures for obtaining the cutscores for each level, and (c) associated cutscores (Brandon, 2005). Performance standards have been used for years in the licensure and certification industry within the United States (US) and Canada (Meara, Hambleton, & Sireci, 2001). However, setting standards at this level is relatively easier because there are usually only two levels—pass or fail—for any licensee. By contrast, education tests require multiple proficiency levels, sometimes as many as six. As a result, the process of standard setting in education is more complex and susceptible to error, with some popular methods harder to implement. Most psychometricians now agree that the current set of standard-setting procedures are rational, well documented, scientific approaches to developing reasonable standards for educational performances (Cizek, 1993; Zieky, 1997).

It is important to distinguish between the “examination standards” in public examinations and the performance standards constructed in national assessments. The method of setting standards in both large-scale assessments is quite different, with different conceptions of rigour and defensibility. In public examinations like the CXC or GCE O Levels, the process of setting grade boundaries is organized very differently and makes use of very different protocols. For example, in the setting of grade boundaries in public examinations, there is usually a central management role for the accountable officer in the examination agency (Baird, 2007; Tomlinson, 2002). In one specific procedure described for the UK Assessment and Qualifications Alliance (AQA), an awarding committee of four to eight members scrutinizes the work of judgemental grades only, with the remaining grade boundaries determined by calculation (called arithmetic grades). Professional judgements by members of this awarding committee will be based on the quality of work, the comparability of students’ work, and statistical data from the current and preceding years. Each member of the awarding committee independently reviews a different sample of scripts within the grade boundaries initially recommended by the principal examiner (Meyer, 2005).

In standard setting for national assessments of educational achievement, there is no formal role for any single individual equivalent to the accountable officer of an examination board; instead, decisions are made by a large panel of judges. Secondly, the validity and defensibility of the performance standard is closely tied to following a well-known and published protocol. While defensibility is a legal and technical benchmark, there can be no true standard, and different procedures will give vastly different results (Cizek, 1995). Therefore, defensibility emphasizes procedural and technical accuracy and rigour. Cizek (1993) considered a defensible standard to be “the proper following of a prescribed, rational system of rules or procedures resulting in the assignment of a number to differentiate between two or more conceivable states or degrees of performance” (p. 100). Current best practice in most procedures requires (a) many high quality judges, (b) well-written descriptors, and (c) large number of work samples.

Designing the Standard-Setting Plan

To establish a comprehensive standard-setting system for national assessments of educational achievement, five components are required: (1) the administrative arm, consisting of the implementation team and an authority to set policy; (2) a content domain; (3) selection of persons (judges) to make judgements about desired levels of performance; (4) a methodology for collecting judgements and estimating standards; and (5) some means for reporting the results (Reckase, 2000). A standard-setting plan for Trinidad and Tobago would have to fit the social and cultural milieu and take into account possible barriers to implementation. It was hypothesized that possible barriers to implementation might include: (1) lack of familiarity with standard setting in the educational community, (2) lack of expertise by teachers in both standard-setting procedures and the content areas assessed, (3) variability in the outcomes and procedures of current standard-setting methods, and (4) the absence of Item Response Theory (IRT) item analysis. The last meant that some standard-setting procedures, such as item mapping and the bookmark method, were automatically excluded. The standard-setting plan would also need to consider the (a) length and cost of the process, (b) administrative capacity of the MOE, and (c) the number of potential qualified judges in the system (Greaney, 1996).

Table 1. Preliminary Labels and Performance Levels Descriptions for the 2006 Primary School National Achievement Tests of Trinidad and Tobago

DERE Performance Labels	Alternative Performance Labels Used in the Literature	Description of Performance Level
Level 4	Advanced, Exemplary, Accelerated, Distinguished Exceed Standards	(Exceeds the Overall Standard of Work required at this Level): Superior academic performance indicating an in-depth understanding and exemplary display of the skills required.
Level 3	Proficient, Competent, Mastery, Satisfactory, Meets Standards	(Meets the Overall Standard of Work required at this Level): Satisfactory academic performance indicating a solid understanding and adequate display of the skills required.
Level 2	Partially Proficient, Nearing Proficiency, Nearly Meets Standards, Borderline, Approaching Basic, Just Below Standards	(Nearly Meets the Standard of Work required at this level): Marginal academic performance, work approaching, but not yet reaching, satisfactory performance. Performance indicates a partial understanding and limited display of the skills required.
Level 1	Below Basic, Beginner, Novice, Emergent, Minimal, Developing, Well Below Standards	(Well Below the Standard of Work required at this level): Inadequate academic performance that indicates little understanding and minimal display of the skills required. There is a major need for additional instructional opportunities, remedial assistance, and/or increased student academic commitment to achieve at the Proficient Level.

An ad-hoc implementation committee was constituted consisting of members of the DERE and the consultant. A proposal was put forward detailing the plan for standard setting. The plan included a framework that suggested four to six performance levels and accompanying labels and descriptions (T&T. DERE, 2005). The administrative arm finally agreed to four performance levels with accompanying descriptions, illustrated in Table 1. While the implementation team believed that the 2004 data pattern might have supported a five-level specification, it was decided to begin with a simplified four-level system that would enable judges and users to easily identify passing (Levels 4 and 3) and failing students (Level 1 and 2). For this paper, the preferred performance level labels are (Level 4) *Exceeds Standards*; (Level 3) *Meets Standards*; (Level 2) *Nearly Meets Standards*; and (Level 1) *Critically Below Standards*.

The next step was to select a set of procedures to ensure defensibility and transparency. The implementation committee considered these two attributes critical to eventual acceptance of the standards by policy makers, teachers, and other key stakeholders. Defensibility, the most important attribute, called for appropriate attention to both substantive and procedural issues. However, choosing between the range of standard-setting methods was a daunting task. Berk (1986) listed as many as 38 different methods and, more recently, Cizek and Bunch (2007) listed 12 popular procedures, each with many variants. The traditional approach has been to classify methods as either test- or examinee-centred (Kane, 1998). Test-centred methods, such as the Angoff and bookmark, are focused on items or the test, while examinee-centred methods, such as the contrasting group and borderline method, are focused on the performance of the candidates. The holistic methods, such as the body of work and booklet classification, are best classified as performance-centred methods rather than test-centred because they make use of student work samples.

Standard-setting methods for national assessments of educational achievement differ in cognitive complexity, reproducibility, precision, and appropriateness (Reckase, 2000). This might be one argument against using a single method, such as the Angoff. While the Angoff has high precision and reproducibility, its higher cognitive complexity might result in difficulty for novice panellists. Indeed, there is debate over the

viability of some procedures used in the Angoff within the education setting (Hambleton et al., 2000; Zieky, 1997). The final decision, therefore, was to use three separate methods and a final synthesis decision. The methods chosen were two variants of the Angoff—whole booklet classification and the contrasting group method. This multiple-method synthesis procedure was originally employed by the Kentucky Department of Education and is further documented in the academic literature by Green, Trimble, Scott, and Lewis (2003).

The DERE synthesis strategy is summarized in Table 2. Each method is identified by the locus and reference of the individual judgement task. As shown, the chosen methods in the triangulated design differ in the nature of panellists' judgement. The Angoff procedure requires judges to estimate the probability of a minimally competent candidate at any achievement level answering the question correctly. There are multiple variants of the method, including some protocols that require judges to estimate the most likely score of students at each level. The latter variant is used for constructed response questions. Whereas the Angoff variants required judgement on expected performances on items, in the whole booklet classification procedure, panellists made a judgement on the entire body of work in the completed test. This post-hoc holistic judgement required judges to classify student performance in one of 12 levels. The variant of this method followed the published procedure in Jaeger and Mills (2001). The 12 levels were also used by panellists in classifying student performances in school on the contrasting group method. Details of each procedure used in 2005 were first published in a workshop manual (De Lisle, 2005) and further adjustments were made in 2006.

Evaluating the Standard-Setting Process

Evaluating the standard-setting process involves gathering evidence for the validity of performance standards. Assuming that there is already sufficient credible evidence for validity of the test instrument, evidence is also needed to support the conceptual basis of the standards and its operationalization into cutscores. There were a number of useful evaluation frameworks in the literature. For example, Schafer (2005), using a sponsor's perspective, proposed an institutional perspective, which subsumed legal, psychometric, and definitional foci.

Table 2. Standard-Setting Strategy for the 2005 and 2006 Primary School National Achievement Tests

Traditional Classification of Procedures	Test-Centred Methods		Examinee-Centred Method	Combination
Standard-Setting Procedures	Probability & Modified Angoff	Whole Booklet	Contrasting Group	Synthesis
<i>Classification Based on Nature of Decision-Making Process (Locus & Procedure)/Characteristics</i>	<i>Binary/Probability Judgements & Estimated Proficiency on Item</i>	<i>Binary Judgement & Proficiency on Set of Items</i>	<i>Binary Judgement & Proficiency in all School Settings</i>	<i>Impact Data and Comparison of Cutscores</i>
1. Iterative process	Yes	No	No	No
2. Review of student performance records in classes	No	No	Yes	No
3. Examined examination scripts	Round 2	No	No	No
4. Examined normative data, including relative difficulty of items	Yes	Not required	No	No
5. Knowledge of minimum pass levels for each item or cutscores set by other teachers	Yes	No	No	Yes
6. Made use of relative impact data ^o	No	No	No	Yes

^o Provided in Synthesis only

Schafer successfully applied the 20 evaluation criteria developed by Hambleton (2001) to this institutional perspective. Haertel and Lorie (2004) developed five validity arguments covering content standards, alignment, accuracy and precision, performance standards, and the cutscore. Gomez and Padilla (2004) offered a sixth possible validity argument for the consequences associated with implementing performance standards. Earlier, Kane (1994) had developed a comprehensive framework for validating performance standards, based on three distinct areas: procedural, internal, and external. Hambleton and Pitoniak (2006) only recently summarized each component area, with procedural validity involving explicitness, practicality, implementation, panellist feedback, and documentation.

Significantly, all evaluation frameworks consider procedural validity as critical to legal and technical defensibility and a defence to questions that might be raised about the arbitrariness of the standards. US courts of law consider “arbitrariness” in standard setting as unreasonable and capricious decision making (Carson, 2001). Thus, this evaluation study is restricted to those validity arguments focused on procedural propriety. Traditionally, this information is obtained from panellists’ feedback using quantitative methods. For example, Hambleton (2001) developed a standard-setting evaluation questionnaire to measure the adequacy of training, activities, confidence in the process, and outcomes. This evaluation questionnaire includes a mixture of closed and open-ended items and provides primarily retrospective, recognition information when administered at the end of the event (Skorupski & Hambleton, 2005). While useful, this type of data provides only weak support for procedural validity and gives little insight into the cognitive processes of panellists (McGinty, 2005). It is therefore prudent to also collect qualitative data that provides insight into panellists’ thinking when making judgements (Skorupski & Hambleton).

Qualitative evaluation data collected concurrently with the judging process will provide greater insight into the actual cognitive decision-making processes of judges (Giraud, Impara, & Plake, 2005; Hartz & Auerbach, 2003). This type of data will be useful in evaluations that are focused on improving the programmes being evaluated and that are intended to provide in-depth contextualized information on practices (Greene, 2007). Since judges are the key to improving the quality of the

process, the panellists' perspective is a critical source of information in monitoring. There is certainly need for more information on how judges make decisions in standards setting. For example, Brandon (2004) noted that, despite its popularity, the Angoff was a virtual black box when it comes to understanding how panellists make judgements. Plake (2008) also highlighted the need to understand the importance of the orientation and training components. These arguments hold doubly true for standard setting in Trinidad and Tobago, with the lack of standard-setting experience among judges.

Evaluation Methodology

The evaluation study employed multiple methods, gathering different types of data using a questionnaire and an unstructured diary/journal. This approach was designed to provide both concurrent recall and retrospective recognition data. The diary/journal was especially valuable as a data collection tool because it allowed in-depth reflection of participants and an exploration of evolving insight over the duration of the process and the event (Hiemstra, 2001). Panellists were expected to keep an account of their daily experiences as well as to recount their personal understandings, thoughts, and feelings about the process. In 2005, 49 questionnaires and 38 journals were collected. In 2006, 65 questionnaires and 51 diaries/journals were returned. Since the standard-setting procedure was an intensive one-week process, the 2006 return rate of 92% for the questionnaires and 69% for the journals and diaries can be considered adequate.

The questionnaire had a slightly different design for the two administrations. In 2006 only, items from Part 1 of the questionnaire were based entirely on the evaluation framework of Hambleton (2001). The instrument included closed and open-ended questions targeting retrospective recognition of key standard-setting activities, such as construction of descriptors, efficacy of training, and overall confidence in the outcomes. Part 2 of the 2005 and 2006 questionnaire consisted of eight open-ended questions originally designed by the consultant for use in a focus group interview. In 2005, the questions were:

1. Did you understand the standard-setting process?
2. What did you like or didn't like about the process?

3. What were your personal understandings of the performance levels and did they match what you were required to do?
4. Which standard-setting methods were most difficult in terms of understanding what was required and what you were required to do?
5. Which standard-setting method in your mind best captures the levels of student performance?
6. To what extent do you agree with the process and the standards derived?
7. Did you benefit from the discussion within your group and how did that discussion change your individual ratings?

In 2006, a number of additional questions were added, including:

1. Explain the strategy you used to assign students to the different categories.
2. Were there any problems that might have influenced the way you categorized students?
3. To what extent do you agree with the process and the standards derived?

All descriptive quantitative data were collated and presented in a tabular format. Qualitative data from the two years were transcribed, coded, organized, and reviewed by two independent researchers. Data analysis focused on exploring the themes that emerged from the data rather than applying a coding scheme based on Hambleton and Pitoniak's (2006) specifications for procedural validity. Content analysis was first used to identify the recurring themes, which were then coded. This preliminary coding was constructed after an initial reading of the text. Themes and sub-themes were then re-grouped and organized into tables. The final coding sets of both reviewers were then reconciled. The reconciled tables listed themes, codes, and sample statements constructed. The entire coding scheme was then independently reapplied to the text data and the tables further refined. Further reconciliation was attempted before constructing the final sample statements, codes, and themes. Following the finalization of the themes and codes, the narrative was written by decoding each of the tables.

Evidence for Procedural Validity

Findings From the Questionnaire

Table 3 provides self-report data from the 2006 panellists on their perceived success on different standard-setting tasks. As shown, the majority of panellists reported success on all seven standard-setting tasks. The highest rate of perceived success (successful and very successful) was on completing the test (98.4%), but judges also felt successful on the integrated training and practice exercises. However, 22.2% of the panellists felt they had achieved only partial success setting the descriptors. Likewise, for the Angoff procedure, 17% of the panellists reported only partial success on the Round 1 Angoff task. This figure, however, dropped to 10.9% in Round 2 and 4.7% in Round 3. This suggests that the ability to make judgements improved in later rounds, possibly due to greater opportunities for practice. Table 4 shows panellists' perception of the performance standards produced. Overall, panellists felt that the quality of the standards generated was adequate, with 96% giving an adequacy rating of 3 to 5 for Level 4; 98.2% for Level 3; 98.2 for Level 2; and 96.6% for Level 4. It appeared, then, that panellists were more confident about the integrity of performance levels at the extreme ends of the scale, including Level 4 (40.8%) and Level 1 (43.1%). However, they were comparatively less confident about the quality of the standards in the middle categories, Levels 2 (31.6%) and 3 (28.1%).

Table 5 displays data on panellists' confidence in the outcomes and the different standard-setting methods used. Supporting the data presented in Table 4, panellists were most confident about the standards in Level 4 and 1, with 36.9% indicating that they had high confidence in Level 4 and 43.1% in Level 1. However, 21.5% of the panellists expressed a moderate level of confidence in Level 3 and 24.6% in Level 2. When responding to the close-ended items, panellists considered the contrasting group and whole booklet classification methods to be the best procedures for setting cutscores, with only 14.5% of the respondents answering in the affirmative for the Angoff. Table 6 listed seven possible factors that influenced panellists' judgements. The factors were ranked in order of importance, with descriptors (65.1%), classroom experience (58.5%), and student responses (53.8%) seen as most critical. The least important

factors appear to be first round judgements and judgements made from the Angoff.

Table 3. 2006 Panellists' Perceptions of Success on Different Standard-Setting Tasks (N = 60–65)

Standard-Setting Activities	Not Successful	Partially Successful	Successful	Very Successful
Completing the Test	0	1.6	41.3	57.1
Angoff Round 3	0	4.7	57.8	37.5
Developing Descriptors	0	22.2	49.2	28.6
Angoff Round 2	0	10.9	62.5	26.6
Angoff Round 1	0	17.2	65.6	17.2
Contrasting Group	0	8.1	75.7	16.2
Integrated Training and Practice	0	8.3	83.3	8.3

Table 4. 2006 Panellists' Perceptions of Quality in the Performance Standards Generated (N = 49–58)

Performance Level	Ratings				Totally Adequate
	Totally Inadequate	1	2	3	
Level 4 (Mastery)	2.0	2.0	12.2	42.9	40.8
Level 3 (Proficient)	1.8	0	17.5	56.2	28.1
Level 2 (Borderline Proficient)	1.8	0	19.3	47.4	31.6
Level 1 (Remedial)	3.4	0	8.6	44.8	43.1

Table 5. 2006 Panellists' Confidence in and Comparison of Standard-Setting Methods, Outcomes, and Procedures (N = 65)

	Confidence			
	Very High	High	Medium	Low
Level 4 Classification	36.9	53.8	9.2	0
Level 3 Classification	16.9	61.5	21.5	0
Level 2 Classification	20.0	55.4	24.6	0
Level 1 Classification	43.1	47.7	9.2	0
Overall Procedure	12.3	69.2	15.4	0
Method	Percentages indicating particular method as "best" procedure			
Angoff	14.5			
Whole Booklet	41.8			
Contrasting Group	43.6			

Table 6. 2006 Factors Used by Panellists in Classifying Student Performances (N = 65)

Factors	Not Important	Somewhat Important	Important	Very Important
Descriptors at each level	0	3.2	31.7	65.1
Own classroom experience	1.5	7.7	32.3	58.5
Student responses	1.5	4.6	40.0	53.8
Perception of test difficulty	3.1	7.7	40.0	49.2
Discussion in group	0	20.0	36.9	43.1
First round judgements	11.9	33.9	33.9	20.3
Judgements from the Angoff	6.3	25.0	51.6	17.2

The Qualitative Enquiry

The qualitative data included retrospective recall and recognition information from the open-ended part of the questionnaires and concurrent recall information from the five-day journal. Analysis of the text data established five main themes in the respondents' comments. In order of frequency, these were: (1) outcomes related to the performance standards; (2) organization of the process, including group discussion; (3) levels of the performance standards; (4) cognitive challenge of the process; and (5) advantages and challenges of specific standard-setting methods. Focusing first upon outcomes, panellists often reflected on the value and use of the standards for the education community, as the following comment by one panellist illustrates:

This process clearly demonstrates or assists me in understanding how important our role as stakeholders are in terms of coming together with dedication and commitment to formulating standards that would serve the needs of all children across Trinidad and Tobago. [Panellist-Diary]

Many teacher panellists felt that the standard-setting process helped them gain deeper insight into student performances within the classroom. These teachers considered that they might be used to modify expectations of students. For example, one teacher noted that performance standards were useful as benchmarks within her classroom:

The process, when completed, painted a clear picture about the students' performance. After making judgments, it was easy to see the level of our children and to be well equipped to make informed decisions about going back to the classroom and putting my house in order. [Teacher Panellist 1]

An often-repeated theme was the reality check experienced when panellists confronted students' actual work performances in the test. This revelation occurred after panellists had examined students' scripts during the whole booklet classification method or after being presented with test statistics during Round 2 of the Angoff. Most judges were surprised by the poor quality of student responses in the scripts and low mean performance on some test items. This reality check was a critical part in the decision-making process, often leading panellists to lower their initial

expectations. More significantly, this reality check resulted in some panellists re-evaluating the quality of teaching and learning within the nation's schools; a notable point revealed in the following 2006 journal entries of two panellists:

We completed the 3rd round of the Angoff today, using the statistics from the test. It was amazing that the results were so low when compared. There were also a lot of children getting zero in items that should be simple for them. [Panellist 1]

Today, the groups were involved in marking the sample scripts using the whole-book method. We marked STD3 in 4 rounds and Std 1 in 5 rounds. There were not many students found in the upper ranges of the level 4 but quite a few in the lower ranges of level 1. This is good information to be used in better understanding the way children think and learn. . . . There are many children at risk and we must act now. [Panellist 2]

However, to be sure, not all comments about outcomes of the standard-setting process were positive. For example, some panellists felt that skills identified in the standard neither fully matched nor extrapolated to the outcome domain. Additionally, some panellists believed that none of the existing procedures captured fully the range, variety, and complexity of student performance in schools and classrooms. These panellists often argued that, in reality, there were too many factors influencing student competence. While these concerns were often based upon the narrowness of the test and limitations of content standards, some of these panellists were also concerned with limitations in the standard-setting procedures. Indeed, quite a few of this group felt that there was a mismatch between expectations in the standard and real-world student performances, as revealed in the following questionnaire response:

The process was done without any information of school, teacher, or pupil. There are many [other] variables to consider [such as]: (1) completion of syllabus, (2) mental disposition of students on a one-day test, (3) socioeconomic and psychosocial issues affecting children, and (4) children's ages and promotion by attainment. [Panellist-Questionnaire]

In the 2005 standard-setting process, a number of panellists commented negatively on the overall organization of the event. This perceived disorganization was discomfoting, leading to frustration and confusion during the judgement process. While administrative delays and alterations might have been relatively infrequent, unsatisfactory arrangements were especially burdensome because they made a difficult and challenging task even more arduous, impacting upon the quality of judgements. For example, one panellist noted that having groups and individuals shifted about constantly added to the confusion over the process:

I did not like the way the group was shifted around constantly with no clear indication of what was to be done. This phase was a bit confusing. However, I understand that since this is the first time that this exercise is being done, adjustments were being made as we went along. [Panellist, 2005]

The situation was much improved in 2006, with better accommodation, quality meals, and additional administrative staff. Thus, in the second year of the process, panellists were more likely to commend the MOE for the physical setting and better organization. For example, one panellist expressed her feelings about the new venue for the workshop in 2006:

This morning I set forth for ____ auditorium in _____ for the standard setting of the National Tests 2006. As I entered I was amazed at what befell my eyes, the setting was pleasing to the eye. . . . This was the first time I was attending a workshop/session held by the Ministry of Education in such style. What a difference! Finally, teachers are being treated as the professionals they are! [Panellist, 2006]

In this exercise, group discussion at the end of each Angoff round appeared to be a critical component of the judgement process, ensuring that different ideas about student performance were disseminated and unrealistic or biased expectations moderated. Without prompting, panellists often noted this aspect in their journals, highlighting the value of group discussion in generating deeper insight into the process of judging. Significantly, in the 2006 standard-setting exercise, a number of panellists also indicated that the group process enhanced their sense of belonging, which made the overall task easier by ensuring quicker social

adjustment to the unfamiliar surroundings. Asked specifically about the role and value of group discussion, panellists indicated five main advantages provided by these discussions: (1) useful information for altering initial expectations and judgements ratings; (2) deeper, thought-provoking insight into student performance for panellists who were new teachers; (3) credible arguments for cementing and making judgements; (4) added insight into students' thinking when responding to written questions; and (5) alternative perspectives on issues and responses. The perceived value of group discussions in the decision-making process is clearly noted in the statements of the following two panellists:

The discussions were extremely healthy. Some people had strong views, but provided credible arguments to substantiate those views. In some instances, it caused me to change my ratings while in others it didn't since I am also one who would argue strongly for anything that I truly believe in. [Panellist 1]

Group discussions, helped me to expand my way of thinking about students at a Standard 1 level, since my beliefs were based entirely upon interaction with students in my class. The group discussions shed light on very dark matters, so, yes I did benefit from these discussions. The number of years experience in the teaching profession placed me at a slight disadvantage from the other seasoned teachers in my group. I was therefore, very open and willing to accept any thoughts or ideas that they were willing to share. [Panellist 2]

Panellists often commented at length on the nature and role of performance standards, highlighting seven main areas: (1) the normative aspect, (2) the substantive aspect; (3) the measurement instrument; (4) the standard-setting methods employed; (5) differences in their own understanding and expectations; (6) the need for challenging expectations; and (7) outcomes, such as the workability of standards and procedures. On recall, panellists most often emphasized either the substantive element, such as the value of the criterion-referenced descriptors, or the normative element, such as the nature of differences among students at the various proficiency levels. Generally, panellists felt that the real value of the system was the ability to judge student performance using criterion-referenced statements. They also believed that teachers within their classrooms could directly apply both the

substantive and normative elements. This application might prove critical to improving learning outcomes, as indicated by the following two panellists' comments:

I liked the formulation of the descriptors for the various content strands. These descriptors gave a clear indication what one is looking for when doing the analysis of the test items. However, what I did not like, or should I say [was] uncomfortable with, was that I felt my judgments might be too subjective. [Panellist 1]

The process, when completed painted a clear picture about the students' performance. After making judgments, it was easy to see the level of our children and to be well equipped to make informed decisions about going back to the classroom and 'putting my house in order'. [Panellist 2]

Although most teachers placed a high value on establishing challenging expectations for student performance, some panellists did question the measurement process, including the quality of the test and even the particular standard-setting methods employed. To some extent, this reflected a mature view of the assessment process, since deficiencies in the test and measurement process would invalidate the performance standards, as one panellist observed in her journal:

The [whole booklet classification] method does not seem to be a fair process as the results are based on your perception [rather than] taking into consideration the child. The Angoff method uses statistics and the descriptor also does not focus on the child, as there are the exceptions in both cases. The standard setting process cannot be based on one test. [Panellist-Journal Entry]

It was expected that the standard-setting process would be cognitively challenging for most panellists. Although many possessed at least a first degree in a content area or education, few had prior knowledge or experience with standard setting. Therefore, most of the content and methodology were new and unfamiliar to participants. This added greatly to the challenge, as one experience test scorer noted in an entry on the first day of her journal entry:

Today was quite interesting and confusing. Many concepts introduced were quite new. It seemed like quite a task to complete. I am quite interested in understanding the standard setting as the Angoff, [Whole Booklet Classification] and the Contrasting Group. This exercise interested me because after correcting the National Test scripts, I was eager to learn how it was going to be standardized. The process of determining descriptors for the various levels was a tedious one, since agreement was hard to reach at times. However, it was a very good method of knowing the descriptors. . . . At the end of the day I felt I did not have a clue about the Angoff method and I hoped that the next day would assist with this. [Panellist-Journal Entry]

The exercise was also regarded as especially difficult because respondents had to participate in multiple tasks and use information from multiple sources, all at the same time. For example, in making judgements about student performance, panellists were required to construct and use descriptors, examine data, observe student work samples, while discussing their own judgements with those of other panellists. Some processes also demanded skills that panellists may not have acquired. To illustrate, one panellist commented on the difficulty of dealing with the item analysis data presented in Round 2 of the Angoff:

The afternoon session, with the introduction of the statistical evidence, in terms of percentages, brought an even deeper level of reasoning for me. I now had to look at the percentages and determine the difficulty level of the item, look at the rubric and the item itself and determine my round 3 cutscores for the standard 1 Maths. This was a very draining exercise. I do not believe that thinking about anything over my holidays so far drained me that much. I just hope that I performed the exercise well. I dread to think of doing it again tomorrow for Standard 3 Maths. [Panellist-Journal Entry]

The degree of challenge was also high because judges had to learn and employ three different procedures at roughly the same time. While journal entries indicated that panellists' knowledge evolved rapidly, this sharp learning curve sometimes resulted in frustration. The integrated design of the exercise also made it difficult to manage time and some respondents believed they had insufficient training and practice.

When asked directly which standard-setting method was most difficult, in the 2006 questionnaire, 39 of 53 panellists identified the Angoff. The Angoff was considered especially difficult because it was abstract, required interpretation of statistical data, lacked authenticity, and employed the concept of a borderline or minimally competent student, which was difficult for some to conceptualize. Only seven panellists regarded the whole booklet classification method as the most difficult. These panellists gave a variety of reasons such as tediousness, lack of fairness, and difficulty in making a compensatory, holistic judgement. However, when asked directly which method best captured student performance levels, in 2005, 18 out of 41 said the Angoff, and in 2006, 30 out of 56. This compared with 23 out of 45 for the booklet classification method in 2005 and 26 out of 56 in 2006. Thus, surprisingly, although the majority of judges regarded the Angoff as the most difficult, it was also the most highly valued procedure. Panellists who favoured the Angoff pointed to the fact that it was item-based and made consistent use of the descriptors. Many panellists therefore considered this approach as more “objective” because their professional judgement was focused on expectations of performance for each item. Panellists who favoured the whole booklet classification method felt that (a) it better accommodated the multiple factors that could influence student achievement, and (b) was more authentic because it was based on real students’ responses.

An analysis of the journal comments revealed some additional reasons for panellists placing high value on the Angoff, despite its cognitive complexity. It seemed that in many cases, panellists regarded judgements on the whole booklet classification method as simply too subjective. Some panellists also reported that it was very difficult to find scripts within a folder that matched the entire ranges of performances; although the scripts in every folder were always graded by percentile scores. In addition, having to make judgements across 12 different performance levels often proved difficult, as noted in the following comment:

The [whole booklet classification] method for standard setting seems to me to be more difficult to assess. The range between the descriptors is hard to decipher. For example what makes the difference between a LEVEL 1-M and a LEVEL 1-H? I did the exercise of evaluating the test

scripts (Whole Booklet Method) but I wasn't comfortable with it.
[Panellist-Questionnaire]

Another issue that contributed to the dislike of the whole booklet classification method was the difficulty in making compensatory holistic judgements, a problem that occurred especially in the Language Arts paper. This was because the tasks in this paper varied greatly in the nature of the assessment tasks and length of responses. Thus, differences in performance were not easily resolved, with some students writing excellent essays but doing poorly on some shorter prompts. For example, in a journal entry, one panellist explained her difficulty in making compensatory judgements:

The [whole booklet classification] method is too subjective. . . .in any particular booklet the student showed strength in one area and weaknesses in others and this had to be weighed [against each other in order] to make a good judgment. For e.g., a pupil [might solve] high order questions in problem solving, but showing no performance with basic number or measurement etc. This may lead to bias where students may make the same total score but because of where they showed their strength they are rated according to certain considerations. [Panellist-Journal Entry]

Discussion

This study documented and evaluated the design and reporting of performance standards for national assessments within the primary school system of Trinidad and Tobago. It was argued that one advantage of standards-referenced measurement over capricious or norm-referenced systems was the ability to provide a direct answer to the question of “*how good is good enough?*” This was achieved by the reporting of performance standards, which categorize student achievement into multiple levels based on the judgements of informed panellists. Information on satisfactory and unsatisfactory performances provided in the standards allows a judgement in terms of whether those standards were met across schools and educational districts. This kind of information is necessary for identifying and resolving patterns of inequality. Identification of achievement gaps is the first step in developing a system policy that ensures that all students have equal

opportunities to learn. Such interventions are especially important in Latin America and the Caribbean because inequality remains a pervasive problem despite the expansion in schooling (Perry, Arias, Lopez, Maloney, & Serven, 2006). Perhaps this issue has been neglected in the past because of the lack of high-quality monitoring data.

The quantitative data revealed that the majority of panellists were very confident about the performance standards constructed. They believed that they had successfully performed most of the key tasks. As to the quality of the standards, while respondents were generally positive, some believed that standards at the extreme levels (mastery and remedial) were superior to the levels in the middle (proficient and borderline proficient). During the process of making judgements, panellists relied on a variety of factors, but were mostly guided by written descriptors and their own classroom experiences. This suggests that an important qualification for judges should be the length and variety of experience in the classroom. Data from the open-ended instrument and the journals suggested that some panellists were negatively affected by the limited time for training. Others were concerned about the organization of the process and the increase in cognitive challenge presented. These are significant issues, which impinge on procedural validity and call for better management and organization coupled with increased training in the future. With hindsight, an integrated design, although cost saving, might not be the best approach in this context, considering that most panellists had little standard-setting experience.

As expected, no single method perfectly captured all the nuances of the standard. Thus, the synthesis procedure was an important aspect of the design. However, the DERE might now want to experiment with alternative approaches such as item mapping, the bookmark, or the analytical judgement method (Cizek & Bunch, 2007). Even if a new method is implemented, group discussion must be retained, as it appears critical to the quality of the decision-making process (Hurtz & Auerbach, 2003). This study adds to the international debate over the utility of different methods of standard setting (Hambleton et al., 2000; Plake & Impara, 2001; Zieky, 1997). While most respondents regarded the Angoff as the most difficult method, they also believed that it produced “better” cutscores. One reason may have been that they perceived the Angoff to be less subjective than the whole booklet classification and

contrasting group methods. It might be that panellists did not value methods that relied solely upon human judgement of work samples, or they might have regarded the broad compensatory judgements as unreliable, difficult, or impossible. Improving the process might be a costly exercise, valuable only if costs can be recovered through effective data use.

Of course, improving standard setting must never come at the expense of improving the quality of the test itself (Downing & Haladyna, 2006). Improving the test requires research into test development, with ongoing validation and alignment studies (Bhola, Impara, & Buckendahl, 2003; Haladyna, 2006). It was of concern, therefore, that a few panellists expressed concern over some items and aspects of test scoring. Technical legitimacy of the test is the foundation upon which all other processes and policies are built. Therefore, performance standards can only be valid if the entire assessment system provides scores that allow useful and meaningful inferences. This does not mean, however, that implementing performance standards now is an entirely futile exercise because validity is not an all or nothing characteristic. What is required is a system of continuous improvement at all test development stages (Greaney, 1996). Perhaps, then, the greatest weakness of the current system is that while there is national testing, a national educational quality evaluation system has not been developed (Gvirtz & Larripa, 2004; Olivares, 1996; Wolff, 2004). Thus, to date, the focus has been solely on collecting and reporting data, and not on using this information to diagnose weaknesses, evaluate education reforms, and create systemic changes. This uninformed use of data increases the potential for error and misuse of the tests. Indeed, policy can be constructed that allows a test to become a weapon in pursuit of some vague political, technocratic, or social agenda (Crespo, Soares, & de Mello e Souza, 2000; Heubert & Hauser, 1999).

The Way Forward: Towards Better Use of the Data

Little can be done in the absence of an overall policy on assessment and evaluation. This requires that the MOE constitute a broad-based technical committee to develop a framework for assessment and data use within a national educational quality evaluation system (Greaney & Kellaghan, 1996). This policy might specify exactly the role and purpose of different

assessments, specifying ways in which data are to be used. The policy should also specify structures, resources, and training to be put in place, including teacher training and public education programmes designed to improve assessment and data for all stakeholders. Turning towards the role of national assessments of education achievement, the critical question that emerges is, “*What policies are to be put in place to make best use of national test data?*” The answer to this question lies in examining current practice within Latin America and becoming an active borrower of international assessment strategy (Sebatane, 2000). The evaluation systems of Latin America provide the best guide for Caribbean national assessment systems because these countries share similar education problems, including the plague of persistent inequality (Ferrer, 2006; Winkler, 2000). Interestingly, Chile’s system is older than that of the US (Greaney & Kellaghan, 2007).

Still, there is also much to be learnt about appropriate data use from accountability systems in the more developed countries, such as the US and the UK (Herman & Haertel, 2005). On the positive side, one lesson is that while data might be used to ensure that schools and teachers are held accountable for student learning, publicly disseminated rankings of schools based on weak or inappropriate criteria will do more damage than good (Heubert & Hauser, 1999). Therefore, there are important choices to be made between competing agendas when developing data use policy. The most notable choices relate to decisions about the nature of the test, use of data, and administration of national assessments. The first decision relates to nature of the test, with the choice between high stakes, summative or low stakes, formative. The second decision relates to policy, with a choice between accountability and compensatory emphases. The third decision relates to authority and control, with the choice between centralization versus decentralized data use. These three decisions relate to the core function of national assessments, which is to bring to the notice of policy makers and the public the need for more effective education, to justify the reallocation of discretionary resources with more efficient resource allocation, and to ensure management efficiency (Kellaghan & Greaney, 2004).

Test Nature: High Stakes, Summative vs. Low Stakes, Formative

Some consequences are always attached to any high-stakes test, and, in theory, this might be a reward or sanction. In contrast, a low-stakes test is primarily used to provide information on the system (McDonnell, 2005). However, a national assessment system might also be considered low stakes if the information is used primarily to help low-performing schools improve (Ravela, 2005). Admittedly, there is tremendous pressure on administrators to increase the stakes associated with national assessments (Heubert & Hauser, 1999). However, it is more than likely that a high stakes national assessment will impact negatively upon schools or students. In the case of students, these negative consequences might include grade retention or placement in remedial tracks, whereas for schools and teachers, embarrassment or sanctions might be the possible outcomes. An increase in test stakes will also occur if the focus is to ensure that teachers teach certain topics. Thus, while washback can be positive, the evidence suggests that the impact of any high-stakes assessment is more likely to be negative (Broadfoot, 2002). There is the likelihood, even in the case of national assessments, that teachers will simply teach to the test (Firestone, Schorr, & Monfils, 2004).

It is possible, then, that the associated prestige of reporting high-achieving students in a high-stakes system might lead some schools to neglect low-achieving students. In the primary school, this will impede organizational arrangements designed to ensure early remediation. Nevertheless, a low-stakes programme by itself will not lead to improved student learning instead, unless the policy encourages data use for formative purposes. This approach implies a focus on assisting the teaching and learning process. There are numerous examples of such practice within the Latin American region (Ferrer, 2006). For example, Ravela (2005) documented the low-stakes, formative approach in Uruguay, where results from national achievement tests are used to inform and direct in-service teacher training programmes linked to performance in the examinations and the overall social context.

Use of Data Policy: Accountability vs. Compensatory

The formative approach to national assessment data use is evident in compensatory education policies, which are designed to help

underperforming schools succeed. In Latin America, best practice is found in Chile, Uruguay, and Mexico. There is a dire need for such an approach throughout the region because inequality remains a significant educational and social issue. Contrary to what some might think, educational inequality is also a critical issue in Trinidad and Tobago (Perry et al, 2006; World Bank, 1995). Not only is the link between education and poverty notable, but if the most disadvantaged students cannot receive a high-quality education, that link becomes unbreakable (Perry et al.). Indeed, without quality education, disadvantaged students cannot escape the poverty that characterizes many disadvantaged communities (Shapiro & Trevino, 2004). A compensatory education policy provides an efficient mechanism for directing scarce resources to schools and communities in need of help. Winkler (2000) has classified the different compensatory systems in Latin America using type of intervention and targeting mechanism. Interventions may be either supply or demand and targeting mechanism may be geographic, group, or self.

In Chile, the P900 is a supply, group targeting education policy that facilitates the direct link between national achievement test scores and compensatory funding in the form of materials and technical assistance channelled to low-performing schools. Data from national assessments are one of the starting points of a system diagnosis leading to a comprehensive strategy for improving the performance of low-achieving schools. Notable interventions include teacher professional development, targeted support for students at-risk, pedagogic counselling and guidance, and distribution of educational materials. Mexico also has a supply side geographic targeting compensatory education policy called the Program to Abate Educational Lag [PARE] (*Programa Para Abatir el Rezago Educativo*) operated by the National Council of Education Promotion [CONAFE] (*Consejo Nacional de Fomento Educativo*). The programme is designed to provide extra resources to schools that enrol disadvantaged students (Shapiro & Trevino, 2004). Compared to the P900 programme in Chile, this policy places a greater focus on curriculum materials and local decision making, which might increase the cost-effectiveness and impact of the programme.

Locally, it is common for schools to blame low student performance on factors outside the control of the school, with the pretence that

institutions can do little to change what the home has constructed. Thus, one might argue that in Trinidad and Tobago any compensatory policy must work hand in hand with increased accountability. To be sure, systemic results-driven accountability systems have become a worldwide trend (Anderson, 2005). These are built upon objectives, assessments, instructions, resources, and rewards or sanctions, with the general premise that educators are held accountable for student learning (McDonnell, 2005). In theory, a workable accountability system will be structured to transform schools, teachers, and classroom environments, and under such a policy, student learning failures will be attributed to weaknesses in programmes and practices. Ultimately, though, even under a compensatory policy, school personnel must understand that they are *professionally accountable* for student learning. Such accountability, then, becomes a moral endeavour rather than a legal or policy requirement; it is a philosophy of action captured explicitly in the mission and goals of empowered schools and districts.

Data use in compensatory and accountability systems represents a significant conflict in philosophical and political values (Benveniste, 2002; McDonnell, 2005). Compared to compensatory policies, an accountability system, with public reporting of results and sanctions, will increase the stakes associated with national assessments. As in Chile, an accountability system might also work hand in hand with quasi-market school systems to encourage competition and choice between schools (Benveniste). Certainly, public ranking of schools, as already practised by some local educational divisions, might lead to stakeholders demanding impossible quick fixes. In turn, this might make some schools, especially those in exceptionally challenging and difficult circumstances, much more vulnerable (Muijs, Harris, Chapman, Stoll, & Russ 2004). These are rather unpleasant side effects, which suggest that the best approach at this time might be to emphasize the formative use of assessment data. Such an approach synchronizes with the State's current commitment to poverty reduction and enhanced educational opportunity.

Administration: Centralized Control vs. Empowering Schools and Teachers

In Trinidad and Tobago, national assessments are currently administered by the DERE. As discussed by Greaney & Kellaghan (1996), there are

disadvantages to employing an agency within the MOE to manage a national assessment system. One of these disadvantages within a plural society such as Trinidad and Tobago is the difficulty in fully disaggregating data. Failure to disaggregate data fully means that some achievement gaps remain unidentified. Perhaps a more significant problem, however, is the persistent and dysfunctional focus of school communities on the MOE, pointing to a general lack of ownership and powerlessness. Harvey (1981) referred to such a pervasive sense of powerlessness and argued that it might be exemplified in the tendency to personify and scapegoat the “Ministry” for all the system’s ills. These weaknesses may be attributed to the governance structure for education in Trinidad and Tobago and the tendency to centralize key educational processes, including assessment and evaluation. Existing arrangements for national assessments only serve to confirm and validate the role and importance of the centralized structure.

Therefore, such a strongly centralized model might lead to further powerlessness and lack of ownership, further limiting effective data use by schools and teachers. It is notable that in Mexico, the PARE compensatory programme includes a separate school-based management policy called Support to School Management (Apoyo a la Gestión Escolar) [AGE]. This programme includes monetary support and training to parent associations, with consequent involvement of parents in decision making relating to school processes. In the local context, the problem-solving and intervention capacity of school personnel must be enhanced, allowing them to use data to solve site issues of academic underachievement. This approach might include strategies to improve assessment literacy among administrators and teachers. Enacted policy should also provide training and resources for contextualized decision making and problem solving (Earl & Katz, 2006). Such approaches will lead to the empowerment of schools and communities and will lessen the focus on “the Ministry” as the sole source and solution of every local problem.

The current pattern of results from national and international assessments suggests a very wide disparity between rural and urban educational districts, private and public schools, selected public schools, and boys and girls. A recent study also suggests that schools facing difficult circumstances do much less well (De Lisle et al., 2007). The central

administrative office of the MOE certainly does not have the capability to redress the weaknesses of every individual school or district. In any case, even if it had the needed human and physical resources, it would not want to intervene in such a way, because such over-involvement might lead to further disempowerment. Greater collaboration is certainly necessary between different Divisions of the MOE. Certainly, planning for systemic reform requires joint planning by the DCD, the Student Support Services, the Division for School Supervision, and the DERE. The requirement for collaborative work is a major hurdle, but without it, compensatory education reforms are difficult to implement.

Empowering teachers to use data better requires building assessment and data literacy. Currently, the low levels of teacher assessment literacy will lead to poor classroom assessment practice. Moreover, it portends possible data misuse. An example of this problem is the tendency for teachers to mimic large-scale pencil and paper tests rather than to use performance standards to guide and develop assessments that promote classroom learning. Therefore, while the introduction of performance standards might be a significant advance in the evolution of national assessment in Trinidad and Tobago, it will not by itself lead to systemic improvement in the education system. While performance standards are necessary for evaluating quality, only a comprehensive and enacted policy on data use will ensure systemic improvement. Therefore, the MOE must move towards developing a national evaluation system with clear policies for assessments and data use. Successful implementation of such a system would require cooperation of different arms of the MOE working together to tackle achievement gaps that now exist and those that might emerge (Porter, 2005).

Declaration of Competing Interests

The author is the standard-setting consultant for the primary school national assessments of educational achievement (2004–2007).

Acknowledgements

The author wishes to acknowledge the independent work of officers of the DERE in developing and implementing the national assessments of educational achievement. The officers have always been keen to provide feedback and are intimately involved in the management of the standard-setting exercise. The author wishes to single out Mr. Peter Smith, Mr. Meryvn Sambucharan, and Mr. Harrilal Seecharan for their assistance.

References

- Anderson, J. A. (2005). *Accountability in education* (Education Policy Studies). Paris: The International Institute for Educational Planning; and Brussels: International Academy of Education.
- Baird, J-A. (2007) Alternative conceptions of comparability. In P. Newton, J-A Baird, H. Goldstein, H. Patrick, & P. Tymms (Eds.), *Techniques for monitoring the comparability of examination standards* (pp. 124–165). London: Qualifications and Curriculum Authority.
- Benveniste, L. (2002). The political structuration of assessment: Negotiating state power and legitimacy. *Comparative Education Review*, 46(1), 89–118.
- Berk, R. A. (1986). A consumer's guide to setting performance standards on criterion-referenced tests. *Review of Educational Research*, 56(1), 137–172.
- Bhola, D., Impara, J., & Buckendahl, C. (2003). Aligning tests with states' content standards: Methods and issues. *Educational Measurement: Issues & Practice*, 22(3), 21–29.
- Braithwaite, R. H. E. (1981). Plus ca change. *Trinidad and Tobago Education Forum*, 2(1), 5–12.
- Brandon, P. R. (2004). Conclusions about frequently studied modified Angoff standard-setting topics. *Applied Measurement in Education*, 17(1), 59–88.
- Brandon, P. R. (2005). Using test standard-setting methods in educational program evaluation: Addressing the issue of how good is good enough. *Journal of Multidisciplinary Evaluation*, 2(3), 1–29.
- Braun, H. & Kanjee, A. (2006). Using assessment to improve education in developing nations. In H. Braun, A. Kanjee, E Bettinger, & M. Kremer (Eds.), *Improving education through assessment, innovation, and evaluation* (pp. 1–46). Cambridge, MA: American Academy of Arts and Sciences.
- Broadfoot, P. (2002). Beware the consequences of assessment! [Editorial]. *Assessment in Education: Principles, Policy and Practice*, 9(3): 285–288.
- Broadfoot, P., & Black, P. (2004). Redefining assessment? The first ten years of assessment in education. *Assessment in Education: Principles, Policy and Practice*, 11(1), 7–26.
- Carson, J. D. (2001). Legal issues in standard setting for licensure and certification. In G. J. Cizek (Ed.), *Setting performance standards; Concepts, methods, and perspectives* (pp. 427–444). Mahwah, NJ: Lawrence Erlbaum.
- Chapman, D. W., & Snyder, C. W. (2000). Can high stakes national testing improve instruction: Reexamining conventional wisdom. *International Journal of Educational Development*, 20(6), 457–474.
- Cizek, G. J. (1993). Reconsidering standards and criteria. *Journal of Educational Measurement*, 30(2), 93–106.

- Cizek, G. J. (1995, April). *Standard setting as psychometric due process: Going a little further down an uncertain road*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco, CA.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.
- Crespo, M., Soares, J. F., & de Mello e Souza, A. (2000). The Brazilian national evaluation system of basic education: Context, process and impact. *Studies in Educational Evaluation*, 26, 105–125.
- De Lisle, J. (2005). *Standard setting manual for the Trinidad and Tobago Ministry of Education's standard setting exercise for the 2005 national assessments*. Unpublished manuscript, St. Augustine, Trinidad: Centre for Medical Sciences Education, Faculty of Medical Sciences, UWI.
- De Lisle, J., Lewis, Y., Mc David, P., Smith, P., Keller, C., Jules, C., Lochan, S., & Seunariningsingh, K. (2007). *In the context of Trinidad and Tobago, how do we identify schools that are succeeding or failing amidst complex and challenging circumstances?* Paper presented at the Biennial Conference of the UWI Schools of Education: Reconceptualising the Agenda for Education in the Caribbean, 23–27 April, 2007 at the School of Education, Faculty of Humanities & Education, UWI, St. Augustine
- Downing, S. M., & Haladyna, T. M. (Eds.) (2006). *Handbook of test development*. Mahwah, NJ: Lawrence Erlbaum.
- Earl, L. M., & Katz, S. (2006). *Leading schools in a data-rich world: Harnessing data for school improvement*. Thousand Oaks, CA: Corwin.
- Eisemon, T. O. (1990). Examination policies performance to strengthen primary schooling in African countries. *International Journal of Educational Development*, 10, 69–88.
- Elley, W. B. (1992). *How in the world do students read? The IEA Study of Reading Literacy*. Hamburg: International Association for the Evaluation of Educational Achievement.
- Evans, H. (2001). *Inside Jamaican schools*. Mona, Jamaica: University of the West Indies Press.
- Ferrer, G. (2006). *Educational assessment systems in Latin America: Current practice and future challenges*. Washington, DC: Partnership for Educational Revitalization in the Americas.
- Firestone, W. A., Schorr, R. Y., & Monfils, L. (Eds.). (2004). *The ambiguity of teaching to the test*. Mahwah, NJ: Lawrence Erlbaum.
- Fiske, E. (2000). *Education for all: Status and trends 2000: Assessing learning achievement*. Paris: UNESCO.
- Giraud, G., Impara, J. C., & Plake, B. S. (2005). Teachers' conceptions of the target examinee in Angoff standard setting. *Applied Measurement in Education*, 18(3), 223–232.

- Gomez, J., & Padilla, J. L. (2004). The evaluation of consequences in standards-based test score interpretations. *Measurement: Interdisciplinary Research & Perspectives*, 2(2), 104–128.
- Grant, N. (2006). Teacher resistance in the United Kingdom. *Rethinking Schools Online*, 20(4). Retrieved February 21, 2006, from http://www.rethinkingschools.org/archive/20_04/uk204.shtml
- Greaney, V. (1996). Stages in national assessment. In P. Murphy, V. Greaney, M. E. Lockheed, & C. Rojas (Eds.), *National assessments: Testing the system* (pp. 111–128). Washington, DC: World Bank.
- Greaney, V., & Kellaghan, T. (1996). *Monitoring the learning outcomes of education systems*. Washington, DC: World Bank.
- Greaney, V., & Kellaghan, T. (2007). *Assessing national achievement levels in education*. Washington, DC: World Bank.
- Green, D. R., Trimble, C. Scott, & Lewis, D. M. (2003). Interpreting the results of three different standard-setting procedures. *Educational Measurement: Issues & Practice*, 22(1), 22–32.
- Greene, J. C. (2007). *Mixed methods in social inquiry*. San Francisco, CA: Jossey-Bass.
- Gvirtz, S., & Larripa, S. (2004). National evaluation system in Argentina: Problematic present and uncertain future. *Assessment in Education: Principles, Policy and Practice*, 11(3), 349–364.
- Haertel, E. H., & Lorie, W. A. (2004). Validating standards-based score interpretations. *Measurement: Interdisciplinary Research and Perspectives*, 2(2), 61–103.
- Haladyna, T. M. (2006). Roles and importance of validity studies in test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 739–758). Mahwah, NJ: Lawrence Erlbaum.
- Hambleton, R. (2001). Setting performance standards on educational assessments and criteria for evaluating the process. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 89–116). Mahwah, NJ: Lawrence Erlbaum.
- Hambleton, R. K., Brennan, R. L., Brown, W., Dodd, B., Forsyth, R. A., Mehrens, W. A., Nellhaus, J., Reckase, M., Rindone, D., van der Linden, W. J., & Zwick, R. (2000). A response to “Setting reasonable and useful performance standards” in the National Academy of Sciences’ “Grading the nation’s report card.” *Educational Measurement: Issues and Practice*, 19(2), 5–14.
- Hambleton, R. K., & Pitoniak, M. J. (2006). Setting performance standards. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 433–470). Washington, DC: American Council on Education.
- Hansche, L. N. (1998). *Handbook for the development of performance standards: Meeting the requirements of Title I*. Washington, DC: United States Department of Education.

- Hanushek, E. A., & Wobmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *Economic Journal*, 116(510), C63–C76.
- Harvey, C. (1981). *Practitioner's perceptions of an innovative school system in a developing country: A qualitative analysis*. Unpublished doctoral dissertation, University of Toronto, Canada.
- Hazel backs down, SEA exam here to stay. (2006, July 9). *Trinidad & Tobago Express*. Retrieved February 21, 2006, from http://www.trinidadexpress.com/index.pl/article_archive?id=160980190
- Herman, J. L., & Haertel, E. H. (Eds.). (2005). *Uses and misuses of data for educational accountability and improvement*. Chicago, IL: National Society for the Study of Education.
- Heubert, J. P., & Hauser, R. M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academies Press.
- Heyneman, S. P. (1987). Uses of examinations in developing countries: Selection, research, and education sector management. *International Journal of Educational Development*, 7(4), 251–263.
- Hiemstra, R. (2001). Uses and benefits of journal writing. In L. M. English & M. A. Gillen, (Eds.), *Promoting journal writing in adult education* (New Directions for Adult and Continuing Education, No. 90; pp. 19–26). San Francisco, CA: Jossey-Bass.
- Hurtz, G. M., & Auerbach, M. A. (2003). A meta-analysis of the effects of modifications to the Angoff method on cutoff scores and judgment consensus. *Educational and Psychological Measurement*, 63(4), 584–601.
- Jaeger, R. M., & Mills, C. (2001). An integrated judgment procedure for setting standards on complex, large-scale assessments. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 313–318). Mahwah, NJ: Lawrence Erlbaum.
- Jones, L. V. (2001). Assessing achievement versus high stakes testing: A crucial contrast. *Educational Assessment*, 7(1), 21–28.
- Kane, M. T. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research*, 64, 425–461.
- Kane, M. T. (1998). Choosing between examinee-centered and test-centered standard setting methods. *Educational Assessment*, 5(3), 129–145.
- Kane, M. T. (2001). So much remains the same: Conception and status of validation in setting standards. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 53–88). Mahwah, NJ: Lawrence Erlbaum.
- Kellaghan, T., & Greaney, V. (2001). The globalisation of assessment in the 20th century. *Assessment in Education*, 8(1), 87–102.
- Kellaghan, T., & Greaney, V. (2004). *Monitoring performance: Assessment and examinations in Africa*. Paris: Association for the Development of Education in Africa; Washington DC: World Bank.

- Lewis, S. (2005). Issues related to disaggregating data in a new accountability. In C. A. Dwyer (Ed.), *Measurement and research in the accountability era*. Mahwah, NJ: Lawrence Erlbaum.
- Linn, R. (2005). Issues in the design of accountability systems. In J. Herman & E. Haertel (Eds.), *Uses and misuses of data for educational accountability and improvement* (National Society for the Study of Education Yearbook, Vol. 104, Pt. 2: pp. 78–98). Chicago: NSSE.
- Manning, H. (2004). *Senate contribution on the 2003-2004 Appropriation Bill*. Retrieved February 21, 2006, from <http://www.ttparliament.org/hansard/senate/2004/hs20041025.pdf>
- Mayo, C. (2005) Testing resistance: Busno-cratic power, standardized tests, and care of self. *Educational Philosophy and Theory*, 37(3), 357–363.
- McDonnell, L. M. (2005). Assessment and accountability from the policymakers' perspective. In J. L. Herman & E. H. Haertel (Eds.), *Uses and misuses of data in accountability testing* (Yearbook of the National Society for the Study of Education, Vol. 104, Pt. 1; pp. 35–54). Boston, MA: Blackwell.
- McGinty, D. (2005). Illuminating the “black box” of standards setting: An exploratory qualitative study. *Applied Measurement in Education*, 18(3), 269–287.
- Meara, K. C., Hambleton, R. K., & Sireci, S. G. (2001). Setting and validating standards on licensure and certification exams: A survey of current practices. *CLEAR Exam Review*, 7(2), 17–23.
- Meyer, L. (2005). *A basic guide to standard setting*. London: AQA. Retrieved October 21, 2006, from <http://www.aqa.org.uk/over/pdf/guidetostandardsetting.pdf>
- Miyazaki, I. (1976). *China's examination hell: The civil service examinations of Imperial China*. New York: Weatherhill.
- Muijs, D., Harris, A., Chapman, C., Sroll, L., & Russ, J. (2004). Improving schools in socioeconomically disadvantaged areas – A review of research evidence. *School Effectiveness and School Improvement*, 15(2), 149–175.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 International Report: IEA's Progress in International Reading Literacy Study in Primary Schools in 40 Countries*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Nitko, A. J. (1980). Distinguishing the many varieties of criterion-referenced tests. *Review of Educational Research*, 50(3), 461–485.
- Oakland, T. (1995). Test use with children and youth internationally: Current status and future directions. In T. Oakland & R. Hambleton (Eds.), *International perspectives on academic assessment* (pp. 1–24). Boston, MA: Kluwer.

- Olivares, J. (1996). Inclusive national testing: Chile's 'quality of education assessment system.' In A. Little & A. Wolf (Eds), *Assessment in transition: Learning, monitoring and selection in international perspective* (pp. 118–133). Oxford, Pergamon.
- Payne, M. A., & Barker, D. (1986). Still preparing children for the 11+: Perceptions of parental behaviour in Barbados. *Educational Studies*, 12(3), 313–325.
- Perry, G. E., Arias, O. S., Lopez, J. H., Maloney, W. F., & Serven, L. (2006). *Poverty reduction and growth: Virtuous and vicious circles*. Washington, DC: World Bank.
- Plake, B. S. (2008). Standard setters: Stand up and take a stand! *Educational Measurement: Issues and Practice*, 27(1), 3–9.
- Plake, B. S., & Impara, J. C. (2001). Ability of panelists to estimate item performance for a target group of candidates: An issue in judgmental standard setting. *Educational Assessment*, 7(2), 87–97.
- Porter, A. C. (2005). Prospects for school reform and closing the achievement gap. In C. A. Dwyer (Ed.), *Measurement and research in the accountability era* (pp. 59–95). Mahway, NJ: Lawrence Erlbaum.
- Ravela, P. (2005) A formative approach to national assessments: The case of Uruguay. *Prospects*, 35(1), 21–43.
- Reckase, M. (2000). A survey and evaluation of recently developed procedures for setting standards on educational tests. In M. L. Bourque & S. Byrd (Eds.), *Student performance standards on the National Assessment of Educational Progress: Affirmation and improvements* (pp. 41–70). Washington, DC: National Assessment Governing Board.
- Schafer, W. D. (2005). Criteria for standard setting from the sponsor's perspective. *Applied Measurement in Education*, 18(1), 61–81.
- Sebatane, M. (2000). International transfers of assessment: Recent trends and strategies. *Assessment in Education: Principles, Policy and Practice*, 7(3), 401–416.
- Shapiro, J., & Trevino, J. M. (2004). *Compensatory education for disadvantaged Mexican students: An impact evaluation using propensity score matching* (World Bank Policy Research Working Paper No. 3334). Washington DC: World Bank.
- Skorupski, W., & Hambleton, R. K. (2005). What are panelists thinking when they participate in standard setting studies? *Applied Measurement in Education*, 18(3), 233–256.
- Tomlinson, M. (2002). *Inquiry into A Level standards – final report*. London: HMSO.
- Trinidad and Tobago Chamber of Industry and Commerce. (2006, July 13). Let's harness our nation's intelligentsia. *Business Guardian*, p. 29.
- Trinidad and Tobago. Ministry of Education. (2005). *SEA report: Secondary Entrance Assessment progress report*. Port of Spain, Trinidad: Author.

- Trinidad and Tobago. Ministry of Education. Division of Educational Research and Evaluation. (2005). *Proposal for standard setting: The 2005 national test*. Port of Spain, Trinidad: Author.
- Trinidad and Tobago. Ministry of Education. Division of Educational Research and Evaluation. (2006). *National test report, 2005*. Port of Spain, Trinidad: Author
- Wall, D. (2005). *The impact of high-stakes testing on classroom teaching: A case study using insights from testing and innovation theory* (Studies in Language Testing, volume 22). Cambridge, UK: Cambridge University Press.
- Wilbrink, Ben (1997). Assessment in historical perspective. *Studies in Educational Evaluation*, 23(1), 31–48.
- Winkler, D. (2000). Educating the poor in Latin America and the Caribbean: Examples of compensatory education. In F. Reimers (Ed.), *Unequal schools, unequal chances: The challenges to equal opportunity in the Americas* (pp. 113–132). Cambridge, MA: David Rockefeller Center for Latin American Studies, Harvard University.
- Wolff, L. (2004). Educational assessments in Latin America: The state of the art. *Applied Psychology: An International Review*, 53(2), 192–214.
- World Bank. (1993). *Caribbean region: Access, quality, and efficiency in education*. Washington, DC: Author.
- World Bank. (1995). *Trinidad and Tobago: Poverty and unemployment in an oil based economy* (Report No. 14382-TR). Washington, DC: Author.
- Zieky, M. (1997). Is the Angoff method really fundamentally flawed? *CLEAR Exam Review*, 8(2), 30–33.

**WHAT LEADERSHIP STYLES HAVE TO DO WITH
TEACHER JOB SATISFACTION:
A Review of a British Virgin Islands Study**

Jasmattie Yamraj with Bertram Hudson B. Ross

A review of a study of leadership styles and job satisfaction at a Caribbean high school is reported. Using an adapted LBDQ survey and a researcher-constructed job satisfaction questionnaire, a positive correlation was found between leadership style and the degree of job satisfaction ($R = .70$; $p < .01$). A stronger correlation existed between the consideration “leadership style” and teacher job satisfaction ($R = .70$) than between initiating structure and teacher job satisfaction ($R = .50$). The review explores methodological changes that might give better insights into leadership style and job satisfaction, especially because of the increasing complexity demanded of leaders in the emergent digital age. The method was found appropriate to obtain preliminary data but not the additional data required for triangulation and in-depth analyses that have a futuristic value. A conceptual matrix was recommended to improve data collection, analyses of data, and reporting.

Before and continuing into the 21st century, leadership and leadership styles, or attributes, or traits, or situations, have been the focus of research, debate, and speculation whether in education, nursing, business, the military, or politics. In this information or knowledge age, the styles of high-profile leaders are being documented every day in the electronic media. No longer are books written after leaders have retired or left the scene. Recently, books have been written about prominent public figures like the former Mayor of New York, Rudolph Guiliani; Defense Secretary, Donald Rumsfeld; and Senator Hillary Clinton, among others, as leadership experts try to unravel the variables that contribute to the success of modern leaders.

At a different pace but applying similar methods, the research about education leaders continues. In the fast-changing world, leaders of all disciplines need similar tools to succeed and survive and should consider

all research relevant to their field. Whatever the uncertainty that might still exist within the literature, it is crystal clear that leadership is critical to the success of organizations (Bennis & Nanus, 1985).

The leaders of tomorrow are in today's schools. It is a major imperative to ensure that those who make decisions about schools have access to research that enhances their capacity to make things work. Concerns about teacher quality and performance continue, as well as the need to improve teacher satisfaction if they are to remain in their jobs. Undoubtedly, effective leadership promotes student learning, teacher performance, collaboration among teachers, and teacher job satisfaction.

Studies about how teacher satisfaction and leadership styles are related have been completed in the United States (US) and elsewhere, but only one other study about teacher satisfaction and leadership styles has been found that has originated in the Caribbean region (see Buckley-Jones, 1988). Now, more than ever, Caribbean communities focus on the quality of education that students receive in schools. Yet, inadequate attention has been given to all of the factors that enhance teacher performance. Evidence exists that the satisfaction teachers derive from their work is related to the style of the school's leadership (Yamraj, 2003).

This paper reviews a study conducted by Yamraj (2003) at the British Virgin Islands High School (BVIHS) in order to determine the nature of the relationship between the leadership styles of the school's senior administration and teacher job satisfaction. The British Virgin Islands (BVI) is a group of approximately 50 islands, cays, and rocks located about 60 miles south east of Puerto Rico in the Caribbean. The BVI provides secondary schooling in three public schools. The largest, the BVIHS, is located on the main island of Tortola and administered by a principal and three assistant principals (senior administration).

The BVIHS, not unlike other schools in the Caribbean and worldwide, continues to recruit a considerable proportion of its teachers from outside its shores. An understanding of what promotes teacher satisfaction is important if education systems such as the BVI are to continue to attract and retain good teachers. Teachers leave the profession for a range of reasons such as increased accountability, increased paperwork, negative

and discouraging student attitudes, inadequate salary, teaching load, lack of discipline, lack of respect by students, non-teaching duties, stress, student behaviour, lack of parental support, and conflict with, or dissatisfaction with, the school's leadership and the lack of administrative support (Marlow, Inman, & Betancourt-Smith, 1996; Tye & O'Brien, 2002). The review of the study gives further insights that might minimize the leadership conundrum that persists in the Caribbean, and elsewhere.

Leadership Styles

“Scholarly attempts to analyze leadership have resulted in many diverse definitions, theories, models and applications” (Razik & Swanston, 2001, p. 60). Four dominant dimensions of leadership theories can be identified: trait, behaviour, contingency, and the transformational/transactional leader. Whereas trait studies focused on who the leaders were, behavioural studies focused on what leaders did; hence their leadership style. In order to investigate leadership style it is appropriate to look first to the research on leadership behaviour.

Leadership behaviour studies were conducted at Ohio State University, University of Michigan, and Harvard University (Hoy & Miskel, 1996), focusing on two dimensions of leadership style in each case: initiating structure and consideration at Ohio State University, production-oriented and employee-centred at University of Michigan, and task leader and social leader at Harvard University. In each study, the two opposing leadership styles are compared, either as a means of demonstrating the virtues of one over the other, or the trend of reversing the traditional dictatorial style to accommodate the emerging modern participative style typology.

For the purposes of the study, it was felt that the Ohio State University studies provided the ideal variables that mirrored each of the dichotomies of leadership styles that have been identified. The Ohio State Leadership Studies (Hoy & Miskel, 1996) examined two behavioural leadership styles: initiating (task) structure and consideration. This dichotomy of leadership styles provided the foundation for the leadership aspects of the study. Cunningham and Cordiero (2000) informed us that leaders who display initiating structure define the members' goals, assign tasks, plan

ahead, establish work methods, push for productivity, emphasize deadlines, and encourage adherence to procedures. On the other hand, the considerate leader stresses participation in decision making; encourages communication, staff development, and independent thinking; and provides timely feedback to the staff about the quality of their output.

Studies correlating leadership styles and job satisfaction have determined that the greater the scope for decision making involvement by faculty, the higher the level of job satisfaction (Kauffman, 1980, as cited in McKee, 1990; Lipham, Dunstan, & Rankin, 1981). Patrick (1995) concurred with Bidwell that teacher satisfaction is largely related to teachers' expectations of their administrator as well as their perception of their administrator's behaviour. Similarly, the attitude of the principal towards teachers and informal peer groups, and the amount of freedom that teachers are given to plan work, affect teacher satisfaction.

Job Satisfaction

Klecker and Loadman (1997) reported that interaction with the principal is one of the key variables related to job satisfaction. An ideal environment is fostered when the school's leadership (a) promotes a strong instructional climate, (b) efficiently manages the curriculum and defines the goals, and (c) interacts continually with teachers (Anderman, Belzer, & Smith, 1991). This is not dissimilar to the views expressed by Covey (2004), who promoted servant leadership and enabling others to find their voices as two key characteristics of great leaders. According to Jorde (1984), the quality of relationships between teachers and administrators, and the value and recognition that teachers receive from their colleagues affect teacher's self-reports of job satisfaction. Indeed, teachers report greater job satisfaction when they perceive the principal as someone who shares information with others, delegates, and keeps open channels of communication (Bogler, 1999). School climate, professional autonomy, and working conditions were additional school conditions deemed important to teacher satisfaction. In addition, intrinsic rewards play a significant role in teacher job satisfaction (Chung, 1970; Lantham, 1998; Lipham et al., 1981; Marlow et al., 1996).

Purpose of the Review

Overall, the review will show that the findings of Yamraj's study (Yamraj, 2003; Yamraj, Ross, & Greene, 2004) are in concurrence with the literature. But were there phenomena that were left unexamined due to the utilization of quantitative measures alone? The purpose of the review is to report the study, then to determine how the methodology might be revised in order to create a model for future studies. It is important to (a) determine whether there were variables that might have been obscured in the attempt to relate the leadership styles to job satisfaction, and (b) propose how the methodological framework might be adapted for future follow-up studies on a larger scale.

In recent studies about leaders in business and politics, extensive interviews (Guiliani, 2002; Welch & Byrne, 2001), observations, document data, and record data have been collected. Yamraj's (2003) study followed the trend of applying questionnaires to determine leadership style based on past performance of the leader. However, there is a need to project and predict future expectations of leaders (Drucker, 2002). Such projections require more than the data obtained by questionnaires alone. If leaders are to successfully survive in the ever-changing and more complex post-September 11 environment, their training should include modelling futuristic situations, based on extensive data analyses. Similarly, should job satisfaction be determined only by the analysis of a questionnaire on job satisfaction? With the Internet as a medium for advertising job vacancies and for the posting of information about most preferred places to work, more detail about job satisfaction is necessary.

In Yamraj's (2003) study, data were collected through questionnaires. Certainly, the data obtained by the questionnaires are valid and reliable. However, according to Yin (1993), an investigation that covers "both a particular phenomenon and the context within which the phenomenon is occurring" (p. 31) requires extensive qualitative data as well. The combination of quantitative and qualitative data has long been recommended. Greene (1994) proposes that "rather than maintaining the long-standing conflicts between the quantitative and qualitative, the objective and the purportedly subjective, educational researchers might turn to metaphor and imagination...to the end of recasting old opposition

and, perhaps to link theory and experience together in new and dynamic ways” (p. 457).

Used in mathematics, a matrix is a rectangular table of elements or entries that may be numbers or any abstract quantities that can be added or multiplied. The variables used for correlation analysis can be subjected to more rigorous scrutiny and thorough conceptualization if all the related and interrelated concepts can be identified and processed to provide additional rich and relevant data. Accordingly, a conceptual matrix (Peterson & Bean, 1998; Ross, 1995) might be appropriate to guide further data collection and analysis. In the conceptual matrix, qualitative data either replace or add to the quantitative data found in a traditional matrix. Qualitative data may be obtained through interviews, direct observations, and by examination of documents and materials. Ross (1995) used a conceptual matrix to link two different case studies used to determine how the US community college system might be transferable to the Organisation of Eastern Caribbean States. Peterson and Bean used a conceptual matrix to organize a course in Economics to “improve student understanding....it focuses on meanings, provides a framework for processing information, promotes critical thinking, and provides problematic topics for writing and discussion” (p. 262).

The review of the study is opportune because only one school was involved and both questionnaires had a high level of reliability and validity despite the small sample (n=35). The review will be guided by three questions:

1. Are additional data required to inform characteristics of leadership style?
2. Are additional data required to inform teacher job satisfaction?
3. What design will facilitate data collection for large-scale studies?

The Study

Method and Instrumentation

Two Likert-type surveys were prepared before selecting the sample of BVIHS teachers for the study. In the first, teachers were to respond to a 40-item researcher-constructed Job Satisfaction Questionnaire about the degree of their job satisfaction. For the second, a 30-item Leadership

Styles Questionnaire (Bryant, 2002), adapted from Halpin's (1957) 40-item version of the Leadership Behavior Description Questionnaire (LBDQ), they were asked for perceptions of the leadership styles of the senior administration.

A pilot study was conducted with selected senior faculty to elicit information concerning the validity of the items on both instruments. Subsequently, a few items were deleted or added before the modified instruments were distributed to the sample (n=35).

The 10 dimensions of Job Satisfaction were: personal information, professional autonomy, supervision, policy and administration, working conditions, salary, work, professional development and growth, interpersonal relations, and recognition/responsibility. The Leadership Styles questionnaire was divided into two dimensions: consideration and initiating (task) structure. Consideration leadership style included leader behaviour that demonstrated friendship, trust, warmth, and respect in the relationship of leader and members of the work group. Initiating structure leadership style included behaviours that described the relationship between the leader and the subordinates, while establishing definite patterns of organization, channels of communication, and procedures (Halprin, 1966, as cited in Hoy & Miskel, 1996). The questionnaire was constructed to produce two scores—one for consideration and the other for the initiating structure (task) leadership style.

Population and sample. The population of 106 teachers who had worked at the BVIHS for at least one academic year included full-timers, part-timers, trainees, department heads, and year heads. The sample was procured through a combination of simple random and systematic sampling.

Data analysis. The data were analysed using the Pearson correlation moment and a multiple linear regression technique to determine the nature and strength of the relationship between leadership styles of the senior administration and teacher job satisfaction.

Results

Job satisfaction. Sixty percent of the teachers were either slightly dissatisfied or dissatisfied with their jobs. Of those who expressed satisfaction, none reported that they were very satisfied with the job.

Leadership style. The data revealed that both consideration and initiating structure leadership styles were displayed by the senior administration; however, initiating structure ($M=52.65$) was considerably higher than the consideration leadership style ($M=22.47$). Overall, teachers perceived that the senior administration displayed more initiating (task) structure than consideration.

Correlation. The multiple linear regression analysis produced a correlation coefficient (R) of .70, indicating a strong and positive relationship between leadership styles and teacher job satisfaction at the BVIHS, with consideration leadership style contributing more to total teacher job satisfaction.

The results in Table 1 indicate a positive correlation between leadership styles and the dimensions of job satisfaction. A stronger correlation existed between the consideration leadership style and teacher job satisfaction ($R = .70$) than initiating structure and teacher job satisfaction ($R = .50$). There was a positive relationship between the consideration and initiating structure subscales ($R = .65$) of leadership styles, confirming both leadership styles were exhibited by the leadership at the BVIHS.

Table 2 arrays the correlation coefficients to indicate the degree of the relationship, and the coefficients of determination to measure the strength of the relationship between each leadership style and each of the dimensions that contribute to total teacher job satisfaction.

The results point to the consideration leadership style as the stronger determinant in 70% of the dimensions of teacher job satisfaction, and total job satisfaction had a strong positive relationship with consideration leadership style ($R = .70$). The coefficient of determination (R^2) was .49, indicating that 49% of teacher job satisfaction could be explained by the

aspects of consideration leadership style displayed by the senior administration at BVIHS.

Table 1. Correlation Between Variables (n=35)

Statistical Tool	Variables	Dependent Variable	Independent Variables	
		Job Satisfaction	Consideration	Initiating Structure
Pearson Correlation	Job Satisfaction	1.00	.70	.50
	Consideration	.70	1.00	.65
	Initiating Structure	.50	.65	1.00

$p < .05$

Table 2. Correlation (R) and Coefficient of Determination (R^2) Between Leadership Styles and Dimensions of Job Satisfaction (n=35)

Job Satisfaction	Leadership Styles			
	Consideration		Initiating Structure	
	R	R^2	R	R^2
Personal	.48**	.23**	.07	.00
Professional Autonomy	.55**	.30**	.53**	.28**
Supervision	.60**	.36**	.38*	.14*
Policy and Administration	.59**	.35**	.43*	.18*
Working Conditions	.53**	.28**	.37*	.14*
Salary	.43*	.19*	.46**	.21**
Work	.46**	.21**	.49**	.24**
Professional Dev. and Growth	.23	.05	.24	.06
Interpersonal Relations	.71**	.50**	.42*	.18*
Recognition/Responsibility	.62**	.38**	.43*	.19*
Total	.70**	.49**	.50**	.25**

* $p < .05$; ** $p < .01$

The highest correlation between the dimension of teacher job satisfaction and consideration leadership style occurred with interpersonal relations ($R = .71$). The lowest correlation ($R = .23$) was between professional development and growth and consideration leadership style. The remaining dimensions of job satisfaction yielded values that ranged from .49 to .62. All other correlations between the combined leadership styles and the different dimensions of job satisfaction were significant at a 95% confidence level.

Only 25% of teacher job satisfaction could be attributed to initiating structure leadership style. Salary, work, and professional development and growth were the three dimensions in which the correlation between initiating structure and job satisfaction was higher than consideration.

Both consideration and initiating structure leadership styles produced positive relationships between leadership styles and teacher job satisfaction. However, the correlation between consideration leadership style and teacher job satisfaction was higher than with initiating structure leadership style.

The teachers at the BVIHS reported satisfaction with 30% of the dimensions of job satisfaction: personal, interpersonal relations, and supervision aspects of their jobs. Most of the respondents were dissatisfied with their jobs at the BVIHS because of the low levels of professional autonomy and limited input into decisions about the curriculum. As well, they were also dissatisfied with the working conditions, salary, responsibility and recognition, and the work itself. The greatest amount of dissatisfaction occurred with the policy and administration dimension of job satisfaction, as well as opportunities for professional development and growth.

Conclusion

The statistically significant ($p < .01$) relationship between the leadership styles of the senior administration and teacher job satisfaction, and the high correlation coefficient for consideration style of leadership ($R = .70$) point to the advantages of the consideration leadership style in maintaining high teacher satisfaction. The initiating structure was the preferred leadership style of the BVIHS senior administration. Since the

leaders displayed consideration there was some evidence of warmth and trust given to the teachers at the BVIHS.

Review

The findings were not dissimilar to other studies documented in the literature. However, the literature provides examples of leaders who displayed high performance by displaying both consideration and initiating structure, thereby promoting harmony within groups and satisfied teachers. Due to the predominance of the initiating structure leadership style, more than half of the teachers in the study under review were dissatisfied with their jobs at the BVIHS.

The theoretical implication is that two leadership styles may be investigated in future studies in other Caribbean schools to determine teacher job satisfaction. The results would provide administrators with valuable information about teachers' expectations about their jobs, school, and the working conditions.

The practical implications are twofold. First, the senior administration (principal and assistant principals) should have had prior access to research about the importance of teacher job satisfaction and the variables that predict it. Second, leaders need to continually reflect on their leadership styles and decide under what conditions they remain effective.

1. Are additional data required to inform characteristics of leadership style?

According to one expert: "what a leader stands for is more important than what he or she does" (Sergiovanni, 1984, p. 106). This raises the concern about whether we can know what the leader stands for based only on the questionnaires given to teachers. The LBDQ instrument has been shown to be useful in the collection of data to determine leadership style. Other models, including the Fiedler's contingency and House's goal-path theory, among others, have been reviewed by Covey (2004) and Hoy and Miskel (1996), and shown to be similar in approach as they examine the dichotomy of leadership attributes.

However, the world has changed since the LBDQ has made its mark. Leaders are now placed under the microscope more frequently. With the requirement for increased accountability, the advent of the information age, and the rapid development of digital technology, cellular phones, and the Internet, among others, information is easily circulated about organizational difficulties, improprieties, and anomalies to the competition and the press. So, too, leadership positions have become more tenuous and complex with change often coming swiftly and suddenly, whether in business, politics, or education. Leaders are required to be multifaceted and make decisions about many issues over short time spans.

The changing of the leadership landscape has been compared to the replacement of the analog systems with digital systems. For example, the DVD and the CD have replaced the cassette and LP as the preferred music storage systems; the calculator has replaced the slide rule; e-mail has replaced the typewritten circular, and the teleconference has replaced the urgently scheduled meeting. In addition, the globalization of products; the intense competition for resources, including human resources; the 9/11 scenario; and the tsunami event have all produced a demand for a new type of leader, who Fisk (2002) calls a digital leader. This new leadership style is not too surprising when the question has been raised whether human life is digital or analog.

Our review reveals the inadequacy of the initiating structure-consideration dichotomy to collect data about digital leaders. Data about modern leaders are collected on a case-by-case basis, using in-depth interviews, examination of written records, and direct observations. For example, in a focus on the War General, Colin Powell, Harari (2002) scoured documents with speeches and other data, and drew information from an autobiography. So, too, for the study of the *World's Greatest CEO*, Jack Welch of General Electric, data was obtained from document research, coupled with observations made by colleagues, and triangulation of the data with interviews (Slater, 2003). In a similar way, Krames (2002) conducted an analysis of Rumsfeld, a politician and Secretary of Defense of the US.

Neff and Citrin (2005) documented cases about leaders and their first 100 days, and Spears (1998) used cases to explore issues of servant

leadership, among others. Marcinko (1996), a commando, provided insights into his leadership secrets by detailing how he ensured success on the battlefield by leading by example.

In the study under review, no attempt was made to directly collect data from the school's leadership. Although only one school was involved, the respondents provided information that related only to statements about the initiating structure and consideration leadership styles. No accommodation was made for digital leaders (Fisk, 2002).

Among the tenets of digital leadership espoused by Fisk (2002) are:

- an instinctive ability to make sense of increasingly complex markets;
- harnessing the potential of technology inside their organizations;
- constantly transforming themselves and their organizations;
- reconfiguring themselves and their networks to create competitive advantage for their customers and shareholders.

A thematic review of the research about digital leadership would guide the construction of questions that would drive interviews, observations, the examination of records, and their data analyses.

2. Are additional data required to inform teacher job satisfaction?

There are two challenges to the efficacy of the job satisfaction questionnaire. The first is that the linkage of job satisfaction to job performance has not always been supported by empirical evidence. Blumberg and Pringle (1982) cited Greene and Craft (1979) who alluded to the uncertainty about the causal nature of job performance. However, we continue with the premise of the linkage because it is overwhelmingly supported by the research.

The second challenge is that job satisfaction will cause teachers to remain in their jobs. There exists evidence to the contrary. Drucker (2002) provided evidence that more employees will continue to leave and change the organizations they work for with temporaries emerging as the way of the future.

Notwithstanding the two concerns, the 10 dimensions of job satisfaction remain relevant into the digital age. The 10 dimensions of Job Satisfaction were: personal information, professional autonomy, supervision, policy and administration, working conditions, salary, work, professional development and growth, interpersonal relations, and recognition/responsibility. Even if teacher performance did not hinge heavily on job satisfaction, the study is about job satisfaction not performance or retention, so the 10 dimensions remain relevant. There is no reason to change the job satisfaction dimensions.

3. *What design will facilitate data collection for large-scale studies?*

The inclusion of interview, observation, and document data requires longer preparation and a larger budget for data collection, analysis, and subsequent reporting. However, the efficient use of technology should facilitate the timely analysis of quantitative data or qualitative data.

A conceptual matrix (Table 3) will guide the planning for the collection, analysis, and synthesis of the data (Ross, 1995). A matrix that includes the 3 dimensions of leadership and the 10 dimensions of job satisfaction is proposed to guide the collection of both quantitative and qualitative data. Wedermeyer (1973) used a conceptual matrix to collect qualitative data because “each of the realities is, in fact a variable, because in no specific place or region where such a program is established will the situation be exactly the same” (p. 7). The matrix projects the dimensions of each variable and shows the simple and complex ways in which they interact with each other.

Research will reveal whether there is overlap among the three leadership styles, or if there is a continuum. So, too, some of the dimensions of job satisfaction might be eliminated, if the change from the traditional workforce to the knowledge workforce is accompanied by the change from an employee serving the system, to the system serving the worker (Drucker, 2002).

In either case, the research is better utilized if it can predict how leaders ought to respond to new and ever-changing situations. Models and simulations must be used for leaders to project how organizations might

respond to new situations in the same way young surgeons today learn new surgical techniques using virtual reality without endangering the lives of their patients.

According to Drucker (2002), “a CEO’s job will be much more like the most complex job I know, which is running an opera” (p. 90). Further research should be conducted by schools at primary and secondary level, in the BVI and other Caribbean islands, to determine the effectiveness of the leaders and to institute reforms that would positively impact teacher job satisfaction. This imperative is urgent since Caribbean islands continue to have high teacher turnover rates, and in many cases there are shortages of qualified teachers.

Table 3. Conceptual Matrix for Leadership Styles and Job Satisfaction

Job Satisfaction	Leadership Styles		
	Initiating Structure	Consideration	Digital
Professional Autonomy			
Supervision			
Policy and Administration			
Working Conditions			
Salary			
Work			
Professional Development and Growth			
Interpersonal Relations			
Recognition/Responsibility			

References

- Anderman, E. M., Belzer, S., & Smith, J. (1991, April). *Teacher commitment and job satisfaction: The role of school culture and principal leadership*. Paper presented at the meeting of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service No. ED375497)
- Bennis, W., & Nanus, B. (1985). *Leaders: the strategies for taking charge*. New York: Harper & Row.
- Blumberg, M., & Pringle, C. D. (1982). The missing opportunity in organizational research: Some implications for a theory of work performance. *The Academy of Management Review*, 7(4), 560–569.
- Bogler, R. (1999, April). *Reassessing the behavior of principals as a multiple-factor in teachers' job satisfaction*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada. (ERIC Document Reproduction Service No. ED432800)
- Bryant, M. (2002). *Leader Behavior Description Questionnaire*. Retrieved October 2, 2002, from <http://tc.unl.edu/mbryant/LeaderQuestionnaire.htm>
- Buckley-Jones, B. (1988). *Principals' leadership style, its expression and effect on teachers' job satisfaction*. Unpublished doctoral dissertation, The University of the West Indies, Mona.
- Chung, K. (1970, March). *Teacher-centered management style of public school principals and job satisfaction of teachers*. Paper presented at the Annual Meeting of the American Educational Research Association, Minneapolis, MN. (ERIC Document Reproduction Service No. ED042259)
- Covey, S. (2004). *The 8th habit: From effectiveness to greatness*. New York: Free Press.
- Cunningham, W. G., & Cordeiro, P. A. (2000). *Educational administration: A problem-based approach*. Boston, MA: Allyn & Bacon.
- Drucker, P. F. (2002). *Managing in the next society*. New York: St. Martin's.
- Fisk, P. (2002). The making of a digital leader. *Business Strategy Review*, 13(1), 43–50. Retrieved March 3, 2006, from http://www.paconsulting.com/news/by_pa/2002/by_pa_200205130.htm
- Giuliani, R. W. (2002). *Leadership*. New York: Hyperion.
- Greene, M. (1994). Epistemology and educational research: The influence of recent approaches to knowledge. In L. Darling-Hammond (Ed.), *Review of Research in Education*, 20, 423–464. (Washington, DC: American Education Research Association)
- Harari, O. (2002). *The leadership secrets of Colin Powell*. New York: McGraw Hill.
- Hoy, W. K., & Miskel, C. G. (1996). *Educational administration: Theory, research, and practice* (5th ed.). New York: McGraw Hill.

- Jorde, P. (1984, November). *Teacher job satisfaction: A framework for analysis*. Paper presented at the Annual Meeting of the National Association for the Education of Young Children, Los Angeles, CA. (ERIC Document Reproduction Service No. ED253304)
- Klecker, B., & Loadman, W. E. (1997, October). *Exploring teacher job satisfaction, across years of teaching experience*. Paper presented at the Annual Meeting of the Mid-Western Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service No. ED413316)
- Krames, J. A. (2002). *The Rumsfeld way: Leadership wisdom of a battle-hardened maverick*. New York: McGraw-Hill.
- Latham, A. (1998). Teacher satisfaction. *Educational Leadership*, 55(5), 82–84.
- Lipham, J., Dunstan, J., & Rankin, R. (1981). *The relationship of decision involvement and principals' leadership to teacher job satisfaction in selected secondary schools* (Report No. WRDCIS-TR-571). Madison, WI: Research and Development Center for Individualized Schooling. (ERIC Document Reproduction Service No. ED207129)
- Marcinko, R. (1996). *Leadership secrets of the rogue warrior: A commando's guide to success*. New York: Simon & Schuster.
- Marlow, L., Inman, D., & Betancourt-Smith, M. (1996). *Teacher job satisfaction*. (ERIC Document Reproduction Service No. ED393802)
- McKee, J. G. (1990, February). *Relationship between community college presidents' leadership styles and faculty job satisfaction*. Paper presented at the Annual Conference of the Eastern Educational Research Association, Clearwater Beach: FL. (ERIC Document Reproduction Service No. ED 316295)
- Neff, T. J., & Citrin, J. M. (2005). *You're in charge – Now what? The 8-point plan*. New York: Crown Business.
- Patrick, J. E. (1995). *Correlation between administrative style and school climate*. (ERIC Document Reproduction Service No. ED387853)
- Peterson, D., & Bean, J. B. (1998). Using a conceptual matrix to organize a course in the history of economic thought. *Journal of Economic Education*, 29(3), 262–273.
- Razik, T. A., & Swanson A. D. (2001). *Fundamental concepts of educational leadership* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Ross B. H. (1995). *The applicability of the American two-year college model to the Organization of Eastern Caribbean States*. Unpublished doctoral dissertation, Ohio University, Athens.
- Sergiovanni, T. J. (1984). Leadership as cultural expression. In T. J. Sergiovanni & J. E. Corbally (Eds.), *Leadership and organizational culture: New perspectives on administrative theory and practice* (pp. 105–114). Champaign, IL: University of Illinois Press.
- Slater, R. (2003). *29 Leadership secrets from Jack Welch*. New York: McGraw Hill.

- Spears, L. (1998). *Insights on leadership: Service, stewardship, spirit and servant leadership*. New York: Wiley.
- Tye, B. B., & O'Brien, L. (2002). Why are experienced teachers leaving the profession? *Phi Delta Kappan*, 84(1), 24–32.
- Welch, J., & Byrne, J. (2001). *Jack: Straight from the gut*. New York: Warner Books.
- Wedermeyer, C. A. (1973, June). *Institution building and evaluation*. Paper presented at Conference and Workshop on Correspondence Education, Nairobi, Kenya.
- Yamraj, J. (2003). *Leadership styles and job satisfaction: The case of the BVI High School*. Unpublished master's thesis, University of the Virgin Islands, St. Thomas.
- Yamraj, J., Ross, B., & Greene, S. (2004, April). *Leadership styles and job satisfaction: The case of the BVI High School*. Paper presented at the Annual Meeting of the American Education Research Association, San Diego, CA.
- Yin, R. K. (1993). *Applications of case study research*. Newbury Park, CA: Sage.

THE POTENTIALS OF EMBODIED AND ETHNOMATHEMATICAL PERSPECTIVES FOR CARIBBEAN MATHEMATICS EDUCATION

Steven Khan

In this paper, I explore the potential value of adopting embodied and ethnomathematical perspectives in the specific context of mathematical education in Trinidad and Tobago. I suggest that ideas representing overlaps among the domains of embodied cognition and ethnomathematics are manifested in the reform-oriented mathematics curriculum documents of the Trinidad and Tobago Secondary Education Modernization Programme (SEMP), and provide a simple example of one area where such an investigation might potentially begin.

Introduction

Something interesting is happening in mathematics education and the now-adolescent discipline of mathematics education research. My primary aim in this paper is to identify and explore two of these domains of inquiry that are contributing to creating an Archimedean lever, resting on the fulcrum that is mathematics (education), which have the potential to move the world of mathematics curriculum here in the Caribbean. My secondary aim is to describe what this might mean for a small twin-island Republic of less than 1.5 million inhabitants at the southernmost end of the Caribbean archipelago.

The first people of Trinidad and Tobago were the Amerindians, and the country was subsequently colonized by the Spanish and the British. In addition, other Europeans such as the French settlers arrived. Africans were brought as slaves and, later, East Indians as indentured labourers to serve on agricultural estates. Other groups such as Chinese and Syrians arrived over the course of our history as well. The mathematics curriculum of Trinidad and Tobago, like other states in the Caribbean, has until fairly recently continued to be dominated by a “Eurocentric” view of mathematics. This is exemplified in terminal examinations that

model the English school-leaving examinations, for which they are a replacement; mathematics textbooks that give little insight into the socio-historical development of mathematics and which offer no connections to the various histories that students bring with them to school; and pedagogy that is aimed primarily at helping students to pass examinations. In this situation, mathematics education divorces mathematics from its socio-historical roots, resulting in a homogenously unpalatable discipline. Such is the enduring legacy of colonialism. In this situation, Trinidad, like other small countries in an increasingly smaller world, is concerned about the mathematical competence of its population, given high failure rates in these terminal examinations and a pervasive attitude among the population that while mathematics is important it cannot be done by most people (Trinidad and Tobago. Ministry of Education. Secondary Education Modernization Programme [SEMP], 2002). Perspectives such as ethnomathematics and embodied mathematics offer a means of beginning to redress some of these issues.

In the sections below I describe these two perspectives in detail.

Embodied Mathematics

For Mercy has a human heart,
Pity a human face,
And Love, the human form divine,
And Peace, the human dress.

“The Divine Image,” William Blake, 1789.

According to Wilson (2002), the theoretical starting point of embodied perspectives is “not a mind working on abstract problems, but a body that requires a mind to make it function” (p. 625). Embodied perspectives seek to determine the extent to which sensorimotor processing is implicated in cognition and holds the view that human cognition is body-based. Anderson (2003) argues that it is specifically this concern for the physical grounding of cognition in a body that distinguishes embodied cognition from related situated perspectives.

Lakoff and Nunez (2000) describe what they consider to be a general romance of (Western) mathematics:

- Mathematics is abstract and disembodied – yet it is real.
- Mathematics has an objective existence...

- Human mathematics is just a part of abstract, transcendent mathematics.
- ...mathematical proof allows us to discover transcendent truths of the universe.
- Mathematics is part of the physical universe and provides rational structure to it.
- Mathematics even characterizes logic, and hence structures reason itself...
- To learn mathematics is therefore to learn the language of nature...
- Because mathematics is disembodied and reason is a form of mathematical logic, reason itself is disembodied. (p. xv)

This particular view of mathematics gives rise to conceptions of mathematics that privilege symbolic forms of reasoning over other types of reasoning. Such a view fits with the cognitive representationist view of cognition in which thinking is essentially the manipulation of internal “representations” of an external objective reality. Such a view does not require that mathematics have a human heart, human face, or human dress; indeed, from this view, mathematics becomes potentially disembodied. Such a view emerges from a Cartesian dualism between mind and body. Embodied Cognition, and specifically Embodied Mathematics, is a response to this cognitivist paradigm (Anderson, 2003) and rejects the basic Cartesian dualism in favour of viewing mind and body as co-implicated systems.

Six of the most prominent claims about embodied cognition are examined by Wilson (2002). For example, she examines the claim that “We off load cognitive work onto the environment” (Claim 3). The argument put forward here is that we reduce cognitive workload by strategically leaving information in the environment rather than fully encoding it, and can alter the environment in order to reduce cognitive work further. She considers the following example that:

counting on one’s fingers, drawing Venn diagrams, and doing math with pencil and paper...are both situated and spatial ... The advantage is that by doing actual, physical manipulation, rather than computing a solution in our heads, we save cognitive work. (p. 629)

She goes on to state that when such activities become decoupled from their embodied situations they can be “used to represent abstract, non-spatial domains of thought” (p. 629).

Another claim that is analysed is that “Cognition is for action,” which considers the adaptive role of cognition. She gives the example of memory, citing Glenberg (1997, p. 1) who “argues that the traditional approach to memory as ‘for memorizing’ needs to be replaced by a view of memory as ‘the encoding of patterns of possible physical interaction with a three-dimensional world’” (Wilson, 2002, p. 631). This makes sense, she argues, since we “conceptualize objects and situations in terms of their functional relevance to us, rather than neutrally, or as they really are” (p. 631). Such a strategy provides problem-solving flexibility, which is an adaptive behavior.

The final claim that is evaluated is that “Off-line cognition is body based.” Here she considers the proposition that cognitive structures that evolved for perceptual or motor tasks can appear to be “co-opted and run off-line, decoupled from the physical inputs and outputs that were their original purpose, to assist in thinking and knowing” (Wilson, 2002, p. 633). She gives examples from research on mental imagery, working memory, episodic memory, implicit memory, and reasoning and problem solving. She finds strong evidence for this particular claim from the domains listed previously and concludes that “it appears that off-line embodied cognition is a widespread phenomenon in the human mind...[that may reflect] a very general underlying principle of cognition” (p. 635).

Lakoff and Nunez (2000), working from the embodied perspective, bring the body back into mathematics. Their essential claim is that “mathematics as we know it is limited and structured by the human brain and human mental capacities. The only mathematics we know or can know is a brain-and-mind based mathematics” (p. 1). The mechanism by which they seek to explain the origins of complex, abstract mathematical ideas in sensorimotor activity is the cognitive linguistic device of conceptual metaphor. According to the authors, conceptual metaphor is a fundamental:

cognitive mechanism for allowing us to reason about one kind of thing as if it were another....It is a grounded, inference-

preserving cross-domain mapping – a neural mechanism that allows us to use the inferential structure of one conceptual domain to reason about another. (p. 6)

They provide plausible mechanisms by which concepts in mathematics could be developed based on metaphorical relations. However, their approach is criticized by Schiralli and Sinclair (2003) who state that, “by focusing almost exclusively on metaphor, *Where Mathematics Comes From* undermines, or at least fails to explain, the acts of understanding needed to move towards conceptual mathematics” (p. 88). The important contribution, though, of Lakoff and Nunez’s work, I think, is the explicit attention and focus it brings to the question of the role of the human body and evolutionary mechanisms in the process of mathematicizing. I agree with their statement that “mathematical ideas...are often grounded in everyday experience. Many mathematical ideas are ways of mathematicizing ordinary ideas” (p. 26) and that such ideas in many instances are rooted in sensory experiences.

Drodge and Reid (2000) also use embodied cognition to frame their argument for a mathematical emotional orientation. Their view draws attention not so much to cognitive action but to the ways in which cognition and emotioning (which is body-based) are intertwined. According to Picard et al. (2004), investigations of the way in which the body and cognition interact is a growing area of research. For example, they claim that through using the Logo Turtle:

children learn important geometric ideas in a more ‘body syntonic’ way, imagining themselves as the turtle as it draws out geometric patterns, and thus leveraging their intuitions and experiences of their own bodies into more formal knowledge and into a more personal relationship with mathematics. (p. 262)

They refer to other projects at the MIT Media Lab, which “have contributed to expanding the range of ways in which the body can be morphed into mathematics [including] knot-tying, piano playing, jiggling, skiing, and dance...in which the body in motion can support intuitive, emotionally deeply interconnected conceptual realms” (p. 262). Such research demonstrates, at a very practical level, the importance of the body in doing and learning mathematics.

But what else does drawing attention to the body in mathematics offer to Caribbean educators? A recognition of the important role of the body in mathematical cognition serves to re-establish a sensuality of mathematics that “cold” cognitive approaches eschew. Another benefit is the potential awareness that comes from asking the question of whose body is actually represented or not represented in mathematical discourse, and how such presence or absence is felt and experienced by real bodies, a perspective that resonates with postmodern critiques of mathematics education (Walshaw, 2004) and issues of social justice in mathematics education (Burton, 2003). This perspective also draws attention to the fact that much of human activity which might not be seen as formal academic mathematics probably involves implicit mathematical ideas, for example, recursion and pattern formation. As such, these body and sensory experiences, which are already familiar to learners, could be used to establish connections to more formal mathematical ideas. The Divine Image of mathematics may not be transcendent but might have a human form after all.

Ethnomathematics

Selin (2000), in the introduction to the book *Mathematics Across Cultures*, states that “every culture has mathematics,” which she defines as “the study of measurements, forms, patterns, variability and change” (p. xvii). Barton (1996), however, reminds us to be cognizant of the fact that “the category mathematics is not common to all cultures” (p. 216). The origin of activities and questions that we might consider mathematical under this definition of mathematics is related to the practical necessities of everyday life in different societies at different times, as well as religious and ritualistic connections; what D’Ambrosio (2000) refers to as the “needs of survival and transcendence” (p. 80). In this section I will try to answer the questions of “What is ethnomathematics?”; “How does one ‘do’ ethnomathematics?”; “What are the main debates and controversies in this area of inquiry?”; and “What are the potential implications for teaching, learning, and society?”

What is Ethnomathematics?

The difficulty in providing a simple answer to the question, “What is Ethnomathematics (EM)?” identified by Rowlands and Carson (2002), is

a reflection not merely of the differences of opinion among the many advocates of EM but, more importantly, also represents a healthy diversity of interests, foci, methodologies, and agendas of the various researchers and practitioners.

Barton (1996), for example, examines developmentally the evolving perspectives on EM of D'Ambrosio, Gerdes, and Ascher and posits the following definition of EM:

Ethnomathematics is a research programme of the way in which cultural groups understand, articulate and use the concepts and practices which we describe as mathematical, whether or not the cultural group has a concept of mathematics. (p. 214)

From which he concludes that:

Ethnomathematics does not consist of the mathematical ideas of other cultures, nor is it the representation of these ideas within mathematics...ethnomathematics is an attempt to describe and understand the ways in which ideas which the ethnomathematician calls mathematical are understood, articulated and used by other people who do not share the same conception 'mathematics'...like anthropology, one of the difficulties of ethnomathematics is to describe another person's world with one's own codes, language and concepts. (p. 215)

Eglash (1997) adds the following observations:

Ethnomathematics is typically defined as the study of mathematical concepts in small-scale or indigenous cultures. (p. 79)

Its epistemological basis is not restricted to methods of direct translation...but also includes the types of pattern analysis used in the modeling approach...Unlike mathematical anthropology, however, this research generally strives to include conscious intent as an important component of the analysis. (p. 81)

Ethnomathematics emphasi[zes] the possibilities for indigenous intentionality in mathematical patterns. (p. 85)

...ethnomathematics differs epistemologically from non-Western mathematics by not limiting itself to direct translations of

Western forms and instead remaining open to any mathematical pattern discernable to the researcher. (p. 87)

Rowlands and Carson (2002) suggest that:

The emphasis in EM seems to be on the ‘doing’ of mathematics, in the sense of cultural groups and peoples creating their own mathematics out of their everyday lives, rather than the teaching/learning of mathematics as a formal academic discipline. (p. 84)

Adam, Alangui, and Barton (2003) write that:

ethnomathematics recognizes the uniqueness of traditional cultures by highlighting aspects of their complex knowledge systems and showing them to be living and dynamic, and valuable and valid in their own terms and context.

The (mathematical) ideas of (traditional) peoples are not static but develop through time. Such knowledge may provide us with new concepts and problems in mathematics. (p. 328)

EM is not a philosophy, much less a ‘pedagogic philosophy’. Rather it is a lens through which mathematics itself can be viewed. (p. 329)

Doing Ethnomathematics

Barton (1996) lists and describes four activities considered relevant to EM projects:

1. **Descriptive Activity** — The first task of an ethnomathematician is to describe those practices or conceptions which are under consideration...a description which is...within the context of the other culture (p. 222)
2. **Archaeological Activity** — One way to bring out mathematical aspects is to trace backwards in time to uncover the mathematics which lies behind the current practice or conception. (p. 222)
3. **Mathematising Activity** — A second way of exposing the mathematical aspects in an EM study is by mathematising, i.e., by translating the cultural material into mathematical terminology, and relating it to existing mathematical concepts....As well as interpretative

mathematising, it is possible to work with the interpreted mathematics and extend it in a mathematical way. (p. 223)

4. **Analytic Activity** — Having described and developed mathematical ideas from other cultures, researchers seek to find out why the practices are like they are (p. 223)

Several good examples of this process can be found in the works of Ascher (2002), Eglash (1997), and Gerdes (1999).

Mathematizing or Re-mathematizing?

Pimm (2001) asks a very important and relevant question that any ethnomathematical study should consider:

in what sense are we mathematising the work and in what way...are we remathematising it? In other words to what extent did mathematics consciously play a role in the creation of the piece in the first place; to what extent is mathematics designed in?" (p. 32)

Though Pimm asks these questions of art, given the fact that the first reaction to much of what is considered ethnomathematics is an aesthetic and affective response to the objects as a whole and not as a mathematical work, his comments are well suited. For example, in Ascher's (2002) report of the kolam tradition of the Tamil Nadu region of India, we observe, for example, in one piece (Figure 1) that a 90° rotational symmetry is not an interesting aesthetic artefact but the result of a recursive embodied process in which the individual repeats a pattern after moving her body. Here, our observation of the symmetry is a re-mathematization of the process of construction. However, the observation that the figure is an Eulerian graph is a mathematizing activity since nowhere in the drawer's mind is this intended. This is unlike the African sand drawings described by Gerdes (1999), where the intent is specifically to produce patterns that have the Eulerian characteristic.

Thus, ethnomathematical activity can involve both mathematization and re-mathematization of the artefacts and practices of different cultures, though it is the latter, the discovery or creation of correspondences

between the original intent of diverse cultures of peoples with our own mathematical language and culture, that respects the integrity of both sets of traditions.



Figure 1. Pulli kolam (adapted from Ascher, 2002). Used with permission of Sabrina McMillan Solomon.

Debates and Controversies

There is considerable debate within the ethnomathematical field itself and among ethnomathematicians, mathematicians, and mathematics educators. Two of the issues have already been raised, namely, the inherent problems of meaning in defining “mathematics” and “culture,” and the question of whether or to what extent what is being done or observed in diverse cultures is a mathematization or a re-mathematization.

D’Ambrosio (1997) identifies another issue; that “ethnomathematics is frequently treated as pedagogical revisionism” (p. 13). It is this issue that partially occupies Rowlands and Carson’s (2002, 2004) critique of ethnomathematics. They ask the question, “What would an

ethnomathematics curriculum look like and where would formal academic mathematics fit in such a curriculum?” (Rowlands & Carson, 2002, p. 80), and posit four possibilities, which they then interrogate. Their conclusion that “ethnomathematics runs the risk of attempting to equalize everything down to the poverty of the ‘builders and well-diggers and shack-raisers in the slums’” (p. 98) and that “formal, academic mathematics would not exist in a curriculum informed by ethnomathematics” (p. 91) is challenged by Adam, Alangui, and Barton (2003) in a response to the previous article. Adam, Alangui, and Barton state emphatically that

Ethnomathematics is not out to displace or replace mathematics. At best, ethnomathematics forces us to reflect on our practice as mathematics educators, to reflect on our discipline, and to be aware of how it has contributed to a culture of “intolerance, discrimination, inequity, bigotry, and hatred. (D’Ambrosio, 2001)”(p.329)

They offer five pedagogical possibilities for an ethnomathematically informed curriculum of which their support is largely for:

an integration of the mathematical concepts and practices originating in the learners’ culture with those of conventional, formal, academic mathematics. The mathematical experiences from the learner’s culture are used to understand how mathematical ideas are formulated and applied. This...is then used to introduce conventional mathematics in such a way that it is better understood, its power, beauty and utility are better appreciated, and its relationship to familiar practices and concepts made explicit...a curriculum of this type allows learners to become more aware of how people mathematize and use this awareness to learn about a more encompassing mathematics. (p. 332)

Given that education in general and mathematics and science education in particular are steeped in socio-political struggles (Eglash, 1997), this debate on the pedagogical position that EM can and should take will probably not be soon resolved.

One of the issues around which debate focuses is that of the pre-eminent place given to Greek rationality in modern mathematics. According to

D'Ambrosio (1997), "much of the research in Ethnomathematics today has been directed at uncovering small achievements and practices in non-Western cultures that resemble Western mathematics. Western mathematics remains the standard of rationality..." (p. 15). This is taken up by Rowlands and Carson (2002), who see an attempt among some ethnomathematicians to downplay the Greek contribution to mathematics and rationality. This claim is refuted by Adam et al. (2003), who instead point out that:

ethnomathematics recognizes the pre-eminence of Greek rationality in modern mathematics and seeks to understand this. One way to do so is to see that Greek rationality is only one form of rationality and that the particular form of mathematics that traces its trajectory through the Greek tradition (and a few others) serves particular functions and has particular consequences. (p. 330)

The role of rationality is discussed by Turnbull (2000) as he argues for an understanding of knowledge systems as inherently local as opposed to universal. He states that

rationality is a deeply problematic concept. It is profoundly embedded in the hidden assumptions of late twentieth century Occidentalism about what it is to be a knowing, moral, sane individual...but there are no universal criteria for rationality....Rationality consists in the application of locally agreed criteria in particular contexts. (p. 46)

Like the problems of definition and of pedagogy, the centrality of rationality, of the sort espoused in the Greek tradition, is one that both mathematics and ethnomathematics must continue to wrestle with. However, what can be gained by relaxing the constraint of rationality (as is espoused and practised in the Western mathematical and scientific traditions) is the potential for the creation "of a shared knowledge space in which equivalencies and connections between differing forms of rationality can be constructed" (Turnbull, 2000, p. 53).

There is also a concern, voiced by Rowlands and Carson (2002), that despite the noble aim of EM in attempting to provide equity for all students, "ethnomathematics cannot guarantee any fundamental change

in terms of equity and so if the status quo remains then this notion of 'equity' is misplaced and can only serve to maintain the divisions that exist in society" (p. 86). However, the question can be turned on itself and we might ask, "To what extent does traditional mathematics guarantee changes in terms of equity?" The main point of many in the ethnomathematical field seems to be that traditional mathematics as it is practised ignores important and essential human needs, and indeed continues to serve as a de-historicizing, de-humanizing, colonizing agent. Arguments about guarantees of equity in this form serve as a red herring to deeper issues. Neither EM nor formal academic mathematics can guarantee changes in equity. What can be achieved by EM is an appreciation of what is mathematical about other people's (and one's own) practices.

Does EM empower or disempower students (Rowlands & Carson, 2004)? This is another fundamental issue in the EM debate. Evidence presented by Adam et al. (2003) suggests that "students who have been taught using such an ethnomathematical curriculum perform better on conventional mathematics tests...[and] that learning in context does make mathematics more meaningful to learners" (p. 333). However, what Rowland and Carson are concerned with is whether or not EM denies students access to the "cultures of power" (Delpitt, 1988). They write, "what [will] this will do for learners, especially when their peers in more conventional educational settings continue to undergo the more customary induction into the world's now ubiquitous formalized cultural systems..." (p. 86). This position, to an extent, belittles EM as an educational project, and the position is rendered less potent with the studies referred to by Adam et al. (2003). Of course, additional studies that are able to make the case that EM does allow students to gain access to the cultures of power are needed.

The question, "To what extent is formal academic mathematics disempowering?" should also be asked. The fact that "many people who study geometry, trigonometry, algebra, calculus, and so forth seldom use these branches of mathematics professionally, yet the effect of this learning is regarded by many as a benefit in and of itself" (Rowland & Carson, 2004, p. 336) is not sufficient grounds to claim a privileged position for formal academic mathematics.

The final issue is one of presentation. There is a concern among some (e.g., Eglash, 1997) that some EM pedagogical projects, by singling out minority students, may increase their alienation or sense of Otherness from formal mathematics or may present aspects of a culture in a trivialized or romantic fashion. Within this, we might situate Rowland and Carson's (2002) concern that an ethnomathematical curriculum might be merely an "excursion into geometrical aesthetics" (p. 92). However, given the diversity of EM projects this seems unlikely.

Implications for Teaching, Learning, and Society

Some of the implications of EM for teaching and learning have been suggested above. However, I wish in this section to explore as well the wider implications of this endeavour. First, though, I situate my own opinion with that of Adam et al. (2003), who state that "ethnomathematics is not a pedagogic philosophy. Rather it is a lens through which mathematics itself can be viewed" (p. 329), and extend it to include D'Ambrosio's (1997) view that such a lens must also be used to look on other sociocultural constructions in general.

The EM literature says very little explicitly about instruction. Rowlands and Carson (2002), in their defence of formal academic mathematics, suggest and investigate four possibilities for an EM influenced curriculum. These are, EM as a substitute for formal academic mathematics; EM as a supplement in order to promote an appreciation of other human cultures; EM as an entry point to formal mathematical ideas; and EM as a consideration in the design and preparation of learning situations. They do, however, state that:

ethnomathematics would make a valuable contribution to the curriculum if it demonstrated how mathematical ideas grew out of the needs of various peoples. However, the emphasis would need to be on the historical and cultural development of mathematics rather than the ideas of various peoples. (p. 91)

This position is criticized by Adam et al. (2003) as an attempt to "tak[e] mathematics in its present form as a given, and ask merely that we be sensitive to cultural difference as mathematicians and mathematics educators (p. 334). They suggest, instead, five possibilities for an EM influenced curriculum, "premised on the belief that the cultural aspects

of the students' milieu should be infused in the learning environment in a holistic manner" (p. 331). Their particular preference has been stated previously.

What EM appears to offer educators is, firstly, a sensitizing concept that at the least offers the promise of greater cultural sensitivity and respect for cultural diversity among students. Secondly, it offers mathematics educators new perspectives from which to view their discipline.

With respect to learning, Adam et al. (2003) have noted that EM has already contributed to new mathematical knowledge. For example, Ascher (2002) reports on the contributions of the study of kolam traditions to computer science. This substantiates the claim of Barton (1997) that "ethnomathematics includes a dialogue between the ideas of another culture and the conventional concepts of mathematics. This dialogue is likely to lead both to new areas of application for mathematics, and to new mathematics through adaptation to new ideas" (p. 217). Thus, what is to be learnt involves not only new knowledge, but new ways of looking at and new perspectives on knowledge itself.

The larger EM project is a socio-political one involving a postmodern critique of both the social construction of science, mathematics in particular, and the consequences of the particular sociocultural construction that dominates most of the discourse in the West. According to Adam et al. (2003), "there is more to ethnomathematics than recognizing the failings of mathematics as a discipline. Its philosophy also implies that we recognize its potential to turn an unjust situation around" (p. 329). The ultimate aim of the EM project is to contribute to a new ethics of "respect, solidarity and responsibility" (D'Ambrosio, 2001) and, ultimately, to a more peaceful coexistence for all of humanity.

At a very fundamental level, EM aims to challenge hegemonic conceptions of what counts or should count as mathematical. Barton (1997), for example, states that "part of the purpose of ethnomathematics is to challenge the perceived universal nature of mathematics, and to expose different mathematical conceptions" (p. 216). To this end, then, he states that

mathematics must be changing...It must admit the possibility of other mathematical concepts which are not subsumable by

existing ones, or by some new, overarching generalization....Unless mathematics can change in a radical way, there is no point to examining the way other people view things which we call mathematical. If there is only one view of mathematical phenomena, then why try and find another one? (pp. 218–219)

For example, Wood (2000) examines the changing views on mathematical proof;

modern mathematicians are in considerable conflict about the status and usefulness of proof....Much has been made of the Greek concept of proof as the basis of modern mathematics to the extent that many writers have disparaged the Indians for their supposed haphazard ideas of proof....Indian mathematics does prove theorems though not in the same way as Greek deductions from axioms. (p. 8)

D'Ambrosio (1997) identifies the central role that EM has to play in this challenge when he states that “ethnomathematics places us in an advantageous position to look into the nature of mathematical knowledge, the questions about which cannot be resolved within the framework of Western mathematics itself” (p. 14). Ethnomathematics' first challenge is to one of its parental domains, mathematics itself.

The second challenge that emerges from EM is a historiographical one (D'Ambrosio, 1997, 2000, 2001). According to Selin (2000):

until the past few decades histories of mathematics have virtually ignored the mathematics of non-European cultures....This neglect grew from the colonial mentality, which ignored or devalued contributions of the colonized peoples as part of the rationale for subjugation and dominance. (p. xvii)

This is a situation that is being partially rectified by ethnomathematics and other ethnosciences.

Finally, the most ambitious goal of EM is a contribution to solving what Rowlands and Carson (2004) call the most pressing issue on the planet

today, “that of figuring out how peoples of varied cultural allegiances are to get along” (p. 329). D’Ambrosio (2001) states that:

I see mathematics playing an important role in achieving the high humanitarian ideals of a new civilization, with equity, justice and dignity for the entire human species....But this will depend on our understanding how deeply related are mathematics and human behavior. (p. 327)

Though I disagree with his statement that “the *most universal problem* – that is survival with dignity – must have something to do with the most universal *mode of human thought* – that is mathematics” (p. 328), which seems to set up a false analogy, I do agree with his statement that it is, “the failure to reestablish the lost interconnection of the sciences, technology and human values is causing unavoidable conflicts. This is apparently true with mathematics, in which the acknowledgement of human attributes is conspicuously absent from its discourse” (p. 329). It is precisely because such discussions of mathematics, ethics, and values are seldom classed together that there is a need to engage in this discussion. It is precisely because our conceptions of mathematics are limited that the ethnomathematical project is necessary. It is precisely because the stakes are so high that such re-connections are urgent.

The SEMP Mathematics Curriculum of Trinidad and Tobago

Some of these ideas representing the overlapping of the domains of embodiment and ethnomathematics are manifested in the reform-oriented mathematics curriculum documents of the Trinidad and Tobago Secondary Education Modernization Programme (SEMP). For example, some of the stated goals of education outlined in the SEMP curriculum guide include:

- Providing opportunities for all students to develop spiritually, morally, emotionally, intellectually and physically;
- Providing opportunities for all students to develop an understanding and appreciation of the diversity of our culture; and
- Providing opportunities for all students to develop an appreciation for beauty and human achievement in the visual and performing arts. (SEMP, 2002, Section 1:5-6)

These goals are further elaborated in six essential learning outcomes, three of which are Aesthetic Expression, Citizenship, and Problem Solving. In achieving these learning outcomes, eight core subjects have been identified, of which mathematics and visual and performing arts are two. The stated goals of the mathematics curriculum include:

- To make mathematics relevant to the interests and experiences of the students and to prepare students for the use of mathematics in further studies
- To cultivate creativity and critical thinking in applying mathematical knowledge and concepts to solve routine and non-routine problems
- To develop skills in inquiry by the use of mathematics to explain phenomena, and by recognition of the influence of mathematics in the advancement of civilization
- To promote appreciation of the role of mathematics in the aesthetics and to make mathematics fun
- To encourage collaboration among students and to promote positive attitudes and values in students through the completion of mathematical tasks
- To provide opportunities for students to experience the structure of mathematics and to appreciate the elegance and power of mathematics
- To provide students with a range of knowledge, skills and techniques relating to number, geometry (space and shape), algebra, measurement, relations, functions, and statistics in a manner relevant to the technological advancements of the 21st century. (SEMP, 2002, Section 2:3)

Some connections to the other core curriculum areas are suggested in Section 2:5-7 of the curriculum document. For example, suggested connections between Mathematics and Visual and Performing Arts include:

- Understanding timing and sound in music, which are based on basic mathematical concepts in number and trigonometry
- Observing the beauty in art and nature, which is based on the concept of symmetry in geometry
- Drawing, designing and dancing, which are dependent on acquiring skills in geometry and fundamental mathematics

- Producing art and craft which require the use of calculations, spatial sense and fundamental concepts in mathematics
- Sequencing of dance steps and patterns in dance, which are dependent on geometrical and number concepts and skills

From the above excerpts, one sees an explicit desire to expand the scope of what is considered relevant to mathematical and, indeed, humanistic instruction. In the next section, I will suggest some potential areas for study where some of the ideas explored regarding ethnomathematics and embodied perspectives might fruitfully intersect with domains or aspects of culture of Trinidad and Tobago, and discuss some of the potential implications.

Potential Areas for Curriculum Study

What I am suggesting is, emerging from the perspectives on ethnomathematics and embodied mathematics, that aspects of the culture of Trinidad and Tobago and, indeed, the cultures of the wider Caribbean, become a focus for mathematical inquiry both among students and researchers. Such a project connects the needs of today with the history of yesterday. It is my hope that such a project might lead to a greater valuing of aspects of “West Indian” culture and a recognition of the importance of reclaiming one’s historio-mathematical heritage while simultaneously recognizing how one goes about continuing to create such a heritage.

I limit my initial proposal to aspects of life in Trinidad that are socioculturally situated, though recognizing that they could be extended to other areas such as the natural environment. This limiting should not deter readers from other Caribbean territories and elsewhere from applying the ideas to their own cultural situations, but it is meant to focus the discussion in the one culture with which I am familiar.

Culturally, Trinidad and Tobago is as diverse as its population. Each ethnic and religious group has retained some of its practices. However, of greater interest to me are art forms that are endemic hybrids of the coincidences of proximity and history of diverse groups. As such, my initial proposal is further limited to aspects of the arts that might be considered local. Two areas of inquiry are thus immediately suggested:

the Carnival and the steelband. The latter is discussed in Khan (2008). In this paper, I restrict the discussion to that of the Carnival.

Carnival

Carnival is a two-day street festival involving costuming, music, and dancing. Each of these areas offers possibilities for mathematical inquiries. For example, from casual observation there appears to be a lot of relevant mathematics in the process of creating costumes. Some of this is related to patterns and geometric motifs, symmetry, and tessellations. Creating a costume is not a matter of knowing academic mathematics, but it is an embodied experience of bending wood and wire to the will to create beauty. With respect to choreography, this too might be seen as a type of embodied mathematics—how human movements are enhanced and imbued with mathematical meaning through the elaboration of costuming. One might consider the trajectories, or the surfaces that can be shaped by different types of costumes as the masquerader inhabits the ‘mas’ and moves through space. Consider the following quote on the Calaloo Company’s (a performance art group) website, which speaks to the embodied and ethnomathematical frameworks:

The challenge of mobility in mas is to transmit the movement of the performer to his apparel, to magnify it, and see it articulated in the far reaches of the structure, yards away from the body. As mas is performed to music, the essential movement of the masquerader is his dance. The greatest kinetic potential, then, is to base the mobility of the costume on that movement, so that the mas expresses the dance, and the rhythm of the music can be read high in the air.

In addition to using mathematics in the design of costumes, one might deliberately design costumes to allow for the exploration, display, and embodiment of mathematical ideas. This might include some of the ethnomathematical studies of Ascher (2002) and Gerdes (1999).

Carnival is but one area where the culture of Trinidad and Tobago might productively meet with the goals of mathematics education outlined previously. Other areas could involve a more traditional exposition linking the diverse contributions of Indian, African, Asian, and European mathematicians to mathematical knowledge. Given the multicultural

nature of Trinidadian classrooms and the still colonial view of mathematics, such a project, I believe, could assist in developing a wider appreciation of what is mathematical about one's own sometimes taken-for-granted cultural situation. The mathematics of Carnival could be part of a larger project, for example, involving a study of mathematics in the major festivals including Hosay, Divali, and Phagwa. Such studies, I believe, might contribute positively to fostering an understanding and appreciation of the diversity of the culture of Trinidad and Tobago, allow students to develop an appreciation for beauty and human achievement in the visual and performing arts, and make mathematics more relevant to local interests and experiences. In this regard, it is possible that a uniquely Trinidadian mathematical identity might begin to be fashioned, one in which the words of the national anthem, "every creed and race find an equal place" ring more true than they do at present.

References

- Adam, S., Alangui, W., & Barton, B. (2003). A comment on: Rowlands and Carson "Where would formal, academic mathematics stand in a curriculum informed by ethnomathematics? A critical review." *Educational Studies in Mathematics*, 52, 327–335.
- Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, 149, 91–130.
- Ascher, M. (2002). *Mathematics elsewhere: An exploration of ideas across cultures*. Princeton, NJ: Princeton University Press
- Barton, B. (1996). Making sense of ethnomathematics: Ethnomathematics is making sense. *Educational Studies in Mathematics*, 31, 201–233.
- Blake, W. (1789). *Songs of innocence*.
- Burton, L. (Ed.) (2003). *Which way social justice in mathematics education?* Westport, CT: Praeger.
- D' Ambrosio, U. (1997). Where does ethnomathematics stand nowadays? *For the Learning of Mathematics*, 17(2), 13–17.
- D'Ambrosio, U. (2000). A historiographical proposal for non-western mathematics. In H. Selin (Ed.), *Mathematics across cultures: The history of non-western mathematics* (pp. 79–92). Dordrecht: Kluwer Academic.
- D'Ambrosio, U. (2001). Mathematics and peace: A reflection on the basis of Western civilization. *Leonardo*, 34(4), 327–332.
- Delpitt, L. D. (1988). The silenced dialogue: Power and pedagogy in educating other people's children. *Harvard Educational Review*, 58(3), 280–298.
- Dodge, E. N., & Reid, D. A. (2000). Embodied cognition and the mathematical emotional orientation. *Mathematical Thinking and Learning*, 2(4), 249–267.

- Eglash, R. (1997). When math worlds collide: Intention and invention in ethnomathematics. *Science, Technology and Human Values*, 22(1), 79–97.
- Gerdes, P. (1999). *Geometry from Africa: Mathematical and educational explorations*. Washington, DC: Mathematical Association of America.
- Khan, S. K. (2008). Mathematics and the steel drum. *For the Learning of Mathematics*, 28(2), 28–37.
- Lakoff, G., & Nunez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.
- Picard, R. W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., Machover, T., Resnick, M., Roy, D., & Strohecker, C. (2004). Affective learning – a manifesto. *BT Technology Journal*, 22(4), 253–268.
- Pimm, D. (2001). Some notes on Theo van Doesburg (1883–1931) and his Arithmetic Composition I. *For the Learning of Mathematics*, 21(2), 31–36.
- Trinidad and Tobago. Ministry of Education. Secondary Education Modernization Programme. (2002). *Secondary school curriculum: Form one mathematics* (revised draft). Port of Spain: Author.
- Rowlands, S., & Carson, R. (2002). Where would formal, academic mathematics stand in a curriculum informed by ethnomathematics? A critical review of ethnomathematics. *Educational Studies in Mathematics*, 50(1), 79–102.
- Rowlands, S., & Carson, R. (2004). Our response to Adam, Alanguí and Barton’s “A comment on Rowlands and Carson ‘Where would formal, academic mathematics stand in a curriculum informed by ethnomathematics? A critical review.’” *Educational Studies in Mathematics*, 56(2–3), 329–342.
- Schiralli, M., & Sinclair, N. (2003). A constructive response to where mathematics comes from. *Educational Studies in Mathematics*, 52(1), 79–91.
- Selin, H. (2000) Introduction. In H. Selin (Ed.), *Mathematics across cultures: The history of non-western mathematics* (pp. xvii–xx). Dordrecht: Kluwer Academic.
- Turnbull, D. (2000). Rationality and the disunity of the sciences. In H. Selin (Ed.), *Mathematics across cultures: The history of non-western mathematics* (pp. 37–54). Dordrecht: Kluwer Academic.
- Walshaw, M. (Ed.). (2004). *Mathematics education within the postmodern*. Greenwich, CT: Information Age Pub.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin Review*, 9(4), 625–636.
- Wood, L. N. (2000). Communicating mathematics across culture and time. In H. Selin (Ed.), *Mathematics across cultures: The history of non-western mathematics* (pp. 1–12). Dordrecht: Kluwer Academic.

NOTES ON CONTRIBUTORS

Dorian Barrow is a Lecturer in Science Education at the School of Education, UWI, St. Augustine.

Jerome De Lisle is currently a Lecturer in Educational Administration at the School of Education, UWI, St. Augustine. He has worked in measurement and evaluation in the past, and his current research interests are in schools experiencing challenge, national assessments of educational achievement, and in the impact and consequences of high stakes assessments at eleven plus.

Anthony Griffith is a Senior Lecturer in Social Studies in the School of Education at the Cave Hill Campus of The University of the West Indies (UWI). His major interests are in social studies and geographic education, teacher education, and education for development.

Steven Khan is a former teacher of mathematics and biology at secondary level in Trinidad. His areas of interest include the intersections between mathematics, aesthetics, and ethics in secondary mathematics education and teacher preparation, qualitative research methods, complexity theory and education and, pedagogy of (children's) popular culture. He currently lectures in Mathematics Education at UWI, St. Augustine.

Samuel Lochan is a Lecturer in Social Studies Education at the School of Education, UWI, St. Augustine, Trinidad.

Rawatee Maharaj-Sharma is a Lecturer in Science Education (Physics) at the School of Education, The University of the West Indies (UWI), St. Augustine. She previously lectured at York University, Canada for 5 years in Physics and Astronomy.

Bertram Ross is a former high school and college principal with extensive teaching and administrative experience. A pioneering examiner with CXC, Dr. Ross was a member of the OECS Education Reform Strategy Group. He was employed as Principal Lecturer in Teacher Education and Science at the HLSCC College in the BVI for six years and served as an adjunct Professor at the University of the Virgin Islands.

Sonia St. Hill is a Social Studies teacher at the Lodge School in Barbados. She obtained a Dip.Ed. and an M.Ed. from The University of the West Indies (UWI).

Jasmattie Yamraj is a doctoral candidate at Queen's University, Canada. She was employed for 10 years at the BVI High School as a senior Science teacher.

SALE OF PUBLICATIONS

OCCASIONAL PAPERS

Author	Title	Price \$TT
Peter Kutnick	Insights into the structure and personnel of primary schools: A demographic and biographic study of a representative sample of primary schools in Trinidad and Tobago.	10.00
Peter Kutnick	A survey of primary school teachers' understanding and implementation of moral education in Trinidad and Tobago.	10.00
Betty Ann Rohlehr	Cognitive style, graphicacy skills of CXC geography students	15.00
Maria Byron & V. Mohan Ram	Some teachers' perceptions of teaching and learning mathematics at the primary level in Trinidad and Tobago.	15.00
Glenda Prime	Broad categorizing and achievement in biology.	15.00
Marie-Louise Brown	Annotated bibliography of teacher research in reading in Trinidad and Tobago 1982-1991	10.00

OTHER PUBLICATIONS

Caribbean Curriculum	Vol. 1, No. 1, 1985	5.00
	Vol. 2, No. 1, 1987	5.00
	Vol. 3, No. 1, 1993	10.00
	Vol. 3, No. 2, 1993	10.00
	Vol. 4, No. 1, 1994	15.00
	Vol. 5, No. 1, 1995	15.00
	Vol. 5, No. 2, 1995	15.00
	Vol. 6, No. 1, 1996	15.00
	Vol. 6, No. 2, 1996	15.00
	Vol. 7, No. 1, 1999	15.00
	Vol. 7, No. 2, 2000	15.00
	Vol. 8, 2001	20.00
	Vol. 9, 2002	20.00
	Vol. 10, 2003	20.00
Vol. 11, 2004	20.00	
Vol. 12, 2005	20.00	
Vol. 13, 2006	20.00	
Vol. 14, 2007	20.00	

Kelvin Jarvis & Sonia Manjoo	A bibliography of staff publications 1965-1985. Faculty of Education Library, 1986, 48 pp.	15.00
E.P. Brandon & P.N. Nissen (Comps.)	Proceedings of the 1990 Cross-Campus Conference on Education, 3-6 April, 1990. Kingston, Jamaica. Jamaica: Faculty of Education, 1991, 312 pp.	55.00
Bernard Heydorn, Kerl Crichlow, & Andrew Ballantyne	A study of reading education in the primary schools of Trinidad and Tobago.	10.00
Olabisi Kuboni (Ed.)	Literacy in the Modern World: Proceedings of the Symposium.	20.00
Social Studies Education	"New Dimensions in Social Studies," Vol. 26, 1994. (<i>Back issues available at reduced rates</i>).	10.00
Lynda Quamina-Aiyejina	Undergraduate Research on Education at the University of the West Indies. St. Augustine. Caribbean Studies 1965-1994. CERIS Biographical Series No. 1.	10.00
Lynda Quamina-Aiyejina	Guide to published statistical data on education in the English-speaking Caribbean. 213 p.	20.00
Margaret Cain, Anna-May Edwards- Henry, & Joycelyn Rampersad	A Portfolio Training Manual: Using the Portfolio as a Model for Training	30.00
Margaret Cain, Anna-May Edwards- Henry, & Joycelyn Rampersad	Developing portfolios for integrating teaching, learning, and assessment	50.00

INSTRUCTIONS FOR ORDERING

Orders by post should be accompanied by cheque or money order made payable to "**Publications, School of Education.**" Send orders to: *Librarian, School of Education, The University of the West Indies, St. Augustine, Republic of Trinidad & Tobago, West Indies. FAX: (868) 662-6615 Email: Janet.Fullerton-Rawlins@sta.uwi.edu.*

Cash purchases can be made from the *School of Education Library.*

ORDER FORM

Enclosed is my cheque or money order for:

Title ordered:

.....
.....
.....

No. of copies:

Total sum enclosed \$

Name:.....

Address:.....

.....
.....
.....

E-mail:..... Tel:..... Fax:.....

Information for Contributors

- (1) Manuscripts for publication that have not been published elsewhere and all correspondence regarding articles should be submitted to *The Librarian/Documentalist, School of Education, The University of the West Indies, St. Augustine, Trinidad, West Indies.*
- (2) Articles from all relevant subject areas will be considered by the editorial committee and priority will be given to those that, in addition to showing acceptable scholarship, fulfil one or more of the following criteria:
 - (a) describe or evaluate experiments in curriculum intervention that have relevance for Caribbean school system;
 - (b) discuss theoretical matters that can favourably affect thought on education in the Caribbean;
 - (c) promote continuous curriculum renewal in the Caribbean region.
- (3) Manuscripts should be typed double-spaced (including quotations, footnotes, and references) on standard quarto paper (8 1/2" x 11"), with ample margins. The author's name and affiliation should appear on a separate cover page, and only on this page, to ensure anonymity in the reviewing process. Include, also on separate pages, an abstract of 100-200 words, and a short biographical profile of the author (and all co-authors, if any). The APA Style Manual (5th edition) should be followed.
- (4) The list of references should be arranged alphabetically by authors' last names, and placed at the end of the manuscript. Reference citations are not to be used as footnotes.
- (5) Endnotes for the purpose of clarifying the text are permitted and should be denoted by superscript numbers in the text. The numbered endnotes should be typed double-spaced and placed before the list of references.
- (6) Tables and figures should be numbered, captioned, and inserted at appropriate points in the text.
- (7) Kindly send three copies of each manuscript for editorial review, excluding originals. All manuscripts will be refereed.
- (8) Manuscripts that do not conform to all the above specifications will be returned immediately to the author for conformity.
- (9) Authors will receive five (5) copies of their articles without charge.