

## EDITORIAL ADVISORY GROUP

Dr. B. Brereton  
Prof. M. Cain  
Ms. S. Craig-James  
Mr. M. Morrissey  
Dr. V. Pollard  
Prof. G. Rohlehr

Mr. E. Brandon  
Dr. D. Broomes  
Dr. D. Clarke  
Dr. J. Glasgow  
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, 2007

ISSN 1017-5636

# CARIBBEAN CURRICULUM

Volume 14

2007

## CONTENTS

Impacting Struggling Adolescent Readers: A Socio-Psycholinguistic Study of Junior Secondary Students in Trinidad	1
<i>Cynthia James</i>	
Students' Attitudes to Science in Urban and Rural Schools in Trinidad and Tobago	31
<i>Rawatee Maharaj-Sharma</i>	
Differential Item Functioning and Male-Female Differences in a Large-Scale Mathematics Assessment in Trinidad and Tobago: An Examination of Standard 1 Mathematics Assessment	49
<i>Launcelot I. Brown and Gibbs Y. Kanyongo</i>	
The Promotion of Thinking in Selected Lower Secondary Science Classrooms in Trinidad and Tobago: Implications for Teachers' Education	73
<i>Susan Herbert and Joycelyn Rampersad</i>	

Notes on Contributors

Information for Contributors



**IMPACTING STRUGGLING ADOLESCENT READERS:  
A Socio-Psycholinguistic Study of  
Junior Secondary Students in Trinidad**

*Cynthia James*

This paper analyses the classroom reading behaviours and literacy culture of 19 low SES students, aged 11–14, who scored between 0–30% on their primary school exit examinations, and were placed in an English-and-Math-intensive Form 1 Remedial class at a junior secondary school in Trinidad. Although there were large differences in reading abilities within the group, poor phonics, syntactic, and semantic competencies impeded decoding and meaning-making for the majority of the students. A reading interview, the students' reading journals, and a variety of authentic reading stimuli formed a backdrop for surveying six issues. An exploratory analysis of 8 two-hour sessions over a period of eight weeks suggests that isolating the students in a concentrated English-and-Math-intensive programme would increase their stigmatization and encourage reading fossilization. A blend of interactive, transactional, sociocultural, and engagement approaches seems likely to promote the best outcomes. It is also suggested that given the students' lack of access to the potential literacy benefits of high-end technology, that they be at least exposed to a wider range of sign system literacies for making meaning from text.

**Introduction**

The introduction of universal secondary education in Trinidad and Tobago in 2000 has brought challenges to the secondary school system, a major one being in the area of academic literacy. The system has dealt with this challenge by instituting, at Form 1 level, a remedial programme for students who score between 0–30% on the primary school exit assessment.<sup>1</sup> This paper focuses on one aspect of this academic literacy challenge—that of reading—among a Form 1 group in this programme at an urban junior secondary school in Trinidad. The exploratory study

focuses on reading in the English language arts, and takes its impetus from my interest in the increasing difficulties secondary school teachers of English report. I interfaced, as teacher-researcher, with Form 1-M<sup>2</sup> once a week from September to November 2005 during one of their reading sessions, for a total of eight sessions. The period dated from the students' third week of entry into secondary school to two weeks prior to their end-of-term examinations.

### **Description of group**

Form 1-M was a class of 19 junior secondary school students between 11–14 years old. They came from working-class homes at the lower end of the socio-economic stratum. This meant that although they belonged to the Internet age, neither at school nor in their homes did they have access to computers or Internet technology. Their first language (L1) was Trinidadian Creole (TC), which has a predominantly English language lexicon. The group was composed of 8 girls and 11 boys who had graduated from primary school in July 2005. They were considered struggling readers and had been put in an English-and-Math-intensive programme for the academic year. The class was taught by an English teacher and a Math teacher (females), both experienced, retired teachers employed specifically to give struggling students an opportunity to catch up.

The English teacher, as the person most responsible for the group, was given free rein with the students' curriculum, and she included storytelling, values, and outdoor recreation sessions periodically.<sup>3</sup> However, her remit was to improve the students' English, which she did along traditional lines, using a variety of skills-based texts and a basal reader (Ferguson-Briggs, 2002). The class did Art with their Math teacher twice a week for one hour each session, when, by their report, they drew mainly shapes. The students had all their classes in their classroom in the Home Economics section of the school.

Although the students were ranked at the bottom of the system, 0–30 percentile scores meant that there was a wide mix of abilities in the class. According to the English teacher's evaluation, two could not decipher or decode (males), and by middle-to-upper primary school standards, 3 read at frustration level, 10 at instructional level, and 4 at independent level.

Among the independent readers there was only one male. Determination of their reading level had been done at the beginning of the semester, using Dolch reading lists and comprehension passages. The class was being geared to be mainstreamed into Form 2<sup>4</sup> in September 2006 at the end of the academic year.

### **Questions**

A socio-psycholinguistic approach to the literacy situation described, as it pertained to reading, suggested the following broad lines of inquiry:

1. How did the students perceive reading?
2. What would their specific areas of difficulty be with secondary school reading?
3. To what extent was their L1 affecting their reading?
4. How did grouping them as a unit impact their literacy?
5. How would the students respond to a mix of literacy stimuli?
6. What interventions would be most effective in impacting their literacy development?

### **Literature Review**

The most influential perspectives on reading agree that reading is an act of meaning-making. However, differing positions on the process by which readers make meaning have given rise to reading models with various emphases, with four prominent overlapping strands being the interactive, the transactional, the sociocultural, and the motivational.

Dechant (1991), for instance, endorses the interactive model of reading:

The interactive model (McClelland, 1986; Perfetti, 1985; Perfetti & Roth, 1981; Rumelhart, 1976, 1980; Rumelhart & McClelland, 1981, 1982; Stanovich, 1980) . . . suggests that meaning comes from many sources, that the reader simultaneously uses all levels of processing, that any one source of meaning can be primary at a given time, that utilizing information from one source often depends on utilizing information from the others, and that the reader constructs meaning by the selective use of information from all sources of meaning without adherence to any set order. (pp. 26–27)

Dechant lists 10 sources of meaning: logographic, graphemic, phonological, orthographic, morphemic, lexical, semantic, syntactic, schematic knowledge, and grapheme/phoneme correspondence. He aligns this conceptualization to the schema model of reading, rooted in “the cognitive base, which the reader draws upon to match new incoming information with prior information stored in memory” (p. 114).

Kenneth and Yetta Goodman, along with their associates Burke and Watson, are well-known whole language proponents. The whole-language approach to reading also sees reading as meaning-making, but it places less emphasis on “the word.” For Kenneth Goodman (cited in Marek & Goodman, 1996), reading involves the combination of semantic, syntactic, and graphophonemic systems, as readers construct personal texts from given texts. Different from strategies that pay much attention to letter-word correspondence, word recognition, and word attack skills, Goodman’s model of reading instruction focuses on metacognitive, inquiry-based, and constructive elements such as prediction, inference, self-correction, and self-monitoring (pp. 22–23). He eschews the “diagnosis and remediation metaphor” of reading (Goodman, 1996, p. 13) in favour of the concept of readers “revaluing” their reading practice. Revaluation is described as a transactional process between readers and texts, which positions readers to analyse and reflect on their reading, thus giving them a better understanding of themselves as readers. A central understanding is that “in the construction of meaning, all readers make miscues” (cited in Marek & Goodman, 1996, p. 23). The typical Goodman miscue analysis chart codes readers’ miscues for syntactic and semantic acceptability in meaning construction. It also monitors the graphic and sound similarities of readers’ utterances alongside their self-corrections (Goodman, Watson, & Burke, 2005, p. 78). Based on taping and replaying readings in personal study or in consultation with a tutor toward evaluating miscues, retrospective miscue analysis (RMA) is aimed at moving reading from the control of the teacher, while putting it in the control of readers themselves.

Sociocultural models of reading emphasize broader concepts of literacy development and have visible legacies in the work of Paulo Freire. According to Siegel and Fernandez (2002), Freire regarded illiteracy “not as an individual failing but as a historically constructed product of a

society structured to produce inequality” (p. 70). Thus, as Au (1997), a contemporary who comes to reading from a sociocultural perspective, explains:

Learning to read cannot logically be separated from the particular milieu in which it takes place. When children learn to read, or fail to learn to read, they do so in a particular social, cultural and historical environment. Their success or failure in reading cannot be understood apart from that environment. (p. 184)

Victoria Purcell-Gates (2002), who also espouses a sociocultural perspective, points to cultural differences that impact the reading competence of groups entering formal education with different literacy preparedness from that required to learn from the print culture of school. In her study, based on a poor White Appalachian community, she reminds:

To learn about written language, to learn that ‘print says.’ [sic] To learn that written stories sound different from the way people talk . . . to learn all of these basic concepts [of print] requires extensive experience with people using print, with people reading and writing around you and to you and for you and allowing you to try your hand at reading and writing. (p. 125)

Like Purcell-Gates, Brice Heath (1994) notes that the linguistic and cultural experiences of children from societies with predominantly oral traditions are in no way deficient. A socio-psycholinguistic approach to reading and to literacy respects, in equal measure, the oral language experience of the learner and the print language that he or she has to treat with.

However, socio-psycholinguistic factors surrounding reading and literacy are more complex than the respect-versus-deficit hypothesis that these scholars raise. With English as a Second Language (ESL) learners whose first language is an English-based Creole, for instance, the fact that their first language shares the same lexicon as the target language often leads the speakers to assume that the language they produce and English are one and the same. Then, too, some Creole speakers resist “white Western” language, deliberately adopting their L1 as an identity

marker—a purposefully adopted socio-political emblem to indicate their opposition to perceived past and present hegemonies. These cultural identification positions are not incompatible with the acquisition of English as a second language (L2). Current research in the academy indicates that L1 competence can be used as a very valuable vehicle for L2 acquisition. Wheeler and Swords (2004) cite Rickford (1999) on the failure of traditional language teaching methods to achieve English language competence among African Americans, for instance (p. 471). In doing so, they point to the rich cultural and linguistic heritage that children of minorities and marginalized groups can call upon as they strategize for L2 literacy.

The importance of valuing L2 learners' identities, respecting their culture, and honouring their rich language heritage in the teaching-learning process bears directly on Krashen's Affective Filter Hypothesis (Piper, 2003, pp. 147–149). The implications of the affective filter hypothesis are that motivation, self-confidence, and anxiety-reduction over language processes need to be catered for in the teaching/learning situation. This means that attitudes sensitive to learners' psychosocial well-being, even beyond those of cultural identification raised above, are important; every facet of social relationships that impacts on the quality of teacher-student interaction and peer-group relationships needs to be taken into account. For instance, whether so-called remedial readers should be mainstreamed or put in immersion programmes, and for how long, are decisions that need to be considered in setting up literacy interventions.

In the context of reading motivation and engagement, current early 21<sup>st</sup> century literacy research with a particular focus on aliterate and at-risk adolescents has tended to take a technological perspective on the concept of “remedial.” For instance, citing Luke and Elkins (2000), Alvermann (2003) believes that instead of trying to “‘fix’ learners’ ... educators should be ‘fixing’ or ‘re/mediating’ the instructional conditions in which they learn” (Section 2, para. 2). Continuing to cite Luke and Elkins, she defines “re/mediating” as “refashioning curricular and instructional conditions so that they incorporate multiple forms of media (e.g., trade books, textbooks, magazines, newspapers, visual images, videos, CD-ROMs, music lyrics, and the Internet)” (Section 2, para. 3).

Although children from working-class backgrounds, such as 1-M students, are exposed to Alvermann's re/mediated world, most of them do not have access to the multiplicity of media she names, either at home or at school. Their access to re/mediation is on the more modest sociocultural scale outlined by Berghoff, Egawa, Harste, and Hoonan (2000), who in their own way question whether reading and writing that disadvantage and demotivate those with the least command of the formal language of school should be the predominant sign systems for meaning-making used for determining literacy. Thus they put forward a sign systems approach, which puts a wider spectrum of the arts at the centre of ways of knowing. In their view "language, art, music, drama, mathematics and movement are sign systems. They represent ways humans have learned to mediate the world in an attempt to make and share meaning" (p. 10). A sign systems approach to literacy is based on collaborative inquiry and democratic principles within the learning community in which it is practised (Berghoff et al., p. ix). The social environment in which transmediation of texts is produced encourages open-ended expression, critical thinking, and creativity.

The models discussed above outline the current position on reading. In a useful summary, Stahl (1997) notes that the various models and approaches have arisen from solutions to different problems (pp. 1–8). It is worth keeping Stahl's overview in mind in light of the polarization that still plagues reading instruction.

### **Methodology**

Because of the exploratory nature of this study, a multi-research design was employed. It was principally socio-linguistic, ethnographic, and discourse analysis in focus (Kamil, Mosenthal, Pearson, & Barr, pp. 80–84, 124–125). On these three fronts, the study examined the students and their reading from the point of view of "the people and process that create the data" (Piper, 2003, p. 25). Further, as a study that sought to "make sense of classroom worlds" (Baumann & Duffy-Hester, 2002), procedures undertaken involved aspects germane to teacher-research, in that they were conducted by a teacher-researcher, were theoretically driven, and were aimed at clarifying student and classroom realities through illustrative means, using interviews and students' work as artifacts (Kamil et al., pp. 7–15). The primary paper-based artifact was a

reading journal that the students kept, which included pages for drawing. Their assignments were open-ended, and after a reading they could express their thoughts in any way they chose.

A miscue analysis and two reading interviews were undertaken. For the miscue analysis, a modified version of the Goodman, Watson, and Burke (2005, pp. 63–76) marking procedures was employed, in which only the students' substitutions, non-word substitutions, omissions, and insertions were marked on the typescript. Repetitions and other complex miscues such as intonation shifts accruing from dialect variations were not included (see Fig. 1). This means that only the words the readers actually produced were marked. The reading on which the miscue analysis is based took place at the first class session using a narrative passage, and was tape-recorded.

Of the two reading interviews, the first, a modified version of the Burke Reading Interview (Goodman, Watson, & Burke, 2005), took place in the third class session (Appendix A). The second, a closing interview of the teacher-researcher's own making, took place in the penultimate session (Appendix B). In addition, students kept a reading journal and the teacher-researcher kept an observation journal.

One possible limitation of the study is that the reading and closing interviews were not conducted individually. The teacher-researcher wrote the questions on the blackboard one by one and read them to the students, who wrote their answers in their journals. The interviews were done in this way because it would have taken quite a few sessions to conduct the interviews orally. Additionally, because of the restive nature of the class, it was the best way to avoid major disruptions while the teacher-researcher was engaged with individual interviews.

Students who are not strong in reading usually also have writing difficulties. However, as their journal responses to the questions show, the students of Form 1-M were undeterred, using invented spelling to express themselves. The teacher-researcher monitored the two reading interview sessions to help with comprehension of the questions. However, in cases where a student did not want to answer a particular question, the student was not pressed. Wherever the number of responses recorded is less than 19, it follows that either students did not wish to

answer the relevant question or that they were absent on that day. All responses for any given session are included. Spelling emendations have been made to the reading and closing interviews. In all other samples of students' work, only punctuation has been inserted as a syntactic aid for readers of the following data analysis.

Each session had a planned reading focus, followed by free-writing, which students did in their reading journals. The purpose of the writing was to encourage students to see their own experiences as important texts, and to engage them as makers of texts. Paired reading and individual readings with the teacher-researcher were the two main forms of reading employed, but whole-class read-alouds were also done, and some sessions devolved into small-group round-robin reading. Reading was done in a variety of genres. For example, one session was devoted to reading and responding to consumer-oriented advertisement flyers. A library session allowed book choice and incorporated drawing, and, in addition to a session when the daily newspaper was the stimulus, other sessions included one-on-one reading with the teacher-researcher from realistic narratives and folktales. The students' prescribed class reader also provided expository and narrative readings, from which an emphasis on meaning-making and comprehension drove mini-lessons on phonics, spelling, word study, syntax, dictation, and editing.

An obvious problem was that a group of 19 mixed-ability struggling readers was too large for any one person to handle. However, having arranged with the school and the teacher to interface with the group in their two-and-a-half-hour Thursday morning reading session for the semester, and having seen their positive response on the first day to being taught by a stranger, it was felt that dividing the class would generate disappointment. Another alternative was to have their regular teacher help at each session. However, the students seemed to be looking forward to a change in normal routines. Therefore, the teacher-researcher worked with the group alone. In this study, the term *respondents* refers to students who were present at any given session. On an average there were 15 students present at each session.

The neurolinguistic implications of working with struggling adolescent readers lie outside the competence of this researcher, and so do not factor into the following analysis of data.

## **Data Analysis**

### **Question 1: How did the students of Form 1-M perceive reading?**

The discussion in this section is based on an analysis of students' responses to a modified version of the Burke Reading Interview (Appendix A). For purposes of readability, Standard English spelling emendations have been made to the responses.

The majority of the students saw reading as task-based and associated with "learning" school work. Two associated it with meaning-making while one wrote "I think reading is fun." Their responses to "What do you think reading is" included "helping you to learn" and "something you learn." The tendency to associate reading with mechanical contact with individual words was also evident in responses such as "learning new words," "learning to read and spell words," "a lesson to help you read and spell better." Connections with coherence and thought were not very evident. Reading seemed academically driven and extrinsically imposed.

The answers to the second question on the reading interview: "What do people do when they read," corroborated the impression that the students held a task-oriented, school-driven perspective of reading. Responses to this question also suggested that they thought that being able to read well improved a person's standing or redounded to countable gain: When people read they "get very intelligent," "they get better passes at exams," they "learn how to call and spell big words." This is not difficult to understand in a society where education is the prime source of social mobility. One student's response betrayed perhaps her own motivational issues with reading; when people read: they "sleep, sometimes pay attention."

Thus it was not surprising that students' strategies for dealing with words they did not know suggested interruptions in meaning-making and coherence. Most indicated that they wrestled with decoding and deciphering: "try to sound it out," "break it into syllables," "take my time

and break it up into syllables,” “spell.” Equally favoured was asking someone.

In response to the question to name good readers that they knew, the students named themselves, their classmates, and the teacher as good readers in equal proportion (5 responses each). The fact that five cited their classmates as good readers suggests that, as a group, they may have spent many of their school years in remedial classes where they would have known few competent readers. Also, only two indicated that people become good readers with practice: “they read plenty,” “they read often.” Most of the others associated reading competence with skills-based strategies such as spelling and sounding out words. Two other yardsticks of good reading that the students cited were vibrancy and good articulation in public performance: reading “loud and clear for everyone to hear,” and “pronounc[ing].” Additionally, becoming a good reader was associated with “learning” to read in the childhood stage: “she learnt her work when she was small,” “she went to school and learnt how to read.”

In the students’ opinion, good readers used reading strategies similar to theirs: “break up the word,” and “try to spell.” Only two students cited strategies such as “figure it out,” suggesting that, on the whole, students had not been exposed to metacognitive appraisals of the act of reading or themselves as readers.

The strategies they recommended for helping someone having difficulty reading are a compendium of frustration techniques, suggesting the ways in which they have been taught. These are:

- Tell them to learn to read. Help them
- Try harder
- Teach them to read and sound out
- Spell and sound out the words
- Help them or laugh at them accordingly
- Try to help them; read more than three times
- Give them to spell; let the teacher help them
- Tell them, or ask somebody
- Help them say it, or break it up
- Make them read everyday; help them with their work

- Tell them the word; ask teacher first; let them read it by themselves to me
- Show them or ask someone to help them
- Make them read it many times; make them spell the word they do not know

**Question 2: What would their specific areas of difficulty be with secondary school reading?**

A modified Burke Reading Interview (Goodman, Watson, & Burke, 2005) miscue marking procedure was used to analyse students' readings from the first session. An excerpt from a story entitled "Sunday with a Difference" (Forde, 1980), taken from an anthology "designed to encourage students entering secondary schools to develop an interest and appreciation for the short story" (Narinesingh & Narinesingh, 1980, p. v), and widely used in secondary schools, was used as the reading material. The excerpt was chosen for its likely authentic appeal, since it portrayed a beach scene to students who live on an island and are familiar with the atmosphere of a beach picnic. It could be anticipated that the students would have had difficulty with the reading since they were known to be struggling readers. However, since cultural background and contextual prior knowledge for the story were in place, their readings would provide insights into the links they made between decoding and authentic experience. More than this, it was important to get an insight into the kinds of difficulties they would have with the reading that they would be expected to do in the Language Arts, because after a year they would be mainstreamed into Form 1.

Figure 1 shows the readings of a group of five male students. The reading became a round-robin reading, because of the high interactivity of the class and their tendency to involve themselves in each other's affairs. Since it was the first class session with the teacher-researcher, small-group collaboration was allowed to give them security, and also because it revealed more about their individual and group literacy practices. The section each student read is numerically marked. Their actual words as sounded are inserted in the lines above the original. Repetitions and pauses are not reflected on the diagram. However, as an indication of their poor reading rate, it should be noted that the reading of the paragraph took three minutes and fifteen seconds.

Reader 1 showed a better syntactic grasp of sense-clusters and phrases than his four other classmates. He substituted the words “blue” for “dull,” and “set” for “sheet,” but tried to maintain the coherence of the clusters to which they belonged. The other four readers showed more severe levels of graphophonic, syntactic, and semantic difficulty. For instance, during one of his long pauses, Reader 2 accepted “calmness” for “crinkles” and “sprayed” for “separate” from Reader 1. He substituted both “will” and “were” for “would,” without making sense of the text, which he attacked word by word, showing sight word deficiencies. Reader 3 showed collocation inadequacies with “huff and puff” of the staple *Three Little Pigs* story, suggesting phonic deficiencies as well as a history of limited exposure to print. But, more significantly, his substitutions of short vowel words for long vowels words that would have made more sense in context (“bite,” and “hidings”) suggested a need for him to read in context, rather than in the isolation of short-vowel/long-vowel drill. Readers 4 and 5 showed a similar tendency to call words rather than to make sense of them, indicating a need for a blend of word study and meaning-making strategies.

As a group, the readers tended to have more problems with medial and final syllables than with initial syllables. The substitution of “careful” for “careless,” and the addition of the morpheme “ing” to “play” in Reader 4’s effort, also suggested the tendency to guess words rather than to use critical judgment and observation. Miscues such as the omission of word endings and other Creole features of pronunciation did not significantly impair meaning-making, but showed a need for greater attention to features of formal English.

The difficulties that students experienced with this figurative excerpt indicated that they had not had much exposure to narrative, the predominant genre of the early years of life and school. In discussing inadequate reading skills among communities with little exposure to print, Brice Heath (1994) notes that: “teachers (and researchers alike) have not recognized that ways of taking from books are as much a part of learned behavior as are ways of eating, sitting, playing games, and building houses” (p. 73). In genre-specific terms, the reading difficulties implied that most of Form 1-M did not have adequate schema for narrative, even at the literal level of comprehension. Additionally, their

lack of schema for the elements of setting in narrative spoke of inadequate exposure to strategies such as the writers' workshop, readers' theatre, and the author's chair.

Sundays  
**Sunday with a Difference**

The sun had not yet risen and in the dull morning light the sea was <sup>blue</sup> across  
of dark colour, it [sic] <sup>cole</sup> greens and <sup>green</sup> blues <sup>blue mering</sup> merging in a grey sheet of water. <sup>set</sup> ①

In an hour's time, I knew, everything would be changed. The sun would <sup>will</sup> <sup>changding</sup> <sup>were</sup>  
set studs of light in the crinkles of the waves, and the sea would separate <sup>sterd</sup> <sup>calmness</sup> <sup>sprayed</sup>

the shades of green from the shades of blue. The breakers would <sup>sheds</sup> <sup>sheds</sup> ② huff and  
puff against the rocks or <sup>bit</sup> bite into the sand or <sup>such</sup> <sup>holler</sup> search the holes and

hidings of little fish and crabs and water spiders. <sup>spayches</sup> ③ It would be very  
invisiting playing ground <sup>careful</sup> <sup>indeed</sup>  
inviting: a playground for the young and careless and idle. And so it

always was, every Sunday <sup>wary</sup> <sup>package</sup> ④ weather permitting. But this was a <sup>sunny</sup> Sunday  
<sup>the</sup> ⑤  
with a difference.

Figure 1. Modified miscue marking of students' reading.

In terms of reading behaviours, there was a tendency toward finger-tracing among the majority of boys. Readings took a long time with many hesitations and repetitions. Punctuation was ignored in the struggle to get past each word. Students read in a questioning tone as if rhetorically asking about the accuracy of the word being tackled. Phrasing was not in evidence, except in the case of two of the girls (sisters) who needed to be placed in more stimulating environments.

During reading, however, there was high interactivity within the class. The classroom furniture accommodated two to a desk, and students depended on their pair-partners for prompting with problematic words. But even outside of these pairs, students rallied around the reader, supplying words with great assurance and conviction. The person who was reading accepted the word uttered by the peer, whether it made sense or not. When that person was asked to read, however, his or her intonation changed to one of questioning, and words were supplied in turn whenever he or she paused. Peer-partnerships can provide useful support in reading activities; however, indiscriminate use can lead to dependency and can retard reading development, as was evident with Form 1-M.

### **Question 3: To what extent did the students' L1 affect their reading?**

As speakers whose L1 is a West Indian English-based Creole, the students tended to gloss final consonants /d/, /t/, and /s/ in their reading. /Th/ was also pronounced /d/. These transfers from Trinidadian Creole (TC) to their production of Trinidadian Standard English (TSE) did not impact meaning-making and construction of texts to any great extent, and so are not considered miscues in this paper.

However, the students' written personal texts suggested that their L1 impacted their syntactic coherence, which determined how they made and took meaning. In introducing herself, for instance, one of the more competent readers wrote:

My name is Lisa. I am thirteen year old. I live El Carmen Rode No: 1.  
My mother name is Juanita Valencia. My father name is Terrence Paul.  
I love hip hop music me and my friend. Her name is Monica. [Full-stops inserted]

The omission of the final /s/ in ‘years’ does not complicate meaning. But as the student’s ideas got more complex, the structure of TC syntax deepened. TC possession, indicated by the juxtaposition of two nouns (mother name; father name), became evident; and the elliptical intransitive TC “live” resulted in the oral cadence “I live El Carmen Rode.” Further, although the positioning of the subject “me and my friend” at the end of the sentence, and not in the normal SE frontal position, can be considered a TSE emphatic appositive, the use of “me and my friend” and not “my friend and I” are signals of the former being TC syntax, since the TC end-position appositive is usually used in speech in a fragmentary way to clarify, not emphasize, an idea previously stated. When asked to read what she wrote in a subsequent session, the student had difficulty, indicating that the orality which had initiated the writing was distant from her mind.

Beyond L1 orality, the students’ written personal texts suggested that speech impediments, and perhaps hearing disorders, also impacted the meanings they made from texts. For instance, in saying why traffic signs should be obeyed, Candice wrote “beaus you mant get your senf kill” for “because you might get yourself killed.” Without doubt, phonemic awareness is evident in this text, but although she was able to read what she wrote on the day that it was written, she had difficulty reading it in a later session. There are other areas of word recognition that bespoke complex relationships among pronunciation, rote-learning, and exposure to print for some students. Joel, for instance, could spell “children,” and “house” but had difficulty reading them contextually in print.

Differences between TSE and TC pronunciations sometimes impacted on responses to texts. For instance, students had difficulty reading the newspaper headline “3 Teens Steal Vault with Gun, Ammo” because of their unfamiliarity with “vault” in print. Since a prompt to use the news photo proved futile, the teacher-researcher gave the pronunciation, whereupon they scoffed, saying they knew what the word was all along, insisting that the correct pronunciation was “volt.” They subsequently

read the newspaper article with good comprehension, saying “volt” whenever they met the word.

Overall, the students had little difficulty expressing their understanding of most readings orally in TC, after the texts had been read. However, their lack of exposure to a wide variety of cultural experiences often shut them out of making meaning from print, even after they had conquered graphophonic and syntactic elements. This happened even with texts depicting local and Caribbean scenarios. Therefore, it is not adequate to argue that their reading material should be more culturally relevant. They live in the world, and so they need to widen their bases beyond their local culture, if only because the electronic media that they also have to read, literally and figuratively, contain large quantities of print.

The observations made in this section are exploratory. No deficit is implied. The deep structure of the students’ TC personal texts showed them to be no less adequate conveyers of sophisticated meaning for their lack of reading competence. Nevertheless, since the print they have to negotiate is coded in formal English, their attention needs to be drawn to cultural differences in meaning between L1 and L2 texts that share the same words, and sometimes share similar structures.

#### **Question 4: How did grouping the students as a unit impact their literacy?**

Discussion in this section focuses on the students’ behaviour as it relates to (1) reading engagement and (2) the class as a literacy unit. Data from the Closing Interview also inform the analysis (Appendix B).

Good peer relationships existed in the class, but there were a few loners among the boys. Reference has already been made to the pair partnerships that had developed between students who sat at double-desks in same-gender pairs. The girls were the stronger readers and, with the exception of one girl, did not mind moving to work with the boys. However, the boys were reluctant to work in cross-gender pairs, so that the only literacy stretching they could get in the classroom was denied them. Also, the way in which pair partners rallied around each other to share each other’s burden suggested that language fossilization had taken place for most of the students. In light of this dependency, it would have

been better if they were not streamed as they were for the entire day, so that they could be exposed to more challenging literacies.

Placement in a remedial class had initially brought feelings of low self-esteem. Nine of the 13 who responded to Question 2 of the Closing Interview about their feelings on being put in the class said they felt badly, sad, awful, or jealous when they found out that they were placed in Form 1-M. Five of the 13 had come to terms with their lot: “sad at first, but then normal” or “I feel bad, but I learn more to read and understand better.” One girl was vociferous about the treatment by one of the teachers and about the stigmatization of the class within the school community (“they trest us like dog”). One of the boys wanted to be integrated into Form 2 right away. On the whole, it was the stronger readers and writers who expressed frustration with their lot. All the same, after almost a semester as a unit, they had bonded and 9 out of 13 respondents said that they liked their class.

The biggest peeve of the girls was that their curriculum was limited to English and Math, and that their schooling was largely confined to their classroom, which they sarcastically referred to as the “office.” Their class teacher also felt that the group needed exposure to a broader curriculum, including at least the arts and physical education. The boys did not express strong feelings on the issue, perhaps because they had interests such as track and field and fishing that they pursued outside of school. One boy played football for the school.

However, there was a high desire within the class as a whole to read. They were aware that they had been singled out and one could sense that they would have liked to become anonymous within the mainstream. Most of students were not hesitant to try, and sometimes initiated reading on their own, aloud, when the session began. They impersonated the competence that they thought they should be showing. This initiating of reading seemed linked, simultaneously, to a desire to overcome as well as a denial that they had problems with print. Barry, for example, jumped the gun on the first day, showing much confidence, but his voice petered out after a minute or so. All of the students tried very hard. However, anxiety and confusion about what to do to sustain their efforts soon caused them to participate *sotto voce*.

In spite of these indications of a potential for reading engagement, students manifested negative behaviours imbibed from a legacy of learned classroom codes. They were very restive and their attention wandered when they were not receiving direct attention from the teacher. As a group, Form 1-M dictated the proceedings of their literacy encounters adversely. Very much like “the bottom group” of McDermott and Gospodinoff’s study (as cited in Au, 1997), much reading time was wasted “in positionings called ‘getting a turn to read’ and ‘waiting for the teacher’”(p. 185). The dynamics of their shuttered environment and learned behaviours forced the teacher-researcher to stop the lesson often “to deal with problems.” The students were “attuned to changes in the position of the reading group, and knew that they could get the teacher’s attention by interrupting during the lessons” (p. 185). Thus they ended up lessening the time that they spent reading.

**Question 5: How did the students respond to a mix of literacy stimuli?**

In this section, the students’ reactions to three kinds of reading stimuli are compared: (1) their basal reader, (2) copies of a daily newspaper, and (3) their choice of books from a fairly varied range, when they visited their disused school library<sup>5</sup>. Comments are also made on a session at which consumer-oriented advertisement flyers were used as the reading stimulus. Students’ reading journals provide the data for much of the analysis. The limited range of stimuli used is glaring, in light of current reading engagement theory such as that of Gee (2003) and Alvermann (2003), which notes the salutary impact of electronic technologies on students’ learning. Nevertheless, within the context of the realistic means available to Form 1-M, the analysis presented here provides insights into basic, literacy-instruction practices that can be implemented to make a difference in the students’ reading habits and, as a consequence, on their academic literacy.

The four readings done from the class reader produced the most cloned responses in students’ reading journals. Students showed a tendency to repeat phrases that were remembered from the reading. Few launched into connections with meaningful and/or personal experiences that would amplify the text. However, where individual students had strong feelings or interests, their writing was more expressive. On the reading about

sports heroes, Lenora, who had written previously about her disappointment at not being able to continue playing football at school, wrote about her admiration for Dwight Yorke, her local football hero, and continued:

I love football just as he do. But I just don't [know] when I will play again be[cause] in my school [Name omitted] Junior Secondary they have football only for the boys. I dont no why. . . I can\* wait to fine out when they will have football for the girls in this school.

\*TC Creole for 'can't'

Similarly, on the narrative reading entitled "The fishing trip," Kerry, who in introducing himself on the first day wrote, "I like to go river and cech fish," showed an eye for fisheries conservation. He also showed his observation of "fish" used as a national symbol in the Caribbean:

I think that the fishermen know mor about fishes [T]hey know that some fish are not to be catch like the flying fish. The flying fish is agenst the law in Barbados [I]f you look at the nasinal flag you will see it.

(capitalization inserted in brackets)

Apart from producing little extrapolation, in the four sessions in which the class reader was used, students tended to engage in distracting each other and in calling their friends outside. It was difficult to keep them on task. For much of the time they chatted or got into play-fights with their reading partners, while the teacher-researcher was engaged with individual students.

By contrast, the use of the newspapers was a highly successful literacy event. Reading the front page headlines whetted appetites, and students were eager to try out the articles of their choice. There was sustained attention to task and building of joint-personal texts out of intrigue. A wide variety of readings was chosen. Some girls' groups read four articles instead of the traditional single item that was customary for the basal reader sessions. In the basal reader sessions they expected to be directed. However, in the newspaper session they went through the material, choosing their own texts. Sensational articles emanating from familiar local scenarios drove reading interest. The chatter during reading tended to be extrapolation, as inferential thinking led to guessing words in context, based on prediction and speculation. The articles that caught

the attention of most groups were: “3 Teens Steal Vault with Gun, Ammo,” “Dangerous Sidewalk,” “Police Stations in the Dark for Divali,” and “Murder Accused Tries to Hang Self” about which Tricia wrote: “I think if you have a had [hard] time don’t hurt yourself or your family. . . . If I were him I will make I trust someone and tell them. He is very much so wrong for tring to hang himself.”

The session at which the consumer-oriented advertisement flyers were used proved the most unsuccessful. Most of the flyers promoted supermarket and furniture store Christmas bargains or advertised pizzerias. There were also fitness and health brochures, and pamphlets advertising banking services. Students looked at the flyers but argued there was nothing to read, although when asked to tell what the brightly coloured slogans said, such as “Let me Loose” and “Gift Certificate,” they seemed surprised to find words amidst the colour. With the exception of the four independent readers most could not figure out the words. When asked to order a delivery from the pizzeria using the menu, Kerry had problems. He could not recognize items such as “cheese,” “topping,” “orange juice,” and “soft drink.”

There was indeed much to read on the flyers, especially in the smaller print that the students glossed over; so that when pointed to limitations on attractive furniture such as “Fabric Subject to Change,” they balked in surprise. Needless to say they could not figure out what the words said.

The unsuccessfulness of this literacy event holds many implications. First, perhaps the teacher-researcher did not prepare the students adequately for the literacy event. It is also possible there was lack of motivation, since the advertisements were adult material that was likely to interest parents more than adolescents. However, their response to the session gave insights into what they thought reading was in terms of type of text, quantity of words, format, purpose, and seriousness. Their lack of engagement in the discussion prompts also indicated their lack of metacognitive awareness about reading in terms of their sensitivity to the world around them, and their political citizenship as consumers with regard to meaning-making, on levels that would affect their well-being as a group susceptible to exploitation.

Finally, although the full effect of Berghoff et al.'s (2000) sign systems approach to literacy cannot be gauged from this study, it was tried at one of the sessions. The sign systems approach to literacy does not privilege language as a medium of expression; it encourages students to transmediate text through other sign systems such as art, music, drama, mathematics, and movement. Because of limited resources, drawing was the medium used for the session, which took place in the disused school library. Students could choose from an assortment of texts that were made available by the librarian.

This session was very successful and suggests great potential for all-round literacy development, if done in a sustained way. The benefits were: (1) the motivation of students to interact with reading materials from a wide variety of genres, (2) students' attention to task and collaboration with each other in discussing their chosen texts and drawings, (3) the quantity of work produced, and (4) the variety of captioned interpretations expressed in art.

At the session, guidelines for the students were that they could choose any book they liked and draw or write about it in any way they wished. They could collaborate on pieces. They were to put the title of the book on their representations to give an idea of what the representations were about. Among the books that elicited interest were: a short biography of Jennifer Lopez, books on Caribbean folk tales, an illustrated book on Jamaica, a book on a South African family, pamphlet-sized story books, illustrated books on cricket, and books on animals. Some students browsed but did not choose. They engaged their own ideas, labelling their pictorial interpretations in their own way.

For example, Anselm drew and labelled warning signs: "Smoking is Death. It is Bad. It Kill us. Guns are Bad."

Henry's abstract pictorial interpretation was captioned in this way: "The name of the book I am read is Welcome to South Africa! It is a good place to live and for children to grow up. I love you."

Trevor, whose journal entry on the second day of class read, "I fiel vie Jreat at Reading two Mi friend. [I feel very great at reading to my friend.]," drew a perfect crab, modelled on the cover picture of the book

that he had selected. He inserted the book title: *Caribbean short stories for children*.

It is not the intention of this study to endorse any one approach to literacy, but the teacher who wants to engage the sign systems approach must be willing to shed traditional ideas of class operations. At the session, the students took a while to settle down. The boys, who sat mainly by themselves, flipped pages at first, just looking at the books. Peer influence was leading them in the direction of David, who said he could not draw and he did not want to write. But quite soon interest built up with the teacher-researcher's encouragement and praise. Gradually, students began to browse through the books they had chosen, interpreting them and chatting while they did so.

A wider range of expression and reading inquiry is encouraged by the sign systems approach to literacy. For, it is evident that Form 1-M has been exposed to skills-based training—Trevor's "fiel" and "two" are just two indications. What the students seem to lack in order to read and write better is adequate exposure to language in authentic and personalized usage. When Trevor copied out the title of the story to encapsulate his crab drawing, for instance, spelling and dictation were not enforced out of a prescribed reading for the day, in which he may not have been interested. The astute teacher can encourage peer-centred interest in reading and writing generated by students themselves. Interest in books can fuel study of texts, creating schemas and propelling vocabulary development that a class reader used in the confines of a classroom without resources cannot adequately build.

It is pathetic to see 14-year-olds who are normal in most respects being left out of the benefits of literacy because of language prejudice. The sign systems approach offered students ways of communicating on a multiplicity of integrated levels of literacy development, reading being only one of them.

## **Conclusion**

### **Question 6: What interventions will most likely impact the literacy development of the students?**

It is not surprising that 11 of the 15 students present at the closing interview (Appendix B) said that their reading had improved over the eight sessions. Three did not respond and one said “I am good in reading,” indicating that she knew how to read all along. The more important testimony to the usefulness of (1) the emphasis on meaning-making, (2) a variety of stimulus material, and (3) wider forms of expression on the students’ literacy development comes from their answers to the question on how they knew that their reading had improved.

By comparison with their first interview, in which the majority of questions on the reading process provoked skills-based answers such as “sound it out” and “break it up in syllables,” none of the 12 responses to this question mentioned skills-based yardsticks. Three students did not respond. However, the 12 other responses to how they knew they had improved either spoke of reading in a holistic context or in terms of the affective: “I read different books in the library,” “Because I can read at home,” “I can read much better,” “I see for myself,” “Because I did not believe in myself. But now I do.” While the students did not expand on their responses, it is clear that they did not come away from the sessions with the view that the spelling, dictation, and syllabification that were done as a subset of the readings were what the readings were about. These observations are instructive in light of reading research that endorses phonics instruction for the junior years of primary schooling, but is silent on the effects of phonics instruction among adolescents. Of interest, too, is the impact of the affective.

Au (1997) corroborates these essential understandings that should be borne in mind for dealing with struggling readers, and indeed for students generally:

Research on school literacy learning conducted from a sociocultural perspective proceeds on the assumption that students need to engage in authentic literacy activities, not

activities contrived for practice. School literacy activities should involve the full process of reading and writing, not the practices of skills in isolation. (p. 183)

Au also refers to the work of Vygotsky; in particular, his concept of the zone of proximal development, which reflects on the dangers of “the use of fixed ability groups for reading instruction” (p. 183).

Students of Form 1-M are unlikely in the near future to have the technological tools to practise the full range of alternative literacy strategies required to fast-track their academic careers. However, respected reading research suggests that the concepts of reading under which they currently operate must be revamped appreciably to get them moving. For instance, Au also suggests a programme that concentrates less on reading aloud and more on comprehension, scaffolding discussion, discussion of unfamiliar vocabulary, meaning-making, and “literacy instruction organized in a manner responsive to and accepting of students’ home culture and language” (pp. 188–189).

Needless to say, too, the student-teacher ratio needs to be lessened if students in Form 1-M are to be helped. McCormick (2003) cites in detail Eldridge’s successful, three-year, intensive intervention, one-on-one for two hours a day, with “a reader, who previously had profound difficulties” (p. 424). Although a one-to-one ratio for Form 1-M may not be feasible because of cost, the teacher-student ratio will have to be reduced in order to achieve the best results.

In order to give greater prominence to an analysis of the students’ difficulties, the wider politics of literacy has been marginalized in this exploratory paper; however, the impact of the politics of literacy on children’s lives cannot be ignored. In this regard, advocacy for students, particularly at the stigmatized end of the secondary school system, is sorely needed. An ombudsperson is perhaps needed to intervene for those who have no voice and whose parents are unable to operate in their best interest, because they lack enough knowledge of the system to ask the right questions. The range of 0–30% encompasses a wide span of abilities. The experience with Form 1-M indicates that the literacy of children designated as “remedial” should be constantly reviewed in pull-out programmes that operate on student need, in order to prevent

demotivation, loss of self-esteem, and wastage of the crucial formative years of adolescence. It should not be left up to individual schools to set policy for the children designated as remedial. As well-intentioned as the efforts of these schools may be, strategies such as giving those who are weak in literacy and numeracy, English and Math to do all day are counter to education and human development.

Additionally, advocacy is imperative for this group because both public and private opinion within the teaching system in Trinidad and Tobago are split about whether students who have scored so low on primary school exit assessment should be “promoted” to secondary school. Keeping the students in a primary system that has failed them seems hardly the issue. A more relevant consideration seems to lie in an examination of the quality of teacher preparation for reading that occurs throughout the education system. Also, attention needs to be paid to sourcing neurolinguists, reading specialists, and speech pathologists. However, generations of students cannot be allowed to pass through the system, waiting for these personnel, before new conceptualizations can begin.

One of the more important points that this exploratory paper has brought out is that reading is a subset of wider literacy development practices, such as visual and aesthetic appreciation from a variety of sign systems. In this regard, low-end technology that is within the reach of Form 1-M, such as recording devices, needs to be exploited. This needs to be said in a country where reading in an electronic age still means, by and large, reading from a paper book; where reading very often still means being given the next few sentences to prepare for homework; where reading is still conceived by many parents and teachers as they knew it in their childhood. It is now accepted that there are many ways to read. Schools must be compelled to open the wide doors of literacy to all.

## Notes

1. English or Language Arts in Form 1 Remedial classes at secondary schools is taught by teachers with some form of reading experience or training. The teachers are contracted by the Ministry of Education and are provided with teacher packages. Many of them are retired primary school teachers. The cadre is also composed of people who have had training in adult literacy teaching. Originally, Form 1 Remedial classes were called Form 1 Special (1S) and a draft syllabus was written for their use (Trinidad and Tobago. Ministry of Education, 2001). The summary of work for 1Rs is also included as the last section of the new Form 1 Secondary Education Modernization Programme (SEMP) syllabus. After this article was accepted for publication, the Minister of Education announced that students scoring 0–30% on the primary school exit exam would repeat the year at their respective elementary schools (Connelly, 2007). The teachers' union welcomed the move but expressed concern over the timing of the Minister's release, since some students who would have expected to move on to secondary school in the new academic year beginning September 2007 would no longer do so.
2. No real names are used in this study.
3. Informal interviews done between September to November 2005.
4. Later checks revealed that the students were put in Form 1.
5. The school library was not being used for its intended purpose during the period of the study. Special arrangements were made with the librarian for use of the facility by the researcher for the reading sessions.

## References

- Alvermann, D. E. (2003). *Seeing themselves as capable and engaged readers: Adolescents and re/mediated instruction*. Naperville, IL: Learning Point Associates. Retrieved January 25, 2007, from <http://www.learningpt.org/pdfs/literacy/readers.pdf>.
- Au, K. H. (1997). A sociocultural model of reading instruction: The Kamehameha Elementary Education Program. In S. A. Stahl & D. A. Hayes (Eds.), *Instructional models in reading* (pp. 181–202). Mahwah, NJ: Lawrence Erlbaum Associates.
- Baumann, J. F., & Duffy-Hester, A. M. (2002). Making sense of classroom worlds: Methodology in teacher research. In M. Kamil, P. Mosenthal, D. Pearson, & R. Barr (Eds.), *Methods of literacy research* (pp. 1–22). Mahwah, NJ: Lawrence Erlbaum Associates.

- Berghoff, B., Egawa, K. A., Harste, J. C., & Hoonan, B. T. (2000). *Beyond reading and writing: Inquiry, curriculum, and multiple ways of knowing*. Urbana, IL: National Council of Teachers of English.
- Brice Heath, S. (1994). What no bedtime story means: Narrative skills at home and school. In J. Maybin (Ed.), *Language and literacy in social practice* (pp. 73–95). Milton Keynes, England: Multilingual Matters.
- Connelly, C. (2007, April 6). Poor SEA performers to repeat exam...*Trinidad Guardian*, p. 6.
- Dechant, E. (1991). *Understanding and teaching reading: An interactive model*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ferguson-Briggs, G. (2002). *Gateway to reading*. Chaguanas, Trinidad: Caribbean Educational Publishers.
- Forde, A. N. (1980). Sunday with a difference. In R. Narinesingh & C. Narinesingh, *Sunshine and shadow: An anthology of short stories* (pp. 81–85). London: Cassell.
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.
- Goodman, K. (1996). *On reading*. Portsmouth, NH: Heinemann.
- Goodman, Y. M., & Marek, A. M. (Eds.). (1996). *Retrospective miscue analysis: Revaluing readers and reading*. Katonah, NY: Richard C. Owen.
- Goodman, Y. M., Watson, D. J., & Burke, C. L. (2005). *Reading miscue inventory: From evaluation to instruction* (2<sup>nd</sup> ed.). Katonah, NY: Richard C. Owen.
- Kamil, M. L., Mosenthal, P. B., Pearson, P. D., & Barr, R. (Eds.). (2002). *Methods of literacy research*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Marek, A. M., & Goodman, Y. M. (1996). Understanding the reading process. In Y. M. Goodman and A. M. Marek (Eds.), *Retrospective miscue analysis: Revaluing readers and reading* (pp. 21–38). Katonah, NY: Richard C. Owen.
- McCormick, S. (2003). *Instructing students who have literacy problems* (4<sup>th</sup> ed.). Upper Saddle River, NJ: Pearson Education.
- Mohammed, S. (2005, October 27). 3 teens steal vault with gun, ammo. *Newsday*, p. 3.
- Narinesingh, R., & Narinesingh, C. (1980). *Sunshine and shadow: An anthology of short stories*. London: Cassell.
- Piper, T. (2003). *Language and learning: The home and school years* (3<sup>rd</sup> ed.). Upper Saddle River, NJ: Pearson Education.
- Purcell-Gates, V. (2002). “. . . As soon as she opened her mouth!”: Issues of language, literacy, and power. In L. Delpit & J. Dowdy (Eds.), *The skin that we speak: Thoughts on language and culture in the classroom* (pp. 121–141). New York: New Press.

- Siegel, M., & Fernandez, S. F. (2002). Critical approaches. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Methods of literacy research* (pp. 65–76). Mahwah, NJ: Lawrence Erlbaum Associates.
- Stahl, S. A. (1997). Instructional models in reading: An introduction. In S. A. Stahl & D. A. Hayes (Eds.), *Instructional models in reading* (pp. 1–29). Mahwah, NJ: Lawrence Erlbaum Associates.
- Trinidad and Tobago. Ministry of Education. (2000). Secondary Education Modernization Programme (*SEMP*) draft secondary school curriculum - *Form One: Language arts*. Port of Spain: Author.
- Trinidad and Tobago. Ministry of Education. (2001). *Draft operational framework and guidelines for teachers working with students in Form 1 Special*. Port of Spain: Author..
- Wheeler, R. S., & Swords, R. (2004). Codeswitching: Tools of language and culture transform the dialectally diverse classroom. *Language Arts* (81)6, 470–480.

## **Appendix A**

### **Modified Burke Reading Interview**

1. What do you think reading is?
2. What do people do when they read?
3. When you come to a word you do not know what do you do?
4. Who is a good reader that you know?
5. What makes them a good reader?
6. Do they come to a word or something while reading that they do not know?
7. What do they do?
8. What would you do to help someone having difficulty reading?

## **Appendix B**

### **Closing Interview**

1. Do you read out of class? What books do you wish you could read?
2. How did you feel when you found out that you were put in Form 1-M?
3. How do other children in the school treat you? Do you like your class?
4. Since we began reading, do you find that you are improving?
5. How do you know?

## **STUDENTS' ATTITUDES TO SCIENCE IN URBAN AND RURAL SCHOOLS IN TRINIDAD AND TOBAGO**

*Rawatee Maharaj-Sharma*

This study is a pilot project that attempts to analyse, on a small scale, students' attitudes towards science in two different localities—one urban and one rural—in Trinidad and Tobago. Though the sample was small, the study suggests that students attending urban schools generally display a more positive attitude toward science—in terms of its perceived relevance and usefulness as well as its interest and challenge level—than students attending rural schools. Furthermore, the project revealed that the attitudes displayed by boys were different from those displayed by girls within each locality, and that the gender difference was more pronounced in the rural schools. Additionally, it was found that urban girls had the most positive attitudes towards science, while more than 80% of rural girls perceived science as being difficult, with many of them indicating that it was not a subject they were inclined to pursue at higher levels.

### **Introduction**

In Trinidad and Tobago, great significance is placed on terminal examinations, which are usually external to the school and which are the predominant focus of both teachers and students. Teachers are hard-pressed by public opinion and school administrators to produce good examination results, as they are made to feel that their competence is reflected in these results. There is a perception among many teachers and administrators that students' interests and desirable attitudes would be automatically nurtured through the application of the "correct" pedagogical principles of teaching for the test. Thus, teaching has been geared to ensuring that content is covered, while the long-term attitude development is left unattended (Ibrahim, 1984). In the local context, this neglect of emphasis on the development of positive attitudes has resulted in many students expressing sentiments of displeasure and unwillingness to pursue certain subjects on the national curriculum. Expressions of this

nature are very common among science students, many of whom have suggested that science is “too abstract,” “difficult,” and “no fun.” The result is that many students pursue science learning by rote, and many only because it is a subject requirement for a future course of study that they intend to undertake. It is not surprising therefore that many students clinically experience the science teaching/learning process but do not develop desirable attitudes towards science (Hu, 2003).

This project attempts to investigate the attitudes to science of students in two different localities—one urban and one rural—in Trinidad and Tobago. It is a pilot project as it is the first phase of a more comprehensive study that will comparatively explore students’ attitudes towards science in both localities on a larger scale by: (a) surveying more schools, and therefore more students, in each locality; (b) examining other aspects of attitudes that may emerge based on the findings of this project; and (c) making any necessary amendments to the design and instrumentation of this project to gain further insights into students’ attitudes towards science. The larger study will probe deeper into the underlying beliefs and assumptions that students may have developed through interactions in the home environment, the school environment, and in the social arena, which may have influenced their attitudes towards science.

The participants surveyed for this project comprised students who had undergone two years of Integrated Science or General Science in six New-Sector schools in Trinidad and Tobago. These schools have been described as:

government schools, [which] emerged during the 1970s and 1980s as a result of the government’s efforts to expand the secondary school intake and broaden the curriculum. The New Sector consists mainly of a two-tiered system incorporating the three-year junior secondary schools, with transitions into the two-year or four-year senior secondary/senior comprehensive schools. It also encompasses the five-year composite schools. Schools in this sector offer courses in both academic and technical/vocational subjects. (Quamina-Aiyejina, L., et al., 2001, p. 2)

Three urban schools and three rural schools were selected. It is important to note that the Trinidad and Tobago context is somewhat unique in that any typical urban school will have a small but not negligible percentage of suburban and rural students. The converse is also true for some rural schools. In this project, the influence of this variability among the students was significantly minimized by analysing the biographical data of the students and subsequently eliminating those students who lived in localities different to the locality of the school they attended. Since this is a pilot project, only two aspects of students' attitudes were assessed:

1. Interest in and/or enjoyment of science.
2. Awareness of the importance of science to social and economic life.

Interest, in this context, refers to liking and having a positive view and personal desire to pursue science learning, while awareness refers to the knowledge and understanding of the relevance and usefulness of science to one's well-being in everyday life and for future plans. For this research, the term *attitude* will be used to describe the combined effect of students' interest and awareness as described above, which is defined, according to Bednar and Levie (1993), as affective responses, cognitions, behavioural intentions, and behaviours. Other definitions include "a manner of acting, feeling or thinking that shows one's disposition and or opinions" (Penguin Webster Handy College Dictionary, 2003). Zimbardo and Leippe (1991) suggest that "attitudes are learned or established predispositions" that vary in direction (positive or negative), degree (amount of positive or negative feeling), and intensity (the level of commitment the individual has to the position). In this project, the latter interpretation will be implied in the reporting.

### **Research Questions**

Three research questions tailored the approach adopted in this project:

1. Do urban students display attitudes different to those of rural students towards science?
2. Do boys and girls display different attitudes towards science?
3. How do students' attitudes toward science relate to their view of its importance to their lives?

## **Literature Review**

Fishbein and Ajzen (1975) suggested that a person's attitude towards an object is a function of his or her beliefs about the object and the implicit evaluative responses associated with those beliefs, while Bednar and Levie (1993) described attitudes by saying that they are not directly observable, but the actions and behaviours to which they contribute may be observed. These behavioural descriptions arising out of beliefs and responses, in large part, govern an individual's perception. With respect to students, Osborne, Simon, and Collins (2003) found that students' attitudes had a significant impact on their views of subject matter, particularly in the area of science, when constructs such as value of science and usefulness of science were used to gauge their attitudes.

Furthermore, attitudinal differences towards science between boys and girls were noted in several studies (e.g., Leventhal & Brooks-Gunn, 2000; Marsh, 1992; Weinburgh, 1995). The first study found that the neighbourhood in which a child lives has a significant impact on his or her attitude towards science learning, to the extent that urban students generally had more positive attitudes towards science than students schooled in rural settings. Roscigno and Crowley (2001) reported that urban boys displayed a more positive attitude toward science, indicating that it would be useful to them in the future, while urban girls reported that science was fun to learn. DeYoung (1987) reported that, generally, rural boys thought that science was relevant and enjoyable and that rural girls indicated that science was a difficult subject; while Jones, Howe, and Rua (2000) found that, in general, across urban, suburban, and rural communities, significantly more females than males reported that science was difficult to understand, but that more males were interested in pursuing science as they saw it as necessary for a future career. It is clear, therefore, that the existing literature suggests that there are differences in attitudes toward science between urban and rural students, as well as differences between boys and girls in each of the two localities.

## **The Sample**

The sampling unit was a school. A rural school is defined in this project as a school that is located outside a radius of 12 km from any major city, borough, or town in Trinidad and Tobago. Given the local context in which the project was done, this definition of a rural school was arrived at based on the defining limit as described by Leventhal and Brooks-Gunn (2000). The three rural schools selected were located at distances of 13, 17, and 25 km respectively from the closest town, and were purposively chosen from the north, central, and south regions of the island respectively.

Similarly, as described by Leventhal and Brooks-Gunn (2000), in the local context, an urban school is defined as a school that is located within a 12 km radius of any major city, borough, or town. Two urban schools were chosen from the north and one from the south. Because this was a pilot project that will be used to inform a larger, more comprehensive study, six schools selected from among the eight educational districts provided a fair representation of students (though on a very small scale) across the country.

The students surveyed were in their third year of secondary school and were all aged between 14–16 years. These schools practise ability grouping or streaming, and so in order to arrive at a random sample, a stratified random sampling was carried out in all the third-year science classes in each of the six schools selected.

### **Urban sampling**

Most of the classes in these schools were small, with less than 20 students in a class on any given day. A total of 201 urban students volunteered for the survey but only 189 actually completed the opinionnaire, of which 15 were incomplete and were therefore not included in the project. A total of 60 urban students, comprising 31 boys and 29 girls, were selected from the three urban schools. To maintain some degree of uniformity, in terms of the number of students selected from each of the three urban schools, 17, 21, and 22 students were respectively selected from each of the three schools, since in one of these

schools only 17 opinionnaires had intelligible data that was relevant to the project.

### **Rural sampling**

A similar selection criterion was use in the case of the three rural schools, where a total of 181 students from the three rural schools completed the opinionnaire. In this locality, 28 of the opinionnaires were not included in the project. A total of 58 rural students—17, 19, and 22, respectively, selected from each of the three rural schools—were included in the project. The project therefore involved the participation of a total of 118 students—60 urban and 58 rural. Table 1 shows the distribution of the sample by gender and locality.

**Table 1. Distribution of Sample by Gender and Locality**

<b>Locality</b>	<b>Sex</b>		<b>Total</b>
	<b>Male</b>	<b>Female</b>	
Rural	26	32	58
Urban	31	29	60
<b>Total</b>	<b>57</b>	<b>61</b>	<b>118</b>

### **Instrument**

A Likert-type opinionnaire (Trochim, 2002) was used to collect data for this project. An instrument of this type was chosen as it provided an easy way of gauging students’ responses. This was important because many classes in the New Sector schools chosen had students with reading/comprehension difficulties, so that the Likert instrument provided the researcher with an opportunity to easily and accurately manage students’ responses (see Appendix A). A pilot testing of the instrument was carried out several weeks before the data collection phase of the project started. Some items in Section A were reworded to remove ambiguities identified in the pilot-testing phase before the instrument was used in this project.

The data collection process began by reading the statements and response options to all the students (sometimes several times), and then allowing

them to record their responses on the Likert range from *Strongly Agree* to *Strongly Disagree*. Some of the students had reading difficulties and so additional time was spent reading the statements in Section A to these students to ensure that they understood what they were required to respond to. They were subsequently asked to circle their response to each question.

Section A of the opinionnaire consisted of 12 opinion statements that sought to assess students' attitudes towards science. The statements were all negatively skewed to allow for consistency, ease of analysis and interpretation, and to avoid ambiguity for students.

Additionally, though, because it was practically impossible to capture all aspects of students' attitudes from the characteristic closed items of the Likert opinionnaire, three open-ended questions were also included as Section B of the opinionnaire. These questions provided students with further opportunity to explain their attitudes towards science and to suggest behaviours, events, or occurrences that might have shaped their attitudes. This section also allowed them to articulate in writing the extent of their awareness of the role of scientific knowledge in their lives. Some students also had difficulties writing responses to the open-ended questions, and so the questions in this section were also read to them. Their oral responses were audiotaped shortly after they completed section A of the opinionnaire. The audiotapes were subsequently transcribed and coded.

### **The Data**

The data obtained from Section A of the opinionnaires were used to identify codes and relevant labels that captured different aspects of students' attitudes. The written responses to the open-ended questions given by students were analysed against these labels and new/emerging ones were added. The data, including the codes and labels assigned to the written responses, were reviewed several times to ensure that the codes and labels reflected the attitudes expressed by the students. Attitude statements with similar labels were grouped together under a theme that represented the data in a broader sense. In so doing, five broad themes emerged from the data: "relevancy of science," "usefulness of science," "enjoyment in science," "difficulty in science," and "science is not for

me.” In the context of this project, the themes “relevancy of science,” “usefulness of science,” and “enjoyment of science” were interpreted as contributing toward a positive attitude to science, while “difficulty in science” and “science is not for me” were interpreted as contributing toward a negative attitude to science.

The numbers in Table 2 represent the actual number of students (boy or girl; urban or rural) who expressed views on each of these themes. The themes emerged after careful examination and coding of the raw data.

**Table 2. Emerging Themes of Students’ Attitudes to Science**

Attitudes	Urban		Rural	
	Boys {31}	Girls {29}	Boys {26}	Girls {32}
Relevancy of science (everyday understandings)	20	23	15	15
Usefulness (future education/career)	26	24	12	3
Science enjoyment	28	28	18	5
Science is difficult	10	9	16	26
Science is not for me	12	6	20	30

### Findings

The numbers in Table 2 were translated into percentages to comparatively report the findings of this project (see Table 3).

Similar to the findings of Leventhal and Brooks-Gunn (2000), this project revealed that there are differences in the attitudes of urban and rural students towards science. In general, urban students showed greater interest in science learning and were more aware of its importance to the environment and to their livelihood. However, contrary to what was reported by Roscigno and Crowley (2001), this project revealed that girls in the urban locality were slightly more enthusiastic than boys about learning science, with 79% of the girls indicating that science was relevant to them academically and socially. The boys in these schools indicated that they had interest in science and that they saw it as being a

relevant subject worth pursuing, but only 64% of them indicated explicitly that it was relevant to their lives. This suggests that they may have slightly less desirable attitudes to the pursuit of science than urban girls.

As for the usefulness of science, however, 84% of the boys surveyed said “it was useful to their life” but that “the classes were boring.” An almost equal percentage of urban girls (83%) reported that science was useful to their lives. Furthermore, in terms of perceived difficulties, again an almost equal percentage of urban girls and urban boys (32% and 31% respectively) felt that science was difficult.

In the rural schools, however, the effect was slightly different, with only about 58% of the boys in this locality reporting that they had a strong inclination to science learning based on their view of it being relevant. Only 47% of the rural girls opted to pursue science for its relevance to their lives. A total of 46% of the rural boys indicated that science was useful to them but only 9% of rural girls felt that science was useful to their lives; while 62% of rural boys and 81% of rural girls perceived science to be a difficult subject. Despite the perceived difficulties, the responses obtained from 69% of rural boys suggested that they found science to be “interesting and exciting” while 46% thought that it was “important for their future studies.” In contrast to the boys in the rural schools, only 16% of rural girls indicated that they enjoyed science, while many of the other girls in this locality reported that science was “boring” and that it was “no fun learning science... because they don’t understand the meanings of the words used.” In addition, 81% of rural girls also suggested that the “calculations were too hard,” with many of them indicating that they would “not be needing this in their future life.”

There was a unique theme emerging from the data, based on several responses obtained, which suggested that some students felt that science was “not meant for them.” This view seemed to be more common among the students in the rural schools and more specifically among the girls in this locality. Examples of actual verbatim responses suggestive of this include “this is not for me,” “I wasn’t cut-out for science,” and “my mind cannot do science.” The available data indicated that only 21% of the urban girls felt this way as against an overwhelming 94% of the rural

girls. It would be interesting to explore the reasons behind this feeling/belief in the larger study.

In terms of the combined numbers (boys and girls) in both localities, the project revealed that:

1. 88% of urban students indicated that science was relevant to their lives, while only 52% of rural students viewed it as a relevant pursuit;
2. 83% of urban students reported that science was useful to them while only 26% of rural students saw the usefulness of science to their lives;
3. 93% of urban students reported that they enjoyed science but only 40% of rural students indicated that they enjoyed science learning;
4. 32% of urban students felt that science was difficult but 72% of rural students felt that it was difficult; and
5. 30% of urban students indicated that science “was not for them” but 86% of rural students expressed this sentiment.

Table 3 summarizes the findings of this pilot project. Columns 2–8 give the findings with respect to the emerging themes listed in Column 1 in the following categories:

1. The attitudes of **all** the students surveyed across both localities (Column 2)
2. The attitudes of **all** urban students (Column 3)
3. The attitudes of urban boys (Column 4)
4. The attitudes of urban girls (Column 5)
5. The attitudes of **all** rural students (Column 6)
6. The attitudes of rural boys (Column 7)
7. The attitudes of rural girls (Column 8)

**Table 3. Percentage of Students in Each of the Emerging Themes**

Columns							
1	2	3	4	5	6	7	8
Themes	All Students	Urban Students			Rural Students		
		All	Boys	Girls	All	Boys	Girls
Relevancy	62	88	64	79	52	58	47
Usefulness	55	83	84	83	26	46	9
Enjoyment	67	93	90	97	40	69	16
Difficulty	52	32	32	31	72	62	81
Science is not for me	57	30	39	21	86	77	94

### Conclusion

Within recent years, some attempts have been made to include affective learning outcomes in science curriculum documents. However, teachers have indicated that it is difficult for them to incorporate these into their classes, either because they have not been trained to do so or because of the limited time in which to cover the science content of the respective syllabuses. This project suggests that two gaps, in terms of attitudes, exist among the students surveyed. The first is the difference in attitude toward science between urban students and rural students, and the second is the difference in attitude between girls and boys towards the subject. This seems to indicate, therefore, the need for teacher training in areas such as the use of more innovative teaching strategies, lesson planning, and time management skills in an attempt to bridge the gap in both cases.

Among the urban students, only a few more girls than boys found science to be useful, relevant, and enjoyable but a few more boys than girls indicated that science was a difficult and unappealing subject. In general, however, urban boys and urban girls seem to share similar feelings and views about the various aspects of science education explored in this project.

Among the rural students, however, there were large differences in the attitudes expressed by boys and girls, especially with respect to the usefulness, enjoyment, and difficulty of science—more boys than girls expressed positive attitudes in these areas. As for the relevance of science, only a few more rural boys than girls thought that science was a relevant pursuit. In a very general sense, this project revealed the following:

1. Urban students had more positive attitudes to science than rural students.
2. Urban girls had more positive attitudes to science than urban boys.
3. Rural boys had more positive attitudes to science than rural girls.
4. Despite the locality in which students were schooled, those who viewed science as enjoyable, relevant, and useful to their lives had more positive attitudes towards the subject.

### **Discussion**

It is important to remember that this is a pilot project, and that the findings cannot be generalized to all schools in each of the localities identified or even across gender borders. No tests of significance were done, as this project was meant to give an overview, in a broad sense, of some of the differences in attitudes that exist across locality and gender. Indeed, a more detailed study that probes deeper into the various aspects of attitudes, with a much larger sample, is needed to further substantiate the findings obtained in this project.

It is interesting, however, to note that the findings here were similar to that obtained by Leventhal and Brooks-Gunn (2000) and Weinburg (1995), in that more students schooled in urban settings exhibited positive attitudes towards science than those schooled in rural settings. Further, in general, more boys than girls had a positive attitude towards science. However, the findings of this project in the urban context were different to that reported by Hu (2003), who reported that in urban areas more boys than girls had a positive attitude towards science.

In view of the findings of this project, it can be speculated that there is perhaps an unspoken/subtle relationship between students' social, economic, and cultural backgrounds and the way they view/perceive science education. It would be interesting to investigate this supposition in science, in the first instance, and then in other curriculum areas, as this project did not probe into the influence and/or possible relationship/s that these variables may have on students' attitudes towards science learning.

Several aspects of students' economic, social, and cultural backgrounds and experiences (in both localities) will be explored in the larger study in an attempt to gauge the impact that these variables have on students' overall attitude towards science learning. Furthermore, the influence/s of students' cultural contexts may also have a bearing on the way they view science and the attitudes they develop toward the discipline. Consideration of this perspective is an excellent topic for future research.

### References

- Bednar, A., & Levie, W. H. (1993). Attitude-change principles. In M. Fleming & W. H. Levie (Eds.), *Instructional message design: Principles from the behavioral and cognitive sciences* (pp. 283–304). Englewood Cliffs, NJ: Educational Technology Publications.
- DeYoung, A. J. (1987). The status of American rural educational research: An integrated review and commentary. *Review of Educational Research*, 57(2): 122–148.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Hofman, H. H. (1977). An assessment of eight-year-old children's attitudes toward science. *School Science and Mathematics*, 77(8): 662–670.
- Hu, S (April 29, 2003). Educational aspirations and postsecondary access and choice: Students in urban, suburban, and rural schools compared. *Educational Policy Analysis Archives*, 11(14). Retrieved from <http://epaa.asu.edu/epaa/v11n14/>.
- Ibrahim, A. B. (1984). Attitudes to science of students in Sarawak. *Journal of Science and Mathematics Education in Southeast Asia*, 7(1): 10–14.
- Jones, G. M., Howe, A., & Rua, M. J. (2000). Gender differences in students' experiences, interests, and attitudes toward science and scientists. *Science Education*, 84(2), 180–192.
- Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhood they live in: The effects of neighborhood residence on child and adolescent outcomes. *Psychological Bulletin*, 126(2), 309–337.

- Marsh, H. W. (1992). Content specificity of relations between academic achievement and academic self-concept. *Journal of Educational Psychology, 84*(1), 35–42.
- Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25*(9), 1049–1079.
- Penguin Webster Handy College Dictionary* (3<sup>rd</sup> ed.). (2003). New York: Penguin.
- Quamina-Aiyejina, L., Mohammed, J., Rampaul, B., George, J. M., Kallon, M., Keller, C., & Lochan, S. (2001). *A baseline study of the teacher education system in Trinidad and Tobago* (Monograph Series, No. 1). St. Augustine, Trinidad: School of Education, UWI.
- Roscigno, V. J., & Crowley, M. (2001). Rurality, institutional disadvantage and achievement/attainment. *Rural Sociology, 66*(2), 268–292.
- Trochim, W. M. K. (2002). *Research methods knowledge base* (2<sup>nd</sup> ed.). New York: Longman.
- Weinburgh, M. (1995). Gender differences in students' attitudes towards science: A meta-analysis of the literature from 1970–1991. *Journal of Research in Science Teaching, 32*(4), 387–398.
- Zimbardo, P. G., & Leippe, M. R. (1991). *The psychology of attitude change and social influence*. New York: McGraw-Hill.

## Appendix A

### INSTRUMENT TO DETERMINE STUDENTS' ATTITUDES TOWARDS SCIENCE (3 URBAN SCHOOLS AND 3 RURAL SCHOOLS)

Please complete both sections of the questionnaire as completely and honestly as you can. If you are unsure about any question, please raise your hand and ask for assistance.

THIS IS NOT A TEST

#### Section A

	Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	Science is boring					
2	Only bright people can do science					
3	Science is a difficult subject					
4	Science is no fun					
5	Science is not useful to me					
6	Science is too abstract					
7	Science cannot help me solve everyday problems					
8	Science is not interesting					
9	Science is too much mathematics					

10	Science is not important to my future education					
11	I dislike attending science classes					
12	I will skip science class if I can					

**Section B**

1. Write a short description of **ALL** your feelings about the subject Science.

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

2. Do you feel that Science is useful to your everyday life? Why?

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

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

3. What do you believe is the importance of Science to you now and later on in your life?

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

Thank you for your assistance.



**DIFFERENTIAL ITEM FUNCTIONING AND  
MALE-FEMALE DIFFERENCES IN A  
LARGE-SCALE MATHEMATICS ASSESSMENT  
IN TRINIDAD AND TOBAGO**  
**An Examination of Standard 1 Mathematics Assessment**

*Launcelot I. Brown and Gibbs Y. Kanyongo*

This study investigates gender differences and the existence of gender-related differential item functioning (DIF) in a large-scale Standard 1 mathematics assessment in Trinidad and Tobago. Although research consistently shows that mathematics scores for male students are usually higher than for female students at the secondary and tertiary level, the differences are not very clear at the primary level. Actually, results from this study show that female students performed slightly better than male students on this examination. Logistic regression procedure was used to detect DIF items, and the results show that about 17% of the items in the examination displayed gender-related DIF; however, for all DIF items the effect sizes were negligible.

**Gender Differences in Mathematics Achievement**

Research on differences in mathematics performance related to gender, although waning at times, has continued over the years (Fennema, Carpenter, Jacobs, Franke, & Levi, 1998; Friedman, 1989; Hyde, Fennema, & Lamon, 1990; Maccoby & Jacklin, 1974). Performance differences in mathematics have been postulated to be partly due to attitudinal differences regarding mathematics (Duffy, Gunther, & Walters, 1997; Fennema & Sherman, 1977; Forgasz & Leder, 1996; Meyer & Koehler, 1990; Stipek & Gralinski, 1991; Tocci & Engelhard, 1991), and parental expectations that suggest a belief in the mathematical superiority of boys over girls (Blevins-Knabe, Austin, Musun, Eddy, & Jones, 2000). This parental belief has been found to be predictive of students' achievement in mathematics courses (Parsons, Adler, & Kaczala, 1982) and students' belief in their own mathematical abilities (Tiedemann, 2000).

It has also been noted that teachers' attitudes toward students' performance in mathematics classes parallel those of the parents. Girls are seen as successful due to their hard work (Jussim & Eccles, 1992; Siegle & Reis, 1998; Tiedemann, 2000), while boys' success is attributed to their talent (Jussim & Eccles, 1992). With regard to the English-speaking Caribbean, there is some evidence that teachers also hold differing perceptions of students' mathematical ability based on gender. However, these perceptions contrast with those articulated in the existing literature. The limited research conducted in this area has shown that, for various reasons, teachers respond more positively to girls than to boys (Kutnick, 2000; Kutnick, Jules & Layne, 1997; Parry, 2000). Additionally, on mandated exams and on teacher-made tests, girls perform better than boys at all ages between 8 and 16, in all curriculum areas, and across all subjects including subjects within the sciences (Kutnick et al., 1997).

Brown (2005) observed that among seven- and eight-year-olds in Trinidad and Tobago, on the mathematics component of the 2000 Continuous Assessment Programme (CAP), overall, girls performed better than boys, a higher proportion of boys than girls was in the lower tail of the distribution, and non-response on items was higher among boys than girls. However, this general finding did not apply to students at or above the 84th percentile where boys had the higher mean.

Hanna (2003) noted that gender differences have decreased to the point that several countries have in effect achieved gender equity in mathematics. Tapia and Marsh (2004), in their review of the relevant literature, also report the non-existence of measurable differences in the achievement scores of boys and girls on math tests at age nine. Furthermore, differences that began to become apparent at age 13 seem to have all but disappeared since 1994. While these findings are in the desired direction, the Trinidad and Tobago and Caribbean data suggest a continuing differential in favour of girls at the primary level but allow for less definitive interpretations at the secondary level (Brown, 2005).

### **Explaining gender differences in mathematics achievement**

Gallagher et al. (2000) presented a taxonomy of content and cognitive characteristics that attempts to account for gender differences in mathematics. The taxonomy was based on outcomes reported in the educational and psychological literature. According to the Gallagher et al. taxonomy, females should perform better than males on items with contextual characteristics likely to be more familiar or interesting to females, on items that require a high level of verbal skill, and on items that require mastery of mathematical content. Males should perform better than females on items that have contextual characteristics likely to be more familiar or interesting to males, on items that are likely to place heavy demands on spatial skills, and on items that have multiple solution paths.

In their study, Gierl, Bisanz, Bisanz, and Boughton (2003) found that males perform better than females on items that require significant spatial processing. This finding was consistent with previous research (e.g., Casey, Nuttall, Pezaris, & Benbow, 1995; Halpern, 1997). They also found that females outperformed males on items requiring memorization, but the differences were small.

### **Differential Item Functioning (DIF)**

Differential item functioning (DIF) occurs when one group is more likely than another to answer an item correctly when both groups are similar on the trait that is being assessed. In DIF terminology, a sample from the population of interest is referred to as the focal group and the sample from the comparable population is referred to as the reference group. Group similarity is usually established by use of total raw score or some measure of the trait (Meyer, Huynh, & Seaman, 2004). In this study, girls comprise the reference group and boys the focal group.

DIF occurs when individuals in a focal group respond differently to a test item than do individuals in a reference group, even when comparisons are restricted to individuals with similar overall skill levels on the trait in question (Johanson, 1997). Dodeen and Johanson (1997) note that the Educational Testing Service (ETS) classifies DIF into three categories: Category A contains items with negligible DIF; Category B for

intermediate DIF items; and Category C for large DIF items. This classification is usually used for investigating DIF in testing situations.

DIF is important in detecting multidimensionality in a test. Multidimensionality occurs when an item assesses a competency other than that being assessed by the test. The source of the multidimensionality for a differentially functioning item may or may not be relevant to the psychological construct being assessed by a particular test in question. Such items unfairly punish those students who might not be familiar with the competency that is not relevant to the test, while unfairly rewarding those who might be familiar with it. It is therefore important to identify and flag items displaying multidimensionality in a test because all items on a test should be one-dimensional, that is, measuring only one construct.

There are a number of reasons that lead to DIF occurring in a test. The test may have items which have negative wording that is confusing to one group. In addition, some test items present the content in ways that prevent certain groups of examinees from demonstrating their knowledge or traits. The way an item is worded or structured may introduce a source of difficulty that is not relevant to the construct being measured. Thus, the item may be measuring two or more constructs, of which only one is relevant to the purpose of the instrument. The extraneous construct may influence the responses of one group differently. However, how it influences the primary construct may not be explicit (Camilli & Shepard, 1994). For example, on a test of math achievement, a weakness in reading comprehension may prevent a group of students from accurately completing math word problems. In this instance, reading comprehension is not the construct of interest, but the examinees' reading level influences their responses.

### **Test and item bias**

At this juncture, it is important that we distinguish between DIF and test bias. Camilli and Shepard (1994) have defined test bias "as invalidity or systematic error in how a test measures for members of a particular group." They further state that "bias is systematic in the sense that it creates a distortion in test results for members of a particular group" (p.

8). That is, it allows for invalid conclusions to be drawn with regard to the performance of a particular group.

But test bias is as a result of item bias. Items are biased when they exhibit DIF. It is important to note that the existence of substantial mean differences between two groups on a test is not sufficient evidence to consider a test biased (Camilli & Shepard, 1994), because true differences could exist between the groups on the underlying construct being assessed by the instrument. The empirical evidence that leads to the conclusion that an item is biased is the extent to which there is *differential item functioning* (DIF) (Hambleton, Swaminathan, & Rogers, 1991).

### **DIF and gender**

Gender DIF is a major concern on large-scale achievement tests in mathematics because differences between females and males are often found (e.g., Bielinski & Davison, 2001; Boughton, Gierl, & Khaliq, 2000; DeMars, 1998; Garner & Engelhard, 1999; Scheuneman & Grima, 1997; Willingham & Cole, 1997). Over the years, many studies have looked at DIF and gender on mathematics items, for example, Doolittle & Cleary (1987), Ryan & Chiu (2001), Scheuneman & Grima, (1997), and Wang & Lane (1996). Harris and Carlton (1993), controlling on the total test score using the Mantel-Haenszel (M-H) procedure, observed that on the Scholastic Aptitude Test (SAT), there were “identifiable patterns of gender differences” (p. 137) in the overall performance of male and female students. While girls performed differentially better than boys in algebra and on textbook-like items, confirming the previous findings of Doolittle and Cleary, the opposite was found for geometry items, real-life non-textbook-like items, and items requiring higher-level mental processing. With the exception of Wang and Lane, whose sample comprised sixth and seventh grade students (age range 11–12 years), all the other studies examined DIF for items on examinations taken at the senior high school or college level.

### **DIF Detection Procedures**

Some of the commonly used psychometric procedures that detect DIF in dichotomously scored test items are the Mantel-Haenszel (MH), the

Simultaneous Item Bias Test (SIBTEST) and the logistic regression (LR) procedures.

### **The Mantel-Haenszel (MH)**

The MH statistic has been one of the most widely used procedures to evaluate DIF (Clauser & Mazor, 1998). The MH procedure was proposed by Holland and Thayer (1988) as a technique for detecting DIF. Since its proposal, there have been many studies about its efficacy in the identification of items with DIF under different conditions, such as DIF magnitude, DIF type, sample size, and sample ability differences (Hidalgo & Lopez-Pina, 2004). Despite its wide use, results from previous studies have indicated that it lacks power in detecting non-uniform DIF (Swaminathan & Rogers, 1990).

### **SIBTEST**

SIBTEST is a nonparametric procedure that both estimates the amount of DIF in an item and statistically tests whether that amount is different from 0 (Bolt, 2000). SIBTEST uses a regression correction method to match examinees from the two groups at the same latent ability levels so as to compare their performances on the studied items. It requires an initial distinction between two non-overlapping subsets of items in the test: (1) a valid subtest, which contains items that are assumed to measure the ability the test is designed to measure; and (2) a suspect subtest, which contains the items being tested for DIF. Scores on the valid subtest are used to match examinees having the same ability levels across groups in order to test items from the suspect subtest for DIF (Bolt, 2000).

### **Logistic regression procedure**

This is the DIF procedure that is used in this study. In DIF detection with logistic regression, the probability of item score (correct vs. incorrect) is predicted from a constant, total test score, group membership, and the interaction between group membership and total score (Swaminathan & Rogers, 1990). Logistic regression is model-based, therefore, uniform and non-uniform DIF can be modelled in the same equation and coefficients for each can be tested separately (French & Miller, 1996).

The logistic regression equation is expressed in general terms as:

$$p(u = 1/x) = \frac{e^{f(x)}}{1 + e^{f(x)}}$$

where  $p(u = 1/x)$  is the conditional probability of obtaining a correct answer given  $\mathbf{x}$  (vector of independent variables), and  $f(\mathbf{x})$  is the function that defines the linear combination of the predictor variables. In DIF analysis, the dependent variable is the item score, and the independent variables are the group variable  $G$  (usually gender or ethnicity), the observed ability  $\theta$  as a matching criterion (usually the total test score), and the group by ability interaction  $\theta G$ . The logistic regression model is therefore expressed as:

$$f(x) = \tau_0 + \tau_1\theta + \tau_2G + \tau_3\theta G$$

where  $\tau_0$  is the intercept,  $\tau_1$  is the ability regression coefficient,  $\tau_2$  is the coefficient for the group variable, and  $\tau_3$  is the coefficient for the ability\*group interaction parameter. Logistic regression is a highly effective technique for detecting DIF in dichotomous items (French & Miller, 1996; Swaminathan & Rogers, 1990; Zumbo, 1999).

Zumbo and Thomas (1997) proposed  $\Delta R^2$ , a weighted least squares effect size measure when using logistic regression in DIF detection, to quantify the magnitude of uniform or non-uniform DIF. The guidelines they provided were:

Type A items – negligible DIF:  $\Delta R^2 < 0.13$

Type B items – moderate DIF:  $0.13 \leq \Delta R^2 \leq 0.26$

Type C items – large DIF:  $\Delta R^2 > 0.26$

However, Jodoin and Gierl (2001), based on their simulation study, have advised caution when interpreting the effect size based on the Zumbo and Thomas (1997) guidelines, and have instead suggested the following guidelines as more accurately quantifying the magnitude of DIF:

Type A items – negligible DIF:  $\Delta R^2 < .035$

Type B items – moderate DIF:  $.035 \leq \Delta R^2 \leq .070$

Type C items – large DIF:  $\Delta R^2 > .070$

In this study, we use the Jodoin and Gierl criteria in interpreting the effect size of the DIF.

### **Purpose of the Study**

A common feature of most of the previous studies on gender DIF is that none of them were done in the Caribbean context using Caribbean data. It is therefore the objective of the current study to accomplish this.

The National Test is administered annually to students in Standard 1 (Std 1)—age range 7 to 8 years—in public and private primary schools. It is based on the national curriculum, and assesses two curriculum areas—Mathematics and Language Arts. The purpose of the test is diagnostic, but as has happened with many large-scale examinations, in the eyes of the public and many school administrators, the test is seen as rating the performance of the schools and, as a result, classifying schools into good schools and bad schools.

Based on student performance on the tests, schools are assigned various support personnel, but also receive greater critical oversight from school supervisors. Unfortunately, this has created an atmosphere of anxiety and competition around the tests, and has resulted in similar kinds of pressure on students, teachers, and administrators as seen in cases where tests results are used for high-stakes decisions. Because decisions are made based on students' test scores, it is prudent for examination boards to demonstrate that any inferences made on the basis of the tests are valid for examinees of all groups.

Therefore, the purpose of this study is to investigate gender differences and the existence of DIF in the Std. 1 mathematics assessment in Trinidad and Tobago. The need to investigate DIF and gender differences in mathematics achievement was motivated by an ongoing quest to explain gender differences in mathematics achievement in the Caribbean. We felt that it was important to look at gender differences through the lens of test quality, where test quality was measured by the number of DIF items. Of interest is whether items function the same for

boys and girls of comparable ability, and the prevalence of the gender-related DIF. Specifically, this study attempts to answer the following questions:

1. Do male examinees perform significantly different from female examinees on a Standard 1 large-scale mathematics assessment in Trinidad and Tobago?
2. How prevalent is gender-related DIF on a Standard 1 large-scale mathematics assessment in Trinidad and Tobago?

## **Method**

### **Data**

The data used in this study came from the Standard 1 national mathematics test of 2004. The test consisted of 20 items, which are a mixture of objective and structured-type items. Each of the items fell into one of the following categories: *Number*, *Measurement and Money*, *Geometry*, and *Statistics*. Each item tested either knowledge computation (KC), algorithmic thinking (AT), or problem solving (PS). Some items had multiple parts, with each part testing a different skill. For example, item 10 had three parts and it measured all three skills. Overall, students responded to 30 questions. Table 1 shows the structure of the examination, and the number of items under each category and what skill each item measured.

Most items in the examination were dichotomously scored as either 1 for a correct response or 0 for an incorrect response. Only items 14-KC and 17-AT were polytomously scored. A correct response was scored as 1 or 2 (depending on whether a complete answer was provided or not) and an incorrect response as 0. To make responses of these two items suitable for analysis in this study, they were recoded as follows: Response  $k = 0$  as 0; response  $k = 1, 2$  as 1. Since this is a national examination, it was conducted under government regulations by which all Standard 1 students the country take the examination at the same time and date.

**Table 1. Examination Questions (Items) Under Each Category**

<b>Number</b>	<b>Measurement and Money</b>	<b>Geometry</b>	<b>Statistics</b>
1-KC	8-KC	13-KC	17-AT
2-KC	9-KC	14-KC	18-KC
3-AT	10-KC	15-AT	18-PS
4-KC	10-AT	16-KC	19-KC
4-AT	10-PS		20-KC
5-KC	11-KC		20-PS
5-AT	11-PS		
6-KC	12-AT		
6-AT	12-PS		
7-KC			
7-PS			

### **Sample**

The original data contained 16,210 examinees, of whom 1,143 were absent on the day of the examination, or their papers were not scored, or gender was not recorded. These were excluded and this resulted in 15,067 (male = 7,720; female = 7,347) valid cases being considered for this study. DIF requires that groups be comparable with respect to the attribute the items supposedly measure. Because the National Test is curriculum based, we expect all students to be similarly familiar with the math content and, as a result, differences in performance on the exam should be random. Therefore, the total test score was used as a measure of ability, and was the criterion on which boys and girls were matched.

For this study, logistic regression was chosen over the MH and SIB Test procedures because of its superior ability to detect both uniform (the item favours one group over another at every ability level) and non-uniform or crossing DIF (there is an Ability \* Group Membership interaction). With this approach, the logistic regression model is used for predicting the probabilities of correct and incorrect responses to each dichotomously scored item, given an observed total test score and its associated group membership.

## Analysis

To investigate gender differences, descriptive statistics were obtained. The descriptive statistics for the test items and the four categories of the examination for male and female examinees are shown in Tables 1 and 2. The independent t-test was performed to find whether the differences are statistically significant on each of the four categories of the test and on the entire test. Since the t-test is sample size dependent and could yield significant results even when the differences are small, we calculated the effect sizes (*d*) to determine the practical significance or meaningfulness of the findings. The interpretation of *d* is as follows: Trivial:  $d < .20$ ; small:  $d = .20$ ; medium:  $d = .50$ ; large:  $d = .80$ . To investigate DIF, logistic regression analysis was performed.

**Table 2. Descriptive Statistics of the Test Sections for Male and Female Examinees**

<i>Category</i>	<b>Male Examinees</b> <i>(n = 7,720)</i>		<b>Female Examinees</b> <i>(n = 7,347)</i>		<b>Total</b> <i>(n = 15,067)</i>	
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>t*</i>	<i>d</i>
Number (11 items)	.64	.45	.69	.43	-6.98	.11
Measurement and Money (9 items)	.49	.44	.53	.44	-5.58	.09
Geometry (4items)	.60	.47	.68	.45	-10.68	.17
Statistics (6 items)	.45	.47	.53	.45	-10.68	.17
Entire Exam	.55	.46	.61	.44	-8.19	.13

\* All t-values significant at  $p < .001$

## Results

Results are presented according to the research questions investigated in the study. The first research question was: Do male examinees perform significantly different from female examinees on a Standard 1 large-scale mathematics assessment in Trinidad and Tobago? To answer this question, descriptive statistics and an independent t-test were computed. The descriptive statistics are presented in Tables 2 and 3. To interpret the

results in Tables 2 and 3, the means represent the proportion of examinees that get an item correct. For example, Table 2 shows that for the Number category, the mean score for male students was .64 while the mean score for female students was .69 out of a possible score of 1.00. Alternatively, this can be interpreted as 64% of male students getting this category correct while 69% of female students got it correct. A similar interpretation can be made to Table 3. It is clear that the mean scores for female students are higher than the mean scores for male students for each category and item. To investigate whether these differences were statistically significant, an independent t-test was performed and the results are also presented in Table 2.

**Table 3. Descriptive Statistics of the Test Items for Male and Female Examinees**

<b>Item</b>	<b>Male Examinees</b> ( <i>n</i> = 7,720)		<b>Female Examinees</b> ( <i>n</i> = 7,347)		<b>Total</b> ( <i>n</i> = 15,067)	
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>Mean</i> <i>n</i>	<i>SD</i>
<b>Number</b>						
1-KC	.85	.36	.91	.28	.88	.32
2-KC	.44	.50	.47	.50	.46	.50
3-AT	.71	.45	.73	.44	.72	.45
4-KC	.64	.48	.68	.47	.66	.47
4-AT	.66	.47	.71	.45	.68	.47
5-KC	.34	.47	.39	.49	.36	.48
5-AT	.43	.50	.48	.50	.45	.50
6-KC	.75	.43	.82	.38	.79	.41
6-AT	.81	.39	.87	.33	.84	.36
7-KC	.68	.47	.72	.45	.70	.46
7-PS	.74	.44	.79	.41	.76	.43
<b>Measurement and Money</b>						
8-KC	.86	.34	.90	.30	.88	.33
9-KC	.90	.30	.93	.26	.91	.28
10-KC	.28	.45	.32	.47	.30	.46

Item	Male Examinees ( <i>n</i> = 7,720)		Female Examinees ( <i>n</i> = 7,347)		Total ( <i>n</i> = 15,067)	
	Mean	SD	Mean	SD	Mean	SD
10-AT	.51	.50	.55	.50	.53	.50
10-PS	.33	.47	.39	.49	.36	.48
11-KC	.21	.41	.24	.43	.22	.42
11-PS	.37	.48	.44	.50	.40	.49
12-AT	.50	.50	.52	.50	.51	.50
12-PS	.49	.50	.51	.50	.50	.50
<b>Geometry</b>						
13-KC	.59	.49	.67	.47	.63	.48
14-KC	.79	.41	.85	.36	.82	.39
15-AT	.59	.49	.69	.46	.64	.48
16-KC	.44	.50	.51	.50	.47	.50
<b>Statistics</b>						
17-AT	.77	.42	.87	.34	.82	.39
18-KC	.41	.49	.53	.50	.47	.50
18-PS	.42	.49	.53	.50	.47	.50
19-KC	.58	.49	.67	.47	.63	.48
20-KC	.27	.45	.29	.46	.28	.45
20-PS	.27	.45	.29	.46	.28	.45

The differences between male and female students were all significant for the test and for the four categories on the test. For the Number category,  $t(15065) = -6.98$ ,  $p < .001$ ; the Measurement category,  $t(15065) = -5.58$ ,  $p < .001$ ; Geometry category,  $t(15065) = -10.68$ ,  $p < .001$ ; and the Statistics,  $t(15065) = -10.68$ ,  $p < .001$ . For the entire test,  $t(15065) = -8.19$ ,  $p < .001$ . The effect sizes,  $d$ , were trivial indicating that significant differences were more due to the sample size than true differences in the population.

The second question was: How prevalent is gender-related DIF on a Standard 1 large-scale mathematics assessment in Trinidad and Tobago? To answer this question, logistic regression was employed. The analytic strategy in logistic regression is model comparison by successively

adding the terms into the model  $(\theta, G, \theta G)$ . DIF evaluation is carried out by statistically evaluating the incremental contribution of each successive model term  $(\theta, G, \theta G)$ . If the group effect  $G$  is statistically significant ( $\tau_1 \neq 0$ ) and the effect of the interaction  $\theta G$  is not ( $\tau_3 = 0$ ), then the item has uniform DIF. When  $\tau_2 > 0$  DIF favours the reference group, and when  $\tau_2 < 0$  DIF favours the focal group. If the interaction is statistically significant, that is, an item favours the higher-ability members of the reference group and the lower ability members of the focal group or vice versa, the item has non-uniform DIF. While non-uniform DIF does occur, this is not commonly found (Jodoin & Gierl, 2001). Therefore in this study, our focus is on uniform DIF.

The statistic used is the Wald test statistic, which follows a chi-squared ( $\chi^2$ ) distribution with degrees of freedom ( $df$ ) = 1. The results in Table 4 show that item 1-KC displayed non-uniform DIF because the Total Score by Gender ( $\theta G$ ) interaction term was statistically significant,  $\chi^2 = 6.55$ ,  $p < .011$ . For item 13-KC, the group effect (Gender) was statistically significant,  $\chi^2 = 6.74$ ,  $p < .009$ ; but the interaction term was not statistically significant,  $\chi^2 = 2.15$ ,  $p < .142$ , which means that this item displayed uniform DIF. Similarly, items 15-AT, 20-KC, and 20-PS displayed uniform DIF because the interaction terms were not significant, but gender was.

### **Effect size measures**

As previously stated, we compute and interpret  $\Delta R^2$  using the Jodoin and Gierl (2001) guidelines. The statistics summary of the effect sizes for the DIF test is the differences in Nagelkerke's  $R^2$  between the model with absence of DIF and the full model; that is, the additional proportion of variability accounted for as each term is added to the model. These are provided in Table 5. Based on Jodoin and Gierl's criteria for using  $\Delta R^2$ , a weighted least squares effect size measure to quantify the magnitude of uniform or non-uniform DIF, all these items are Type A items—negligible DIF, because  $\Delta R^2 < .035$ . Item 1KC is the only item that

approached moderate DIF. Thus, although these items displayed DIF, the magnitude was negligible as measured by the effect size.

**Table 4. Items Displaying Uniform and Non-Uniform DIF**

<b>Item 1- KC</b>		<i>Wald</i>	<i>df</i>	<i>Sig.</i>
	Total score	265.876	1	.000
	Gender	20.277	1	.000
	Total score by Gender	6.545	1	.011
	Constant	96.025	1	.000
<b>Item 13- KC</b>				
	Total score	341.381	1	.000
	Gender	6.742	1	.009
	Total score by Gender	2.151	1	.142
	Constant	261.961	1	.000
<b>Item 15- AT</b>				
	Total score	339.694	1	.000
	Gender	7.568	1	.006
	Total score by Gender	.640	1	.424
	Constant	279.711	1	.000
<b>Item 20- KC</b>				
	Total score	293.502	1	.000
	Gender	7.956	1	.005
	Total score by Gender	3.024	1	.082
	Constant	292.728	1	.000
<b>Item 20- PS</b>		<i>Wald</i>	<i>df</i>	<i>Sig.</i>
	Total score	301.692	1	.000
	Gender	6.048	1	.014
	Total score by Gender	2.037	1	.154
	Constant	302.711	1	.000

**Table 5. Effect Size Measures for Differential Item Functioning Magnitude**

Item	$\Delta R^2$
Item 1-KC	.03
Item 13-KC	.01
Item 15-AT	.02
Item 20-KC	.02
Item 20- PS	.02

Figures 1 and 2 show two item characteristic curves for items that displayed DIF (item 1-KC) in Figure 1 and an item that did not display DIF (item 6-KC) in Figure 2. The fact that these two figures look similar even when one item has DIF and the other does not, supports results in Table 5 that the magnitude of the effect size is negligible.

### Discussion and Conclusion

Almost all researchers like their findings to be significant and of practical importance; however, sometimes an alternative finding, that is, one that is non-significant, or trivial in its effect, yields information that is just as important. It is tempting to say that this study discounts DIF as a possible source for the gender differences in performance on a test, because the magnitude of the male-female differences on particular items after controlling for group differences in mathematics ability was negligible. Yet, it is important to note Hidalgo and Lopez-Pina's (2004) observation that the statistic  $\Delta R^2$  appeared to be insensitive in classifying variables with moderate DIF and, therefore, one needed to be aware of the limitation of the interpretation criteria.

Additionally, it is necessary to recognize the probability of DIF amplification (Nandakumar, 1993). As Nandakumar explains, while items individually might display negligible DIF, collectively these items may have a significant effect on total test scores. Therefore, 5 of 30 items on a test exhibiting statistically significant DIF in favour of girls could be a cause for concern, because one group is being given an undue advantage over another. It was also noted that even where items did not display significant DIF, on many items the item characteristic curve

(ICC) showed the reference group (girls) to be above the focal group (boys).

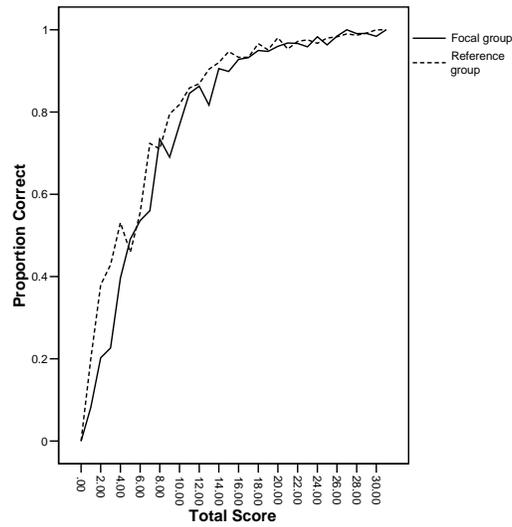


Figure 1. Example of item exhibiting differential item functioning.

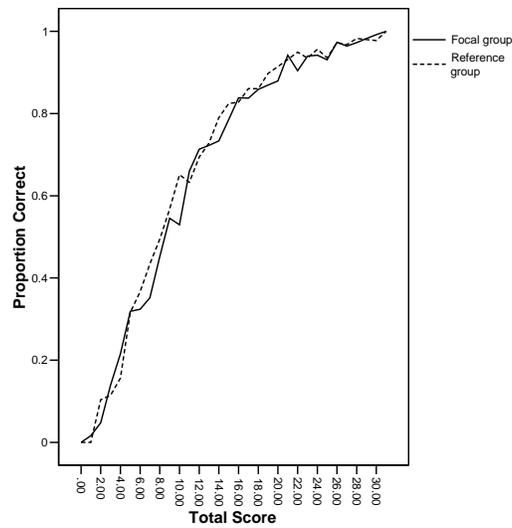


Figure 2. Example of item not exhibiting differential item functioning.

Whether these results are interpreted to mean that DIF is not a problem, or emphasis is placed on possible DIF amplification, in the context of Trinidad and Tobago and the wider English-speaking Caribbean, this finding is important. We have not been able to locate any record of test items ever being examined for DIF. Therefore, this study adds a new dimension to the ongoing debate in the Caribbean about differences in performance between boys and girls. Best practice suggests that items that exhibit DIF should be identified before a test is published. When items show DIF, test developers must decide whether these items should be retained by determining if gender is a relevant dimension in the measurement of mathematics competency.

Ignoring the indication of DIF, the consistency of the mean differences in favour of girls seems to confirm findings of other research on mathematics achievement in the Caribbean that have shown female students doing better than male students, not only in mathematics but in other subjects as well. This evidence strongly suggests that while the examination board addresses the construction of the tests, the research on differences in the academic performance of male and female students also has to focus on sociocultural and socio-psychological factors as explanatory variables. In addition, the school has to become a focus of the research. Caribbean researchers will need to extend the work of Jules and Kutnick (1990) to better understand school and classroom dynamics that impact the sexes differently, and potentially limit one group or enhance another group's performance.

Walstad and Robson, (1997) posit that possible factors which may explain gender differences in achievement are socialization, differential reasoning, instructional practices, or the format of testing. Also, and as already noted (Jussim & Eccles, 1992; Leedy, Lalonde, & Runk, 2003; Siegle & Reis, 1998; Tiedemann, 2000), teacher attitude and beliefs also more than likely reflect the prevailing attitude and beliefs in the wider society. While some of these factors are beyond the control of the school, others are not. It is expected that as schools are made aware of their role in contributing to the disproportion in academic achievement between the sexes, they are more likely to initiate solutions to the problem. This is critical because despite the social and cultural factors that may explain the differential in performance, most issues involving education still have to be addressed in the school and classroom.

Although the gender differences on mathematics achievement were small, they are worth noting and need to be addressed from a policy standpoint. What was interesting in this study was that the female-male differences were consistent across the different sections of the test, with female students consistently outperforming male students in all content areas.

Our findings also indicate that, overall, students performed the worst on statistics, followed by measurement and money. Also, our findings indicate that, generally, the problem-solving skills for most students are not well-developed. These findings have important implications for classroom practice. They are important because they provide evidence of the need to develop effective strategies to teach these concepts. Teaching mathematics is not just about the focus on numbers; rather, it requires teachers to tailor their strategies to specific areas of mathematics. For example, teachers cannot teach statistics the same way they teach measurement or money.

Also, teachers need to be innovative in developing effective teaching methods. For example, the use of real-life examples is one very effective way of teaching statistics or money concepts. Students always want to see how concepts being taught in class relate to their everyday life situations. However, improving classroom practice is not the sole responsibility of the teachers, but of all stakeholders involved. Ultimately, teachers and school administrators are responsible for the implementation of the strategies in the classroom, but policymakers should be willing to provide funding and incentives for teachers to be creative in developing effective teaching strategies.

### **References**

- Bielinski, J., Davison, M. L. (2001). A sex difference by item difficulty interaction in multiple-choice mathematics items administered to national probability samples. *Journal of Educational Measurement*, 38(1), 51–77.
- Blevins-Knabe, B., Austin, A., Musun, L., Eddy, A., & Jones, R. M. (2000). Family home care providers' and parents' beliefs and practices concerning mathematics with young children. *Early Child Development and Care*, 165, 41–58.

- Bolt, D. M. (2000). A SIBTEST approach to testing DIF hypotheses using experimentally designed test items. *Journal of Educational Measurement*, 37(4), 307–328.
- Boughton, K. A., Gierl, M. J., & Khaliq, S. N. (2000, May). *Differential bundle functioning on mathematics and science achievement tests: A small step toward understanding differential performance*. Paper presented at the Annual Meeting of the Canadian Society for Studies in Education (CSSE), Edmonton, Alberta, Canada.
- Brown, L. (2005). Gender and academic achievement in math: An examination of the math performance data on seven to nine year olds in Trinidad and Tobago. *Caribbean Curriculum*, 12, 37–56.
- Camilli, G., & Shepard, L. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage.
- Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology*, 31(4), 697–705.
- Clauser, B. E., & Mazor, K. M. (1998) Using statistical procedures to identify differentially functioning test items. *Educational Measurement: Issues and Practice*, 17(1), 31–44.
- DeMars, C. E. (1998). Gender differences in mathematics and science on a high school proficiency exam: The role of response format. *Applied Measurement in Education*, 11(3), 279–299.
- Dodeen, H., & Johanson, A. G. (2003). An analysis of sex-related differential item functioning in attitude assessment. *Assessment and Evaluation in Higher Education*, 28(2), 129–134.
- Doolittle, A. E., & Cleary, A. T. (1987). Gender-based differential item performance in mathematics achievement items. *Journal of Educational Measurement*, 24(2), 157–166.
- Duffy, J., Gunther, G., & Walters, L. (1997). *Gender and mathematical problem solving*. *Sex Roles*, 37(7–8), 477–494.
- Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). A longitudinal study of gender differences in young children's mathematical thinking. *Educational Researcher*, 27(5), 6–11.
- Fennema, E., & Sherman, J. A. (1977). Sex-related differences in mathematics achievement, spatial visualization and sociocultural factors. *American Educational Research Journal*, 14(1), 51–71.
- Forgasz, H. J., & Leder, G. C. (1996). Mathematics classrooms, gender and affect. *Mathematics Education Research Journal*, 8(2), 153–173.
- French, A. W., & Miller, T. R. (1996). Logistic regression and its use in detecting differential item functioning in polytomous items. *Journal of Educational Measurement*, 33(3), 315–332.

- Friedman, L. (1989). Mathematics and gender gap: A meta-analysis of recent studies on sex differences in mathematical tasks. *Review of Educational Research, 59*, 185–213.
- Gallagher, A. M., De Lisi, R., Holst, P. C., McGillicuddy-De Lisi, A. V., Morely M., & Cahalan C. (2000). Gender differences in advanced mathematical problem solving. *Journal of Experimental Child Psychology, 75*(3), 165–190.
- Garner, M., & Engelhard, G., Jr. (1999). Gender differences in performance on multiple-choice and constructed response mathematics items. *Applied Measurement in Education, 12*(1), 29–51.
- Gierl, M. J., Bisanz, J., Bisanz, G., & Boughton, K. (2003). Identifying content and cognitive skills that produce gender differences in mathematics: A demonstration of the multidimensionality-based DIF analysis paradigm. *Journal of Educational Measurement, 40*(4), 281–306.
- Halpern, D. F. (1997). Sex differences in intelligence: Implications for education. *American Psychologist, 52*(10), 1091–1102.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. H. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hanna, G. (2003). Reaching gender equity in mathematics education. *The Educational Forum, 67*(3), 204–214.
- Harris, A. M., & Carlton, S. T. (1993). Patterns of gender differences on mathematics items on the Scholastic Aptitude Test. *Applied Measurement in Education, 6*(2), 137–151.
- Hidalgo, D. M., & Lopez-Pina, J. A. (2004). Differential item functioning detection and effect size: A comparison between logistic regression and Mantel-Haenszel procedures. *Educational and Psychological Measurement, 64*(6), 903–915.
- Holland, P. W., & Thayer, D. T. (1988). Differential item functioning and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 129–145). Hillsdale, NJ: Erlbaum.
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139–155.
- Jodoin, M. G., & Gierl, M. J. (2001). Evaluating Type 1 error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education, 14*(4), 329–349.
- Johanson, G. A. (1997). Differential item functioning in attitude assessment. *Evaluation Practice, 18*(2), 127–135.
- Jules, V., & Kutnick, P. (1990). Determinants of academic success within classrooms in Trinidad and Tobago: Some personal and systemic variables. *Educational Studies, 16*(3), 217–235.

- Jussim, L., & Eccles, J. S. (1992). Teacher expectations II: Construction and reflection of student achievement. *Journal of Personality and Social Psychology*, 63(6), 947–961.
- Kutnick, P. (2000). Girls, boys and school achievement: Critical comments on who achieves in schools and under what economic and social conditions achievement takes place – A Caribbean perspective. *International Journal of Educational Development*, 20(1), 65–84.
- Kutnick, P., Jules, V., & Layne, A. (1997). *Gender and school achievement in the Caribbean* (Education Research Serial, No. 21). London: Department for International Development.
- Leedy, G. M., Lalonde, D., & Runk, K. (2003). Gender equity in mathematics: Beliefs of students, parents, and teachers. *School Science and Mathematics*, 103(6), 285–292.
- Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.
- Meyer, J. P., Huynh, H., & Seaman, M. A. (2004). Exact small-sample differential item functioning methods for polytomous items with illustration based on an attitude survey. *Journal of Educational Measurement*, 41(4), 331–344.
- Meyer, M. R., & Koehler, M. S. (1990). Internal influences on gender differences in mathematics. In E. Fennema & G. Leder (Eds.), *Mathematics and gender* (pp. 60–95). New York: Teachers College Press.
- Nandakumar, R. (1993). Simultaneous DIF amplification and cancellation: Shealy–Stout’s Test for DIF. *Journal of Educational Measurement*, 30, 293–311.
- Parry, O. (2000). *Male underachievement in high school education in Jamaica, Barbados, and St. Vincent and the Grenadines*. Kingston, Jamaica: Canoe Press.
- Parsons, J. E., Adler, T. F., & Kaczala, C. M. (1982). Socialization of achievement attitudes and beliefs: Parental influences. *Child Development*, 53(2), 310–321.
- Ryan, K. E., & Chiu, S. (2001). An examination of item context effects, DIF, and gender DIF. *Applied Measurement in Education*, 14(1) 73–90.
- Scheuneman, J. D., & Grima, A. (1997). Characteristics of quantitative word items associated with differential performance for female and black examinees. *Applied Measurement in Education*, 10(4), 299–319.
- Siegle, D., & Reis, S. M. (1998). Gender differences in teacher and student perceptions of gifted students' ability and effort. *Gifted Child Quarterly*, 42(1), 39–47.
- Stipek, D. J., & Gralinski, J. H. (1991). Gender differences in children's achievement-related beliefs and emotional responses to success and failure in mathematics. *Journal of Educational Psychology*, 83(3), 361–371.

- Swaminathan, H., & Rogers, J. H. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27(4), 361–370.
- Tapia, M., & Marsh G. E. (2004). The relationship of math anxiety and gender. *Academic Exchange Quarterly*, 8(2), 130–134.
- Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, 92(1), 144–151.
- Tocci, C. M., & Engelhard, G. (1991). Achievement, parental support, and gender differences in attitudes toward mathematics. *Journal of Educational Research*, 84(5), 280–286.
- Walstad, W. B., & Robson, D. (1997). Differential item functioning and male-female differences on multiple-choice tests in economics. *Journal of Economic Education*, 28(2), 155–171.
- Wang, N., & Lane, S. (1996). Detection of gender-related differential item functioning in a mathematics performance assessment. *Applied Measurement in Education*, 9(2), 175–199.
- Willingham, W. W., & Cole, N. S. (1997). *Gender and fair assessment*. Mahwah, NJ: Erlbaum.
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores*. Ottawa, ON: Directorate of Human Resources and Evaluation, Department of National Defense.
- Zumbo, B. D., & Thomas, D. R. (1997). *A measure of effect size for a model-based approach for studying DIF* (Working paper of the Edgeworth Laboratory for Quantitative Behavioral Science). Prince George, Canada: University of Northern British Columbia.



**THE PROMOTION OF THINKING IN SELECTED  
LOWER SECONDARY SCIENCE CLASSROOMS  
IN TRINIDAD AND TOBAGO:  
Implications for Teachers' Education**

*Susan Herbert and Joycelyn Rampersad*

This article reports on a study into the levels of thinking promoted in selected lower secondary science classrooms in Trinidad and Tobago. Twenty-seven teachers, 10 of whom were professionally certified, were observed twice. The levels of thinking were inferred from an analysis of teaching/learning strategies, teachers' and students' questions, the process skills used, and types of classroom interactions. The findings indicate that the teachers, both professionally certified and uncertified, were unable to promote higher-order thinking in their classrooms. There are implications for teacher education programmes.

**Introduction**

The development of students' thinking is an educational goal that is receiving renewed attention. From the United States of America (USA), Feiman-Nemser (2001, p. 1051) reports that "many contemporary reforms call for content-rich, learner-centred teaching, which emphasizes conceptual understanding and gives all students opportunities to think critically, solve problems, and learn things that matter to them and have meaning in the world outside of school." However, the research shows that significant proportions of high school and university students are unable to engage effectively in activities that require higher-order cognitive functioning (see Beyer, 2001a).

In Trinidad and Tobago, the country in which this study was conducted, many teachers, employers, and stakeholders also express the view that the students/graduates of the secondary education system, including science graduates, are unable to think critically. Most seem to operate at low-order cognitive/mental levels (knowledge and comprehension) as identified by Bloom, Englehart, Furst, Hill, and Krathwohl (1956), where

they can recall information and explain concepts and principles, but are unable, for example, to critically examine ideas/concepts, solve problems, and make reasoned arguments. The crucial question is: Are our secondary science teachers, providing the kinds of experiences that would facilitate the development of these higher-order thinking processes/operations?

### **The Context**

Trinidad and Tobago, a twin-island republic, is located in the southern Caribbean. In keeping with international trends, Trinidad and Tobago is also currently involved in educational reform. At the secondary level, the reform initiative is described as the Secondary Education Modernization Programme (SEMP). The SEMP began at the lower secondary level—the first three years of secondary schooling—with the development of curricula in seven core areas, including science. Students at the lower secondary level may range in age from 11 to 17 years. They begin their lower secondary schooling in Form 1, moving on to Forms 2 and 3 over a 3-year period.

Included among the expected learning outcomes of the SEMP science curriculum are that students will “interpret and evaluate data,” “critically reflect on and interpret ideas presented through a variety of media,” (Trinidad and Tobago. Ministry of Education [T&T], 2003, p. 1-9), “reflect critically on ethical and other issues” (p. 1-10) and “develop a critical awareness of the role of science in everyday living” (p. 2-1).

Underpinning these outcomes is the assumption that the experiences provided in the science classroom have the potential to facilitate/promote thinking. This assumption has been the impetus for a number of research activities outside of Trinidad and Tobago (Adey & Shayer, 1994; Zimmerman, Raghavan, & Sartoris, 2003) investigating the kinds of approaches used to develop thinking skills, or the effectiveness of various interventions on the development of thinking. This type of research has not been conducted within Trinidad and Tobago, and, therefore, this survey is the first of its kind within the country that seeks to gather empirical data on the approaches that have the potential to promote higher-order thinking in science classrooms, as required by

SEMP documents and in light of the unique circumstances described below.

In Trinidad and Tobago, the teachers at the secondary level do not all have professional certification. It is not a requirement for entry into the teaching service, as obtains in other countries, for example, the USA and the United Kingdom (UK). It means, therefore, that both professionally certified teachers (hereafter referred to as trained teachers) and uncertified teachers (hereafter referred to as untrained teachers) are expected to enact the curricula, including the science curriculum. The professional certification offered to graduate teachers at the secondary level is the one-year in-service Postgraduate Diploma in Education (Dip.Ed.) programme offered at the St. Augustine Campus of The University of the West Indies (UWI). The programme involves face-to-face sessions at the campus and field visits to schools.

An eclectic model, which draws on aspects of teacher behaviour development (see Koballa, Jr. & Tippins, 2001) and reflective practice, underpins the prospectus for the Dip.Ed. programme. For example, “The programme attempts to ensure that classroom practice is informed by a solid theoretical base in the foundation disciplines, curriculum theory, and methodology” (The University of the West Indies [UWI]. School of Education, 2004, p. 63). The objectives of the programme include the following:

- encourage teachers to give the greatest attention to past and present practices and future possibilities in the teaching of their subjects;
- lead teachers to consider the professional implications of the nature of their occupation and to strive for continued professional growth. (p. 63)

Guided by the stated objectives, the science teacher educators develop the course outline before the programme begins. During face-to-face sessions, theoretical frameworks are presented and discussed, and participants suggest implications for their practice. Teachers are then expected to implement the strategy described and their efforts are monitored during the next school visit or field day. In addition, the teachers are expected to keep a reflective journal, and to submit a

teaching portfolio as part of the assessment requirements. The portfolio serves as a vehicle for self-evaluation, as it allows the teachers to reflect on their philosophies of teaching and on practice, to assess strengths and weaknesses in the professional work accomplished over the period of training, and to highlight significant growth points. They are also required to submit a formal report of an action research project that was conceived and implemented during the year.

A perusal of the prospectus reveals that there are no explicit statements that allude to the preparation of, or assessment of, teachers in the area of teaching critical thinking skills. However, during the course entitled “Curriculum Process,” the science teachers are introduced to evidence-based strategies—selected by the tutors—which can facilitate the development of higher-order thinking skills. These include uses and purposes of questions, the integration of process skills in science teaching, use of practical work involving inductive and deductive approaches to concept development, and active learning strategies. Following the face-to-face sessions conducted at UWI, the teachers are expected to implement these strategies and to reflect on their developing competencies. Tutors and peers provide support and feedback through field experiences.

There is evidence that teachers’ instructional practices are guided by their thinking and beliefs, which in turn affect student thinking (Brickhouse, 1990; Hashweh, 1996; Onosko, 1990). The evidence also suggests that these relationships are not simple, and that there are multiple factors that may obstruct a simple and direct influence (Zohar, 2004). One factor has been described as the ambivalent character of teachers’ conceptions of thinking and teaching science—on the one hand they believe that efficient learning is realized when they transmit knowledge in an effective way, and on the other hand they believe that it is important for students to learn in an active way. Thus, there is a contradiction between teachers’ conscious way of thinking about teaching and learning (which they may have acquired in a teachers’ course) and a less conscious way of processing information that may be influenced by their own experiences as students (De Jong, Korthagen, & Wubbels, 1998, as cited in Zohar, 2004). Another factor that might determine whether teachers act on their beliefs is cultural traditions. These traditions shape students’ perceptions of their roles in the

classroom, and may frustrate teachers in their attempts to engage students in activities that enhance higher-order thinking. For example, students may resist attempts to move them from a passive mode of learning to an active mode that encourages collaboration and interaction (Snell, 1999).

This paper reports on a preliminary survey to determine the kinds of experiences that are provided for stimulating higher-order thinking in selected lower secondary science classrooms. The lower secondary level was chosen because the recently introduced SEMP curriculum reflects a strong thinking skills focus. The findings are intended to serve two purposes. The first is to make some inferences about the levels of thinking promoted based on students' and teachers' behaviours within the science classroom. The second is to provide research evidence to inform the approach for developing strong thinking skills pedagogy in teacher education programmes.

With respect to the first purpose of the paper, the premise is made that some classroom activities/experiences facilitate low-order mental operations (low-level thinking) while others facilitate higher-order mental operations (high-level thinking). The higher-order operations, which are fundamental to scientific activity, are also referred to as "critical thinking" or the "skilled and active interpretation and evaluation of observations and communications, information and argumentation" (Fisher & Scriven, 1997, p. 21, as cited in Fisher, 2001). Watson and Glaser (cited by Bitner, 1991), and Facione (1998) refer to core critical thinking skills, which include inference, recognition of assumptions, deduction, interpretation, analysis, evaluation of arguments, explanation, and self-regulation.

Higher-order mental operations (critical thinking) have informed the development as well as the interrogation of scientific knowledge, and are also inherent in the procedural aspect of scientific enquiry (the science processes). Hence, for the development of higher-order thinking, the activities in the science classroom should reflect similar foci. The literature provides direction as to the kinds of activities/experiences and teacher behaviours that facilitate thinking in the science classroom.

## **Facilitating Student Thinking – A Brief Review of Related Literature**

Inferences about students' thinking can be gleaned from their behaviours (Mislevy, 2004). Such behaviours include the types of questions students ask; the quality of their responses to teachers' questions; their ability to make sense of data from a number of sources, for example, discerning patterns, making generalizations, applying principles to solve problems; and making reasoned judgements based on evidence. The research suggests that there is a relationship between certain instructional strategies and thinking skills outcomes (Beyer, 2001a). For example, teachers can facilitate the development of student thinking by: (a) providing opportunities for process skill development, (b) using questioning appropriately, and (c) including collaborative classroom interaction/activities (Higgins et al., 2004).

### **Process skills and development of thinking**

Prior to the 1980s, the term “process skills approach” was usually associated with science curricula, in which the development of general science processes was the primary goal. During the 1980s, there was much debate about the appropriateness of focusing on the teaching of general process skills with the expectation that the use of these skills would be transferred to specific domains. In a seminal paper published in 1987, Millar and Driver presented critiques of the process approach to science teaching from various perspectives—the philosophy of science, cognitive psychology, and pedagogy. Their main argument was that use of the process skills is context and domain dependent, and should be linked to the teaching of science concepts and constructs.

Certain process skills have been identified in the literature as necessary for the examination and explanation of natural phenomena. In science education, process skills are usually, but not exclusively, developed in association with practical work. By practical work is meant “any teaching and learning activity which involves at some point the students observing or manipulating real objects and materials” (Millar, 2004, p. 2). Chiappetta and Koballa, Jr. (2002) classify the science process skills as basic and integrated. Basic skills include observing, measuring, and making predictions. Some of the integrated skills are designing

experiments, developing models to explain phenomena, and interpreting information. While these skills are not exclusive to science, the integrated skills and advanced reasoning are required for scientific enquiry. All of these processes, to some extent, involve analytical thinking, which is as essential in science as critical thinking skills (Swartz & Fischer, 2001). Therefore, students who are given opportunities to use the basic and integrated process skills to enhance conceptual understanding of science are likely to be engaged in critical and analytical thinking about science content and the nature of science and scientific enquiry.

Millar (2004) cautions that student thinking about scientific enquiry does not advance much unless practical tasks are designed with specific, progressively challenging objectives in mind. It means that teachers must be cognizant of the multiple purposes of practical work. Furthermore, he states that the prevalence of the empiricist/inductivist view of science and the development of scientific knowledge—the belief that “ideas ‘emerge’ automatically from the event itself” (p. 11)—can limit the opportunities that teachers provide for developing thinking in and about science. In addition, students may have limited exposure to the hypothetico-deductive approach to doing science, for which the integrated process skills of hypothesizing and experimenting are required.

### **Questioning**

According to Shakespeare (2003), arguments that support the use of questions in teaching and learning are clear and convincing, and include the roles that questions play in the development of thinking skills and in enhancing students’ understanding of the nature of science. Bloom et al.’s (1956) taxonomy is perhaps the most familiar to science teachers, and many researchers acknowledge that it provides a framework for the development of questions that require varying levels of thinking (Blosser, 1990; Chiappetta & Koballa, Jr., 2002; Cotton, 1988). Costa’s (2001) information processing model for questioning is less familiar, but just as useful.

Costa (2001) is of the view that “questions invite different levels of complexity of thinking” (p. 361). He refers to three levels of questioning.

The first level of questions, which is the input level, is pitched at gathering and recalling information. Questions at this level require students to recall past experiences, information, and concepts. At the second level, classified as the process level, questions should require that students make sense of the information gathered. These questions require students to determine cause-effect relationships, compare, contrast, analyse, summarize, and synthesize data. The third level is described as the output level, requiring students to apply and evaluate actions in novel situations. Questions pitched at the output level require students to think creatively and hypothetically, use their imagination, expose their value systems, and make judgements. While there are some similarities between Costa's levels of questions and Bloom's taxonomy, Costa identifies the thought processes required for each of his levels, thus explicitly relating questioning to levels of thinking.

The research evidence (e.g., Blosser, 1990) shows that teachers tend to pose questions at low levels of Bloom et al.'s (1956) taxonomy, so that students operate mainly at the knowledge and understanding levels (the input level in Costa's model). It also shows that even when teachers attempt to challenge students to engage in higher mental operations by asking probing questions, "wait-time" is usually very short (Chiappetta & Koballa, Jr., 2002; Costa, 2001). The seminal work of Rowe (2003) in the area of questioning has shown that increased wait-time results in an increase in the level of cognitive functioning. There is also evidence that redirection/probing/reinforcement, use of higher-order questions, and lengthening wait-time not only enhance thinking skills, but are also associated with increases in student engaged time and level of participation (see Cotton, 1991).

Increased levels of cognitive functioning, which can be determined by a comparison over time of student performance on tasks designed to challenge their thinking and reasoning, are also facilitated when students ask questions, especially higher-order questions. These higher-order questions, also referred to as "quality questions," are reflective of critical thinking and they also facilitate the development of critical thinking skills (Fisher, 1990, as cited by Alsop, de Jesus, & Watts, n.d.). However, research on classroom questions indicates that teachers do not select the types of strategies that encourage student questions during classroom sessions (Chin, 2004). For example, problem solving has been

shown to elicit more, and a wider range of, students' questions at the higher-order level than teacher-directed activities (Chin, Brown, & Bruce, 2002).

### **Classroom interactions**

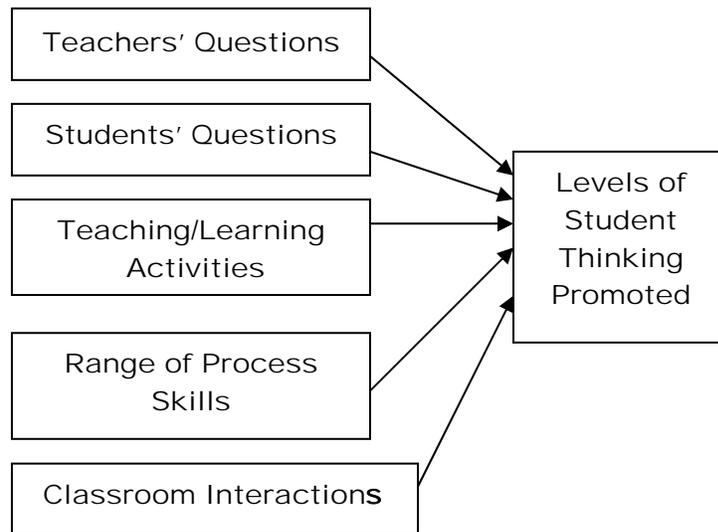
There is some theoretical support for the view that the development of thinking is facilitated by social interaction (Leach & Scott, 2000; Vygotsky, 1978). As applied to the classroom context, the literature reveals that one-way interaction, such as the lecture method, produces the least benefits in terms of acquisition of cognitive skills and strategies. On the other hand, student-centred types of classrooms in which there is a variety of interaction patterns, and where students are, for example, working cooperatively in groups to make decisions, collecting data by hands-on manipulation of material, constructing strategies to solve problems, and validating solutions based on student-generated data, produce the highest output in terms of complex thinking processes (Costa, 2001; Johnson & Johnson, 2001). Furthermore, the literature suggests that classroom interactions should optimize opportunities to develop communities of inquiry, since these can assist in the development of students' thinking and learning through social interaction and reflective inquiry (Kovalainen, Kumpulainen, & Vasama, 2001).

### **Conceptual Framework and Methodology**

Levels of thinking promoted in the lower secondary classroom were examined through an analysis of the teaching/learning strategies, the questions asked by students and teachers, the range of science process skills (basic and integrated) implemented, and the patterns of classroom interactions, (see Figure 1).

The following research question informed the study:

- What levels of student thinking are promoted in the lower secondary science classroom in Trinidad and Tobago?



*Figure 1.* Activities/experiences linked to promotion of student thinking.

### **The sample, data collection, and analysis**

A total of 27 science teachers participated in the study. The participating teachers were a heterogeneous group, which was differentiated by professional certification and years of teaching experience. Of the 27 participating teachers, 10 were trained and the others were untrained. Trained and untrained teachers were purposely included to determine to what extent teachers spontaneously include higher-order thinking experiences in their repertoire. The study was conducted over a 3-month period. The data collection strategy was classroom observations at each of the three lower secondary class levels.

Four researchers, who were science education lecturers on the Dip.Ed. programme at UWI, were involved in data collection. Each science teacher was observed while teaching on two occasions. The lessons were audio taped, transcribed, and summarized to highlight key episodes in terms of teachers' actions and students' actions. The data were further analysed using the following pre-established frameworks—Bloom et al. (1956) and Costa (2001) for analysis of questions, and Chiappetta & Koballa, Jr. (2002) for categorizing science process skills. Open coding

of data, followed by thematic clustering, was used to categorize classroom interactions and teaching/learning activities.

There are two limitations of this study. One was that each teacher was observed on two occasions only. While this small number of lessons observed may not reflect typical teacher behaviour, the observations do give some insights into classroom practices. The second is that all inferences made about students' thinking outcomes were based only on classroom observations, since students' written work was not available for analysis.

## **Findings**

The findings are presented under the following headings: teaching/learning strategies, teachers' and students' questions, and range of process skills developed. Comments on the types of interactions associated with specific strategies are embedded in the analysis. The findings are presented not only in terms of the potential of selected strategies to promote higher-order thinking, but also in relation to the effectiveness of teacher behaviours to facilitate same.

### **Teaching/learning strategies**

Table 1 shows the teaching strategies used by the sample of teachers observed. As illustrated in this table, teaching strategies were categorized according to their potential for promoting higher-order thinking. The results indicate that the proportion of trained teachers using the five strategies that could definitely promote higher-order thinking was high for one category (probing questions), medium for three categories, and low for one category. The proportion of untrained experienced and untrained inexperienced teachers using these strategies was low for four of the five categories, and medium for one. Overall, a higher proportion of the trained teachers used strategies that had the potential to definitely promote higher-order thinking than the untrained teachers.

Six of the strategies were categorized as having the potential for possibly promoting higher-order thinking. The proportion of trained teachers using these strategies was high for two of the categories (small-group activities and practical activities), and medium for four. The strategy

used by a high proportion of untrained experienced and untrained inexperienced teachers was the use of textbooks.

With the exception of the use of textbooks, the proportion of untrained experienced teachers using the other strategies in this category was low. What is noteworthy is that a medium proportion of untrained inexperienced teachers used demonstrations and visual aids, which was not evident with their experienced colleagues. This was an example of untrained inexperienced teachers spontaneously or intuitively using strategies that could possibly promote higher-order thinking.

Conversely, all the strategies categorized as having the potential to marginally promote higher-order thinking were used by a high proportion of untrained teachers in both groups, with the exception of copying notes. A medium proportion of trained teachers used these strategies. It is evident that a higher proportion of untrained teachers selected those strategies that had the potential to marginally promote higher-order thinking.

A critical analysis of how some of the strategies were used, and the extent to which the potential for developing higher-order thinking was realised, follows.

The most frequently used strategy was the quick-paced question and answer sequence (recitation)—that is, initiation, response, feedback sequence, generally referred to as IRF (Sinclair & Coulthard, 1975), which involved two-way interaction. The following examples illustrate:

**T:** What are some examples of organisms that show asexual reproduction?

**S:** Amoeba.

**T:** Amoeba. Very good.

**S:** Algae.

**T:** Some algae, what else?

**Table 1. Strategies used by Lower Secondary Science Teachers for Delivering Instruction**

Strategy	Potential for Promoting Higher-order Thinking	Frequency/ Proportion of Trained N = 10	Frequency/ Proportion of Untrained Experienced N = 7	Frequency/ Proportion of Untrained Inexperienced N = 10	Total Frequency
Use of Ss past experiences	Definitely	6 (medium)	4 (medium)	3 (low)	13
Questioning (probing)		7 (high)	1 (low)	3 (low)	11
Wait-time strategy		5 (medium)	1 (low)	3 (low)	9
Linking new concepts		4 (medium)	1 (low)	3 (low)	8
Analogies		3 (low)	0 (low)	3 (low)	6
Use of textbook	Possibly	5 (medium)	6 (high)	7 (high)	18
Small-group activities		9 (high)	2 (low)	3 (low)	14
Practical activities		9 (high)	2 (low)	2 (low)	13
Demonstration		5 (medium)	2 (low)	3 (low)	10
Use of visual aids		4 (medium)	2 (low)	3 (low)	9
Discussion		5 (medium)	2 (low)	1 (low)	8
Questioning (recitation)	Marginally	6 (medium)	6 (high)	9 (high)	21
Lecturing /teacher presentation		4 (medium)	7 (high)	7 (high)	18
Copying/ reading/ collecting notes		6 (medium)	2 (low)	8 (high)	16

**Key:** For N = 10, 7 - 10 = high, 4 - 6 = medium, 0 - 3 = low  
 For N = 7, 5 - 7 = high, 3 - 4 = medium, 0 - 2 = low

- T:** The last thing we did was blood vessels. You all recall that? What are the three types of blood vessels?
- S:** Arteries, veins, capillaries.
- T:** What do arteries do?
- S:** Carry blood away.
- T:** Most arteries carry blood away from the heart. What about that blood?
- S:** What was the question?
- T:** If you remember the structure of the heart, left and right...

The lecture method was the second most frequently used strategy. This involved one-way interaction in which students were passive recipients. Practical work was the next most common strategy adopted. This involved interaction between students and manipulatives. Of these three dominant strategies, only practical work provided the kinds of experiences that had the potential to possibly promote higher-order thinking.

Practical work was included in 22 of the 54 lessons observed. Six of the trained teachers and four of the untrained experienced teachers implemented verification laboratory activities, either as group activities or as demonstrations. The main intent of these sessions was to confirm principles and concepts that had been discussed in a previous class. For example, after a series of lessons on balancing equations, the mole concept, and neutralization reactions, an untrained, experienced teacher took the students into the laboratory to engage in procedures to verify the concepts. The students were therefore involved in deductive reasoning. In some cases, the teachers attempted to take the exercise further by having students apply the concepts to solve a problem. On these few occasions, the students were given opportunities to think at Bloom et al.'s (1956) level of application or Costa's (2001) process level.

In contrast, only three trained teachers and one untrained experienced teacher involved their students in inductive reasoning for concept development, during laboratory activities that would have required students to make inferences based on a series of observations or data in order to make generalizations. In one case, students were provided with apparatus and materials, and through questioning and analysis of data were led to induce Hooke's law. These are examples of experiences where there was the potential for students to operate at Costa's (2001)

output level. During another laboratory activity, students did calculations of surface area/volume ratios using samples of cheese, and then made inferences with respect to surface area and rate of absorption in the small intestine and in the lungs. These latter are examples of Costa's process level of thinking.

There were a few occasions when teachers used analogies, visual aids/models, concept mapping, and the learning cycle (which involves exploration, introduction, and application of concepts). These less-frequently seen strategies, during which students were operating at Costa's (2001) process level, were used by trained and untrained teachers alike, and provided opportunities for students to make linkages, compare, look for patterns, apply principles, and make inferences. For example, an untrained teacher used students' prior knowledge of pressure and water pumps at the local water authority—the Water and Sewage Authority (WASA)—to make the link between the structure of the heart (muscular walls) and its functions in withstanding pressure. An example of the learning cycle strategy involved the use of a model to represent airflow in a system as a result of a pressure differential. Students were given opportunities to compare the parts of the model with the respiratory system. They then applied the principles behind the operation of the model to the mechanism of breathing in the lungs. In the latter two cases, a high level of thinking was being promoted since students were required to operate at Costa's output level, that is, to look for patterns, derive generalizations, and make judgements based on evidence. These opportunities for thinking at this level were, however, the exception instead of the rule.

### **Teachers'/students' questions**

The lower secondary science teachers asked a variety of questions. These questions were predominantly of the low-level type, and teachers used them in various ways. One use of questions was to find out what students had learnt either during the lesson or from a previous lesson. For example:

*What is a hormone? What are some examples of hormones you know?*

*What kinds of bonding occur between atoms?*

In such cases, students' responses were short, quick, and knowledge-based. For example, in response to the question requesting examples of hormones, students' responses included insulin and adrenaline. They were not required to engage in any extended mental processing before responding, since their responses were immediate and they were simply recalling information held in memory. They were operating at Bloom's level of recall or the level referred to by Costa (2001) as the input level.

The teachers also used questions to identify gaps in students' knowledge, to draw attention to misconceptions, or to link existing prior knowledge to new learning. The examples below from an untrained teacher and a trained teacher illustrate the latter. After a discussion on the structure and function of arteries and veins, the untrained inexperienced teacher wanted the Form 3 students to make an inference about the structure of capillaries based on the pre-established functions. In an attempt to have students operate at Costa's (2001) process level, she asked:

*What do you think the structure of the capillary would look like?*

After the question was posed, the students did not respond immediately. The teacher was then unable to provide the necessary follow-up questions and scaffolds, and instead resorted to giving the information herself. This was an example of ineffective teacher behaviour.

During a lesson on food webs with a Form 1 class, a trained teacher attempted to elicit from students various sources of food for a selected organism in a food chain. He mentioned that the chain was becoming complex or "tangled-up." He then tried to use students' knowledge of ducks' webbed feet to introduce the concept of interrelationships among the chains that comprise a food web. However, the teacher did not use probing questions to have one of the students make her reasoning more explicit, and, hence, was unable to effectively facilitate the use of analogical thinking as intended. The following represents the exchange between the teacher and the student:

**T:** Think of the duck's feet. What's between?

**S:** Toes, phalanges.

**T:** Define the term *food web* in terms of food chains. What does it show us?

**S:** It shows what feeds on what.

**T:** Is there any connection?

A similar pattern of teachers' inability to use probing questions, or to scaffold students' ideas was apparent during a lesson on vegetative propagation and cloning with Form 3 students. The untrained inexperienced teacher presented explanations on different forms of plant clones. A student asked the question: "What is tissue culture?" The teacher then posed this question to the class: "What do you think tissue culture might mean?" The students did not respond after appropriate wait-time, and the teacher gave an explanation of the process. It was evident from her follow-up explanations that she was hoping that the students would have been able to make a link between bacterial cultures growing in a medium (that had been previously discussed) and this new concept of tissue culture. However, the teacher lacked the skills to ask the kinds of probing questions that would enable the students to think at Costa's (2001) process level in order to make this connection, and thus to bridge the gap in understanding of the cloning process. This pattern was also observed with untrained experienced teachers.

The majority of teachers who asked probing questions used wait-time effectively. But even on these few occasions, the teachers were unable to probe effectively to have students make their explanations more robust. The following is an example of how one trained teacher attempted to use probing questions, which would have allowed students to operate at Costa's (2001) output level, that is, to make an evaluation or prediction, or develop a hypothesis during a lesson on the characteristics of a scientist. After giving short narratives on the contributions of some scientists (Democritus, Dalton, Thompson, and Rutherford) to the development of the atomic theory, he asked the following questions:

*What does all that suggest?*

*If we look at the time?*

It is highly likely that his intention here was to have the students synthesize the information in order to look for patterns in the development of the atomic theory. For example, the students could have focused on the use of confirmatory evidence to build the theory, or on evidence that may have disconfirmed the existing theories and resulted in a change of thinking. Rather than continue with this line of questioning after an unexpected response from the student, which was "It had

scientists in Jesus' time," he proceeded to change the line of questioning as follows:

**T:** What did it take to become a scientist?

**S:** Hard work.

**T:** Why was it hard?

**S:** Because they had to search for answers.

**T:** And what does that involve?

**S:** Being able to know what the answer is.

**T:** Yes. But in doing that what do you need?

**S:** Patience, time, investigation, instruments.

The line of questioning did not really challenge the student to make any inferences about the qualities of scientists, for example, being open-minded (willing to change ideas and theories in the light of new evidence), or persevering (willing to take tasks to a conclusion), or curious (a disposition to investigate). In this example, the second question does not provide a scaffold for the student to extend his understanding of the concept of *hard work* as it applies to a scientist. Instead, after the student's response to the first question, the teacher might have asked "Can you give some examples of ways in which any one of the scientists demonstrated that he was engaged in hard work?" This line of questioning would have challenged the student to reflect on his understandings of the nature of scientific work, and to understand more fully why he described scientific work as "hard." The assumption here is that the teacher intended to have the student give examples and then match the examples with evidence from the narrative. In this way, the teacher would have interrogated the students' understandings of the work of scientists, and would have helped them to discern the relationship between science and the qualities of the scientist, and the relationship between evidence and formulation of theory.

Little attention was given to providing a forum for students' ideas about the particular science concepts under examination or to the generation of students' questions in relation to science phenomena. Students' questions, which were significantly fewer than their teachers' (1:10), were pitched mainly at the levels of knowledge and comprehension (Bloom et al., 1956). Examples of low-order questions include "What are cuttings?" in a lesson that dealt with vegetative propagation, and "This is the same displacement?" (Translation: Is this another example of a

displacement reaction?) in a lesson that dealt with types of chemical reactions.

However, students did ask a higher percentage (25%) of higher-order questions, compared to their teachers (14%). In most cases, the higher-order questions posed by students were not handled satisfactorily by the teachers. For example, during a lesson on blood circulation, a male student sought some explanations on what seemed to be a discrepancy. The untrained, inexperienced teacher had referred to wrist slashing as a common method of suicide. This immediately seemed to result in cognitive conflict for the student, who asked: “Miss, and they does cut off people hand?” (Translation: Miss, don’t surgeons amputate people’s hands?). The implied reasoning here seemed to be that since people do not die from amputations then they should not die from wrist slashing. The teacher, however, did not respond to the question; hence, there was no resolution for the student. Most importantly, however, it was a missed opportunity for engaging not only the student but also the entire class in high-level cognitive operations, by use of probing questions that would have allowed for comparing and contrasting the two procedures (what occurs in a controlled environment, such as tying off blood vessels during surgery to prevent blood loss, versus what occurs in an uncontrolled situation), as well as looking for patterns and significant differences (Costa’s (2001) process level).

In summary, it was noted that trained and untrained teachers were unable to use questioning effectively. It was significant that three trained teachers resorted to use of low-level questions only and that trained teachers were unable to manage unexpected responses from students. While the trained teachers asked a higher percentage of higher-order questions than the untrained teachers, this percentage was very low (less than 25% of all the questions that they asked). In addition, untrained teachers did not deal satisfactorily with students’ higher-order questions. As outlined earlier, the literature suggests that higher-order questions posed by both teachers and students play an important role in students’ cognitive engagement and, by extension, in their thinking. It was evident that the teachers observed were not paying sufficient attention to the asking or eliciting of higher-order questions.

### **Range of process skills**

Table 2 shows the range of process skills promoted in the classrooms observed. Chiappetta and Koballa, Jr.'s (2002) framework for classifying process skills as basic and integrated was used. The latter requires higher-order thinking. All of the teachers observed implemented activities that gave students opportunities to use some of the basic process skills such as, observation, use of numbers, inferring, measuring, and predicting.

The skill of observation received attention from a high proportion of trained and untrained experienced teachers, and a medium proportion of untrained inexperienced teachers. The development of this basic process skill was evident in 22 of the 54 science lessons during short activities that formed part of a set induction, a demonstration, or a group activity. For example, an untrained experienced teacher demonstrated the difference in reactivity of metals (magnesium, zinc, and copper) with dilute acids to her Form 3 students. They observed the reaction between the metal and the acid, and measured the volume of gas produced over time to determine the rate as a measure of reactivity of the metal. A trained teacher began his Form 1 integrated science class on the expansion and contraction of substances by asking students to make predictions about the effect of heat on the metal ball. He then demonstrated the effect using the ball and ring apparatus. He further developed the concept by asking his students to make similar predictions about the expansion of liquids and gases.

Unfortunately, the opportunities for students to engage in complex thought processes during these activities were few. This was due either to poor task management or teachers' inability to use probing questions to help students to interpret what they were doing, why they were doing it, and what was the relevance to their everyday lives. Instead, the discussions focused primarily on the data collected and quantitative measures rather than, for example, on the relationship between evidence and explanations described by Millar (2004) as linking the two domains of knowledge: domains of objects and observables with that of ideas. In addition, there were few attempts to have students apply the information gleaned to solve problems in a new setting. Consequently, students were engaged most of the time in low-level thinking.

**Table 2. Process Skills Used in Lower Secondary Science Classrooms**

Process Skills	Frequency/ Proportion of Trained Teachers N = 10	Frequency/ Proportion of Untrained Experienced N = 7	Frequency/ Proportion of Untrained Inexperienced N = 10	Total Frequency
<b>Basic Skills</b>				
Observing	8 (high)	7 (high)	7 (high)	22
Inferring	7 (high)	3 (medium)	1 (low)	11
Measuring	7 (high)	2 (low)	1 (low)	10
Using numbers	5 (medium)	3 (medium)	1 (low)	9
Data gathering	5 (medium)	2 (low)	1 (low)	8
Predicting	3 (low)	1 (low)	0 (low)	4
Classifying	2 (low)	0 (low)	0 (low)	2
Space/time relation	0 (low)	0 (low)	1 (low)	1
<b>Integrated Skills</b>				
Communicating	10 (high)	5 (high)	5 (medium)	20
Experimenting	7 (high)	2 (low)	1 (low)	10
Interpreting data	3 (low)	2 (low)	0 (low)	5
Defining operationally	2 (low)	0 (low)	0 (low)	2
Hypothesizing	1 (low)	0 (low)	0 (low)	1
Formulating models	0 (low)	0 (low)	0 (low)	0
Controlling variables	0 (low)	0 (low)	0 (low)	0

Teaching to facilitate use of integrated skills was sometimes seen (see Table 2). The integrated skill of communication was the only one that was used by a high or medium proportion of all categories of teachers. The only other integrated skill that received attention from a high proportion of trained teachers was experimenting. The proportion of all categories of teachers who included the other integrated skills was low. These integrated skills require the use of complex processes that characterize higher-order thinking. Unfortunately, teachers did not focus sufficiently on including experiences requiring the use of integrated

process skills, which provide opportunities for students to give explanations, to ask questions, to adopt a new idea, or to highlight a discrepancy, all of which are examples of higher-order thinking behaviours.

### **Conclusion/Discussion**

This paper was intended to serve two purposes—to make inferences about the levels of thinking promoted in the science classrooms observed, and to provide research evidence to inform the approach for developing strong thinking skills pedagogy in teacher education programmes. The discussion addresses each purpose in turn.

With regard to the levels of thinking promoted, the findings indicate that, in general, higher proportions of the trained teachers selected strategies that either definitely or possibly had the potential to promote higher-order thinking than their untrained colleagues. Conversely, higher proportions of untrained teachers selected strategies that marginally had the potential to promote higher-order thinking than their trained colleagues. However, when all of the data on teachers' and students' behaviours are combined, it is evident that the lower secondary science classrooms observed were characterized by low levels of thinking. In sum, the patterns of interactions, the levels of teachers' questions, the paucity of students' questions, the choice of teaching strategies, as well as the limited range of process skills utilized all give rise to concerns.

The interactions in the classrooms of both trained and untrained teachers generally did not reflect the types of collaboration that could facilitate critical and creative thinking as well as reflective inquiry, as suggested by Kovalainen, Kumpulainen, and Vasama (2001). There was also a predominance of low-level teachers' questions that required thinking at Costa's (2001) input level. The teachers observed did not optimize the use of questions to probe students' responses and extend their thinking to higher cognitive levels. In addition, they did not use strategies that could encourage students' questions. For example, there was little evidence of problem solving—an activity that facilitates students' questioning (Chin, Brown, & Bruce, 2002). The literature reports that questioning is a complex activity which requires teachers to “monitor their own questions, pose questions that intentionally challenge students' intellect

and imagination and purposely draw forth students' awareness and employment of thinking skills, cognitive tasks and dispositions" (Costa, 2001, p. 360). There was little evidence that teachers' questioning techniques had been developed at the level of mastery that would facilitate attainment of these higher-order outcomes.

The teachers observed focused primarily on the development of basic process skills. The literature suggests that in spite of exposure to training which encourages the use of hypothetico-deductive reasoning, many teachers retain an empiricist/inductivist view of science based on their own beliefs about the nature of science (Millar, 2004). This may explain the findings of this study that teachers focused their practical work on the processes of gathering empirical data, and that students were rarely exposed to practical work that required the use of integrated skills such as hypothesizing and experimenting, which develop higher-order thinking.

There was some evidence that teachers, either intuitively or based on prior knowledge and experiences/training, attempted to use strategies that could take their students to higher levels of cognitive engagement. However, both trained and untrained teachers lacked the skills and expertise to effectively use these strategies. On a few occasions, teachers were able to facilitate students' thinking up to Costa's (2001) process level, but facilitation of thinking at the output level was rare. The research evidence points to the important role of the teacher in placing pedagogical emphasis on collaborative group work, effective patterns of classroom talk and interactions, and in eliciting pupils' responses as approaches to developing thinking and making reasoning explicit (Higgins et al., 2004). For the teachers surveyed, the implementation of these principles was either not evident or ineffective.

The untrained teachers did not or could not, as a rule, select strategies that had the potential to definitely promote high levels of thinking. It may be inferred that their practical teacher knowledge was insufficient to facilitate higher-order cognitive engagement in the classroom. While the trained teachers attempted to include some of these strategies, their inability to utilize such strategies effectively is consistent with findings on the behaviours of trained teachers reported in the literature (Beyer, 2001b; Zohar, 2004). The actual teacher behaviours observed in this

study were therefore significantly different from the behaviours needed to promote higher-order thinking. The findings therefore indicate the need for a stronger thinking skills pedagogy in the teacher education programme, as discussed below.

When the findings are examined within the Trinidad and Tobago context, there may be two plausible explanations for the behaviours of the trained teachers. Firstly, at the macro level, the model of teacher education that informs the science curriculum process does not focus sufficiently on the eliciting of teachers' prior knowledge within the cultural context in which teaching occurs. The findings perhaps illustrate the weaknesses of the teacher behaviour development model reported in the literature (Koballa, Jr. & Tippins, 2001). This top-down approach, in which issues related to the development of student thinking (e.g., questioning in the science classroom) are selected by "experts," may not have allowed teachers to reconcile what might be perceived as problematic by experts with their own perceptions about problems that exist in the classroom.

Secondly, the specific micro-level strategies used by teacher educators on the Dip.Ed. programme should be revisited. It is clear that more attention must be paid to specific sub-skills such as: (1) teachers' use of probing questions to have students (a) strengthen and extend their explanations, (b) compare and contrast, and (c) discern patterns and relationships; (2) skills to scaffold students' ideas to extend their understanding; and (3) skills for eliciting students' questions.

The uncovering of the deficiencies noted during the post-training observation of the teachers is significant. It has implications for the type of model used in training programmes to develop teachers' capability to use strategies that could enhance thinking, and for the types of strategies that the science teacher educators use to assist in building teachers' competencies. The success of most interventions aimed at improving thinking is usually based on the evaluation of the outcomes, in terms of pupils' attainment. However, monitoring the effectiveness of the processes used during the implementation, as done in this study, can also provide valuable feedback that may improve outcomes.

After training, it is difficult for teachers to display behaviours that contradict entrenched cultural patterns. Beck and Kosnik (2006) describe

individuals who attempt such changes on their own as playing Don Quixote, suggesting that such change “goes nowhere” (p. 41). Over the years, a culture has developed that has socialized many teachers and students into a style of teaching in which teachers adopt the role of transmitters and students function as passive recipients. In addition, teachers who have been socialized into this traditional mode of teaching are often resistant to change (Rudduck & Flutter, 2000), even in the face of curriculum reform initiatives that include goals related to the development of critical and creative thinking.

How then can our science teachers be empowered to provide the kinds of experiences needed to promote the development of competent thinkers? A three-pronged approach is suggested. While this approach is directly applicable the Trinidad and Tobago context, there are lessons for teacher development programmes in general.

Firstly, there is need for a re-examination of the model of teacher development on which the Dip.Ed. programme is based. A re-orientation towards the social reconstructivist model (see Beck & Kosnik, 2006) of teacher development is suggested. The teachers and teacher educators should work together to design the course content by reflecting upon teacher behaviours in the science classroom, and on their concerns about levels of thinking promoted. Engaging in reflection would provide opportunities for teachers to deconstruct their prior knowledge, beliefs, and attitudes about teaching thinking, to comprehend how these understandings evolved, and to explore the effects they have on actions and behaviour (Abdal-Haqq, 1998; Zohar, 2004). The reflective process should allow teachers and teacher educators to engage in collaborative inquiry about levels of student thinking, while still keeping the focus on building teacher capacity to address the deficiencies identified. Teachers and tutors could also collaborate to develop complex tasks that demand a high level of cognitive engagement at Costa’s (2001) output level. A data base of such tasks could be developed for general use.

Secondly, a thinking skills approach as a direct intervention for building teacher capacity (skills and dispositions) might be included in the programme. The literature suggests that the use of thinking skills approaches as a direct intervention has a positive impact on pupils’ attainment (Higgins et al., 2004). Tishman, Jay, and Perkins (1992) also

suggest that certain dispositions, such as the tendency to explore, to inquire, to seek clarity, to take intellectual risks, and to think critically and imaginatively, contribute to good thinking. Training that focuses on teaching thinking ought to cultivate these dispositions.

Thirdly, there should be site-based interventions for untrained teachers. Such interventions could draw on approaches used elsewhere, for example, the *Thinking in the Science Classroom* (TSC) project (Zohar, 2004), and also build on some of the positive aspects of the practical knowledge that was apparent in the study. The interventions should be flexible enough to monitor both processes and outcomes of the training. Such interventions are important because the model of teaching that guides teachers' practice on entry to the service is often based on: (a) experiences as students, (b) knowledge and beliefs about teaching and learning, (c) observation of peers, (d) advice from a senior teacher/mentor, or (e) a combination of these (see Rampersad & Herbert, 2005), and this model does not necessarily promote the development of thinking.

This study provided some insights into levels of thinking into selected science classrooms. It should serve as a springboard for more in-depth research to better understand the cultures in which science teachers operate, and the challenges that they face as they attempt to enact science curricula. Such research should also reveal the kinds of interventions and resources that are necessary to support the development of teachers who can think critically, and who can in turn promote the development of higher-order thinking in the science classroom.

### References

- Abdal-Haqq, I. (1998). *Constructivism in teacher education: Considerations for those who would link practice to theory*. Washington, DC: ERIC Clearinghouse on Teaching and Teacher Education. Retrieved September 7, 2006, from <http://www.ericdigests.org/1999-3/theory.htm>.
- Adey, P., & Shayer, M. (1994). *Really raising standards*. London: Routledge.
- Alsop, S., de Jesus, H. P., & Watts, M. (n.d.). *Student-generated questions in the learning and teaching of chemistry*. Retrieved January 15, 2005, from [http://www.roehampton.ac.uk/staff/mikewatts/docs/aera\\_submission.doc](http://www.roehampton.ac.uk/staff/mikewatts/docs/aera_submission.doc).
- Beck, C., & Kosnik, C. (2006). *Innovations in teacher education: A social constructivist approach*. Albany, NY: State University of New York Press.

- Beyer, B. (2001a). What research says about teaching thinking skills. In A. L. Costa (Ed.), *Developing minds: A resource book for teaching thinking* (3<sup>rd</sup> ed., pp. 275–282). Alexandria, VA: Association for Supervision and Curriculum Development.
- Beyer, B. (2001b). Teaching thinking skills: Defining the problem. In A. L. Costa (Ed.), *Developing minds: A resource book for teaching thinking*, (3<sup>rd</sup> ed., pp. 35–40). Alexandria, VA: Association for Supervision and Curriculum Development.
- Bitner, B. (1991). Formal operational reasoning modes: Predictors of critical thinking abilities and grades assigned by teachers in science and mathematics for students in grades nine through twelve. *Journal of Research in Science Teaching*, 28(3), 265–274.
- Bloom, B. S., Englehart, M. D., Furst, E. J. Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives, Handbook I: Cognitive domain*. New York: Longmans, Green.
- Blosser, P. E. (1990, March 1). Using questions in science classrooms. *Research matters – to the Science Teacher*, No. 9001. Retrieved November 27, 2004, from <http://www.narst.org/publications/research/question.htm>.
- Brickhouse, N. W. (1990). Teachers' beliefs about the nature of science and their relationship to classroom practice. *Journal of Teacher Education*, 41(3), 53–62.
- Chiappetta, E. L., & Koballa, Jr., T. R. (2002). *Science instruction in the middle and secondary schools*. Columbus, OH: Merrill/Prentice Hall.
- Chin, C. (2004). Students' questions: Fostering a culture of inquisitiveness in science classrooms. *School Science Review*, 86(314), 107–112.
- Chin, C., Brown, D. E., & Bruce, B. C. (2002). Student-generated questions: A meaningful aspect of learning in science. *International Journal of Science Education*, 24(5), 521–549.
- Costa, A. L. (2001). Teacher behaviors that enable student thinking. In A. L. Costa (Ed.), *Developing minds: A resource book for teaching thinking* (3<sup>rd</sup> ed., pp. 359–369). Alexandria, VA: Association for Supervision and Curriculum Development.
- Cotton, K. (1988). *Close-Up #5: Classroom questioning*. Retrieved June 19, 2004, from Northwest Regional Educational Laboratory's School Improvement Research Series Web site: <http://www.nwrel.org/scpd/sirs/3/cu5.html>.
- Cotton, K. (1991). *Close-Up # 11: Teaching thinking skills*, Retrieved June 18, 2004, from Northwest Regional Educational Laboratory's School Improvement Research Series Web site: <http://www.nwrel.org/scpd/sirs/6/cu11.html>.
- Facione, P. A. (1998). *Critical thinking: What it is and why it counts?* Retrieved June 9, 2004, from <http://www.insightassessment.com/>.

- Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record*, 103(6), 1013–1055
- Fisher, A. (2001). *Critical thinking: An introduction*. Cambridge: Cambridge University Press.
- Hashweh, M. Z. (1996). Effects of science teachers' epistemological beliefs in teaching. *Journal of Research in Science teaching*, 33(1), 47–63.
- Higgins, S., Baumfield, V., Lin, M., Moseley, D., Butterworth, M., Downey, G., Gregson, M., Oberski, I., Rockett, M., & Thacker, D. (2004). Thinking skills approaches to effective teaching and learning: What is the evidence for impact on learners. In *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.
- Johnson, D. W., & Johnson, R. T. (2001). Cooperation and conflict: Effects on cognition and metacognition. In A. L. Costa (Ed.), *Developing minds: A resource book for teaching* (3<sup>rd</sup> ed., pp. 445–458). Alexandria, VA: Association for Supervision and Curriculum Development.
- Koballa, Jr., T. R., & Tippins, D. J. (2001). Portraits of professional development models in science teacher education: A synthesis of perspectives and issues. In D. R. Lavoie & W-M. Roth (Eds.), *Models for science teacher preparation: Theory into practice* (pp 213–224). Dordrecht, Netherlands: Kluwer.
- Kovalainen, M., Kumpulainen, K., & Vasama, S. (2001). Orchestrating classroom interaction in a community of inquiry: Modes of teacher participation. *Journal of Classroom Interaction*, 36(2)/37(1), 17–28.
- Leach, J., & Scott, P. (2000). Children's thinking, learning, teaching and constructivism. In M. Monk & J. Osborne (Eds.), *Good practice in science teaching: What research has to say* (pp. 41–56). Buckingham, UK: Open University Press.
- Millar, R. (2004, June). *The role of practical work in the teaching and learning of science*. Paper prepared for the Meeting: High School Science Laboratories: Role and Vision, National Academy of Sciences, Washington, DC.
- Millar, R., & Driver, R. (1987). Beyond processes. *Studies in Science Education*, 14(9), 33–62.
- Mislevy, R. J. (2004). *The case for an integrated design framework for assessing science inquiry* (CSE report 638). Los Angeles, CA: Center for the Study of Evaluation (CSE), National Center for Research on Evaluation, Standards, and Student Testing (CRESST)
- Onosko, J. J. (1990). Comparing teachers' instruction to promote students' thinking. *Journal of Curriculum Studies*, 22(5), 443–461.
- Rampersad, J., & Herbert, S. (2005). Reflections of science teachers in an in-service teacher education programme. *Caribbean Curriculum*, 12, 1–23.

- Rowe, M. B. (2003). Wait-time and rewards as instructional variables: Their influence on language, logic, and fate control: Part 1- Wait-time (Reprinted article) *Journal of Research in Science Teaching*, 40(S1), S19–S32. (Original 1974, 11(2), 81–94).
- Rudduck, J., & Flutter, J. (2000). Pupil participation and pupil perspective: ‘Carving a new order of experience.’ *Cambridge Journal of Education*, 30(1), 75–89
- Shakespeare, D. (2003). Starting an argument in science lessons. *School Science Review*, 8 (311), 103–108.
- Sinclair, J., & Coulthard, M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. London: Oxford University Press.
- Snell, J. (1999). *Improving teacher-student interaction in the EFL classroom: An action research report*. Retrieved November 26, 2004, from <http://iteslj.org/Articles/Snell-Interaction.html>.
- Swartz, R. J., & Fischer, S. D. (2001). *Teaching thinking in science*. In A. L. Costa (Ed.), *Developing minds: A resource book for teaching thinking*, (3<sup>rd</sup> ed., pp. 303–309). Alexandria, VA: Association for Supervision and Curriculum Development.
- Swift, J. N., Gooding, C. T., & Swift, P. R. (1996, October). Using research to improve the quality of classroom discussions. *Research Matters - to the Science Teacher*, No. 9601.
- Tishman, S., Jay, E., & Perkins, D. N. (1992). *Teaching thinking dispositions: From transmission to enculturation*. Retrieved April 1, 2006, from <http://learnweb.harvard.edu/ALPS.thinking/docs/article2.html>.
- Trinidad and Tobago. Ministry of Education. (2003). *Secondary Education Modernization Programme, Secondary school curriculum. Form Three. Science*. Port of Spain: Author.
- The University of the West Indies.. School of Education. (2004). *Regulations & syllabuses 2004–2005*. St. Augustine, Trinidad: Author.
- Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Zimmerman, C., Raghavan, K., & Sartoris, M. L. (2003). The impact of the MARS curriculum on students’ ability to coordinate theory and evidence. *International Journal of Science Education*, 25(10), 1247–1271.
- Zohar, A. (2004). *Higher order thinking in science classrooms: Students’ learning and teachers’ professional development*. Dordrecht, Netherlands: Kluwer.
- Zohar, A., Weinberger, Y., & Tamir, P. (1994). The effect of the biology critical thinking project on the development of critical thinking. *Journal of Research in Science Teaching*, 31(2), 183–196.
- Zoller, U. (1987). The fostering of question-asking capability. *Journal of Chemical Education*, 64(6), 510–512.

### Notes on Contributors

**Launcelot I. Brown** is a Lecturer at the School of Education, Duquesne University, Pittsburgh, USA. His current research investigates the role of school leadership and other organizational and wider systemic factors that impact the effectiveness level of the school. The ultimate aim of the research is to develop models of school leadership relevant to the Caribbean context. Related to his interest in school effectiveness are issues of gender differentials in academic performance and the various factors that maintain the differential in student achievement.

**Susan Herbert** is a Lecturer in Science Education at the School of Education, The University of the West Indies (UWI), St. Augustine. She taught chemistry and general science at the secondary level for 15 years.

**Cynthia James** is a Lecturer in Education (English Curriculum) at the School of Education, UWI, St. Augustine, who has also lectured in the Department of Liberal Arts, UWI, St. Augustine. Dr. James is a graduate of Howard University, Washington, D.C., and taught for many years in the public school system in Trinidad.

**Gibbs Y. Kanyongo** is an Assistant Professor of Educational Statistics and Research Methods at Duquesne University, Pittsburgh, USA. His research interests are in multivariate statistics and measurement issues in large-scale assessment tests. He is also interested in research on mathematics education.

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

**Joycelyn Rampersad** is a Lecturer in Science Education at the School of Education, UWI, St. Augustine. She has over 35 years of experience in education at both the secondary and tertiary levels. Her research interests are reflective teaching, classroom practice, and assessment. She is currently pursuing a Ph.D. in Education.

## 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. 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. 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
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: *jfullerton-rawlins@fhe.uwi.tt.*

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

<b>ORDER FORM</b>
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 (5<sup>th</sup> 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.