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*.

ISSN 1017-5636

CARIBBEAN CURRICULUM

Volume 13

2006

CONTENTS

ETHNIC AND GENDER DIFFERENCES IN SELF-REPORTED ACHIEVEMENT AND ACHIEVEMENT-RELATED ATTITUDES IN SECONDARY SCHOOL STUDENTS IN TRINIDAD[1]

Frank C. Worrell[2]

In this study, self-reported achievement, achievement-related behaviours, and achievement-related attitudes of 1,434 students attending secondary schools in Trinidad were examined. Females reported higher achievement than males, and males reported cutting class more than females, and both of these differences yielded medium effect sizes. Females also reported completing homework more frequently and higher academic perceived life chances than males. East Indian students reported higher achievement, homework completion, time on schoolwork, and academic perceived life chances than their Black and Mixed counterparts, as well as spending less time with friends during the week and lower rates of cutting class. However, all of the ethnic comparisons yielded low effect sizes. Given the differences found and the potential for achievement differences to increase over time, more research on gender and ethnic group differences is recommended.

Although no consistent gender differences have been found in measures of general cognitive ability (Burstein, Bank, & Jarvik, 1980; Hyde, 2005; Lezak, 1976; Maccoby & Jacklin, 1974), consistent gender differences have been reported in specific domains (Denno, 1982; Moore & Smith, 1987). For example, researchers have reported that females obtain higher scores than males on verbal ability measures (e.g., Burstein et al.; Cohen & Wilkie, 1979; Maccoby & Jacklin; Matarazzo, 1972), and males obtain higher scores than females on spatial tasks from adolescence onward (e.g., Cohen & Wilkie). Similar patterns have been reported with

[1] This study was conducted under the auspices of the School Leadership Center of Trinidad and Tobago.

[2] Correspondence concerning this article should be addressed to Frank C. Worrell, 4511 Tolman Hall, Berkeley, CA 94720-1670. Electronic mail can be sent to frankc@berkeley.edu.

regard to self-concept and motivational variables, with males obtaining higher scores on mathematical self-concept, females obtaining higher scores on verbal self-concept (e.g., Marsh, Parker, & Barnes, 1985), and neither gender consistently higher on global academic self-concept (e.g., Battle, 1976; Mullis, Mullis, & Normandin, 1992; Worrell, Roth, & Gabelko, 1998).

With regard to ethnicity, studies in the US have indicated that African Americans generally obtain lower scores on both cognitive (Neisser et al., 1996; Williams & Ceci, 1997) and achievement measures (Singham, 2003) than their White and Asian American counterparts, and though trends indicate that the gap has diminished since the 1970s, there are still substantial differences between the groups. In this paper, I examined gender and ethnic group differences in self-reported achievement and variables related to achievement in adolescents in Trinidad.

Gender Differences in Achievement

In spite of the apparently consistent findings in earlier research in this area (Burstein et al., 1980; Cohen & Wilkie, 1979), more recent research evidence has questioned the generality of the findings on gender differences in the US (e.g., Kohr, Masters, Coldiron, Blust, & Skiffington, 1989; Sabers, Cushing, & Sabers, 1987; Skaalvik & Rankin, 1994), as well as the magnitude or meaningfulness of these differences when they occur (e.g., Hogrebe, Nist, & Newman, 1985). For example, in reporting on a meta-analysis of 165 studies that examined gender differences in verbal ability, Hyde and Linn (1988) concluded that gender differences in verbal ability are so small, they "can effectively be considered zero" (p. 64). A meta-analysis of spatial ability tasks led to a more nuanced conclusion about gender differences. Linn and Petersen (1986) found consistent gender differences favouring males on spatial perception and mental rotation. However, on spatial visualization tasks, which make up "the largest category of spatial ability tests,…no significant or meaningful gender differences were found" (p. 74).

Hogrebe et al. (1985) examined gender differences in reading achievement using an extremely large sample from the High School and Beyond national survey (N = 48,040). Not only were some of the relationships between gender and reading achievement not significant in

spite of the enormous sample size, gender also accounted for less than 1% of the variance in reading achievement. These researchers concluded that the null hypothesis of no relationship between gender and reading achievement was supported.

In a recent article, Hyde (2005) summarized meta-analytic studies on gender differences across a variety of constructs and domains, including cognitive variables, communication, social and personality variables, psychological well-being, motor behaviours, moral reasoning, charting behaviours and attitudes, job preference attitudes, and computer use and self-efficacy. Only 28 (22%) of the 124 effect sizes examined were in the moderate to large range—"that is, 78% of gender differences are small or close to zero" (pp. 582-586). The largest and most consistent differences favoured males in motor performance, casual attitudes about sex, and aggression. Differences in achievement that yielded moderate to large effect sizes were spelling and language favouring females, and mechanical reasoning, spatial perception, and mental rotation favouring males. In sum, the research on gender differences in achievement indicates that males and females are more alike than different on most variables.

Ethnic Differences in Achievement

Research on Black-White differences in the US has indicated that although the magnitude of some of the differences declined through the 1970s and 1980s, they stabilized in the 1990s and gaps still remain (e.g., Lee, 2002). For example, African Americans have scored about one standard deviation (SD) lower than Whites (Neisser et al., 1996) on early editions of IQ tests—i.e., 16 points on the Stanford-Binet and 15 points on the Wechsler scales. However, the Black-White difference on the fourth edition of the Stanford-Binet test ranged from 10 to 13 points (Thorndike, Hagen, & Sattler, 1986), and the difference on the most recent edition of the Wechsler Intelligence Scale for Children-Fourth Edition, which was published in 2003, is 11.5 points (Sattler & Dumont, 2004). Information on racial and ethnic group differences is not included in the technical manual for the fifth edition of the Stanford-Binet (Roid, 2003).

The decrease in the IQ difference is also reflected in reading and mathematics achievement scores (e.g., Lee, 2002; Williams & Ceci, 1997). For example, the average difference in mathematics achievement for 17-year-olds dropped from 1.1 SDs in 1978 to .65 SDs in 1990 (Neisser et al.). Although the differences are smaller, they are still substantial, and African Americans still graduate from high school and enrol in college at lower rates than their White counterparts (Frederick D. Patterson Research Institute, 1997a, 1997b). Thus, unlike the gender differences, ethnic differences in achievement in the US are robust and reported consistently in the literature.

The Cultural Context of the Study

Before proceeding further, it is important to provide the cultural context for the current study. Trinidad and Tobago is a twin-island nation and the southernmost country in the Caribbean chain. The country is less than 2,000 square miles in area, but it is one of the most prosperous nations in the Caribbean, due to large oil and natural gas reserves (Brereton, 1981). The country is also one of the most cosmopolitan of the Caribbean nations. Based on the 2000 census, the country's population of 1.1 million includes people of African (37.5%), East Indian (40%), and Mixed (20.5%) descent, as well as smaller numbers of Chinese, European, and Arabic descent (1.2%), and an even smaller (0.8%) unspecified group. Trinidad and Tobago boasts one of the highest reported literacy rates in the world at 98% (Central Intelligence Agency, 2006[1]), although the functional literacy rate may only be about 55% (Health Ministry Official, personal communication, January 3, 2003).

Economic prosperity notwithstanding, Trinidad and Tobago (T&T) does not yet have universal education. The primary school attendance rate is only 93% (United Nations Development Programme[2] [UNDP], 2001), and this figure decreases at the beginning of secondary school by at least three percent (Trinidad and Tobago Central Statistical Office [CSO], 1998). Additionally, as in the US, there are increasing complaints from the business community about students completing their primary and secondary education with low literacy and numeracy skills. The claims of poor educational outcomes are complemented by dire crime statistics for adolescents and young adults. The largest age group in prison falls "between the ages of 17 and 21" (UNDP, 2001, p. 77), and "the average

age of the vast majority of violent perpetrators [is] 18 to 25" (Marajh, 2001, p. 10).

Gender and Ethnic Differences in Achievement in T&T

The problems in T&T noted above, particularly in the areas of crime and education, are often presented in specific ethnic and gender terms. In 1997, Gypsy, a leading calypsonian, won the Calypso Monarch title with a song entitled, *Little Black Boy*. In this calypso, Gypsy argued that Black boys study music and clothing in school, rather than reading and mathematics, and concluded that the fate of these Black boys, who are unemployable, is stealing and dying from a police bullet through the skull. To the extent that calypsoes reflect the zeitgeist of the culture, this calypso presents a bleak description of young Black males in T&T society.

Nor is Gypsy alone in his analysis. In 1998, Noguera and Worrell commented on the increased use of terms like *endangered*, *at-risk*, *marginal*, and *in crisis*, to discuss Black males in the Caribbean region (e.g., see Miller, 1992, 1994). Subsequently, Worrell and Noguera (2000) reported that Black male secondary school students in T&T agreed that Black males in T&T are in crisis, and they defined the crisis in terms of personal responsibility, much as Gypsy defined it. In another line of research on gender, Parry (1996) discussed the patriarchal view in Caribbean society of education as effeminate and, consequently, inferior, leading to an anti-academic orientation on the part of males in schools.

The concerns about gender are also evident in the crime statistics, and declining male enrolment at The University of the West Indies (UWI), alongside increasing female enrolment. The concern about the gender gap was reflected in the theme chosen by the School Leadership Center of Trinidad and Tobago for its 2002 institute: "Boys to men: The challenges of engaging boys academically and emotionally in the primary and secondary school system" (see Worrell, 2003a, 2003b). Traditionally, little empirical data have been available in the less-developed countries, as these governments do not provide substantial funding for educational and social science research as in the more-developed countries. However, there are a number of sources of data in

Trinidad and Tobago that address gender or ethnic differences on educational variables.

Gender differences

Government statistics indicate that males are being held back a grade in greater numbers at all of the primary grade levels, except Standards 5 and 6—the years of transition to the secondary school system (CSO, 1998). Moreover, primary school dropout rates for the school years for 1993/1994, 1995/1996, and 1996/1997 indicate that males are also leaving school in greater numbers than females (UNDP, 2001). Additionally, the percentage of females enrolled in all secondary schools is higher than the percentage of males enrolled, with the difference being greatest in the rural counties (UNDP), although boys are being placed in junior secondary schools in greater numbers than females. In sum, these data indicate that boys are more likely to leave the school system before the secondary school years, and are more likely to be placed in the lowest-tier schools when they do attend secondary school. The higher placement rates in junior secondary schools also suggest that boys may be less likely to repeat Standard 5 to attain the test scores necessary to get into a higher-tier secondary school (Worrell, 2003b).

Kutnick, Jules, and Layne (1997) also examined data on gender differences in achievement in the early secondary school years. These researchers examined gender differences in student achievement at the end of the first and third years of secondary school for a group of males and females whose secondary school entrance examination scores did not differ significantly. Their results indicated that one and three years after entering secondary school, males were a substantial majority in the lower third of the achievement scores and were a minority in the upper third of achievement scores.

More recently, Worrell and his colleagues completed several projects involving primary and secondary school students in Trinidad and Tobago (Hall, Watkins, & Worrell, 2002; Watkins, Worrell, & Hall, 2002; Worrell, Hall, & Watkins, 2002; Worrell, Watkins, Runge, & Hall, 2002). In one study, Worrell, Watkins, et al. (2002) examined differences in scores on a phonemic awareness measure in a sample of over 4,000 students attending the first three years of school. They found no

differences between boys and girls on this measure of pre-reading skills. These findings were replicated in a second study using a representative sample of the school population (Hall et al., 2002).

Worrell, Hall, et al. (2002) examined gender differences on teacher ratings of student behaviour. The students who were rated (N = 700) were randomly sampled from all seven years of the primary school, and males and females were equally represented. Teachers completed the Learning Behaviors Scale (LBS, McDermott, Green, Francis, & Stott, 1999) and the Adjustment Scales for Children and Adolescents (ASCA, McDermott, 1994; McDermott, Marston, & Stott, 1993). Factor analyses yielded two LBS factors (Attitude Toward Learning and Strategy Flexibility) and three ASCA syndromes (Attention-Deficit Hyperactive, Conduct Problems, and Underactivity) that were invariant across gender and ethnicity (Worrell, Hall, et al.). Reliability estimates for subscale scores by gender grouping ranged from .80 to .91. No gender differences were found in Attitude Toward Learning or Strategy Flexibility. However, males were rated significantly higher on the Attention-Deficit Hyperactive factor, a common finding in the US (Havey, Olson, McCormick, & Cates, 2005).

At the secondary level, Watkins et al. (2002) collected self-report data from 897 Trinidadian students on depression and anxiety symptoms; verbal, mathematics, and global self-concept; and student fears. The instruments included the Reynolds Adolescent Depression Scale (RADS, W. M. Reynolds, 1987), the Revised Child Manifest Anxiety Scale (RCMAS, C. R. Reynolds & Richmond, 1985), questions from the Self-Description Questionnaire-II (SDQ-II, Marsh, 1990), and the Fear Survey Schedule for Children and Adolescents-II (FSSC-II, Burnham & Gullone, 1997). Construct validity analyses, including internal consistency estimates and factor analysis, indicated that the scales' scores were valid in the Trinidad sample (Watkins et al.). Adolescent females reported significantly more depression symptoms and fears than their male counterparts. Females also reported higher verbal self-concepts and lower math self-concepts. Effect sizes were only in the low to moderate range for the self-concept variables, but were in the moderate and high range for depression and fear, respectively.

In the most recent examination of gender differences in Trinidad and Tobago, Brown (2005) reported that in the CXC examinations, girls outperformed boys in all subjects except mathematics, and the gender differences in obtaining the highest mathematics grade favoured boys by less than 3%. Based on an examination of mathematics performance on the Continuous Assessment of Progress (CAP) data for Standards 1 to 3, Brown also found that females were outperforming males. Females omitted fewer items on the examination, were over-represented in the upper tail of the distribution, and under-represented in the lower tail, and had significantly higher mean scores than did males. In almost all cases, however, effect sizes were small.

Ethnic differences

Since differences in attainment among T&T's ethnic groups are a source of controversy, much like Black-White differences in the US, these differences are seldom discussed openly, and there are limited data available on them. Indeed, much of the data reported by the T&T Government is not disaggregated by ethnicity. Hall et al. (2002) examined ethnic group differences in phonemic awareness in 300 students from the first three years of school. They found no differences among Black (38.4%), East Indian (36.7%), and Mixed (21.3%) students on this pre-reading skill.

Worrell, Hall, et al. (2002) also found no differences on teacher ratings of learning behaviours in a nationally representative primary school sample, but Black students were rated higher on Attention-Deficit Hyperactive and Conduct Problems than East Indian students. At the secondary level, East Indian students reported higher math self-concept scores than Black students (Watkins et al., 2002). Brown (2005) found that students in Hindu and Muslim schools, in which students are predominantly East Indian, had higher mathematics scores than students in other schools.

The Current Study

As the review of the literature suggests, there are gender and ethnic differences in T&T for some variables at some ages. In this study, I examined differences in self-reported achievement, as well as attitudes

and behaviours related to achievement, by gender and ethnicity. The public perception in, and extant data about, Trinidad and Tobago would suggest that females will report higher scores than males, and that East Indians will report higher scores than Black Trinidadians. The contributions of the attitudes and behaviours to academic achievement were also examined across ethnic and gender groupings.

Method

Participants

Participants consisted of 1,434 students from 15 secondary schools in Trinidad. All schools were members of the School Leadership Center network. Ten of the schools were located in northwest Trinidad, one was from Central Trinidad, two were in the northeast, and two were in the south. Fifty-eight percent of the participants were male, 40% were female, and 2% did not report their gender. Ethnic designations were in keeping with the Trinidad and Tobago Census, which uses four categories: Indian descent (East Indian), African descent (Black), Mixed, and Other. Self-designated ethnic group representation in the sample included 26.3% African descent, 40.4% East Indian descent, and 33.3% Mixed descent. Based on 2000 census data, these figures indicated an 11% under-representation of participants of African descent and 13% over-representation of participants of Mixed descent. Students ranged in age from 11 to 19 (M = 14.5, SD = 1.71), and Forms 1 through Upper Six were represented, although representation from the two examination years—Form 5 (4.2%) and Upper 6 (2%)—each constituted less than 10% of the sample.

About a third of the students (32.6%) did not report their fathers' education level, and about a quarter (26%) did not report that information for mothers. However, those who did provide the information indicated that 20% of fathers had some high school or less, 24% had completed 'O' Level or CXC examinations, 16% had completed 'A' Levels, and 40% had at least some college. The equivalent percentages for mothers were 18%, 31%, 19%, and 32%, respectively. Gender representation did not differ across the ethnic groups, χ^2 (2) = 2.06, $p > .05$, and males and females did not differ by age, $F(1, 1394) = .76$, $p > .05$. However, Black

students were older than the East Indian students, $F(2, 1394) = 10.14$, $p < .001$, but the effect size was small (Cohen's $d = .30$).

Measures

In addition to demographic data (e.g., age, gender, parental education level), students were asked to report on six single item variables: (a) how well they were doing in school (6-point Likert scale; 1 = *Below 40%*, 6 = *80% or higher*); (b) how often they completed homework in the previous year (5-point Likert scale; 1 = *Never*, 5 = *Always*); (c) how often were they unsupervised when not in school (5-point Likert scale; 1 = *Never*, 5 = *Always*);, (d) how many days *during the school week* they *limed* with friends (from zero to 5); (e) how many hours of schoolwork they completed outside of class per week; and (f) how many times did they miss class periods in school without permission (4-point Likert scale; 0 = *Never*, 3 = *1 – 2 times per week*).

Students also completed two self-report scales—the Almost Perfect Scale-Revised (APSR, Slaney, Rice, Mobley, Trippi, & Ashby, 2001) and the Measure of Perceived Life Chances (MPLC, Jessor, Donovan, & Costa, 1990). The APSR consists of three subscales: Order (4 items), High Standards (6 items), and Discrepancy (12 items). Order questions assess students' level of organization and neatness, and High Standards questions assess students' striving for excellence in their work. The Order and High Standards subscales represent adaptive or healthy perfectionism. The Discrepancy subscale assesses students' failure to achieve the goal standards that they have set for themselves, and represents maladaptive perfectionism. Internal consistency estimates for the three subscales were in the moderate to high range (see Table 1), with two estimates for Order scores falling below .70.

The MPLC (Jessor et al., 1990) is a 10-item scale that assesses respondents' global perceptions of their future. The scale's two academic items asking about graduating from high school and college were replaced for this study with three questions that asked respondents about getting at least 5 CXC passes, going on to "A" levels, and going to university or getting professional qualifications. An exploratory factor analysis of scores on the modified 11-item measure yielded two factors: a four-item academic perceived life chances and a seven-item general

perceived life chances, and reliability estimates for these scores were in the moderate range (see Table 1).

Table 1. Internal Consistency Estimates for Subscales by Group

	Black		East Indian		Mixed	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**
Almost Perfect Scale Revised						
Order (4)	.66	.78	.79	.78	.67	.83
High Standards (6)	.82	.65	.71	.77	.72	.77
Discrepancy (12)	.85	.83	.86	.90	.85	.86
Perceived Life Chances – TT						
Academic (4)	.77	.73	.79	.79	.80	.74
General (8)	.80	.74	.80	.85	.82	.81

Note. Number of items on subscale in parentheses.

Procedure

Students completed the measures in their classrooms, and were supervised by teachers who were provided with an administrator version of the questionnaire. Teachers answered any questions that students had about the study. The measures were completed anonymously and were turned in upon completion. The study was approved by the Institutional Review Board at the Pennsylvania State University.

Results

Preliminary Analyses

Table 2 contains the means and standard deviations for the major variables in the study. Students reported Marks in the 4 to 5 range, indicating average achievement scores in the 60% to 79% range. Students also indicated that, on average, they completed homework *sometimes*, they were *rarely* supervised by adults when not in school, and they spent time with friends about *one to two days* during the week. Reports of cutting class were very low, with all groups averaging less

than one, indicating that most students *never* cut class. Maladaptive perfectionism scores (Discrepancy) were generally lower than adaptive perfectionism scores (Order, High Standards), and both Academic and General PLC scores were near the top of the distribution with means between four and five.

Correlations among the variables were generally low (see Table 3). Of the 55 correlations, only 9 attained values of at least .30 (i.e., accounting for at least 9% shared variance), and the directions of these correlations made theoretical sense. For example, Marks were positively correlated with Homework Completion and Academic Perceived Life Chances. Homework Completion also had positive relationships with High Standards and Order, and negative relationships with Time Spent with Friends and Cutting Class. High Standards was positively correlated with Order and General Perceived Life Chances, and Academic and General Perceived Life Chances were also positively correlated. The nine correlations were significant at .001, the adjusted alpha level for these analyses.

Major Analyses

Group differences

Two (gender) by three (ethnic group) univariate ANOVAS were used to examine group differences, and the adjusted alpha rate (Bonferroni correction) for analyses was .001. Females reported significantly higher scores on Marks ($F[1, 1329] = 72.58$, $p < .001$, Cohen's $d = .50$), Homework Completion ($F[1, 1392] = 12.46$, $p < .001$, Cohen's $d = .18$), Hours Unsupervised ($F[1, 1377] = 24.8$, $p < .001$, Cohen's $d = .28$), and Academic Perceived Life Chances ($F[1, 1396] = 22.12$, $p < .001$, Cohen's $d = .26$), and males reported significantly higher scores on Time Spent with Friends ($F[1, 1377] = 34.42$, $p < .001$, Cohen's $d = -.32$) and Cutting Class ($F[1, 1374] = 53.19$, $p < .001$, Cohen's $d = -.41$). However, only the Marks and Cutting Class differences had effect sizes in the medium range. There were no significant differences on Order, Discrepancy, High Standards, or General Perceived Life Chances. To compare our results to those reported by Kutnick et al., 1997, we graphed the six achievement levels of Marks by gender, and found that males outnumbered females in the four lower levels and females outnumbered males in the two upper levels of achievement (see Figure 1).

Table 2. Descriptive Statistics for Major Variables in the Study

Variables	Black				East Indian				Mixed			
	Male		Female		Male		Female		Male		Female	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age	14.83	1.68	14.61	1.80	14.33	1.75	14.11	1.56	14.45	1.67	14.65	1.78
Grades Earned (1-6)	3.99	1.16	4.48	1.06	4.26	1.00	4.79	0.90	4.07	1.15	4.56	0.93
Homework Completion (1-4)	2.79	0.94	3.13	0.85	3.15	0.88	3.28	0.76	3.01	0.92	3.05	0.83
Hours Unsupervised (1-4)	2.12	1.24	2.52	1.08	2.36	1.07	2.73	1.11	2.24	1.09	2.41	1.20
Time with Friends (1-5)	1.85	1.75	1.23	1.48	1.33	1.58	0.87	1.31	1.86	1.81	1.37	1.49
Hours on Schoolwork	8.09	7.63	9.78	10.42	11.23	9.58	11.74	6.32	8.27	7.08	10.16	8.10
Cutting Class (0-3)	0.67	0.94	0.29	0.69	0.46	0.84	0.14	0.50	0.64	0.99	0.35	0.70
Discrepancy (1-7)	4.04	1.28	4.06	1.13	4.25	1.21	4.04	1.34	4.04	1.24	4.26	1.21
High Standards (1-7)	5.97	1.03	6.11	0.76	6.00	0.76	6.07	0.75	5.98	0.83	5.91	0.91
Order (1-7)	5.45	1.11	5.41	1.21	5.43	1.20	5.58	1.14	5.39	1.13	5.20	1.28
Academic PLC (1-5)	4.00	0.89	4.27	0.76	4.21	0.74	4.38	0.69	4.05	0.86	4.23	0.79
General PLC (1-5)	4.18	0.56	4.22	0.47	4.13	0.54	4.29	0.55	4.19	0.59	4.13	0.54

Note. PLC = Perceived Life Chances. The ranges of responses for items using Likert-scales are in parentheses.

Table 3. Correlations Among Major Variables in the Study

	1	2	3	4	5	6	7	8	9	10	11
1. Grades Earned	1.00										
2. Homework Completion	.31[a]	1.00									
3. Hours Unsupervised	.11	.28	1.00								
4. Time with Friends	-.19	-.30[a]	-.21	1.00							
5. Hours on Schoolwork	.14	.17	.04	-.13	1.00						
6. Cutting Class	-.20	-.41[a]	-.18	.28	-.04	1.00					
7. Discrepancy	-.23	-.09	-.03	-.01	-.11	.03	1.00				
8. High Standards	.24	.30[a]	.16	-.12	.13	-.16	-.08	1.00			
9. Order	.06	.35[a]	.21	-.16	.04	-.23	.00	.46[a]	1.00		
10. Academic PLC	.39[a]	.20	.03	-.13	.26	-.10	-.21	.29	.06	1.00	
11. General PLC	.09	.18	.15	-.04	.03	-.10	-.13	.38[a]	.26	.34[a]	1.00

Note. Correlations of at least .09 were significant at the .001 level.
[a] Correlations that were interpreted.

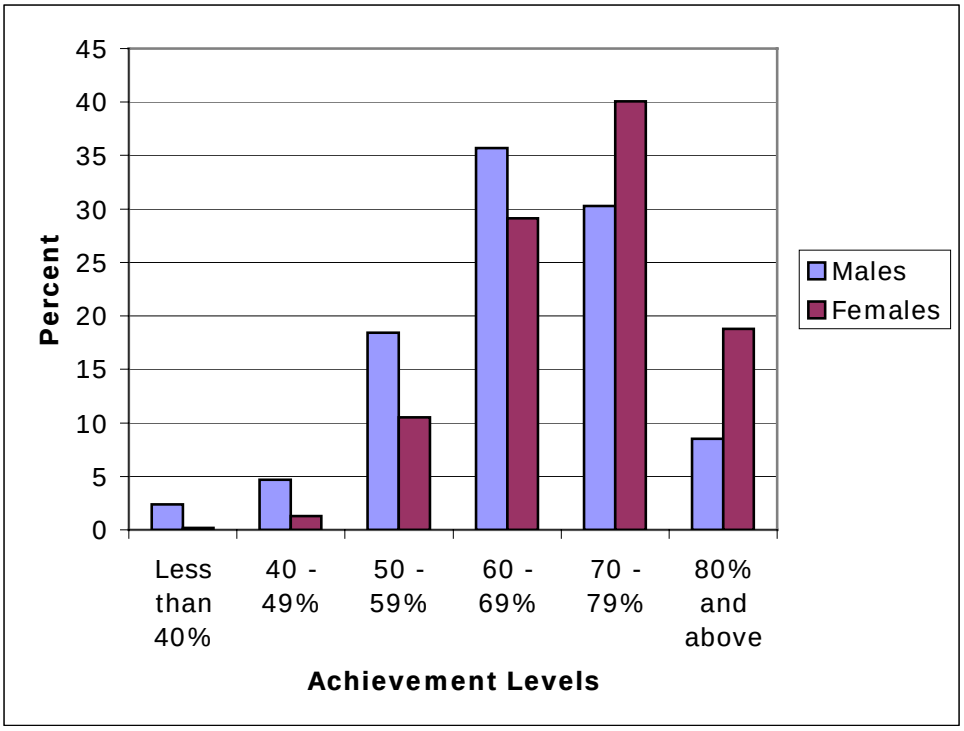

Figure 1. Proportion of males and females across the achievement levels.

The analyses based on ethnicity yielded similar findings. Students of East Indian descent reported significantly higher scores than students of African and Mixed descent on Marks (F[2, 1329] = 9.03, $p < .001$), Homework Completion (F[2, 1392] = 10.42, $p < .001$), Hours Spent on Schoolwork (F[2, 1313] = 11.86, $p < .001$), and Academic Perceived Life Chances (F[2, 1396] = 7.04, $p < .001$), and significantly lower scores than these two groups on Time Spent with Friends (F[2, 1377] = 15.34, $p < .001$), and Cutting Class (F[2, 1374] = 8.74, $p < .001$). Effect sizes (Cohen's d) for these differences were generally small, ranging from |.21| to |.33|.

Predicting achievement

Four regression analyses were computed to examine the contributions of the variables with significant and meaningful relationships (i.e., $r \geq .3$) to student achievement. In the first equation, we examined the contributions of gender, Homework Completion, Hours Unsupervised, Cutting Class, Time with Friends, and Academic PLC on Marks. The variables predicted 24% of the variance in Marks, F (6, 1292) = 67,9, $p < .001$, with significant contributions from gender, Homework Completion, and Academic PLC. The effect size measure for multiple regression ($f^2 = .32$) was large.

The other three regression equations examined the contributions of the same predictors to Marks, with a separate equation for each ethnic group. These results are summarized in Table 4. The equation for Black students accounted for 27% of the variance, F (6, 1292) = 67,9, $p < .001$), and only Academic PLC contributed significantly. The equations for East Indian and Mixed students accounted for 21% and 23% of the variance in Marks, respectively, and gender, Homework Completion, and Academic PLC were significant contributors.

Table 4. Regression Analyses Predicting Student Achievement by Group

	Black		**East Indian**		**Mixed**	
	B	ß	B	ß	B	ß
Gender	.17	.07	.44	.22*	.30	.14
Homework Completion	.20	.16	.22	.18*	.31	.26*
Hours Unsupervised	.05	.05	-.02	-.03	.00	.00
Time with Friends	-.05	-.08	-.02	-.03	-.03	-.05
Cutting Class	-.08	-.06	.02	.02	-.12	-.10
Academic PLC	.50	.37*	.44	.32*	.33	.26*
Adjusted R^2	.27*		.20*		.23*	
f^2	.37[a]		.25[b]		.30[b]	

Note. PLC = Perceived Life Chances.
[a]Large effect size. [b]Medium effect size.
*$p < .001$.

Discussion

In this study, gender and ethnic differences on self-reported achievement, achievement-related behaviours, and achievement-related attitudes were examined in a sample of secondary school students in Trinidad. The contributions of the achievement-related variables to achievement were also examined. Females reported significantly higher scores on achievement and achievement-related behaviours than did males, but in keeping with Hyde's (2005) and Brown's (2005) findings, effect sizes were generally small for all but one variable, Marks. East Indians also reported higher scores on these variables than Blacks, but, as with gender, the effect sizes suggest that the differences are not of practical significance. Moderate amounts of variance in student achievement were predicted by gender, behaviours, and attitudes, but there were differences in significant predictors by subgroup.

The results of this study indicated that, with the exception of self-reported achievement, there were no meaningful differences in attitudes and behaviours by gender. Our findings mirrored findings reported by Kutnick et al. (1997) who found that females outnumbered males in the upper distributions of achievement in Forms 1 and 3, and Brown (2005) who reported similar findings in mathematics for Standards 1 to 3. The findings suggest that the differences between genders should be examined further. For example, although the effect sizes of the differences were small, if the impact of these variables is cumulative over time, they may lead to substantial differences in patterns of achievement and goal-setting across student subgroups.

This hypothesis is supported by recent work on teacher expectancy effects, which suggests that even though teacher expectations are not sizeable in any one classroom, they do accumulate into sizeable effects over a number of school years (Weinstein, 2002). In all three ethnic groups, females reported completing homework more frequently, spending more time on schoolwork, spending less time with friends during the school week, and cutting class less. This pattern of behaviour, both in theory and in the empirical literature, is associated with higher academic achievement in the US. East Indian students reported a similar pattern in comparison to Black and Mixed students. An equally intriguing finding that is in need of more research is the prediction of

school achievement by the attitudinal variables, which differed across ethnic groups. It is difficult to know how to interpret this finding. These findings certainly need to be replicated in other samples before coming to firm conclusions.

Limitations and Conclusion

As with all studies, this study had several limitations. First, even though the sample was large, it was not random and may not be representative of the secondary school population in Trinidad. In fact, prestige schools were over-represented in the sample. Second, a number of variables were single-item variables, which can attenuate the validity of the constructs being measured. Third, the constructs that were measured using subscales have not been examined in this population before, and more analyses will need to be conducted to examine the validity of the instruments' scores in this population.

Limitations notwithstanding, however, the findings are in keeping with the other research in this population on gender (Brown, 2005; Kutnick et al. 1997; UNDP, 2001) and ethnicity (e.g., Brown; Worrell, Hall, et al., 2002). Thus, they cannot be dismissed without some consideration. At the very least, they suggest that gender and ethnic differences on school achievement and attitudes toward school—although currently small—are areas in need of a serious research agenda in Trinidad and Tobago.

Notes

[1] The data in the Central Intelligence Agency factbooks are from local sources such as the Census Bureau and the Central Statistical Office.

[2] This report was compiled by the United Nations Development Programme based in Port of Spain, Trinidad, and is based on data provided by the Trinidad and Tobago Central Statistical Office.

References

Battle, J. (1976). Test-retest reliability of the Canadian Self-Esteem Inventory for Children. *Psychological Reports, 38,* 1343–1345.

Brereton, B. (1981). *A history of modern Trinidad: 1783–1962.* Exeter, NH: Heinemann.

Brown, L. I. (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.

Burnham, J. J., & Gullone, E. (1997). The Fear Survey Schedule for Children-II: A psychometric investigation with American data. *Behaviour Research and Therapy, 35,* 165–173.

Burstein, B., Bank, L., & Jarvik, L. F. (1980). Sex differences in cognitive functioning: Evidence, determinants, implications. *Human Development, 23,* 289–313.

Central Intelligence Agency. (2006). *The world factbook 2006: Trinidad and Tobago.* Retrieved March 11, 2006, from http://www.cia.gov/cia/publications/factbook/geos/td.html

Cohen, D., & Wilkie, F. (1979). Sex-related differences in cognition among the elderly. In M. A. Wittig & A. C. Petersen (Eds.), *Sex-related differences in cognitive functioning: Developmental issues* (pp. 145–157). New York: Academic Press.

Denno, D. (1982). Sex differences in cognition: A review and critique of the longitudinal evidence. *Adolescence, 17,* 779–788.

Frederick D. Patterson Research Institute. (1997a). *The African American education databook, Volume I: Higher and adult education.* Fairfax, VA: Author.

Frederick D. Patterson Research Institute. (1997b). *The African American education databook, Volume III: The transition from school to college and school to work.* Fairfax, VA: Author.

Hall, T. E., Watkins, M. W., & Worrell. F. C. (2002). *Reading assessments for elementary schools.* State College, PA: CEDAR School Psychology Clinic, The Pennsylvania State University.

Havey, J. M., Olson, J. M., McCormick, C., & Cates, G. L. (2005). Teachers' perceptions of the incidence and management of Attention-Deficit Hyperactivity Disorder. *Applied Neuropsychology, 12,* 120–127.

Hogrebe, M. C., Nist, S. L., & Newman, I. (1985). Are there gender differences in reading achievement? An investigation using the High School and Beyond data. *Journal of Educational Psychology, 77,* 716–724.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60,* 581–592.

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104,* 53-69.

Jessor, R., Donovan, J. E., & Costa, F. (1990). Personality, perceived life chances, and adolescent health behavior. In K. Hurrelmann & F Lösel (Eds.), *Health hazards in adolescence* (pp. 25–42). New York: Walter de Gruyter.

Kohr, R. L., Masters, J. R., Coldiron, J. R., Blust, R. S., & Skiffington, E. W. (1989). The relationship of race, class, and gender with mathematics achievement for fifth-, eighth-, and eleventh-grade students in Pennsylvania schools. *Peabody Journal of Education, 66,* 147–171.

Kutnick, P., Jules, V., & Layne, A. (1997). *Gender and school achievement in the Caribbean.* Education Research Paper No. 21. London: Department for International Development.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher, 31,* 3–12.

Lezak, M. D. (1976). *Neuropsychological assessment.* New York; Oxford University Press.

Linn, M. C., & Petersen, A. C. (1986). A meta-analysis of gender differences in spatial ability: Implications for mathematics and science achievement. In J. S. Hyde and M. C. Linn (Eds.), *The psychology of gender* (pp. 67–100). Baltimore, MD: Johns Hopkins University Press.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Marajh, C. (2001, December 23). Chief cop in a quandary. *Sunday Express*, p. 10.

Marsh, H. W. (1990). *Self-Description Questionnaire-II manual.* Sydney, Australia: University of Western Sydney.

Marsh, H. W., Parker, J., & Barnes, J. (1985). Multi-dimensional adolescent self-concepts: Their relationship to age, sex, and academic measures. *American Educational Research Journal, 22,* 422–444.

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Baltimore, MD: Williams and Wilkins.

McDermott, P. A. (1994). *National profiles in youth psychopathology: Manual of Adjustment Scales for Children and Adolescents.* Philadelphia, PA: Edumetric and Clinical Science.

McDermott, P. A., Green, L. F., Francis, J. M., & Stott, D. H. (1999). *Learning Behaviors Scale.* Philadelphia, PA: Edumetric and Clinical Science.

McDermott, P. A., Marston, N.C., & Stott, D. H. (1993). *Adjustment Scales for Children and Adolescents.* Philadelphia, PA: Edumetric and Clinical Science.

Miller, E. (1992). *Men at risk.* Mona, Jamaica: University of the West Indies Press.

Miller, E. (1994). *Marginalization of the Black male* (2nd ed.). Mona, Jamaica: Canoe Press.

Moore, E. G. J., & Smith, A. W. (1987). Sex and ethnic group differences in mathematics achievement: Results from the national longitudinal study. *Journal for Research in Mathematics Education, 18,* 25–36.

Mullis, A. K., Mullis, R. L., & Normandin, D. (1992). Cross-sectional and longitudinal comparisons of adolescent self-esteem. *Adolescence, 27*(105), 51–61.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51,* 77–101.

Noguera, P. A., & Worrell, F. C. (1998). The crisis of the Black male in the English-speaking Caribbean. *Caribbean Perspectives,* 25–31.

Parry, O. (1996). Caribbean masculinities and educational failure: Academic underachievement of Caribbean males. *Caribbean Perspectives,* 8–13.

Reynolds, C. R., & Richmond, B. O. (1985). *Revised Children's Manifest Anxiety Scale (RCMAS) manual.* Los Angeles, CA: Western Psychological Services.

Reynolds, W. M. (1987). *Reynolds Adolescent Depression Scale (RADS) professional manual.* Odessa, FL: Psychological Assessment Resources, Inc.

Roid, G. H. (2003). *Stanford Binet Intelligence Scales, Fifth edition, Technical manual.* Itasca, IL: Riverside Publishing.

Sabers, D., Cushing, K., & Sabers, D. (1987). Sex differences in reading and mathematics achievement for middle school students. *Journal of Early Adolescence, 7,* 117–128.

Sattler, J. M., & Dumont, R. (2004). *Assessment of children: WISC-IV and WPPSI-III supplement.* San Diego, CA: Jerome M. Sattler, Publisher, Inc.

Singham, M. (2003). The achievement gap: Myths and reality. *Phi Delta Kappan, 84,* 586–591.

Skaalvik, E. M., & Rankin, R. J. (1994). Gender differences in mathematics and verbal achievement, self-perception and motivation. *British Journal of Educational Psychology, 64,* 419–428.

Slaney, R. B., Rice, K. G., Mobley, M. Trippi, J., & Ashby, J. S. (2001). The revised Almost Perfect Scale. *Measurement and Evaluation in Counseling and Development, 34,* 130–145.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale* (4th ed.). Chicago, IL: Riverside Publishing.

Trinidad and Tobago. Central Statistical Office. (1998). *Report on Education Statistics 1996/1997.* Port of Spain, Trinidad: Author.

United Nations Development Programme. (2001). *Trinidad and Tobago National Human Development Report 2000: Youth at risk in Trinidad and Tobago.* Port of Spain, Trinidad: Author.

Watkins, M. W., Worrell, F. C., & Hall, T. E. (2002). *Secondary assessment tools.* State College, PA: CEDAR School Psychology Clinic, The Pennsylvania State University.

Weinstein, R. S. (2002). *Reaching higher: The power of expectations in schooling.* Cambridge, MA: Harvard University Press.

Williams, W. M., & Ceci, S. J. (1997). Are Americans becoming more or less alike?: Trends in race, class, and ability differences in intelligence. *American Psychologist, 52,* 1226–1235.

Worrell, F. C. (Ed.). (2003a). *Boys to men: The challenges of engaging boys academically and emotionally in the primary and secondary school system. 2nd Annual Summer Institute, August 6–8, 2002.* Port of Spain, Trinidad: The School Leadership Center of Trinidad and Tobago.

Worrell, F. C. (2003b). What do we know about our male students and how can we use data for decision making and intervening with school problems? In F. C. Worrell (Ed.), *School Leadership Center of Trinidad and Tobago Second Annual Summer Institute. Boys to men: The challenges of engaging boys academically and emotionally in the primary and secondary school system* (pp. 23–33). Port of Spain, Trinidad: The School Leadership Center of Trinidad and Tobago.

Worrell, F. C., Hall, T. E., & Watkins, M. W. (2002). *Behavioural rating scales for elementary schools: Interim manual.* State College, PA: CEDAR School Psychology Clinic, The Pennsylvania State University.

Worrell, F. C., & Noguera, P. A. (2000, April). *The crisis of the Black male: Views of male high school students in Trinidad.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Worrell, F. C., Roth, D. A., & Gabelko, N. H. (1998). Age and gender differences in the self-concepts of academically talented students. *The Journal of Secondary Gifted Education, 9,* 157–162.

Worrell, F. C., Watkins, M. W., Runge, T. J., & Hall, T. E. (2002). Pre-reading skills in Trinidad and Tobago students in the first three years of school. *Caribbean Curriculum, 9,* 1–19.

MEETING PROFESSIONAL LANGUAGE STANDARDS IN THE FL CLASSROOM IN TRINIDAD AND TOBAGO

Jennifer Yamin-Ali

This paper examines the notion of linguistic competence within the context of developing professional standards for foreign language teachers. It focuses on specific language deficiencies recorded during observation of Spanish teachers in the classroom, and provides an initial indication of the need to provide some degree of structured support for competence enhancement in the foreign language. The observation is made that whereas FL teachers may have "latent" knowledge of some fundamental lexical and semantic areas of the language, this knowledge does not always translate into appropriate application. Essentially, the role of the FL teacher as model, resource, and inspirer is diminished when the language used in the classroom shows hesitation, insecurity, inaccuracy, and lack of practice. Finally, some suggestions are made about how teachers can access opportunities for improving their proficiency in the foreign language, at both the personal and the policy level.

Introduction

Much of what we read and learn about foreign language (FL) teaching today is derived from work done in the United States, Europe, and Australia. Current standards and practices derive from policies designed by leading associations in the field such as the American Council for the Teaching of Foreign Languages (ACTFL) and the Council of Europe's Common European Framework of Reference for Languages. Internationally, standards for FL teachers are now evolving.

FL teachers in Trinidad and Tobago currently qualify to teach in the secondary school system with a bachelor's degree in the relevant language. They do not require teacher training. Any such training is in-service and optional. The recognized option is the one-year, part-time, postgraduate Diploma in Education (Dip.Ed.) offered at the School of

Education of the Faculty of Humanities and Education at The University of the West Indies (UWI). If one examines systems elsewhere, for example in the United States, there are entry requirements for FL teachers. Their level of proficiency has to be ascertained in order for them to practise in the classroom. The ever-increasing emphasis on proficiency and meaningful communication signals the important role the teacher has to play in using the target language (TL) in the classroom. Though lesson planning and knowledge of learning theories may be key ingredients in the teacher's portfolio, the ability to comfortably engage with the TL and to inspire students to do the same is one professional requirement that is much downplayed in discussions of classroom practice generally.

Literature Review

The attention paid to FL teacher preparation focuses to a large extent on knowledge of learning theories, language acquisition, and teaching methodologies. Embedded in all of these areas is the understood requirement of teacher proficiency in the TL. Often, this component is not even mentioned or mentioned almost as an afterthought. In proposing five hypotheses on methodology and proficiency in the teaching of foreign languages, Omaggio Hadley (1993) suggests that opportunities must be provided for students to practise using the language in a range of contexts likely to be encountered in the target culture (p. 79). As a corollary to this, she adds that "authentic language should be used in instruction wherever possible" (p. 82). In a final comment she states:

> It is also important to remember that natural language includes the comprehensible input provided by the teacher in everyday exchanges in the instructional setting that are communicative in nature, from giving directions to recounting personal anecdotes in the target language. The proficiency-oriented classroom is one in which such natural acquisition opportunities are exploited as fully as possible. (p. 83)

While not downplaying the evident value of pedagogical skills, educators at all levels need to recognize that there is a common tendency to isolate linguistic competence from overall discussions of teacher competence. There is also the contrasting tendency in which linguistic proficiency is

highlighted at the expense of overall teacher competence. A report compiled by Briguglio and Kirkpatrick (1996, p. 34) found that the main area of concern of teachers interviewed was "linguistic proficiency." The writers' concern was whether researchers and educators were viewing FL teaching holistically enough. But once we decide to focus on proficiency in the language domain, we recognize that the current trend is toward communication and functionality as opposed to discrete-point learning. In FL teacher education programmes, it seems to be the norm to stress that *accuracy* is only *one* component of proficiency so that teachers in the classroom would not overemphasize this element in their teaching as was the norm in "classrooms of old." What is therefore expected of teachers' language use in the classroom?

Even before Omaggio Hadley's (1993) emphasis on the use of the TL in the classroom, the teacher's role as the initiator of interaction in the classroom was recognized (Flanders, 1970; Moskowitz, 1971, 1976 as cited in Brown, 1994). Such use was seen to be combined with all other forms of interaction in the classroom. Characteristics of the "good" FL teacher also point to ease of use of the TL as a model, to inspire its use, and to inspire students' confidence in the teacher. Yet, FL teachers themselves are often concerned about their own linguistic proficiency and may be seen to be reluctant to use the TL communicatively in the classroom. In a report prepared for the National Asian Languages/Studies Strategy for Australian Schools (NALSAS) Task Force by Simpson Norris Pty Limited (1999, p. 5), it is stated that there is currently a perception that language teachers are not as proficient as is desirable. An interviewee stated that "as soon as someone uses the term language teacher proficiency, I think tragedy, sadness, distance, cost, opportunity, lack of confidence, lack of awareness of low levels of proficiency on the part of teachers" (p. 6).

Simpson Norris Pty Limited (1999) quotes Tedick and Tischer (1996, p. 415) who refer to the "constant challenge that second language teachers face in developing and maintaining language proficiency," and suggests that the benefits from a five-week immersion experience for pre- and in-service teachers of French, German, and Spanish may have been related to a more increased comfort level than actual proficiency. Bell (2005) conducted a survey of 457 ACTFL members, of whom more than 94

percent agreed that an effective FL teacher uses the TL competently, and uses it as the predominant means of classroom communication.

Lafayette (1993) suggests that language proficiency is one of the three competencies that must encompass language teachers' subject matter content knowledge, and Vélez-Rendón (2002) argues that it is crucial for effective teaching. She points out that, usually, the appropriate level of proficiency for teachers, based on the ACTFL guidelines, should at a minimum be Advanced Low. In evaluating the professional development of teachers in California, Lozano, Sung, Padilla, & Silva (2002) found that apart from support in classroom management, sequencing lesson plans, and incorporating the *Standards for Foreign Language Learning in the 21st Century* (ACTFL, 1999), teachers needed support in developing language skills. This has implications for the nature of FL teacher preparation—both pre-service and ongoing—and the role of teachers themselves in their own professional development.

In Trinidad and Tobago, though there has been a recent thrust towards the teaching of Spanish as a FL, a parallel preparation of teachers in linguistic proficiency has not occurred. Almost 20 years ago, at the Inaugural Meeting of the Caribbean Language Conference, it was observed that since teaching for communication implies communicative ability on the teacher's part, this may entail the introduction of measures for teacher certification to ensure that teachers bring to the classroom the level of language proficiency and knowledge of the target culture that is essential for effective delivery of a communicative programme. (Morris, 1989, p. 51)

Currently, there is some ongoing discussion at UWI among the Department of Liberal Arts, the Instructional Development Unit, and the Foreign Language Curriculum Department of the School of Education with regard to enhancing the competence of FL teachers in the classroom, both linguistically and technically. However, there is as yet no provision targeted to the specific linguistic needs of the FL language teacher, whether pre-service or in-service.

Research in teacher education, generally, has shown that reflection is a useful tool for self-diagnosis and self-development. There is some evidence from teachers' reflections that the recommended use of the TL can be a challenge to teachers to improve their own linguistic skills

(Morris & Yamin-Ali, 2005). If one bears in mind that teachers' development is not static, but ongoing, and that "the fundamental processes of professional learning are preserved in the journey from student teacher to expert practitioner" (Burroughs-Lange& Lange, 1994, p. 51), then "error, success and refinement" (Shulman, 1987, p. 4) would characterize the professional lives of FL teachers, not only pedagogically, but on the linguistic continuum.

Rationale for Research

FL teacher education in Trinidad and Tobago, though poised to treat with teachers' actual needs in the classroom, is not yet supported by official policy in terms of specific standards for FL teachers. While there is, in theory, a small-scale introduction of clinical supervision, through the appointment of Heads of Departments in some schools, it is insufficient and does not yet translate into a mechanism for ongoing support for the practising teacher.

If we were to be guided by the *Program Standards for the Preparation of Foreign Language Teachers* (ACTFL/NCATE, 2002), we would recognize its policy of emphasizing the development of teachers' oral proficiency as well as ongoing assessment of their oral proficiency together with diagnostic feedback. The standards also advocate the provision of opportunities for teachers to participate in a structured study abroad programme and/or intensive immersion experience in a TL community.

In order to facilitate such development of FL teachers, both FL teacher preparation programme developers and teachers themselves need to be aware of the particular areas of linguistic development that require attention for effective use in the classroom. With this in mind, this examination of TL use in the FL classroom was conducted.

The questions that this study sought to answer were:

1. What are teachers' views of and responses to their use of the Target Language (TL) in the classroom?
2. What areas for improvement are reflected in teachers' use of the TL in the classroom?

Method

The study is situated in the qualitative research paradigm and is an instrumental case study. Analysis is interpretive as it seeks to understand the reflections and practice of FL teachers with regard to linguistic proficiency and use of the TL in the classroom. For the second research question, quantities are used only in order to underscore that it is not the frequency of the error that is being highlighted but the overall nature of the error, and that the fact that it occurs at all is significant to the ensuing discussion.

Due to the context of the research, it was possible and practical to employ the "participant-as-observer" technique since this researcher is a lecturer on the Dip.Ed. programme, with responsibility for the Foreign Language Education component of the programme. Data collection was therefore a natural process and not engineered specifically for the purpose of this research.

The sample used in this study was purposive. The participants were in-service Spanish teachers enrolled in the Dip.Ed. postgraduate programme at the School of Education (SOE) on the St. Augustine campus of UWI. Eight Spanish teachers completed the programme in 2005 when the total intake was nine. The other teacher was a teacher of French. Seven Spanish teachers were supervised by this researcher in 2005/2006, when the total intake, inclusive of French, was 16. They all teach at the secondary level and entered teaching with a bachelor's degree in either Spanish or French, or both. There is no other provision for long-term, in-service training for FL secondary school teachers locally. The small sample therefore represents 100% of the researcher's Spanish students from two consecutive years, totalling 15 students.

Data were collected via observation records and journal entries. Both these sources constitute elements of the Dip.Ed. programme, which requires student teachers to reflect on their experiences and learning during the course of the year (July–May). Classroom practice was observed and recorded by the teacher-educator (this researcher). These records indicated extent of use of the TL, general proficiency level, and types of linguistic errors made. Feedback conference records were also analysed and categorized, focusing on teachers' reactions and

interpretations of their engagement with the TL. In their journal entries, teachers selected the areas they wished to focus on for reflection. From this source, comments pertinent to TL use in the classroom were categorized and interpreted.

Teachers' views and responses were seen as a significant factor in this study because they may be the driving force behind their practice, or they may provide explanations for the strengths or weaknesses observed in the context of their linguistic proficiency. They can be useful in determining the kinds of help or encouragement they need. This is supported by research conducted by Stritikus (2005) in which he proposes that spaces must be created in teacher preparation programmes for discussion of teachers' beliefs and their local settings. In addition to this, research in the field has pointed to the advantages that reflective practice can bring to the classroom (Haynes, 1995; Schon, 1983; Wallace, 1991).

Examination of areas of improvement reflected in the teachers' use of the TL in the classroom would guide the way for planning to meet professional standards in the local FL classroom. Such analysis would indeed call for collaboration among the employer—the Ministry of Education, specifically the Foreign Language Curriculum Officer—the Foreign Language Department of UWI, and the School of Education, UWI in a bid to ensure that FL teachers are well equipped to maximize their students' opportunity for optimum language learning in the classroom.

Perhaps one limitation of this study is that it reflects the practice of teachers whose sensitivity to the use of the TL may be higher than the norm, since it is an element of teaching practice that is highlighted for attention within the programme. Another, clearly, is that the sample cannot be said to represent the body of FL teachers in a general way. It is merely indicative of pockets of practice and attitudes. However, it does point the way for parallel or wider-ranging studies of a similar nature.

Findings

In this section the data will be analysed according to each of the two research questions.

Research Question #1 (Based on comments from teachers' journals)
What are teachers' views of and responses to their use of the target language in the classroom?

Hesitation/Reluctance

In most instances, student teachers reported that they had usually been hesitant to use the TL during lessons for two main reasons. One was that they did not think that their students would be comfortable. Some even reported that their students were frustrated:

> I was very sceptical about this (using the TL), since in my experience at _____ College, any attempt that I make to use purely Spanish is met with blank stares or confused looks from the students.

> Although I use French or Spanish in class and tell my students that they have to use it, I am not consistent with it. Sometimes I do not use it (especially when explaining a grammar point).

> It makes no sense speaking in Spanish to my Form Ones since they will not understand me.

The other reason was that the easier, more manageable way was to resort to the native language, either by conducting the class in English or by translating the TL into English even when it is used. This is illustrated in the reflections of a French teacher on her belief prior to the Dip.Ed. programme:

> I really thought that by using the native language in class I was making life easier for myself and my students.

In an attempt to use the TL consistently in the classroom, a Spanish teacher reported on her students' reactions:

> At first my students seemed perplexed, and in written feedback, many asked for translation.

No doubt students and, admittedly, teachers, resort to their comfort zone, which means functioning in their native language. Students are seen to react negatively to the consistent use of the TL when their prior

experience has not provided them with that exposure. From experience in other FL classrooms locally, where the FL is used consistently by the teacher from the first year of instruction, it has been observed that students are able, at all levels, to function naturally and comfortably in the language. They comprehend and produce it with more ease.

A teacher's estimation of her own competence was seen to be another factor hindering her from using the TL extensively in the classroom:

> I was wary at first, and scared that my own command of Spanish would not measure up.

Generally, early reaction to the suggestion that they should use the TL as much as possible indicates discomfort and some even comment that it would be difficult.

Students' improvement

Despite teachers' hesitancy in using the TL, some reported that they had seen improvement in their students' own use of and response to it:

> I am feeling so happy today…my Form One students are beginning to speak in Spanish in my class….I am giving to them each day what I refer to as "*Frases de hoy*" which, as I give it to them, I begin to use in the classroom while speaking to them in Spanish…. My reservations about using the target language in class are diminishing with each passing day.

Another student teacher reported on the positive effects of using the TL extensively in the lower forms for the first time:

> This is the first time that I have ever used the target language almost exclusively at these levels. It became clear that students at the First Form level were very enthusiastic about hearing Spanish spoken. Most students understood most of what was said….At the Third Form level the results were very similar. I have concluded therefore that the predominant use of the target language is of benefit even at the lower levels. This practice will certainly be continued in my classes.

Self-improvement

Despite their early hesitation due to their feelings of linguistic inadequacy, there was also evidence that some teachers recognized that using the TL consistently in their classrooms was an opportunity for them to practise and to become more confident. In addition to the satisfaction they felt when their students' competence developed as a result of the exposure to the TL, self-improvement was another source of satisfaction for some teachers:

> However, the more Spanish I speak myself, the more fluent I get.

> Additionally, as a teacher using the foreign language in class, I will be challenged to improve my linguistic skills.

Even though not every teacher made the overt observation that self-improvement was an obvious benefit from classroom use, classroom observation has shown that as teachers attempted to use the TL more, they became more comfortable, confident, and used it with more ease.

Recognition of significance of teacher's TL use in the classroom

Modelling has been identified as an important factor in the teaching/learning experience. The recognition that the teachers' behaviour in the classroom is a source of learning for the students is one that has far-reaching consequences for teacher practice. Teachers on the Dip.Ed. programme read and hear about the benefits of the teacher's use of the TL in the classroom. When they enter the programme, they come with their own theories of why they should or should not use it. Some are convinced that their students would be too confused and thus it would be a hindrance. Some find it too tiring mentally, and others feel overwhelmed due to their own lack of proficiency in the language. Once they have begun their sessions in theory and practice, new learnings become evident:

> I realize that my students' confidence in their oral proficiency was linked to my own use of the target language in the classroom. One of my major challenges therefore was to motivate my students to become risk-takers by my own example.

One other teacher also commented on her role as model when she said that:

> The classroom has its limitations so the teacher must make using the target language the norm in order to encourage students to follow suit.

Teacher-as-resource is another element that arose out of the reports of two teachers on the programme. In the case of one of these teachers, the students come from a semi-rural background where the likelihood of practice of the TL outside of the classroom is slim:

> I am realizing that I am possibly the only contact many of them have with spoken Spanish, and as a result I should expose them to as much spoken Spanish as possible.

Inspiring students to use the TL is yet another benefit to be derived from exposure via the teacher. One Spanish teacher spoke about one of the things she remembered about her favourite teacher—her Spanish teacher—who:

> never failed to stop me in the corridor to speak to me in Spanish – everyday!…I always said to myself, "If ever you become a teacher, you must be just like Mr. _____."

The sentiments and observations related above reflect the thoughts of teachers as they were prompted to dwell on their professional and personal development throughout their year of the Dip.Ed. experience. The fact that they chose to remark on the use of the TL points to the significant place it holds in the among the challenges that FL teachers face. It also becomes clear that there are issues to be addressed or re-addressed in formulating programmes that prepare FL teachers for the classroom.

Research Question #2 (Based on observation in the classroom)
What areas for improvement are reflected in teachers' use of the TL in the classroom?

The teachers observed in the context of this study have varying levels of linguistic proficiency in the FL or languages they teach. While this study does not measure their linguistic proficiency holistically via the medium

of a proficiency scale, discrete linguistic details in their use of the FL point to areas of weakness and consequent need for attention. These areas are categorized under the broad headings of *vocabulary*, *grammar*, and *pronunciation.* The details from observation are presented in the following tables and then discussed. The letters A-O have been substituted for the teachers' names.

Vocabulary

Table 1. Vocabulary Errors Made by Teachers in the Classroom

Errors	Intended Meaning	Frequency	Teachers	Correct Version
1. actora	actress	1	A	actriz
2. pretender	to pretend	1	C	fingir
3. acente	accent	1	F	acento
4. custard	custard	1	F	flan
5. encontrar	to meet	2	G, L	conocer (to meet for the first time)
6. tratar de	to treat (someone)	1	L	tratar a (to treat)
7. después de	afterward	1	H	después
8. no más	no longer	1	H	ya no
9. trabajas	jobs	1	H	trabajos
10. pagar atención	to pay attention	1	H	prestar atención
11. revisar	to revise	1	H	repasar
12. saber	to know (someone)	1	J	conocer
13. niños	teenage boys	1	F	chicos

With the exception of three words (*actriz, pretender, flan*), the words in Table 1 can be considered to be high frequency words in Spanish. Whether these errors may be attributed to either nervousness or ignorance, one can propose, with some degree of certainty, the source of the error in most cases.

Some of the errors stem from opting for the first choice that comes to mind—the fabrication of what seems to be a cognate. This seems to be the case for *pretender* and *revisar, which* are in fact Spanish words, but their use is incorrect in the context. In the case of *actora,* a general rule was applied whereas this word is an exception to the rule. *Pagar* is a literal translation of "to pay" and is incorrect in this context, as is the case of *no más.* Both *acente* and *custard* point to ignorance on the speaker's part, as the first seems to have been fabricated by attaching a vowel to the English word and there was no attempt to translate the second one. The ability to distinguish between *encontrar* and *conocer,* and *saber* and *conocer* is not always spontaneous among speakers of Spanish as a second language. Though it is most likely that the teacher knows the distinction in theory, spontaneous use does not reflect this understanding. Lack of practice may account for this deficiency.

The use of both *tratar de* and *después de* seems to suggest that since these are the more common uses to which the user in accustomed—attaching the preposition—spontaneous use includes the preposition *de.* It is likely that when they studied the language, they were required to place emphasis on phrases that required a preposition. There are many textbooks that highlight this aspect of the language and it is an error that is usually emphasized by teachers. These users apparently view the item as a vocabulary item and ignore its grammatical component. *Trabajas* may be a slip on the teacher's part, or minor confusion emanating from the two 'a's that come in the first part of the word. The correct version *trabajos* is far too common to suggest that this was a word not known by the teacher.

The two teachers making most errors were F (5) and H (3). Though they are enthusiastic in the classroom, their own exposure to native speakers has been minimal. Each has been teaching Spanish for over five years and both received their bachelor's degree from the same university. Overall, though, the frequency of errors of most of the teachers is not high; however, the nature of the errors points to a need for more familiarization with common elements of the language.

Grammar

Table 2. Grammatical Errors Made by Teachers in the Classroom

Errors	Frequency	Teachers	Correct Version
Verbs — Instructions			
1. Repiten	2	B, F	Repitan
2. Recorden	1	F	Recuerdan
3. Abren	1	L	Abran
4. Miran	1	L	Miren
5. Escuchan	1	L	Escuchen
6. Buscan	1	I	Busquen
Verbs — Gustar			
1. Me gustan a las chicas	1	C	Me gustan las chicas
2. Les gustan el dibujo	1	D	Les gusta el dibujo
3. Les gustan el pollo	1	F	Les gusta el pollo
Verbs — Agreement			
1. Deseas que ella te acompañes	1	E	Deseas que ella te acompañe
2. Todo el mundo han terminado	1	L	Todo el mundo ha termindao
3. Tú es	1	J	Tú eres
Verbs — Haber			
1. Habrán cinco espectáculos	1	F	Habrá cinco espectáculos
2. ¿Cuántas competiciones habrán?	1	F	¿Cuántas competiciones habrá?
Verbs — Past Participle			
1. Terminadas?	1	K	(Han) terminado?
2. Están interesada	1	L	Están interesadas *or* Les interesa
Verbs — Subjunctive			
1. Quiero que tú responder	1	L	Quiero que tú respondas

Errors	Frequency	Teachers	Correct Version
2. Quiero que cada persona decir	1	L	Quiero que cada persona diga
3. Quiero que tú preguntas	1	L	Quiero que tú preguntes
4. Quiero que los estudiantes arreglan	1	L	Quiero que los estudiantes arreglen
Verbs — Ser vs. Estar			
1. Deben ser en casa	1	J	Deben estar
Verbs — Reflexive			
1. ¿Qué pones?	1	B	¿Qué te pones?
2. Me llamo es	1	L	Me llamo
3. Quedarán	1	H	Se quedarán
Prepositions			
1. (un regalo) por mi padre	1	D	Para mi padre
2. Un espectáculo a la escuela (repeated error 4 times by same person)	1	F	Un espectáculo en la escuela
3. Servir el hombre; ayudando su mamá	1	L	Servir al hombre; ayudando a su mamá
Pronouns			
1. De tu mismo	1	F	De ti mismo
2. *Tú* and *usted* to address same student	1	J	Either one should be used consistently
3. Los estudiantes quien	1	L	Los estudiantes quienes
Nouns			
1. El foto	1	F	La foto
2. El orden	1	J	La orden
3. Puerta España	1	F	Puerto de España
Adjectives			
1. Respuesta rápido	1	F	Respuesta rápida
Adverbs			
1. Mucho difícil	1	F	Muy difícil

Errors	Frequency	Teachers	Correct Version
Numbers			
1. Cien treinta y ocho	1	L	Ciento treinta y ocho
Questions			
1. ¿Qué es la interacción?	1	K	¿Cuál es la interacción?
2. ¿Qué es su nacionalidad?	1	L	¿Cuál es su nacionalidad?
3. ¿Cómo los hombres ven a las mujeres?	1	L	¿Cómo ven los hombres a las mujeres?
4. ¿Quién está hablando de?	1	L	¿De qué está hablando?
5. ¿Por qué X odia a Y?	1	L	¿Por qué odia X a Y?
6. ¿Adónde tendrá lugar?	1	F	¿(En) dónde tendrá lugar?

As seen in Table 2, teachers made a wide variety of grammatical errors. In most cases the error was made once by one person, but overall they represent a wide range of the fundamentals of Spanish syntax. In four instances, the same teacher made the same type of error more than once, for example, teacher L made four errors in the use of the subjunctive and four errors in the formation of questions, and teacher F made two errors with the use of *haber*.

Bearing in mind that some teachers were observed on four occasions over at least eight teaching periods, and some on two occasions over an average of three to four lessons, the errors appear to be too many. An important factor to be considered is that, except for two types, these errors occurred during speech and not in the written form. The two types that did appear in a handout used by the teacher were the incorrect use of *haber* and the use of *a* instead of *en* for location. These two errors were used together within the same context. On none of these occasions did the teacher self-correct or stumble. Since none of the above items were actual teaching points within any of the teacher's lessons, it is not known

whether the teacher has control of the items in another format. That is to say, if the teacher had to actually teach a lesson presenting the item formally and giving students oral or written structured practice, there is no evidence available to suggest that the teacher's content knowledge, which, especially in a linguistic situation, is far removed from its application.

Again, it is apparent that teacher L is among those making the most grammatical errors (17), in addition to F (12), and to a lesser extent J (4). F's errors span a number of grammatical areas, totalling 9 out of 15 categories presented here. L's errors also span nine of the categories, but are more frequent in question formation, use of the subjunctive, and using the imperative form of the verb.

Table 3. Pronunciation Errors Made by Teachers in the Classroom

Pronunciation	Frequency	Teachers	Correct Version
Asist*i*ré	1	H	asistir*é*
Viaj*a*ré	1	H	viajar*é*
Trabaj*a*ré	1	H	trabajar*é*
Univ*er*sidad	1	H	universid*ad*
Ag*us*tín	1	H	Agust*í*n
Nec*es*ito	1	H	neces*i*to

Pronunciation errors were not frequent. In Table 3 all the errors were made by the same teacher. One can surmise that the first three errors are related to the sound of the infinitive from which the form comes. In the infinitive form the stress goes on the syllable containing the last 'r.' The last three errors appear to stem from native language interference, as the stress in the Spanish word was placed exactly where it is in the English word.

There were three teachers who did not appear in the data as there were no errors recorded during their classroom visits. Two of them have been teaching for over 20 years and demonstrated mastery of grammar and vocabulary when using Spanish in the classroom. They were also among those teachers who spoke Spanish most frequently during their teaching.

The third is a comparatively junior teacher teaching for less than 10 years, but who excelled at the language both at the secondary and tertiary levels, during which she had extensive exposure to the TL and culture through frequent visits to Latin America.

Discussion and Recommendations

The data presented show that there are areas of deficiency in Spanish teachers' vocabulary, grammar, and pronunciation. This does not mean to say that they are less proficient than their counterparts in any other part of the world. After all, to varying degrees, they have all been successful at internationally recognized examinations. In Trinidad and Tobago, FL teachers have only recently begun to attempt to prepare their students to be communicative in the TL. However, it must be remembered that most of the teachers themselves have not benefited from such an approach. Many of them are still nervous about using the TL, especially when being observed in the classroom.

It is also important to note that some of the errors noted are common errors for the average non-native speaker, and may even be deemed "high frequency" and be easily forgiven by the native speaker. Yet one has to consider that if professional standards are a target for educators and policy makers, then those standards have to be spelt out in accordance with the needs of the society and the education system. One would assume that teachers of the language must be at a higher level of proficiency than the students being taught, at least Advanced Low according to the ACTFL proficiency scale.

The Personal Level

I have attempted to account for some errors, and while errors are a step towards refinement (Shulman, 1987), there remains the reality that teachers, in their preparation for teaching—pre-service and in-service—need much more opportunities for oral practice, both in and out of the classroom setting.

My principal suggestion is that teachers become the monitors of their own proficiency, thereby maximizing every opportunity to enhance their use of the TL. Formal learning settings will not necessarily furnish them

with the skills, knowledge, and attitudes they require. "Attitudes" is an important component of this preparation package. The attitude must be one of a burning desire to reach "educated native speaker" proficiency. With such an ambitious goal, the learner would always be striving. FL teachers need to become proactive if they are to rise above feelings of inadequacy in the classroom and in other settings. This would assume a lifestyle that would incorporate constant exposure to the language in the written and aural forms in authentic cultural settings. This obviously suggests travel to countries where the language is spoken. As some teachers in the system already do, travel to TL destinations should become the obvious holiday option. Some teachers learn along with their students on school trips. Another suggestion is that they offer their services to travel agencies as travel guides in exchange for travel expenses.

It must be noted that a professional association of Spanish teachers does not exist in Trinidad and Tobago. Such an association has to be formed through the impetus of teachers themselves. Conversation groups in Spanish and sharing of audio-visual material to enhance their own learning is one of the major benefits to be derived from such an initiative.

The Policy Level

The fact that there are no professional standards set for FL teachers speaks to the need to develop a policy that would effect such. University students graduate with varying degrees of competencies, which may not always ensure that they are effective teachers in terms of content and aptitude for teaching. The Ministry of Education would need to ascertain its needs in terms of teacher capability in order to ensure that it is placing competent persons in the classrooms.

Apart from the teachers' personal initiative, there are two scenarios that are not impossible. The first is that all degree programmes in the FL should incorporate a compulsory stay in a TL country. The second is that all permanent FL teachers in the system should receive a stipend and official assistance with visas, diplomacy matters, and so on, in order to travel to such countries every three to five years. The alternative to this would be enrolment in compulsory immersion programmes organized by

the Ministry of Education during the vacation period. This would constitute part of on-going professional development.

With the careful and skilful implementation of clinical supervision in classrooms, schools have the capacity to diagnose their individual needs and put ameliorative measures in place. School policy can facilitate more intradepartmental planning and support through the allocation of time for meetings and the structured facilitation of collaborative work, thereby encouraging learning through sharing. Allied to this is the suggestion that teachers be given the opportunity to teach at varying levels of the school, so that the content of each level is mastered through preparation for classes.

If portfolios constituted an element of teachers' ongoing assessment or a part of a performance management system, teachers could record themselves with any form of technology available, for their personal listening, viewing, and self-critique. They would be required to develop a growth plan and path utilizing this self-improvement strategy. This kind of experience has proven to be very enlightening for practitioners.

Since experience has shown that there is a need for practising FL teachers in Trinidad and Tobago to enhance their skills in the TL, a professional training programme such as the Dip.Ed. may have a role to play in the development of teachers' linguistic proficiency. Teaching a FL requires that teachers be the total model. Since linguistic competence is essential to the FL teacher, there are two possibilities. One is that a language module be incorporated into the programme. The other is that there should be specialized preparation courses within the degree programme for those persons who read for a degree in a FL and want to become teachers. To this end, consultation between the Education and Foreign Language departments would be necessary.

The main thrust of this article has been to indicate that a high level of proficiency in the TL has an integral role to play in determining the professional standards of the FL teacher. While this perspective does not overlook the importance of pedagogical skills, it suggests that professional standards for FL teachers cannot discount the teacher's ability to use the language communicatively, inclusive of the varying types of competencies.

References

American Council for the Teaching of Foreign Languages. (1999). *Standards for Foreign Language Learning.* Lawrence, KS: Allen Press.

American Council for the Teaching of Foreign Languages. (2002). *ACTFL/NCATE Program Standards for the Preparation of Foreign Language Teachers*. Yonkers, NY: Author.

Bell, T. R (2005). Behaviors and attitudes of effective foreign language teachers: Results of a questionnaire study. *Foreign Language Annals, 38,* 259–270.

Briguglio, C., & Kirkpatrick, A. (1996). *Language teacher proficiency assessment: Models and options.* Perth, Australia: NLLIA Centre for Literacy, Culture and Language Pedagogy, Curtin University of Technology.

Brown, H. D. (1994). *Teaching by principles: An interactive approach to language pedagogy.* Englewood Cliffs, NJ: Prentice Hall Regents.

Burroughs-Lange, S., Lange, J., Downing, A., Francis, D., & Sellars, N. (1994). Significant episodes in the acquisition of professional knowledge by teachers. In Queensland Board of Teacher Registration, *Knowledge and competence for beginning teaching: Report of a policy development initiative* (pp. 23–64). Toowong, Queensland: Board of Teacher Registration.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.

Haynes, D. (1995). One teacher's experience with National Board assessment. *Educational Leadership, 52,* 58–60.

Lafayette, R. C. (1993). Subject-matter content: What every foreign language teacher needs to know. In G. Guntermann (Ed.), *Developing language teachers for a changing world* (pp. 124–158). Lincolnwood, IL: National Textbook Company.

Lozano, A. S., Sung, H., Padilla, A. M., & Silva, D. M. (2002). Evaluation of professional development for language teachers in California. *Foreign Language Annals. 35*(2), 161–170.

Morris. J. (1989). Goals for foreign language teaching in the Caribbean. In H. Jarvis (Ed.), *Caribbean Language Conference: Inaugural Meeting, A regional perspective* (pp. 44–54). Port-of-Spain, Trinidad. School of Languages, NIHERST.

Morris, J., & Yamin-Ali, J. (2005, July). Shifting gears, towards reflective practice: A study of foreign language teachers on an in-service teacher education programme. Paper presented at 12th International Conference on (Learning Conference 2005), Granada, Spain.

Omaggio Hadley, A. (1993). *Teaching language in context* (2nd ed.). Boston, MA: Heinle & Heinle Publishers.

Schon, D. A. (1983). *The reflective practitioner: How professionals think in action.* New York: Basic Books.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review,* 57(1), 1–22.

Simpson Norris Pty Limited. (1999). *Language teacher proficiency or teacher language proficiency? An environmental scan of information relating to the competencies/qualities/knowledges required to be an effective language teacher*. Canberra, Australia: Commonwealth Department of Education, Training and Youth Affairs.

Stritikus, T. (2005). Nadie me dijo: (Nobody told me): Language policy negotiation and implications for teacher education. *Journal of Teacher Education, 56*(1), 73–87.

Vélez-Rendón, G. (2002). Second language teacher education: A review of the literature. *Foreign Language Annals, 35*(4), 457–467.

Wallace, M. (1991). *Training foreign language teachers: A reflective approach.* Cambridge, UK: Cambridge University Press.

MOTIVATION AND READING IN CHILDREN AND ADOLESCENTS
An Annotated Bibliography

Janet Fullerton-Rawlins

This bibliography attempts to provide an overview of the literature available in the area of motivation and reading, and is not intended to be comprehensive. It focuses on reading as a lifelong habit and not as literacy development. The bibliography is intended as a basic research tool for researchers on the topic of reading and motivation in children and adolescents. The introduction provides a review of the literature on the topic and is followed by the bibliography itself, which is organized by some of the factors that affect motivation in reading—parental role, strategies, the library's role, and the role of technology. The final section identifies some websites that should be useful for Internet research on the topic. The research was done primarily through the search of online databases, online documents, and print journals. The strong North American perspective reflected in the bibliography underscores the paucity of Caribbean research on the topic.

Research over the past 20 years has shown that students' motivation is a primary concern of many teachers, who acknowledge that it is at the root of many of the problems they face in educating children (O'Flavahan, Gambrell, Guthrie, Stahl, & Alvermann as cited in Edmunds and Tancock, 2003). Research has also shown that motivation plays a major role in learning and that it also affects the reading engagement of children and adolescents. In addition, research has established that students who read perform better in school than students who don't, as reading affects learning in a myriad of ways. For example, it affects comprehension of texts, critical thinking skills, breadth of vocabulary, and proficiency of writing skills. According to Beers (2003), "students who struggle with reading know they lack the most important tool for success in school – the ability to read and make sense of texts" (p. 6). And Xue and Meisels (2004) say that "at the core of children's school success is their literacy achievement" (p. 19). So that motivated readers

are generally high achievers and move through the school system with a greater sense of purpose and achievement than their unmotivated counterparts.

Reading and literacy development can be defined in different ways—as reading for academic gains and achievement and as reading as a lifelong activity for pleasure. The key to reading is motivation because reading is an activity that one can choose to do or not to do. Children and adolescents must develop not only the skill but also the will to read. According to Alvermann (2003), aliteracy, which is the capacity to read but electing not to do so, is one of the most disturbing factors in secondary educators today. Alvermann also posits that students with high self-efficacy are more likely to engage in reading than students with low efficacy. She also believes that teachers cannot "fix" learners, but that the instruction conditions in which they learn can be changed and thus effect a change in their attitude to reading. It is important, therefore, that teachers develop and implement strategies for literacy development and for encouraging their students to read.

Luke and Elkins (2000) used the term remediation to describe one such approach, which involves "refashioning curricular and instruction conditions so that they incorporate multiple forms of media, for example, trade books, textbooks, magazines, newspapers, visual images, videos, CD-ROMs, and the Internet" (as cited in Alvermann, 2003, p. 2). This remediation process is one of the ways that can help to make the difference between a motivated or an unmotivated reader. Allen (1995) describes her methods of working with "at risk" students and gives an account of how she motivated her students by charting a new course. She developed a curriculum with no models to draw from, and successfully built an environment in her English classroom that supported literacy and learning. She used creative methods to deal with reluctant readers and her ideas promoted literacy in and out of the classroom. Some of her ideas are now being promoted, for example, read-aloud sessions, using books that are relevant to students' lives, and book discussions. She did not stick to the curriculum, but became innovative and creative in her classroom, using any printed material she could find to engage her students. She noted that what she discovered is that students will read when they are given the opportunity to read.

Hoy and Miskel (2001), in discussing constructivism, quote Vygotsky's perspective that "knowledge is socially constructed and that knowledge is built upon what participants contribute and construct together (p. 66). From an adolescent literacy perspective, this means that meaning making occurs through social interactions or social practices within classrooms. Allen's (1995) strategy of having her students work collaboratively on song lyrics and other available printed material is an example of one such social practice. It illustrates how an entire class engaged in meaning making around common texts. Furthermore, social constructivist research suggests that adolescents do not merely construct meanings about texts, but that they also develop understandings about the value these texts have in situated contexts.

In discussing motivation for reading, sociocultural contexts are also important. as these have a direct influence on literacy development and sustained reading for pleasure. Taylor and Dorsey-Gaines (1988) have shown how poor, black families from an inner city, whose children were not expected to succeed in school, were able to develop the literacy skills of their children in spite of their socio-economic and sociocultural status. By sheer determination and parental supervision, their children were able to perform at unexpected levels at school. This study looked at the learning styles, coping strategies, and social support systems of children living in poverty to understand how literacy in school became meaningful in their everyday lives.

On the personal level, the literature abounds with theories of intrinsic versus extrinsic motivation. Intrinsic motivation can be defined as the self-generated desire to read for pleasure, literacy, and academic gains, while extrinsic motivation can be defined as reading for incentives, for a defined purpose, or for rewards. Intrinsic readers are termed engaged readers, and according to Guthrie and Wigfield (2000), "their devotion to reading spans across time, transfers to a variety of genres and culminates in valued learning outcomes. In contrast, disengaged readers are inactive and inert. They tend to avoid reading and minimize effort" (p. 403). Guthrie and Wigfield also feel that effective literacy addresses the issues of self-efficacy and engagement.

What then are the results of reading motivation? Since reading and literacy go hand in hand, the primary result of reading motivation is

literacy development, which in turn enhances learning. Guthrie and Wigfield (2000) believe that "motivation is characterized as a unidimensional quality of which more is better and less is worse for learning and performing school tasks" (p. 406). Self-efficacy is another aspect of reading motivation. Students become confident that they can and will perform and succeed, and this increases their learning and performance potential (Wigfield, 1997).

Becoming socially interactive is another consequence of reading motivation. Students are motivated to discuss and share information and books with their peers, which they were not previously inclined to do. Gambrell (1996) says "an engaged reader is motivated, knowledgeable, strategic and socially interactive" (p. 16). This is also borne out by Alvermann, Young, Green, and Wisenbaker (2004) who did a study of Read and Talk clubs to inquire into adolescents' literacy experiences in a setting outside of school. It was discovered that these clubs were perceived to be social outlets for young adolescents, especially for those who liked to read.

Motivation to read can also be a means to healing and wholeness. Introducing books to children and adolescents with people, problems, and circumstances with which they can identify can bring things into a whole new perspective for them. Reynolds (2004) says "reading can be a very important journey towards wholeness, or health, or enlightenment, or whatever label we attach to the mystery of intellectual and spiritual growth" (p. 21).

Reading motivation is clearly a desirable aspect of education and school instruction, without which students' progress is minimized and their literacy capabilities lessened. It also, according to Gambrell (2004), activates their capacity for imagination, and for creative and critical thinking.

How do educators, teachers, and librarians respond to the task of motivating students to read? The initial response is that certain variables must be taken into consideration. Reading skills or ability to read is of primary concern. Interests, background, age levels, and gender are significant factors when considering how reading motivation is facilitated. Various writers and researchers have given tips and

techniques for motivating the student reader. Although there are basic guidelines, it must be realized that motivating a child of three will be significantly different from motivating a child of nine or an adolescent. Once that is established, there are certain guidelines by which educators operate in their attempts at motivating students to read, for example, providing the right atmosphere, providing the right kind of books, displays, read-aloud sessions, dramatization of texts, incentives, and using community and parents to support reading programmes.

Another major motivating factor is recognizing that students possess different interests, so that a wide range of books, journals, and magazines that meet the needs of the diversity of ages, gender, abilities, and cultures must be provided. Different genres of reading material must be provided to satisfy different tastes. In addition, there should be different levels of reading material as all children are not necessarily at the same level in terms of reading ability. Ediger (2001) states that a philosophy of sameness for all students in the reading curriculum with a "one-size fits all" set of beliefs should be discouraged. Worthy (2002) points out that like adults, students prefer to choose their own material and their tastes are varied and individual.

Displaying books in the classroom or the library is another way of motivating students, as an eye-catching display or an interesting book cover may attract the attention of the student.

A teacher or librarian can become innovative and creative in his/her approach to introducing books to students. Read-aloud sessions or dramatizations of books can prove to be motivating factors. Kasten and Wilfong (2005) show how providing a bistro-like atmosphere in the classroom increased their students' interest in reading.

A significant aspect of reading motivation is peer group reading where students come together to read, and to share knowledge and interpretations of material. This also has the added benefit of addressing their critical thinking, language, and social interaction skills. It also allows more understanding of issues and improves the ability to argue a point (Snowball, 2005).

Using students as mentors is another way of motivating other students. Garton (2005) describes a project where she used secondary students to read and discuss books with her elementary class. This proved to be a successful motivating factor for her students and it also helped to build self-esteem in the older students.

It is said that children live what they learn. Reading therefore becomes a part of one's life if it is seen as a habit and a pleasurable activity. One way that a child can be motivated to read is by seeing his/her parents reading. Parents must be encouraged to provide a reading atmosphere at home and must also be willing to cooperate with school programmes that foster reading motivation. Along with this, teachers themselves must be enthusiastic about reading and point out or encourage cooperation with role models who read and who could motivate students to read. Moniuszko (1992) has shown that involving the community helped motivate her students. When celebrities from the community were brought to read, this had a great impact. Students became enthusiastic, eager, and willing to read.

In the 1960s, the United States introduced a Sustained Silent Reading (SSR) policy in all public schools. This policy encouraged all students to read for a given period of time each day. The policy was introduced at all grade levels and proved to be highly successful in motivating students to read. Several other countries have since used this policy to generate an interest in reading.

The use of technology can also play an important role in reading motivation. As classrooms become more technology oriented and students are more techno-savvy, the Internet, CD-ROMs, Power Point presentations, software, videos, and e-books can all play an integral part in reading motivation. Coggeshall and Doherty (2004) described how their team focused on technology to help motivate their students and break down their resistance to reading. Other researchers have shown similar results. In addition, the multiliteracies described earlier by Luke and Elkin (as cited by Alvermann, 2003) show that students' literacy ability can be enhanced by technology.

Many researchers and teachers have also recognized the relevance of Gardner's theory of multiple intelligences as regards reading and reading

motivation. Recognizing this theory, reading motivation for each and every student can become a varied, profitable, and rewarding exercise for both students and educators.

The school's role, however, in helping to motivate students, is firstly to have a policy on reading, then to provide an environment in the classroom or library rich with books to satisfy the needs of all its students. Displays, author visits, reading competitions, reading rewards, and all kinds of imaginative and exciting projects can all play a significant role in intrinsic and extrinsic motivation and make lifelong readers of all students.

References

Allen, J. (1995). *It's never too late: Leading adolescents to lifelong literacy*. Portsmouth, NH: Heinemann.

Alvermann, D. E. (2003). *Seeing themselves as capable and engaged readers: Adolescents and re/mediated instruction.* Retrieved October 24, 2005, from North Central Regional Laboratory Website: http://www.ncrel.org/litweb/readers/readers.pdf.

Alvermann, D. E., Young, J. P., Green, C., & Wisenbaker, J. M. (2004). Adolescents' perceptions and negotiations of literacy practices in after-school read and talk clubs. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 870–913). Newark, DE: International Reading Association.

Beers, K. (2003). *When kids can't read: What teachers can do: A guide for teachers 6–12.* Portsmouth, NH: Heinemann.

Coggeshall, K., & Doherty, J. (2004). Technology that powers up learning. *Voices from the Middle, 11*(3), 23–29.

Ediger, M. (2001). *Reading: Intrinsic versus extrinsic motivation.* ERIC Document Reproduction Service ED 458 566.

Edmunds, K. M., & Tancock, S. M. (2003). Incentives: The effects on the reading motivation of fourth-grade students. *Reading Research and Instruction, 42*(2), 17–37.

Gambrell, L. B. (1996). Creating classroom cultures that foster reading motivation. *The Reading Teacher, 50*(1), 14–25.

Gambrell, L. B. (2004). Motivating kids to read. *Instructor, 113*(5), 10–11.

Garton, J. (2005, September 30). Passing on a passion for reading. *TES Teacher,* p. 4.

Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 403–422). Mahwah, NJ: Lawrence Erlbaum.

Hoy, W. K., & Miskel, C. G. (2001). *Educational administration: Theory, research, and practice* (6th ed.). Boston, MA: McGraw Hill.

Kasten, W. C., & Wilfong, L. G. (2005). Encouraging independent reading with ambience: The Book Bistro in middle and secondary school classes. *Journal of Adolescent and Adult Literacy, 48*(8), 656–664.

Moniuszko, L. K. (1992). Motivation: Reaching reluctant readers age 14–17. *Journal of Reading, 36*(1), 32–34.

Reynolds, M. (2004). *I won't read and you can't make me: Reaching reluctant teen readers.* Portsmouth, NH: Heinemann.

Snowball, C. (2005). Teenage reluctant readers and graphic novels. *Young Adult Library Services, 3*(4), 43–45.

Taylor, D., & Dorsey-Gaines, C. (1988). *Growing up literate: Learning from inner-city families.* Portsmouth, NH: Heinemann.

Wigfield, A. (1997). Children's motivation for reading and reading engagement. In J. T. Guthrie & A. Wigfield (Eds.), *Reading engagement: Motivating readers through integrated instruction* (pp. 14–33). Newark, DE: International Reading Association.

Worthy, J. (2002). The intermediate grades. What makes intermediate-grade students want to read? *The Reading Teacher, 55*(6), 568–569.

Xue, Y., & Meisels, S. J. (2004). Early literacy instruction and learning in kindergarten: Evidence from the Early Childhood Longitudinal Study-Kindergarten Class of 1998–1999. *American Educational Research Journal, 41*(1), 191–229.

BIBLIOGRAPHY

Parental Role

Baker, L. (2003). The role of parents in motivating struggling readers. *Reading and Writing Quarterly, 19(*11), 87–106.

This article seeks to identify the latest research about home influences on reading motivation, and to provide teachers with suggestions on how they can enlist the assistance of parents in motivating struggling readers.

Baker, L., & Scher, D. (2002). Beginning readers' motivation for reading in relation to parental beliefs and home reading experiences. *Reading Psychology, 23*(4), 239–269.

This study used 65 six-year-olds and their mothers from different sociocultural backgrounds to examine children's motivation for reading in relation to parental beliefs and home literacy experiences. It assessed each child's interest and competence in reading, and parents' beliefs about reasons for reading. The results showed that motivation crossed the barrier of income level, ethnicity, and gender, and that parents identified pleasure as the children's motivating factor for reading. The results also showed that parents who enjoy reading transfer this enjoyment to their children.

Baker, L., Scher, D., & Mackler, K. (1997). Home and family influences on motivations for reading. *Educational Psychologist, 32*(2), 69–82.

This article reviews some of the literature on home and family influences on children's motivation for reading. It also shows that children who enjoy reading at an early age become frequent readers later on. It explains that shared storybook reading between parents and children plays an important role in promoting reading motivation. The article states that parents who believe that reading is a source of entertainment produce children with more positive views of reading than parents who emphasize the learning aspect of reading.

Chandler, K. (1999). Reading relationships: Parents, adolescents, and popular fiction by Stephen King. *Journal of Adolescent and Adult Literacy, 43*(3), 228–239.

This article reports on a case study of 12 high school juniors who saw themselves as avid readers of fiction by Stephen King. The study found that for 8 out of 12 students, the influence of this choice of fiction was the parents. A survey was then done of parents and their perceptions of their children as readers and their own reading habits with respect to popular fiction. The article also explores how popular fiction became a medium of communication for students and parents in the study, and describes the role that parents play in forming their children's reading habits.

McDowall, H. (2004). Parents as partners in the school's reading drive for excellence. *The Educational Journal of Trinidad and Tobago, 14*(1), 48–57.

This article stresses that parental involvement in children's reading does not preclude teacher instruction and that every effort should be made to empower parents to facilitate and nurture their children's emerging literacy. In so doing, parents would be enabled to assume their role as partners in the school's reading drive. The article also gives hints specifically designed to provide parents with information on promoting their children's literacy development.

Morrow, L. M. (1983). Home and school correlates of early interest in literature. *Journal of Educational Research, 76*(4), 221–230.

This study sought to determine the level of literary interest of kindergarten children from 21 classrooms. Data were collected from parent questionnaires, teacher evaluations, and tests on children's free-time home activities, parents' characteristics and activities, school achievement, and the quality of the in-classroom literary environment. The results showed that children who showed higher interest were read to more often, watched less television, and had more books in their homes. The study also showed that parents and teachers of high-interest children provided supportive literary environments at home and in school.

Morrow, L. M., & Young, J. (1999). A family literacy program connecting school and home: Effects on attitude, motivation and literacy achievement. *Journal of Educational Psychology, 89*(4), 736–742.

In this study, a group of 56 children from first to third grades were studied in a special programme that connected home and school literacy by involving parents with their children's literacy development. The purpose of the programme was to enhance children's achievement and interest in literacy. The family programme was similar to a literature-based school programme that included classroom literacy centres, teacher-modelled literature activities, and literacy centre time. Monthly

meetings were held with teachers, parents, and children. Pre- and post-test data showed that the children in the programme were more motivated and had achieved more.

Shockley, B. (1994). Extending the literate community: Home-to-school and school-to-home. *The Reading Teacher, 47*(6), 500–502.

In this article, a second grade teacher describes an experiment in which she included parents in her drive to develop her students' literacy growth. The experiment was a year-long project, which involved parents reading and discussing books with their children and writing journals. The goal was to connect families with books, their children, and their literacy experiences. The experiment proved to be successful in developing the parent-teacher relationship.

Sonnenschein, S., & Munsterman, K. (2002). The influence of home-based reading interactions on 5-year-olds' reading motivations and early literacy development. *Early Childhood Research Quarterly, 17*(3), 318–337.

A survey of 5-year-olds was undertaken to understand the impact of home-based practices on young children's literacy development. The comments made about books read and the affective quality of interactions with parents reading to their children were recorded. Parents were also interviewed about the frequency with which their children engaged in reading activities at home. The results showed that the affective quality of the reading interaction was the most powerful predictor of children's motivation for reading.

Teale, W. H. (1981). Parents reading to their children: What we know and need to know. *Language Arts, 58*(8), 902–912.

This article states that researchers all agree that reading to preschool children is valuable, as children develop interest and skill in literacy from this activity. It also shows how the research associated with reading to children helps literacy development and concludes by looking at some key areas where more research is needed.

Strategies for Motivating Readers

Alden, K., Lindquist, J. M., & Lubkeman, C. A. (2003). *Using literature to increase reading motivation.* Chicago, IL: Saint Xavier University & Sky Light Professional Development Field-Based Master's Program. (ERIC Document Reproduction Service No. ED 481 442)

The purpose of this project was to increase Grade four students' reading motivation by helping them to develop a broader awareness of literary genres, and through participation in literature-based discussion activities. A literature-rich classroom was created, a reading interest survey was conducted, a variety of genres was introduced, and book buddy discussions were also introduced. The final outcome was that the students' motivation to read increased.

Allen, J. (2001). Eliminating a "yes, but" curriculum. *Principal Leadership, 2*(2), 10–15.

This article states that everyone agrees that literacy is important but that there seems to be barriers to helping secondary school students become more literate. It is felt that the barriers to teaching reading effectively must be examined and practices must change. The writer discusses her research on student literacy and provides advice on how to encourage students to read.

Baker, M. (2002). Reading resistance in middle school: What can be done? *Journal of Adolescent and Adult Literacy, 45*(5), 364–366.

This article provides advice for language arts teachers on overcoming reading resistance. Strategies include providing a greater variety of classroom materials, providing class time for reading and discussions, and examining classroom atmosphere from an adolescent perspective.

Clary, L. M. (1991). Getting adolescents to read. *Journal of Reading, 34*(5), 340–345.

This article lists six strategies for teachers to get adolescents to read: 1) capitalize on interests, 2) make reading material accessible, 3) build a conducive environment, 4) allow time to read in school, 5) provide adult

models, and 6) use motivational techniques such as reading aloud and book discussions.

Collins, N. D. (1996). *Motivating low performing adolescent readers.* Bloomington, IN: ERIC Clearinghouse on Reading, English, and Commmunication. (ERIC Digest ED 396265)

This article provides some guidelines on motivating students to read, for example, by providing appropriate reading material that reflects the interests of adolescents, encouraging parents to become involved in their children's reading programmes, listening to oral and recorded reading, and asking questions. It also states that supplementary material like newspapers and magazines can help to motivate students to read.

Davis, M., & Lyons, S. (2001). Improving reading by … reading: Ideas from two teachers. *Voices from the Middle, 8*(4), 51–57.

In this article, two teachers share their motivational strategies, which were initiated with a home and school programme called "Reading — Exercise Your Mind." In this programme, bags with three books on one of a variety of topics were given to students. The material ranged from "too easy," "too hard," and "just right." Parents were enlisted in this programme. The too-easy books were read by students to parents, the too-hard books were read to students by parents, and the just-right books were read by students with parents. Parents also had to give feedback on the programme to the teachers. This programme proved to be successful since it involved teachers, parents, and children.

Dorion, R. (2003). Motivating the lifelong reading habit through a balanced use of children's information books. *School Libraries Worldwide, 9*(1), 39–49.

Literacy teachers tend to use mostly fiction as the chief source of materials for motivating students to read. In this article, the author shows how the use of information books can encourage and motivate girls and boys to do more independent reading.

Dreher, M. J. (1999). Motivating children to read more nonfiction. *The Reading Teacher, 52*(4), 414–417.

This article states that there is an imbalance between reading fiction and non-fiction in the elementary school. It encourages teachers to assist children to read more material and to extend their reading to non-fiction, and discusses how to motivate children to read more non-fiction.

Dreher, S. (2003). A novel idea: Reading aloud in a high school English classroom. *English Journal, 93*(1), 50–53.

In this article, a high school teacher shares his experience with reading aloud to and with his students, and shows how it improved and motivated their interest in reading.

Ediger, M. 2001. *Reading: Intrinsic versus extrinsic motivation.* (ERIC Document Reproduction Service No. ED 458 566)

This paper seeks to analyse the two points of view about motivating students in reading achievement. It provides techniques to assist teachers in developing both intrinsic and extrinsic motivation, and concludes that most reading teachers will use a combination of both techniques to motivate students.

Fischer, C. (2000). An effective (and affordable) intervention model for at-risk high school readers. *Journal of Adolescent and Adult Literacy, 43*(4), 326–335.

This article describes a programme used by a teacher to help high school students raise their literacy levels. It included tailor-made assignments with appropriate interest material for each student, the use of community volunteers and one-on-one tutors, and regular opportunities for students to read to elementary school children. The programme proved to be highly successful. The most important benefit was that students became efficient at, and enjoyed, reading to younger students.

Gambrell, L. (1996). Creating classroom cultures that foster reading motivation. *The Reading Teacher, 50*(1), 14–25.

This article discusses what research and theory suggest about the role of motivation in reading engagement. It describes six research-based factors

that are related to increased motivation to read. These are: 1) the teacher as a reading model, 2) a book-rich classroom environment, 3) opportunities for choice, 4) opportunities to interact socially with others, 5) opportunities to become familiar with lots of books, and 6) appropriate reading-related incentives.

Gambrell, L. (2004). Motivating kids to read. *Instructor, 113*(5), 10–11.

This article provides strategies that teachers can use to motivate students to read. These strategies include acquiring a large selection of books, displaying books appropriately, endorsing books, and providing students with an introduction to the books offered.

Guth, N., & Heaney, P. (1998). A challenge for school administrators: Motivating adolescents to read. *NASSP Bulletin, 82*(600), 34–40.

This article discusses the importance of literacy and describes literacy programmes used in one high school. It states that each school is unique and must use its available resources. In addition, it advocates that literacy programmes must be a collaborative effort involving staff, students, parents, and the community at large.

Guthrie, J. T., Alao, S., & Rinehart, J. M. (1997). Engagement in reading for young adolescents. *Journal of Adolescent and Adult Literacy, 40*(6), 438–446.

This article reviews research that shows how strategies for reading are linked to motivation. It also shows how motivation and engagement can be increased in the classroom through the Concept-Oriented Reading Instruction programme developed at the National Reading Research Centre, USA. In this programme, teachers emphasize seven principles for creating engaging classrooms: 1) real-world observation, 2) conceptual themes, 3) self-directed learning, 4) strategy instruction, 5) social collaboration, 6) self-expression, and 7) coherence in the curriculum.

Guthrie, J. T., & Cox, K. E. (2001). Classroom conditions for motivation and engagement in reading. *Educational Psychology Review, 13*(3), 283–302.

This study sought to identify ways of increasing long-term reading engagement in classrooms. The article focuses on three questions: 1) How can we increase long-term reading engagement in the classroom? 2) Is our approach for increasing reading engagement and motivation more effective than traditional reading instruction? and 3) What are the critically important features of a classroom context that fosters long term reading engagement?

Heathington, B. S. (1979). What to do about reading motivation in the middle school. *Journal of Reading, 22*(8), 709–713.

This article suggests that middle-grade students differ physically, emotionally, socially, and intellectually from students in other grades, so that a reading programme should bear this in mind. In developing a reading programme for her school, the author conducted a survey of students to determine their reading activities. The survey revealed that their reading activities were indeed influenced by their physical, emotional, social, and intellectual conditions. The students also indicated that: 1) they did not have enough time to read, 2) there were too many interruptions during reading time, 3) there were not enough books on the topics they liked, 4) they could not read well, and 5) they wanted to choose their own books. A reading strategy was developed, which incorporated these variables.

Kasten, W. C., & Wilfong, L. G. (2005). Encouraging independent reading with ambience: The Book Bistro in middle and secondary school classes. *Journal of Adolescent and Adult Literacy, 48*(8), 656–664.

This article shows how teachers used a café atmosphere to promote reading and interaction among adolescent students. This was done by encouraging students to read books on their own, bring books to class for a scheduled event, and linger over books with discussions in a café atmosphere.

Lause, J. (2004). Using reading workshop to inspire lifelong readers. *English Journal, 93*(5), 24–30.

In this article, a teacher describes her frustrations at her high school students' lack of interest in reading and the design of a programme that helped them become better readers. This programme included students' selection of books, discussions on books, and reading at home for a specified time.

Lyons, C. A. (2003). *Teaching struggling readers: How to use brain-based research to maximize learning.* Portsmouth, NH: Heinemann.

This book introduces teachers to the concepts, categories, language, and arguments pertaining to the brain's control of what readers do. It also describes strategies for reaching struggling readers, and provides tips to guide teachers on observing their students and planning instructions more effectively.

Moniuszko, L. K. (1992). Motivation: Reaching reluctant readers age 14–17. *Journal of Reading, 36*(1), 32–34.

In this article, the author shows that partnering with people from the community was the key to motivating her students to read, since it made the connection between the real world and student interest.

Morrow, L. M. (2003). Motivating lifelong voluntary readers. In J. Flood, D. Lapp, J. R. Squire, & J. M. Jensen (Eds.), *Handbook of research on teaching the English language arts* (2nd ed., pp. 857–867). Mahwah, NJ: Lawrence Erlbaum.

This article states that voluntary, recreational, independent reading must be an integral part of the total development programme. It describes the benefits of voluntary reading, the characteristics of voluntary readers, and gives a framework and strategies for providing voluntary reading in schools.

Morrow, L. M. (2004). Motivation: The forgotten factor. *Reading Today, 21*(5), 6–7.

This article provides six suggestions for motivating students to read: 1) create literacy-rich environments in the classroom, 2) establish a literacy centre, 3) provide time for choice and collaboration, 4) read to students,

5) encourage relevant reading and writing, and 6) have high expectations for student success.

Mulholland, R. (2002). Using high-interest materials to engage secondary students in reading. *Reading Online, 6*(3). Retrieved April 28, 2005 from http://www.readingonline.org/articles/art_index.asp?HREF=mulholland/index.html

This article shows how one teacher used plays to motivate her students to read. She linked reading a play with finding newspaper articles and information on the Internet to prepare reports to share with the class. This generated a great deal of interest and, as a result, the students developed additional reading interests.

Norton, D. Y. (1992). Motivating reluctant readers using a reading incentive program. *The Reading Teacher, 46*(3), 271–272.

This article shows how one teacher used her initiative to encourage students to read by setting up an out-of-school reading programme to emphasize the importance of reading for pleasure. Parents were involved and incentives were given. Results showed that students selected books from a wider range of topics, read more, spent more time reading, and were eager to share books with others.

Oldfather, P. (1995). Commentary: What's needed to maintain and extend motivation for literacy in the middle grades. *Journal of Reading, 38*(6), 420–422.

A four-year longitudinal study of students' intrinsic motivation for literacy learning found that when students had opportunities for authentic self-expression as part of their literacy activities, they were more intrinsically engaged in reading.

Reynolds, M. (2004). *I won't read and you can't make me: Reaching reluctant teen readers.* Portsmouth, NH: Heinemann.

This book provides practical and effective strategies that can be used to help motivate students to read. Some topics discussed include the importance of respect for students' attitudes, experiences, perceptions,

and choices regarding reading; tips for motivating reluctant readers; classroom management issues for readers; and student/teacher programme accountability in reading.

Richardson, M. V., & Miller, M. B. (2001). Motivating students to read: Using authors and literature from their home state. *Reading Improvement, 38*(3), 119–124.

This article proposes an idea for resolving the problem of motivating students to read. The idea is that children like to read about familiar places, people, and settings; and they also like to read books written by authors who live or have lived in their home state. Therefore, providing this kind of material is a good incentive to reading.

Sanacore, J. (1997). Promoting lifetime literacy through authentic self-expression and intrinsic motivation. *Journal of Adolescent and Adult Literacy, 40*(7), 568–571.

This article states that students' desire to read may lessen as they move from elementary to secondary school, and that a primary cause may be too much emphasis on teacher-centred approaches to instruction. It states that this approach tends to stifle students' opportunities for authentic self-expression, and this problem appears to diminish students' intrinsic motivation for literacy learning. It advises that teachers' goals should be to stimulate students' internal desire for personal expression while guiding their energies to lifetime and lifelong literacy efforts.

The Role of the Libraries and Librarians

Braxton, B. (2004). Encouraging students to read for pleasure. *Teacher Librarian, 31*(3), 39–40.

This article provides strategies that teacher-librarians can use in the library to encourage students to read. These include eye-catching display of books; encouraging parents to read as an example; use of drama at story time; making books more easily accessible; and book discussions.

Dahlhauser, J., & Purcell, H. (2003). Motivating boys as beginning readers. *Teacher Librarian, 30*(3), 29–31.

Two librarians reveal ways in which they encourage boys to read. These include providing non-fiction books and current and up-to-date material that they can recognize, sharing reading experiences, conducting read-aloud sessions, and having members of the community come in to read.

Fenn, J. (2005). Eight ways your librarian can help promote literacy. *Principal Leadership, 5*(6), 49–51.

This article provides eight suggestions for ways in which the school librarian can help to promote literacy among students. These are: 1) providing books that students want to read, 2) promoting reading for pleasure, 3) recommending related reading to teachers in many subject areas, 4) collaborating with teachers to incorporate literature and technology into the curriculum, 5) offering online resources to foster literacy components in curriculum areas, 6) writing grants that involve teachers and students in literacy-related activities, 7) developing special events and displays tied to reading, and 8) modelling lifelong learning and reading for pleasure in every possible way.

Fitzgibbons, S. (2004). What motivates reading? How library media specialists can contribute to the development of good readers. *School Library Media Activities Monthly, 20*(10), 21–25.

This article reveals various findings from research, which show that motivation to read is an essential ingredient in the development of good readers and lifelong readers. It provides some suggestions for motivating students to read, including greater access to new, attractive, and popular paperbacks; allowing student to select material for libraries, identifying role models who are readers, and giving incentives for reading.

Jones, P., & Fiorelli, D. C. (2003). Overcoming the obstacle course: Teenage boys and reading. *Teacher Librarian, 30*(3), 9–13.

In this article, a teacher/librarian makes the point that teenage boys tend to read less than girls. She gives some suggestions for encouraging boys to read. Some of these suggestions are: 1) plan a programme for boys only with books that should interest them, for example, sports; 2) use book talk sessions with non-fiction books and make use of visual aids in

these sessions; 3) put up posters of male readers in the library; 4) enlist coaches of boys' teams in a reading programme; 5) put books near to the computers on the topics that they generally search on the Internet.

The Role of Technology

Coggeshall, K., & Doherty, J. (2004). Technology that powers up learning. *Voices from the Middle, 11*(3), 23–29.

This article shows how the use of technology with students promoted an interest in reading. Students were encouraged to discuss their books with the use of slide shows and PowerPoint presentations, which generated a lot of interest. The blend of reading and technology was successful in motivating students to read.

Franklin, D., & Ferguson, T. (2005). Technologies to achieve reading success. *Media and Methods, 41*(6), 28–29.

This article shows how one teacher uses technology in his strategies for teaching and motivating reading. These strategies incorporate the use of projectors, scanners, the computer, e-books, and reading software.

Glasgow, J. N. (1997). It's my turn! Part II: Motivating young readers using CD-ROM storybooks. *Learning and Leading with Technology, 24,* 18–22.

This article discusses the use of CD-ROMs to motivate and help young readers. It also states that multimedia can help connect the text to the words children hear and the things they see by combining text, sound, and graphics. This technique helps to keep their interest in reading.

Malloy, J. A., & Gambrell, L. (2006). Approaching the unavoidable: Literacy instruction and the Internet. *The Reading Teacher, 59*(5), 482–484.

This article shows how reading online is becoming popular with students who were not otherwise interested in reading. It notes that teachers should appreciate and prepare themselves for this new trend and find ways to use this tool to enhance instruction that engages students. It also

shows how reading online differs from reading print because of the wealth of information available on the Internet.

Tankersley, J. (2003). Using technology to promote reading. *Multimedia Schools, 10*(3), 40–41.

One teacher describes how he used a project that combined reading and technology to stimulate students to read. Students who had read books for a students' choice book award contest were asked to give a book talk, which was videotaped. This was shown to other classes. Because of this innovative use of technology, other students were eager to read, and give their book talks to be videotaped.

SELECTED WEBSITES

Book Adventure
http://www.bookadventure.com

Book adventure is a free reading motivation programme for children in Grades K–8. Children create their own book lists from over 7,000 recommended titles, take multiple choice quizzes on the books they have read offline, and earn points and prizes for their literacy successes. Book Adventure was created by Sylvan Learning.

Enlighten Me
http://www.enlightenme.com/enlightenme/

Created by VerizonReads and FableVision, this website for children ages 7–12, as well as parents, teachers, and caregivers, features articles, activities, and booklists designed to promote creative thinking and encourage a lifelong love of reading.

Reading is Fundamental
http://www.rif.org

RIF develops and delivers children's and family literacy programmes that help prepare young children for reading and motivate school-age children to read. Through a network of teachers, parents, and volunteers,

RIF programmes provide books and other essential literacy resources to children at no cost to them or their families.

Reading Rockets
http://www.readingrockets.org

This is a free website created for parents and educators to help them ensure that children read.

Spaghetti Book Club
http://www.spaghettibookclub.org/

The Spaghetti Book Club encourages children's love for reading by giving them an opportunity to connect on a personal level with the books they are reading and then share their reactions, thoughts, and opinions with family members, friends, and children around the world.

TEAM TEACHING AT THE PRIMARY LEVEL
Insights into Current Practice in Trinidad and Tobago

Margaret E. Cain and *Iris P. Hewitt-Bradshaw*

Team teaching is an arrangement that is sometimes used to organize instruction in primary schools in Trinidad and Tobago, and its use needs to be documented. This qualitative study used the principle of maximum variation to select four teaching teams and to investigate the factors that determine the introduction of team teaching, the organization and functioning of the teams, and the perceived benefits and challenges. Data gathered from interviews with the principals and team members suggest that team teaching was usually introduced in response to the needs of the institution, and that principals primarily determined how teams were constituted. School cultures that fostered collaboration and collegiality facilitated the introduction of team teaching, and where the allocation of subject was based on teacher preference or expertise, there were benefits for both pupils and teachers. Factors at the level of the institution, the class, and the individual influenced the functioning of teams. The findings also indicate a critical need for ongoing assessment of the impact of team teaching arrangements on teaching and learning.

Introduction

Over time, schools have explored different methods of organizing instruction. Traditionally, at the primary level in Trinidad and Tobago, the class teacher is a generalist——one teacher is assigned to one class with the responsibility for teaching all subjects. This can make extensive demands on the knowledge and skills of the teacher as well as foster teacher individualism and isolation.

Although this generalist approach remains the most common arrangement for instruction, other approaches that require teachers to specialize in specific subject disciplines have been adopted. At the primary level, team teaching is one such arrangement. This term is

commonly used to refer to any arrangement in which subjects in the curriculum are divided between two or more teachers who are given the responsibility of instructing one class. Slater (1993) describes this arrangement as an interdisciplinary team with subject specialization and shared responsibilities for the same students.

Primary schools in Trinidad and Tobago have the autonomy to organize the curriculum within broad guidelines. The factors that influence the decision to introduce team teaching vary in schools that use it, and in the Trinidad and Tobago context, these factors need to analyzed and documented.

The most common arrangement of teaching teams is the assignment of two teachers to one class. There are, however, variations in how teams are organized and how they function. These arrangements need to be examined. The findings of a study of team teaching can foster greater understanding of its use and provide a basis for assessing its effectiveness in the Trinidad and Tobago educational context. Such a study can also provide the impetus for practitioners to document, analyze, and reflect on how team teaching functions in their schools, and modify their programmes if necessary.

This study examined the factors that influenced the introduction of team teaching in three schools in Trinidad and Tobago. The questions that guided the research were:

1. What factors determine the introduction of team teaching in primary schools?
2. How is team teaching organized and how does it function?
3. What are the perceived benefits and challenges of team teaching?

Issues in Team Teaching

The nature of contemporary teaching demands extensive skills and knowledge on the part of the teacher, especially at the primary level. Team teaching is one strategy that has been employed by schools to attempt to meet those demands. As a teaching and learning option, it can promote a culture in which expertise is shared as opposed to one that

fosters isolated, individual, professional problem solving (Brownell & Walther-Thomas, 2002). A study by Evans (1997) of newly qualified teachers in Jamaica, for example, found that isolation and absence of collaboration hindered the development of their professional collegial relations with experienced teachers. Educators who favour team teaching contend that it offers a means of meeting the increasingly diverse needs of pupils (Pugach & Wesson, 1995; Slater, 1993). Brennan et al. (1973) and Miller (1992) also suggest that team teaching is a useful technique for improving learning "efficiency" and bringing about changes in pupil or teacher attitudes. Since team teaching involves varied degrees of subject specialist teaching, it offers a means of meeting the demands that the curriculum makes on the content knowledge of primary school teachers.

In the past, subject specialist teaching at the primary level was more commonly found in the creative arts—music, art, or physical education—where it was assumed that an innate talent was essential (Hall, 2000). However, there is now an increased use of subject specialization in the teaching of academic subjects at the primary level. The main advantage cited for specialist (or semi-specialist) teaching is that the subject specialist brings a high level of subject knowledge to his/her teaching, and it is the lack of such knowledge that is the main weakness in the generalist class teacher model (Hall, 2000). Proponents of the generalist model argue that the generalist teacher has better knowledge of pupils and their individual needs, and can ensure the coherence and balance of their educational experience. A report cited in Thornton's (1998) review suggests that it is advantageous to maintain the generalist system as far as possible; but that it is unreasonable to expect one teacher to cope unsupported with the depth and width of the modern curriculum.

Marilyn Friend, a leading expert in professional collaboration, argues that "there is simply too much for any one educator to know in order to effectively meet the needs of all his or her students" (Brownell & Walther-Thomas, 2002. p. 224). Friend thus advocates collaboration in the delivery of the curriculum. From this perspective, if team teaching is well structured and organized, it has the potential to foster collaboration, cooperation, and collegiality and thus make curriculum delivery more manageable and effective.

Any decision to introduce team teaching at the primary level requires consideration of how teams are to be structured and how they will function. The number of teachers in a team may vary and they may teach a single class or classes that are joined. Teams can be arranged vertically or horizontally with discernible differences in authority structure and degree of specialization (Clifford & Friesen, 1993; Lovell, 1967; McGoogan, 2002). Whether members are of equal professional status may be important to the success of the team. Craig (1998) for example, shows the difficulties that can confront teams comprising expert and novice or teacher in training.

With respect to the functioning of teams, the need for sustained collaboration is highlighted in the literature. Coordination requires that members of the team have a common philosophy so that they will have a similar order of priorities. They must share a vision and work towards a common goal. Teams must therefore plan, prepare, and dialogue about delivery. The relationship between team members is consequently an important issue. Craig's (1998) study concludes that team teaching relationships should be collaborative and non-hierarchical, and that team members need to know the direction in which their class is moving and share responsibility for shaping the class "story." It is difficult, though, for teams to achieve high levels of collaboration in the classroom if the culture of the school is not collegial and collaborative. School administration has a critical role to play in this regard (Brownell & Walther-Thomas, 2002).

Studies invariably highlight the benefits of team teaching for both pupils and teachers (Lovell, 1967; McGoogan, 2002; Miller, 1992; Pugach & Wesson, 1995; Slater, 1993). Teachers combine their expertise to decide on appropriate content. Teacher/s not involved in the actual lesson presentation can individualize content for special or at-risk pupils. Feedback and evaluation can also be assigned to the free teacher/s (Rink, 1998). The level of cooperation displayed by teachers can serve as a model for pupils. Brennan et al. (1973) suggest that team teaching provides an ideal opportunity for teachers to learn from one another as they plan for their learners—a situation that the writers see as particularly advantageous for the new teacher. Williams, Prestage, and Bedward (2001) emphasize the role of collaboration and collegiality in

promoting teacher development and in eliminating professional isolation. Indeed, since teachers usually teach the subject in which they have greater expertise, viewing their peers teach may improve their knowledge in their weaker areas.

Time for "constructive communication" is identified as the greatest challenge to successful collaboration for teaching teams (Brownell & Walther-Thomas, 2002). Difficulties are also created when principals and teachers have mindsets that work against the development of a collaborative school culture. If team members do not trust and respect each other, then the team cannot function efficiently.

Method

A qualitative approach was used in this study to gain insights into the phenomenon of team teaching in the primary school system in Trinidad and Tobago. Some primary schools are government schools and are owned and operated by the State. Others are assisted by the State but are managed by private institutions, usually a denominational body. Children spend seven years in one of these schools at the primary level, at the end of which they write the Secondary Entrance Assessment (SEA) examination for entry into secondary school. For this study, the researchers first identified schools in which team teaching existed and examined how team teaching was organized. This examination revealed varying patterns of organization of instruction. These included vertical arrangements in which two or more teachers delivered instruction to classes at different levels. In addition, we discovered horizontal arrangements in which two or more teachers were involved in the instruction of one or more than one class of pupils at a specific level. In some schools, there were both vertical and horizontal arrangements. Given the scope of the study, we confined our investigation to horizontal arrangements of team teaching.

In order to obtain an in-depth understanding of the subject, four cases were selected for study using a purposive sampling strategy. Purposive sampling facilitates the selection of cases that allow the investigator to discover, understand, and gain insights into the phenomenon (Merriam, 1998). Merriam suggests that maximum variation helps to identify and seek out those cases that represent the widest possible range of

characteristics of interest for the study. We selected cases that exhibited variation in horizontal team teaching arrangements at three sites in the primary school system. The cases selected varied on the basis of the composition of the team and the characteristics of members. These characteristics included the number of teachers on the team, their years of experience as trained teachers, and their gender. There was also variation in the way the teams were organized.

Within each site, we selected one case for in-depth study. However, at one site, which had a long history of team teaching, we decided to study a second team, one member of which was a male teacher in an Infant class. Male teachers assigned to an Infant class are not common in the Trinidad and Tobago school system. All teachers in the cases under study were trained. Data were obtained through audiotaped interviews with the teachers involved in team teaching in the selected classes and the principals of the three schools. The interviews were semi-structured and sought to determine why team teaching was introduced, how it was organized, and the participants' perceptions of the benefits and challenges of team teaching. The interviews were transcribed and the two researchers jointly coded the transcripts. The codes were then organized into broader themes/categories. In the interest of anonymity, all individuals and schools have been given pseudonyms.

The Cases

The three sites selected for study were: Seaview Primary, Tamarind Heights Primary, and Pine Valley Primary. Seaview Primary is located in north Trinidad, and the other two in the south. The three schools varied in a number of ways. Two schools, Seaview Primary and Tamarind Heights Primary, had two teachers assigned to one class, while Pine Valley Primary had an arrangement in which three teachers worked with a group of three classes.

The history of team teaching varied at the three sites. Tamarind Heights Primary had a long tradition (20 years) of team teaching, whereas Seaview Primary and Pine Valley Primary had only recently begun experimenting with this approach to teaching. Seaview Primary had introduced team teaching three years prior to the study, while Pine

Valley Primary had only been using this approach for about one year. The three schools had populations varying between 257 and 415 pupils.

Seaview Primary is a denominational boys' school that was established over 100 years ago. It is a relatively small school located in a semi-urban area. At Seaview Primary, the case studied was a Standard 4 class of 46 pupils. In Standard 3, these pupils were in two separate classes that were merged at the Standard 4 level. The pupils in this class were of mixed ability. Two teachers, both male, shared the class in a single, self-contained classroom. One was an experienced teacher with 19 years of service while the other was a newly qualified teacher. In this study, the term *newly qualified teacher* refers to one who is in the first year of teaching after graduation from teachers' college. The more experienced of the two had prior experience of team teaching at another school. For the newly qualified teacher, it was the first experience of team teaching. Both teachers had only joined the staff at the start of the academic year and had no prior knowledge of the pupils.

Tamarind Heights Primary has been in existence for 42 years. It is a government school for boys in an urban district. At this school, two cases were investigated. One team, Team A, consisted of an experienced teacher working with a newly qualified teacher. They taught a Standard 2 class of 33 pupils of mixed ability, in a single classroom. The senior teacher had been teaching at the school for over 10 years. The other team, Team B, comprised an experienced teacher and a teacher with three years experience who had just joined the staff. They shared a First Year class of 45 pupils in a self-contained classroom. The new teacher had no prior experience teaching at this level. Unlike his team partner, he had never before taught in a team. In both cases, the senior teachers were female, while the less experienced teachers were male.

Pine Valley Primary is a denominational school that was established 50 years ago. It is a boys' school located in a rural district. The team comprised three trained teachers, instructing three classes at the Standard 4 level. The classes were of mixed ability, and were only to be streamed just before the SEA examination. Each class comprised 15 pupils who were taught as a group. This case was unique in that each teacher had a self-contained classroom to which one class went for instruction. There were two female teachers and one male teacher in this team. They were

all experienced and had been at the school for a long time, two of them for over 10 years.

Findings

Factors Determining the Introduction of Team Teaching

At Seaview Primary, team teaching was initiated by the principal, who was experimenting with it for the first time. It was introduced because of limited classroom space and a surplus of teachers. Tamarind Heights Primary had a policy of mentoring for relatively inexperienced teachers, whether they were trained or not. Team teaching also addressed the problem of limited space and large numbers of pupils in classes. Team teaching at Pine Valley Primary was initiated by the teachers involved, in consultation with the principal. Their main consideration was a desire to improve pupil learning. Thus, team teaching was mainly introduced to improve the functioning of the school, either as a policy decision or as an ad hoc measure.

Most team members indicated that a culture of collaboration and collegiality facilitated the introduction of team teaching:

> Mother's day, father's day, graduation, you know, whenever we have something…to celebrate, the entire school will be involved. Contributions from each class…So, collaboration takes place on a school level…even though people may be in separate classrooms. (Mr. Williams, Tamarind Heights Primary [THP])

> …the whole staff is one. All the children too, I think it rub off on them…we have a lot of activities…we bring out our Carnival band; all the teachers take part. We have a vibrant house system. (Ms. Jones, Pine Valley Primary [PVP])

Organization and Functioning

At Seaview Primary, members of the team were selected by the principal based mainly on the perceived personality of the teachers. Subjects to be taught were selected by members of the team on the basis of preference. One teacher taught mathematics, language arts, social studies, art, and religious education, while the other taught science, mental mathematics, physical education, and also reinforced areas of mathematics. The

teachers collaborated in planning lessons, but the more experienced teacher usually taught during the morning session while the newly qualified teacher taught during the afternoon. They also sat in on each other's lessons and, at times, there was joint delivery.

Teaching teams at Tamarind Heights Primary were selected by the principal who had practical experience of team teaching. In Team A, the criterion for allocation of subjects was teacher preference. One teacher taught mathematics, science, and social studies, while the other taught all areas of language arts and creative arts. There was no joint planning and each teacher delivered instruction in the subjects for which he/she was responsible. In Team B, subjects were allocated on the basis of expertise. The senior teacher taught mathematics, science, and composition, while the junior teacher taught social studies, grammar, and composition. There was joint planning and, at times, joint delivery of lessons.

At Pine Valley Primary, the selection of team members was based on experience and expertise. A male teacher was deemed desirable due to concerns about discipline. Three teachers taught three classes, and expertise and preference guided subject allocation. One teacher taught mathematics and science, another taught agricultural science, grammar, and comprehension, while the third teacher taught social studies and vocabulary. Teaching was organized in such a way that each teacher was responsible for allocated subjects with all three classes. However, there was a core of subjects (religious instruction, creative arts, and handwriting) that each teacher taught to an assigned class. The timetable for each class changed from week to week during a three-week cycle. There was joint planning and, usually, individual delivery of lessons, although at times two classes were combined and taught by two teachers. Teacher-assigned classrooms allowed each teacher to organize learning centres in the areas of his/her specialty.

Administrative support was a significant factor in team teaching. This support took different forms. At Tamarind Heights Primary, planning sessions were timetabled and administrators supervised classes. At Pine Valley Primary, the principal established the boundaries and gave the team the freedom to structure the arrangement in a manner they thought appropriate. On the other hand, even when the principal supported the arrangement, the lack of structure inhibited the success of teaming. For

example, at Seaview Primary, the team indicated that their programme of work for the class was sometimes disrupted when the administration assigned a team member to duties outside of the classroom. They suggested that principals should know their staff resources and use them judiciously. Brennan et al. (1973) and Lovell (1967) support this position and suggest that optimum use of the strengths of individual teachers is critical in team teaching.

The status of team members varied. In the majority of cases, a senior teacher acted as a mentor to a less experienced teacher, while in one case, all teachers on the team enjoyed equal professional status. The quality of the relationship depended on the attributes of the members of the team. Recurring themes in the analysis were team members' willingness to learn from each other, work cooperatively, and discuss issues; their ability to fulfil classroom responsibility; and the degree to which partners were open to feedback and criticism. The following comments capture some of these themes:

> It is cooperation. Without that, nothing doing. You have to be dedicated. Know that if you stay home, you don't prepare, or you don't pull your weight the children will suffer. So *that* part of it is very, very important, I think, and you have to be able to … work with your partner. (Ms. Johnson, THP)

> Teachers must be willing to work together. They need to be open and honest about shortcomings. (Mr. Paul, Seaview Primary [SP])

Less recurring themes were partners' willingness to (a) learn to work together and discuss issues, (b) demonstrate a level of maturity, (c) show understanding, and (d) command the respect of the pupils. When asked whether team teaching improved classroom discipline, Ms. David at Tamarind Heights Primary replied:

> Once both teachers are firm, and they maintain the discipline, it works well. But if there's one that they have no respect for, well, he could spoil the whole group.

The quality of the relationship was enhanced when a member of the team displayed an attribute that the partner found to be exemplary, especially when the attribute related to the emotional and physical well-being of pupils. For example, when speaking about her experiences with a peer, one teacher commented:

> He's a real father figure to the little ones in the Infant [class]. (Ms. Johnson, THP)

Generally, teams reported positive experiences of team teaching. Both members of the team at Seaview Primary said that the experience of team teaching was favourable. At Tamarind Heights Primary, both members of Team A expressed mixed views about their experiences of team teaching. However, both teachers of Team B were extremely satisfied with the team teaching arrangement. At Pine Valley Primary, the team members were, so far, pleased with the teaming initiative.

Benefits of Team Teaching

From the perspective of the teachers, team teaching offered the opportunity for more effective organization and management of instruction. All teachers saw benefits in the division of subjects according to preference and/or expertise:

> I started off with my strengths first…English…social studies and things like that. So she [the other teacher] would have handled the maths and the science…also the composition. So we drew upon each other's strengths…So team teaching worked well, as far as my experience was concerned…Afterwards, after handling my strong areas…and getting into my comfort zone, I started to tread into new areas, maths…. (Mr. Williams, THP)

All teachers indicated that the arrangement gave them more time to prepare classwork and assess pupils' progress, especially with large classes. Thus, most teachers found team teaching to be less stressful and time-consuming. Referring to the way the team was organized, one teacher said:

> We decided to teach different areas so no one teacher would be burdened with all the subjects…. Since we have this arrangement, it is much easier on me. (Ms. Jones, PVP)

Teachers also found that improved classroom discipline was another major benefit of the team teaching arrangement. The combined effort of the team was considered to be extremely useful in monitoring patterns of pupils' behavior and responding appropriately:

> If I find a boy is not performing well, I go across by Mr. Charles and tell him something or I go by Ms. Jones and tell her…And then sometimes when we look back at the pattern on that particular day, that child did not perform, so something wasn't right with that child, probably from home or somewhere else. (Ms. Stewart, PVP)

A significant feature of teams in a mentoring relationship was the opportunity it offered the less experienced teacher to learn by observation. One newly qualified teacher stated that college training:

> …gives theory. Real teaching starts when you leave college. (Mr. Thomas, SP)

He therefore thought it fortunate to be associated with his mentor teacher to get real experience.

In particular, most junior teachers found that they improved their skills in communicating with pupils and parents by observing the senior teachers. At the same time, the experienced teachers sometimes felt that they benefited from exposure to new ideas and methods. Mr. Paul, a mentor, described the experience as "refreshing" since newly qualified teachers can bring current ideas into the classroom.

The majority of teachers expressed a greater sense of enthusiasm and motivation that extended beyond the classroom:

> We excited. I lie down in my bed in the night and I…thinking about next morning. I can't wait to get out there. I'm in school very early now. (Ms. Jones, PVP)

The presence of a partner assisted with the management of large classes of mixed ability and provided support during the delivery of the lesson. Team teaching also facilitated pupils' learning and teacher-pupil interaction. All teachers found that the involvement of two or more teachers in the teaching of a group of pupils allowed individualized instruction as well as remediation:

> While he [the partner] is teaching the maths, I would go around to see what they are doing, and if you find they are not understanding, or they didn't grasp the concept, you pull them out and like when he is doing something else or they are doing their enrichment … you take them and

> you explain again, give them examples and that kind of thing. (Ms. Johnson, THP)
>
> You know exactly, this child did not understand this particular concept, and you know exactly well you have to work to help that child build on that particular concept. (Ms. Jones, PVP)

Ongoing dialogue between teachers enabled tracking of pupils' progress. Ms. Jones further explained that the teaching was so organized that if one of the teachers experienced difficulty in getting children to understand a topic in his/her area, another member of the team would teach the topic again to enhance pupils' learning. This strategy allowed pupils to be exposed to different teaching styles and enhanced their understanding:

> If I find I'm not getting through [to the pupils], I'll say to Mr. Charles, "I want you to go over this with those boys." …[to] get it from a different teacher, you know, and he would do that. So I would just quickly ask him, and tell him where I feel the problem is. (Ms. Stewart, PVP)

When a team member was absent, continuity of instruction was also possible. The allocation of subjects based on expertise and preference allowed more comprehensive coverage of the curriculum since teachers had less work to prepare, and most teachers in the study stated that this had a positive effect on student achievement.

A consistent theme was the opportunity that team teaching presented for pupils' selective interaction with teachers, based on the nature of the pupils' concern as well as the gender and perceived personality of the teacher:

> I think the children enjoy [being taught by a team] especially if it's a mixed team…There are some things they will come by Sir for, and some things they will come [to me] for. (Ms. Johnson, THP)
>
> You see what I look at, sometimes some children relate better to one teacher than to others. So … one teacher teaching this child for let us say two years…could have a negative impact on the children. Whereas, [with two or more teachers] you have opportunities to go to another teacher…You find, one teacher teaching a class, if there's a child in that class that cannot understand that one teacher, that is two years of trouble. (Ms. Jones, PVP)

Challenges of Team Teaching

The effectiveness of team teaching seemed to be constrained by several factors. This arrangement required a great deal of planning and collaboration, which was difficult if partners had no joint free time. In one school, time was allocated in the timetable for planning and collaboration. In the other schools, teachers found time to meet at their convenience. All the teams emphasized the need to do so.

The absence of one partner also affected the functioning of the arrangement. Sometimes one partner was given responsibility for another class or for an activity elsewhere, and one team member was left in charge of a class for extended periods. Negative qualities of members of the team, such as irresponsibility in fulfilling duties, severely affected its functioning:

> That person will begin to feel taken for granted … because it's just two of us, I would take the burden … and conflicts might arise. So it could be abused. It depends on a good relationship; it would have to be a good relationship. (Mr. Williams, THP)

Where one partner acted as a mentor, unwillingness of the junior teacher to take advice was identified as a critical factor. However, one junior teacher felt that the arrangement inhibited the use of skills acquired during training:

> When I first came, the idea [of sharing a class with a mentor] wasn't too pleasing and appealing in the sense that at [training] college, you learn so many different things; up-to-date strategies and when you're fresh out, you want to try things out on your own. You want to see how you could develop yourself. (Mr. Andrews, THP)

Since classes were sometimes combined in a team teaching arrangement, the sheer size of the class, coupled with the mixed ability of pupils, posed a challenge. Some of the problems related to assessment and feedback. Often, strategies adopted to address these concerns resulted in one partner's limited involvement in lessons being taught.

Discussion

In the literature, collaboration between members is generally a defining feature of teaching teams (Brownell & Walther-Thomas, 2002; Salend, Gordon, & Lopez-Vona, 2002). Our study revealed varying levels of collaboration between members of teams: high to moderate levels of collaboration, planning, and discussion as opposed to little or none of these between members. The findings further suggest that effective team teaching was facilitated when the school had a culture of collegiality, collaboration, and a history of harmonious staff relationship. Institutional structures, such as working committees and teacher participation in school events, helped to foster congenial relations among members of staff and make team teaching easier to implement. In the absence of this, the chances for effective teamwork to attain educational and institutional goals may be reduced. This relates directly to the important consideration of culture when innovations are introduced into a school setting. It must be recognized that in organizations such as schools, habits of individualism, reinforced by strong, articulate personalities or teachers who are insecure about their pedagogical skills, may prevent the formation of teams or limit their effectiveness (Brownell & Walther-Thomas, 2002; Katzenbach & Smith, 1994).

Even when the school culture provides a scaffold for an arrangement such as team teaching, an important consideration in organizing teams is the personality of the members. Individual factors such as work ethic, willingness to learn, and interpersonal skills are critical to the effective functioning of teams. Given the demands that teamwork makes on individuals on the team and the need for members to share a common vision, the ability to work collaboratively and to fulfil responsibility are crucial to the success of the team.

One striking feature of the cases studied was that team teaching was mainly introduced to promote the efficient functioning of the institution; the needs of pupils were secondary. There was only one exception to this tendency. In most cases, team teaching provided an opportunity for teacher professional growth and development, more specifically, for the mentoring of new members of staff and for less experienced teachers to learn by observing their seniors. Brennan et al. (1973) make a similar observation. Our study noted, too, that team teaching helped to alleviate

the problem of limited space, which today constrains the operations of many primary schools in Trinidad and Tobago. However, this is a situation that is not limited to Trinidad and Tobago since the issue is also addressed by Pugach and Wesson (1995). With the growing trend towards the establishment of school libraries and computer laboratories, adequate classroom space may become even more limited, especially where physical expansion of the existing plant is not possible.

The nature of team teaching demands time for planning and coordination. Given the fact that the primary school teacher in Trinidad and Tobago usually teaches a class for the entire school day, finding time for collaboration and planning is difficult. The administration must therefore allocate time for this, if team teaching is to function efficiently. The benefits of doing so are well documented in the literature (Brownell & Walther-Thomas, 2002; Craig (1998); Pugach & Wesson (1995); Rink (1998). With respect to the structure of teams in classrooms, the traditional structure of team teaching was predominant in the cases studied: two teachers were assigned to one class. However, a different arrangement emerged in one case that was marked by the creative use of the experience and expertise of team members. This team used flexible scheduling as they rotated classes among three teachers. Similar arrangements were described in Hall (2000) and Miller (1992). These can serve as models for other schools seeking to implement new ways of delivering instruction, while maximizing the strengths and preferences of teachers.

The findings further revealed that team teaching positively influenced teacher performance and teacher satisfaction, and this redounded to the benefit of pupils. Team members in our study found that the presence of two or more teachers in an arrangement with a class increased the likelihood of the needs of all being met. Principals and staff interviewed by Lovell (1967) and students interviewed by Pugach and Wesson (1995) confirmed this benefit for learners. Team teaching also allowed greater opportunity for assessment and feedback during and after lesson delivery. Teachers in our study reported that they were able to employ more formative assessment while team teaching. The tracking of the progress of individual pupils was also easier and, consequently, it was possible to implement measures for remediation or enrichment as required. This is a crucial benefit, which is also highlighted in

McGoogan (2002) and Miller (1992). Apart from pupils' academic development, opportunities to cater to their social needs increased in classes that were taught by teams, since more than one teacher monitored pupils. Slater (1993) identifies this as an opportunity for increased teacher involvement in pupils' personal and academic lives. Thus, while team teaching in the cases studied was mainly introduced to serve the needs of the institution, many benefits accrued to pupils instructed within this arrangement, once teachers functioned as a team in a collaborative relationship.

The potential of team teaching to increase teacher motivation was also apparent. Increased collegiality among teachers, the comfort of teaching subjects of preference, and the early gains in pupils' academic progress contributed to teachers' feelings of well-being and a positive attitude to work. The fact that all teachers in the study expressed satisfaction with teaching fewer subjects indicates that the education system may have to explore alternate ways of organizing instruction at the primary level, where each class teacher is generally expected to teach all subject areas. The opportunity to teach subjects of preference was a primary benefit identified by Miller (1992). Furthermore, when team members have individual classrooms, these can be organized as learning centres for the specific subject areas that they teach.

Some issues currently facing the administration of schools were raised in the study. A major concern for all groups was discipline. In one case, it was a factor in the selection of a team. There was general agreement that team teaching had a positive impact on discipline. Schools may wish to consider this option in their search for solutions to problems of indiscipline. In Slater's (1993) study, team teaching facilitated the implementation of reform programmes for at-risk students and minority groups. However, effective teamwork is extremely important.

Generally, evidence from this study suggests that team teaching provides opportunities for enhanced teaching and learning, if it is used effectively. However, there is the view that such gains can also be obtained from conventional arrangements for instruction. Pugach and Wesson (1995), for example, observe that many benefits of team teaching simply represent good teaching, which can be found in any classroom. Lovell (1967) also cautions that team teaching does not automatically assure

better teaching or better pupil learning. Perhaps the critical factor is not the form of organization in itself, but the extent to which the potential for enhanced teaching and learning is utilized.

Challenges to team teaching were linked to: (a) institutional factors, (b) issues at the level of the class, and (c) individual factors. There is an apparent willingness by administrators to experiment with different ways of organizing delivery of the curriculum, and this has given rise to different teaching arrangements. For example, there is an increasing practice of using teachers who specialize in the teaching of one subject, usually in the area of creative arts, computer studies, agricultural science, and science. This has implications for the future of team teaching as it is currently structured. Additionally, unless administrators view the team as a unit, members will always be used conveniently as relief teachers in other classrooms when teachers are absent. This severely reduces the effectiveness of team teaching. However, this reflects the difficulty that confronts administrators as they attempt to ensure that the needs of the pupils, teachers, and the institution are all met.

At the level of the class, the circumstances under which a class is brought together can sometimes impact on successful team teaching. If some pupils have prior experience with one teacher in the team and view the other teacher as a stranger, they may respond in ways that undermine the effectiveness of one member of the team.

The literature and findings from this study suggest that the personal attributes of members are important to the functioning of a team. If team members are not professional in their approach to teamwork, and they do not engender trust and respect, then it would be difficult for the team to function effectively.

Individuals may respond differently to institutional arrangements such as mentoring. The findings indicated that while two newly qualified teachers welcomed the opportunity to work with a senior teacher, one found that it limited the opportunity to try out new strategies.

Future of Team Teaching

In the context of primary schooling in Trinidad and Tobago, team teaching lies on a continuum of approaches to curriculum delivery. The generalist approach with one class that is assigned to one teacher stands at one extreme. Theoretically, specialized instruction, in which one class is assigned a different teacher for each subject, is at the other extreme. Different combinations of teaching teams are possible between these two extremes. These teams can differ with respect to the way they are organized and how they function. It is also possible to have a class instructed by specialist teachers who do not collaborate and integrate instruction and, therefore, cannot be said to function as a team.

There was general consensus among teachers in the study that team teaching as it has been traditionally organized in local schools was on the decline, even though one school had just introduced team teaching. The decline was linked to the increasing use of forms of subject specialization, which are seen as more effective options in the delivery of the curriculum at the primary level. Teachers felt that they could plan more effectively if their teaching was restricted to those subject areas that they preferred, or for which they held special aptitude or expertise. As educational institutions strive for greater efficiency and for improved student performance, there is likely to be movement towards increased subject specialization, with or without the type of collaboration and planning among teachers necessary for effective team teaching. If there are fewer subjects for a class teacher to teach, there is less justification for having two or more of them assigned to one class, and team teaching as traditionally structured may be less frequently used.

However, it is possible for the two arrangements—team teaching and other arrangements of specialized teaching—to coexist in the same school. At the time of our study, team teaching had been recently introduced into a setting where there were specialist teachers in some subjects who taught at all levels of the school. It is perhaps significant that team members in that school were all experienced and operating within a school culture that was described as “communal” by both the teachers and the principal. The structure and functioning of that team did not follow the traditional pattern, and was initiated by the team members

to improve pupils' learning. In the process, these teachers discovered benefits for themselves.

Apart from the increase in forms of subject specialization, the greatest threat to the continued use of team teaching was thought to be the unwillingness of new teachers to take advice and value the wisdom of more experienced teachers. New teachers, though, vary in the extent to which they do so. Ultimately, the growth or demise of team teaching depends on its ability to meet the needs of the institution and prove itself to be a viable option for curriculum delivery.

Suggestions for Future Research

This study elicited the views of administrators and teachers; those of the pupils were not included. Future research can therefore investigate pupils' perspectives of team teaching. This can provide useful insights into pupils' experience of being taught by two or more teachers, and can assist educators in evaluating team teaching arrangements. In addition, teachers often expressed the belief that team teaching positively impacted on pupils' learning. This warrants investigation. Improved student achievement may well be a consequence of team teaching arrangements; however, no documented evidence was offered to support this assertion. Also, the use of different forms of subject specialization in primary schools requires close examination. At present, schools appear to have a great deal of autonomy in the use of specialist teachers. The Ministry of Education perhaps needs to develop a national policy/framework on this issue and, as an initial step, monitor and evaluate the use of specialist teachers in schools throughout the primary system.

With specific reference to team teaching, future research can use classroom observation to investigate the dynamics of interaction between members of a team and between the team and the class. The results of further investigation would undoubtedly enhance our understanding of teaching and learning processes at the primary level of Trinidad and Tobago's education system.

References

Brennan, K., Carter, D. C., Jenkins, E. W., Lawson, K., Leigh, R., & Scott, P. J. (1973). *The teaching of science to pupils of low educational attainment.* Leeds, England: Centre for Studies in Science Education, University of Leeds.

Brownell, M. T., & Walther-Thomas, C. (2002). An interview with Dr. Marilyn Friend. *Intervention in School and Clinic, 37*(4), 223–228.

Clifford, P., & Friesen, S. L. (1993). A curious plan: Managing on the twelfth. *Harvard Educational Review, 63*(3), 339–358.

Craig, C. J. (1998). The influence of context on one teacher's interpretive knowledge of team teaching. *Teaching and Teacher Education, 14*(4), 371–383.

Evans, H. (1997). Making the transition from college to classroom: What knowledge do teachers use and why? *Caribbean Journal of Education, 19*(1), 73–87.

Hall, J. (2000). Initial teacher education: Specialist and generalists – A review of the literature. (Retrieved December 30, 2004, from the Scottish Council for Research in Education Website: http://www.scre.ac.uk/scot-research/hallinitial/index.html

Katzenbach, J. R., & Smith, D. K. (1994). *The wisdom of teams: Creating the high-performance organization*. New York: HarperBusiness.

Lovell, K. (1967). *Team teaching.* Leeds, England: Institute of Education, University of Leeds.

McGoogan, G. (2002). The bear den: An elementary teaching team. *Educational Leadership, 59*(5), 30–32.

Merriam, S. B. (1998). *Qualitative research and case study applications in education* (rev. ed.). San Francisco, CA: Jossey-Bass.

Miller, L. D. (1992) Preparing elementary mathematics-science teaching specialists. *Arithmetic Teacher, 40*(4), *228–231.*

Pugach, M. C., & Wesson, C. L. (1995). Teachers' and students' views of team teaching of general education and learning-disabled students in two fifth-grade classes. *Elementary School Journal, 95*(3), 279–295.

Rink, J. (1998). *Teaching physical education for learning (3rd ed.).* Boston, MA: WCB/McGraw-Hill.

Salend, S. J., Gordon, J., & Lopez-Vona, K. (2002). Evaluating cooperating teaching teams. *Intervention in School and Clinic, 37*(4), 196–200.

Slater, J. K. (1993). Using regional school networks to orchestrate reform in California middle grades. *Elementary School Journal, 93*(5), 481–493.

Thornton, M. (Ed.). (1998). *Subject specialists - Primary schools*. (UCET Occasional Papers, 10.). London: Universities Council for the Education of Teachers.

Williams, A., Prestage, S., & Bedward, J. (2001). Individualism to collaboration: The significance of teacher culture to the induction of newly qualified teachers. *Journal of Education for Teaching, 27*(3), 253–267.

DRAGGING ELEVEN-PLUS MEASUREMENT PRACTICE INTO THE FOURTH QUADRANT

The Trinidad and Tobago SEA as a Gendered Sieve[3]

Jerome De Lisle

> *"Group impact and educational impact are the quintessence of social consequence. They drag measurement practice into that fourth quadrant"*. (Willingham, 2002, p.196)

This paper expands upon concerns expressed earlier in De Lisle & Smith (2004) about the relationship between Eleven-Plus test design and patterns of gendered achievement in Trinidad and Tobago. Following Willingham's (1999), evaluation protocol, it includes (1) a critical analysis of gender fairness issues, (2) an empirical evaluation of gendered impact, and (3) a consideration of proposals for resolving gender fairness issues. Datasets used in the analysis are from the 2001–2003 Secondary Entrance Assessment (SEA) and the 1998–2000 Common Entrance Examination (CEE). The results confirmed that the gender gap was significantly greater for the SEA compared with the CEE, and that both males and females were disadvantaged in different ways by the placement system. The results also confirmed the existence of medium-sized gender differentials across urban-rural educational districts, literacy constructs, and high-low ability groups. A recent proposal to change the way the composite score is calculated did little to reduce the overall female advantage. Moreover, misclassification rates for the current remediation cutscore set at 30% were relatively high. These fairness issues are not easily resolved, but suggest the need for evidence-based test designs, test validation studies, and a re-examination of the need for selection.

[3] For the period of this work, which extended over two years, I acknowledge the assistance of the Division of Educational Research and Evaluation. While I am fully responsible for all statements made, feedback from this department helped to improve the study.

Introduction

High-stakes certifying and selection examinations have always had an important role in the education-oriented societies of the English-speaking Caribbean (London, 1989; Payne & Barker, 1986). Taking pride of place is the Eleven-Plus examination, which is taken by every student at the end of the primary school cycle (Heyneman, 1987; Stanley-Marcano & Alexander, 1998). By governing access to secondary schools of different quality, performance in this examination becomes critical to later success at 16 and 18+ and, ultimately, to opportunities for upward social mobility (Jennings, 2001). The use of selection examinations as early as age 11 may be especially critical in situations where opportunities for educational advancement and social mobility are very limited (MacKenzie, 1989). On the one hand, it is possible that such early selection will enhance system efficiency by ensuring that only those students who are best capable can gain access to further opportunity (Durbrow, Schaefer, & Jimerson, 2002; Heyneman, 1987). On the other hand, sorting and weeding students using high-stakes tests in an elitist school system can achieve at best only pseudo-meritocracy and may limit opportunities for widening access (Gould, 1996; Hickling-Hudson, 2002; Sacks, 2001a; Siegel, 2004).

Examinations used for selection purposes shape education systems and influence institutional structures and outcomes (Drenth, van de Flier, & Omari, 1983; Greaney & Kellaghan, 1995; Kellaghan & Greaney, 1992). This is because the examination acts as a gatekeeper (or gateway), creating differential test-taker outcomes, and governing access to further educational opportunity (Sacks, 2001b; Zwick, 2002). While many school systems have primary gatekeeping mechanisms at 16 to 18+, in the Caribbean, the role of the Eleven-Plus examination as an early gatekeeper has been inherited from the colonial era (Jules, 1994; Payne & Barker, 1986). Nevertheless, its continued retention is only possible in a society completely at ease with concentrating "the most disadvantaged children in the least popular [and worst resourced] schools" (Edwards & Tomlinson, 2002, p. 30). The Common Entrance Examination (CEE) at the end of the primary school cycle acted as gatekeeper by either barring or allowing entry to secondary school. Changing the purpose and design of the examination would not substantially alter this purpose; however, with the implementation of universal secondary education, the emphasis might

become sorting students towards different school types, as Hickling-Hudson (2002, p. 572) has noted in the case of Jamaica.

I would argue that a more useful analogy for a selection examination in this role is a "sieve," which emphasizes the concept of separation rather than debarment (Handy, 1989). The purpose of a simple sieve is to separate constituents; a concept expressed in popular terms, such as "*separating the wheat from the chaff,*" which implies that the constituents have dissimilar value. Using this alternative analogy, initially, the CEE would have acted as both gatekeeper and sieve by (1) barring entry into secondary school for some students, and (2) separating or allocating the remaining students to different school types. On the institution of universal secondary education in 2001, however, only the sieve function was retained in the newly designed Eleven-Plus examination—the Secondary Entrance Assessment (SEA) (Trinidad and Tobago. Task Force for the Removal of the Common Entrance Examination [T&T. Task Force], 1998). Although all students could now access secondary schooling, Eleven-Plus test scores were still used to decide which students received particular opportunities. Arguably, a new layer of separation was added, with students below a cutscore of 30% automatically allocated to remedial classes, thereby further ensuring efficient "separation of wheat from the chaff."

Whether or not opportunities have changed is an issue for education sociologists. Test evaluators are instead concerned with both adequacy of the sieve function and consequences of test use (Kane, 2001). Social consequences are also important to the test evaluator because the technical aspect of testing can never be separated from the social and political functions (Messick, 1988). In this regard, Willingham (1999) proposed a useful three-part protocol for evaluating test fairness, which involves (1) identifying fairness issues, (2) assessing impact, and (3) resolving fairness status. The first stage requires an analysis of the full range of issues associated with differential performance. In this paper, both the sieve analogy and an argument-based approach to test validation are used to identify and analyze critical gender fairness concerns. In the second part of the paper, empirical evidence for gendered impact is gathered from Eleven-Plus data; and in the final section, an attempt is made to resolve the fairness issues.

Of course, dual gatekeeping and sieve roles operate for selection examinations elsewhere. For example, in the US, SAT scores are critical determinants of access to institutions of higher learning (Sacks, 2001b). Even though high school grade point averages may be better predictors of undergraduate performance, a low SAT score will either bar a student or result in placement within a less prestigious institution (Camara & Kimmel, 2005; Zwick, 2002). Still, we might expect the fallibility of an Eleven-Plus gatekeeping system to be significantly higher than that at 18+. Indeed, most gatekeeping systems are inherently fallible because inferences and decisions are usually based upon a single composite score. Limitations can also arise if there is no theoretical rationale for the construct that the test score is supposed to measure or insufficient evidence to support inferences that the test measures the construct well (Gardner & Cowan, 2005; Sacks, 2001b).

Fairness and Precision as Examination Myth

There are important hidden benefits to using an examination for making high-stakes decisions on educational opportunity within small states. For example, only an examination will ensure the successful creation of the myth that selection is fair and precise, even when this is conceptually or technically impossible (Gardner & Cowan, 2005). Indeed, locally, the myths of precision and fairness have so successfully been promulgated that Eleven-Plus scores are often considered sacrosanct. Moreover, such myths can be augmented and sustained indefinitely through a lack of transparency. Consequently, although selection examinations have retained such a critical role within Caribbean education systems, there has been little consideration of issues related to test fairness. Indeed, internationally, there are still only a handful of studies of the Eleven-Plus in the Caribbean region (Durbrow et al., 2002). In the absence of these studies, those considered as illuminators (journalists, assessment evaluators, and other academics who bring light to the issue of test quality and fairness) are well advised not to express strong faith in the validity of the selection process.

Thus the real value of the precision and fairness myths may lie in the negation or suppression of dissent or challenge by stakeholders, including illuminators (Kane, 2002; Linde, 2003). Perhaps, then, it is this astounding absence of illumination coupled with the lack of transparency

that is central to the continued legitimacy of the Eleven-Plus selection system (Gardner & Cowan, 2005; Jules, 1994). London (1989), for example, has argued that considerable sponsorship exists in the Eleven-Plus selection process of Trinidad and Tobago, implying that the system is not as meritocratic as many believe. However, while London focused primarily upon the influence of paid lessons, there are also differences in the quality and extent of test preparation and variations in opportunities to learn across educational districts and schools (Haladyna & Downing, 2004). Indeed, local schools vary greatly in quality, especially across the urban-rural divide (De Lisle, Smith, & Jules, 2005). Thankfully, however, pervasive and institutionalized cheating is not yet a feature of Eleven-Plus selection in Trinidad and Tobago, although it might well become so if high-stakes school-based assessments were to be implemented (Overdorf, 2004). So, while there are still some who might seek to sneak over the wall with the help of the Concordat[1], this paper is concerned only with the gender fairness of the actual Eleven-Plus examination. Put another way, the question is: *To what extent is the eleven-plus gate equally open to boys and girls?*

Designing (and Redesigning) the SEA

It would be a grave error to assume that any test is automatically gender fair and, further, that gender fairness is independent of test design. The belief that "test design does not matter but scores from the test do" remains a significant problem in current analyses of gendered achievement in the Caribbean. Many empirical studies of gendered achievement assume that assessment is a completely neutral procedure and pay little attention to the changing nature of assessments (Gipps & Murphy, 1994; Kutnick, 2000; Layne & Kutnick, 2001). However, changes in assessment design must be considered when evaluating trends across subjects or time (Bailey, 2000; Rampersad, 1999). For example, analyses of 'O' Level results might consider changes in the weighting and nature of school-based assessments because these changes will alter the size of the gender gap (Elwood, 1999, 2005).

The assessment cycle includes the processes of test design, development, scoring, and administration, with the first two steps concerned with the knowledge and skills measured by the examination and the instruments used for measurement. It is at these early stages that critical choices must

be made about the purpose of the test, what constructs are required to effect those purposes, choice of assessment format, particular item types used, scoring methods, and statistical characteristics of items (Willingham & Cole, 1997). It is also significant that these test design and development decisions are interwoven into subsequent stages, thereby enhancing the impact of these choices. There is overwhelming evidence that a change in construct, format, or test blueprint will significantly alter the validity of a test for different subgroups, especially gender (Gipps & Murphy, 1994; Ryan & DeMark, 2002; Willingham & Cole, 1997). Since different choices will result in vastly different outcomes for males and females, test design or development committees must reflect on decisions and make public the predicted consequences of their choices, well before implementing any proposed changes (Willingham & Cole).

Acknowledging that a relationship exists between test design and gendered achievement is not the same as arguing that the increasing female advantage is the result of assessment schemes becoming "feminized" (Elwood, 2005). This argument implies some unholy conspiracy among test developers. Rather, the focus is on the need to develop explicit rules for a fair test. For example, neither construct nor authenticity can be the sole concerns in test design. Indeed, as Gipps and Murphy (1994) admitted, even test designers' conception of authenticity in a subject area is open to challenge. This has certainly proved true in the case of physics, where in the past masculine definitions of acceptable content have led to a relative disadvantage for females (Hildebrand, 1996). Indeed, Willingham (2002) has argued that when alternative test designs are available and one group of examinees is likely to be severely disadvantaged on a specific design, the most valid test is the fairest.

Considering the gender fairness of a chosen test design is the essence of analysing the Eleven-Plus as a gendered sieve. Such a focus is important because the format and structure of this examination has been significantly altered, and further changes have been proposed in an attempt to redress gender inequalities in performance (De Lisle & Smith, 2004; T&T. Division of Educational Research and Evaluation [DERE], 2004; T&T. Task Force, 1998). In 1998, the committee charged with the task of removing the CEE explored the issue of the gender fairness in selection, and concluded that there was grave injustice in any system that attempted to place an equal number of boys and girls. They noted that

"the present situation, if one judges on the basis of performance, is one in which girls who deserve places on the basis of merit are deprived selection in favour of boys who did not perform well" (T&T. Task Force, p. 41). Of course, this statement assumes that this or any other examination can objectively measure "merit" (whatever merit might be), and that scores assigned to the approximately 1,000 disadvantaged girls represent an accurate estimation of "capability to succeed in the secondary school," which seems to be the test construct in question.

The SEA is very different to the CEE in the fundamental design areas of construct, format, and weighting of the composite score (T&T. Task Force, 1998). Whereas the CEE tested knowledge of five subject areas—Mathematics, Language Arts, Creative Writing, Social Studies, and Science—the SEA omits the latter two subjects. The Task Force reasoned that although it should remain an achievement test, the instrument could be redesigned to better measure students' reasoning ability and verbal skills. Thus, compared with the CEE, the SEA composite score is heavily weighted towards linguistic-verbal competence, primarily ways of knowing language taught by the school (Myhill, 2005). From the literature, it is apparent that this choice might put some groups at a great advantage, most likely females and students of higher socio-economic status. Additionally, the test development decision to use constructed response items only would give these groups a further advantage. Thus, major choices in test design and development of the SEA are likely to result in females doing relatively better in the SEA compared with the CEE.

The adequacy of sampling may not have been an important consideration during test development, since the SEA had just over 100 items compared to 190 items in the CEE. The 2003–2004 test specifications for the SEA provide details of the sampling for that period (T&T. DERE, 2002). Of the 50 Mathematics items, 20 measure number, 18 measure measurement and money, 8 measure geometry, and 4 measure statistics. Thus, only two of three content strands are adequately sampled. Similarly, in Language Arts, 22 of the 50 items measure grammar, but only 8 measure vocabulary/spelling, 5 measure punctuation, 10 measure comprehension, and 5 measure graphic representation. Thus, only one Language Arts content strand is adequately sampled. While there is limited information on the scoring of questions, for the Creative Writing component, a holistic

rubric and two markers are used. Each rater gives a score between 0 and 6 using the rubric, which is then combined for a final score, with a maximum of 12.

Associated with the implementation of the 2005–2006 SEA was a redesign proposal intended to reduce the effect of the essay on the composite score (T&T. DERE, 2004). This proposal was based on the belief that the weighting of 1:1:1 for the SEA components gave too much weight to the essay. Mistakenly, then, this proposal viewed the essay as a single item, although in fact this is an extended performance task scored with a six-point rubric. Interestingly, the proposal failed to deal with another more significant scoring problem—low discrimination caused by a sketchy holistic rubric. The problem of inadequate discrimination (most students get between 6 and 12 and differences are small) was first considered by the Task Force, which believed that the best solution would be to increase the essay score to 48. In the 2005–2006 redesign proposals, the decision was made to use a 10-point rubric (maximum score of 20 by two raters). At the same time, the impact of the essay will be reduced by weighting it 2:3 against the Language Arts section. Thus, the overall weighting of the components for the composite score would become 5:3:2.

Although these are significant changes, they are not referenced to the critical issues of validity, reliability, or fairness (T&T. DERE, 2004; T&T. Task Force, 1998). More importantly, in the context of this paper, Willingham and Cole (1997) and Willingham (2002) emphasized that evidence for the impact of alternative test designs must be obtained and compared *before* implementation. Comparative studies using measures of impact such as effect sizes, standard deviation ratios, and female to male ratios on valued outcomes will allow us to make judgements on the fairness of alternative test designs. The gendered impact of the proposed redesign is a significant issue because changes in gender differentials across educational districts might also influence placement opportunities. Therefore, rather than reducing the advantage to a particular group, such changes, when implemented, might result in unpredictable and spotty patterns of winners and losers among both males and females.

Validity as the Context for Evaluating Test Fairness

From the standpoint of an evidence-based assessment design, Mislevy, Steinberg, and Almond (2003) reminded us that measures of student learning are essentially machines "for reasoning about what students know, can do, or have accomplished, based on a handful of things they say, do, or make in particular settings" (p. 4). More importantly, assessments are "embedded in a cultural setting and address social purposes both stated and implicit" and "communicate values, standards, and expectations" (p. 4). This understanding of assessment is critical to making appropriate choices when designing examinations to act as efficient sieves. Firstly, as a machine, an assessment can be put together in several alternative ways, with some designs more efficient for specific purposes. Secondly, all assessments are limited by the representativeness of the sample and, therefore, the adequacy of the sample of items or tasks is a major design issue. Thirdly, the values, standards, and expectations embodied within the social purposes of an assessment cannot be divorced from the technical aspects. Thus, educational and social purposes are not opposing facets of some choice wheel but instead must be balanced. More importantly, the values, standards, and expectations embodied in the assessment must be explicitly stated and considered both in planning and in evaluation.

To put fairness concerns into context, we must focus, then, on the current meaning of test validity, which is the most important criterion for evaluating a test (Gersten & Baker, 2002). The central focus of current conceptions of validity is on test score use and interpretation (Kane, 2001). In the modern era, Messick put forward two important concepts that have influenced the process of obtaining validity evidence. Firstly, he proposed a unitary concept that envisions construct validity as the whole. Secondly, he proposed a four-facet model of evidence, consequences, test interpretation, and test use, in which the fourth quadrant focuses upon the social consequences of test score use (Messick, 1988, 1989). It is this fourth quadrant that requires evaluators to address questions of "unintended side effects" and impact when evaluating tests. This is an important goal because test evaluators are really public scientists working for the public good (Ryan & DeMark, 2002). Thus, evaluators must avoid either a conformationist bias or a tendency to single-mindedly defend

current testing programmes (Haertel, 1999), and instead "speak to a diverse and potentially critical audience" (Cronbach, 1988, p. 4).

Validating test score use and interpretation is a major activity for any test agency involved in high-stakes testing. Messick's (1989) definition of validity is the foundation of the validation process, as defined in the 1999 *Standards on Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). Validation, however, remains a nebulous process because new ideas about validity are not easily applied. Indeed, there are currently no guidelines for determining the relevance or types of validity evidence that should be collected (Kane, 2004a). Recently, in response, some authors have put forward an argument-based approach in which validation becomes the process of constructing and evaluating arguments for and against proposed test use and interpretation (Haertel, 1999). A validity argument will provide an overall evaluation of the plausibility of proposed interpretations and use of test scores. According to Kane (2002, 2004a), the first step is the construction of *an interpretive argument*, which are statements laying out the assumptions and inferences about the test score and its proposed uses. The act of constructing an interpretive argument is critical because test evaluators often set out to judge a test without establishing first what these proposed interpretations might be (Kane, 2004b). Using this approach, test validation involves "obtaining and weighing evidence to support or refute the interpretive argument" (Haertel & Herman, 2005, p. 2).

Since the interpretive argument provides an explicit statement of the proposed interpretations or test use, it demands further evaluation, generalization, and extrapolation. To aid the construction of an interpretive statement, Kane (2002) distinguished between *descriptive* and *decision-based* interpretations. The latter is important when judging test use and consequences and must be justified by gathering a variety of evidence from different sources (Ryan & DeMark, 2002). Kane also extended the framework to include *semantic inferences*, focused upon what the test score means, and *policy inferences*, which involve the adoption of decision rules. Using Kane's (2002) framework, four semantic (1-4) and two policy inferences (5-6) related to SEA as a gendered sieve are listed below:

1. Males and females perform similarly on the tasks outlined in the test blueprint.
2. Performance on the range of test objectives is similar for males and females.
3. Achievement on the SEA test objectives provides a comparable indication of underlying ability for males and females.
4. Achievement on the SEA test objectives provides comparable secondary school placement opportunities for males and females.
5. The implementation of the SEA will lead to increased gender equity.
6. The adjustment in weighting will lead to increased gender parity.

These inferences suggest that a fundamental assumption of the SEA is that males and females should perform similarly on the same set of knowledge and skills, and gender equity and parity remain important policy goals.

The SEA as a Gendered Sieve

Although the concepts of fairness and validity are analogous, fairness issues must be explicitly addressed when evaluating tests (Stobart, 2005; Willingham 2002). Test fairness may apply to either individuals or groups. For the individual, test fairness is directly equivalent to validity whereas group fairness is comparable validity for different sub-groups (Willingham, 1999). When considering the SEA as a *gendered sieve*, the focus is on group fairness, meaning that the SEA might have different predictive validities for males and females. As with the SAT, this implies that the SEA functions differently for males and females and is a better predictor of outcomes for one group (Zwick, 2002; Zwick & Schlemer, 2004). Technically, if construct under-representation and construct irrelevant variance were absent, groups might perform similarly. In practice, however, there are always differences between sub-groups and therefore policymakers must consider *how large these differences are* and *how existing educational policy is altered*, especially for high-stakes tests. When analysing group impact, gender, ethnicity, social class, and geographic location are common focal points because social equity and justice demands fair distribution of opportunities across these dimensions.

It is likely that the changes of construct and format from the CEE to the SEA altered the magnitude of the gender gap and created differential placement opportunities for males and females (Murphy, 2000). Any

judgement of fairness in this context must be tempered by the values and expectations implicitly or explicitly espoused about the test and its purpose, as contained in the semantic and policy inferences. If a society strives for gender equity and excellence, it is hypocritical to choose an Eleven-Plus test design that accentuates gender differences. Likewise, it would be foolhardy to propose or implement changes that might further magnify gendered impact. It is clear, then, that in the Caribbean setting, test designers must always consider gendered impact when making policy decisions (De Lisle & Smith, 2004).

Fairness and the Complexities of Selection at Eleven-Plus

While the test is central to the selection process, there are other important elements to consider when analysing fairness (Messick, 1989). For example, both achievement scores and placement opportunities should be evaluated. Interestingly, the Task Force (1998) believed that any disadvantage to females would be fully resolved on the institution of universal secondary education; however, it did not consider the vagaries of the placement system. It might be that the placement system remains as a unique and separate source of inequity (De Lisle & Smith, 2004; Jules, 1994). This inequity is difficult to evaluate using solely a quantitative approach because different individuals will place different values on similar outcomes. For example, one parent might be pleased with a child passing for a less prestigious school, while another parent, whose child had the same composite score, becomes distraught over the outcome because they expected placement in a more prestigious institution (Gersten & Baker, 2002). However, there are some issues that can be answered through quantitative data analysis, including: *To what extent did the change in Eleven-Plus test design alter placement opportunities for males and females?*

Placement opportunities are partly dependent upon the composite score, choices made by parents, and the availability of schools in different geographical regions (Jules, 1994; T&T. Task Force, 1998). For example, if a parent were to choose only prestigious schools, which are also chosen by parents of other higher-achieving students, then the likelihood of that student being placed is significantly reduced. In turn, parental choice is itself determined by a myriad of factors such as the school's perceived prestige, accessibility, and geographical location (Bagley, Woods, &

Glatter, 2001; Ball, Bowe, & Gewirtz, 1995; Parsons, Chalkley, & Jones, 2000). For example, the current distribution of schools will result in fewer realistic opportunities for most rural students, who are located far away from the urban centres, where most highly prestigious schools are located.

While the perceived prestige of a school is primarily based on academic performance, parents will consider other factors when choosing schools, such as whether the school is single-sex or coeducational, religious affiliation, and sixth form availability (Hospedales, 1982). The 2003 data show that parents selected very different schools for choices 1 to 3. In fact, there are 29 different schools in the 36 possible places, suggesting that different geographic communities consider different schools as viable first, second, or third choice opportunities. Indeed, only one school, Queen's Royal College, was chosen in two choice categories. Additionally, while there are similar numbers of single-sex boy and girl schools in choices 1 and 3, choice 3 is dominated by mixed schools. This pattern might influence placement opportunities outside the first two choices, as would a strategy of parents choosing less prestigious schools as their first choice.

The situation becomes even more complex when we consider that the more prestigious the institution, the greater the number of parents selecting it for first choice. This would significantly decrease opportunities for access by students with lower scores. Table 1 provides data on the students choosing the 12 most prestigious schools and those who actually obtain their preferred first choice placement. While the interquartile range suggests wide variation, the mean scores indicate that most parents choose accurately. Nevertheless, considering that over 700 students chose each of the first four ranked schools, with only about 100 students gaining access to each (including the 20%), the chances of obtaining a first choice placement to a prestigious institution seems rather difficult. The difficulty is illustrated in the case of the top-ranked school, St. Augustine Girls, where the median score of those choosing is 677, just below the minimum entry score of 682. This is likely the score of a student gaining access through the 20% Concordat provision. At the same time, the minimum score of 628 for Naparima, the second ranked school, suggests that the Concordat rule may be applied flexibly and differently across like institutions.

Table 1. ***On the Chances of Obtaining the Students' First Choice*****:**
Characteristics of Students Choosing & Obtaining Placement in the Top 12 Ranked Schools, as Measured by Performance in the 2003 SEA

Rank	Name of School (Gender Composition)	Characteristics of Students											
		Choosing School as First Choice							Obtaining School of Choice				
		No.	Mean	*SD*	Median	IQ range	Max	Min	No.	Mean	*SD*	Max	Min
1 (077)	St. Augustine Girls (G)	869	663.22	56.90	677	73.00	735	426	108	723.11	11.246	735	682
2 (053)	Naparima Girls (G)	782	674.22	51.60	690	66.25	738	460	117	722.86	13.336	738	628
3 (069)	Queen's Royal College (B)	729	620.10	66.50	631	90.50	731	359	107	707.59	9.494	731	696
4 (109)	St. Stephen's College (B)	707	610.55	67.44	619	94.00	728	377	120	704.60	17.709	728	616
5 (061)	Presentation College, Chag. (B)	661	636.73	75.39	654	104.50	739	379	83	722.58	13.983	739	654
6 (029)	Hillview College (B)	641	644.98	60.66	657	82.50	734	343	109	714.06	18.634	734	621
7 (037)	Holy Faith Convent (G)	602	633.11	67.84	645	95.00	734	335	108	713.23	17.325	734	621
8 (113)	North Eastern College (M)	577	550.48	67.31	556	103.00	704	332	148	667.57	11.689	704	650
9 (009)	Bishop Anstey High (G)	561	633.45	62.38	646	79.50	735	365	106	706.83	16.149	735	646
10 (065)	Presentation College, San'do (B)	554	657.01	56.65	668	71.00	737	347	111	718.32	10.937	737	685
11 (049)	Naparima College (B)	527	661.14	55.99	672	68.00	736	369	103	715.29	22.185	736	606
12 (097)	St. Joseph Convent, St. J. (G)	522	656.22	51.84	667	69.00	733	279	114	712.86	17.495	733	629

KEY: B — Boys G — Girls M — Mixed

As expected, the mean scores of those choosing top-ranked schools are lower for males. Therefore, although females may score higher, access to the prestigious single-sex schools may be restricted. Reasonably, any change in the number of single-sex schools for males or females will also significantly alter placement opportunities. Therefore, the patterns reported by Jules (1994) may have changed significantly with the secondary school building programme initiated by the Secondary Education Modernization Programme (SEMP). These complexities will result in inequity, especially in instances where some parents do not have a full understanding of the placement process or make poor choices.

Lessons on Fairness from the Geography of Gendered Inequity

One of the complexities that requires special attention is the differences across geographic regions. In some of the work on gendered achievement, it is often assumed that gender differences are homogenous and differences in geographical location are not important. Even when there is an understanding that only some males are underachieving, little thought is given as to where these males might be located. One common view, for example, is that laddish behaviours and anti-school masculinities are more common in urban areas influencing gender differences in achievement (Archer & Yamashita, 2003; Jackson, 2003). It does seem appealing to argue, even without evidence, that the laddish behaviours of males in some urban areas are the prime determinant of male underachievement. However, while this argument is instinctively attractive, other possibilities exist. For example, it might be that some girls are also significantly underachieving in these urban centres (DeBlase, 2003). Another possibility is that some rural males and females are especially at risk because of poor quality teaching-learning, community expectations, and socialization practices (Chevannes, 2001; De Lisle, Smith, & Jules, 2005). Therefore, determining variations in gendered achievement patterns across geographical locations is important in seeking explanations for why some males and females underachieve.

That large differences exist across education districts is neither a new nor startling revelation. Most recently, data from the 2004 national tests results confirmed that the gender gap is indeed larger in the South-Eastern and North-Eastern educational districts and in Tobago (De Lisle, Smith, & Jules, 2005). Lowered performance in these regions is probably

a consequence of multiple factors, including the ineffectiveness of schools, lack of resources including trained teachers, and the quality of teaching-learning. The geography of gender inequity should considerably lessen our faith in the fairness of selection examinations, especially the Eleven-Plus yardstick. For example, if the gender gap varies so significantly across regions, just how meritocratic can the selection process be? Is it really achievement being measured or just opportunities to learn? In such a circumstance, how efficient is the sieve at separating students of different ability or gender for remediation?

Principles of Fairness

A biased sieve is unfair because it allocates "too many" of a favoured group to valued outcomes and creates "underachievement" in the other. With hindsight, then, it is a great paradox for any society to construct general principles of fairness without first re-evaluating its use of Eleven-Plus selection (Schrag, 2004). Again, implicit in such an approach is the belief that examinations are automatically a fair selection tool. However, while no examination is perfectly fair, all assessments can be made fairer (Stobart, 2005). Indeed, Gipps and Murphy (1994) urged test designers to work continuously towards the goal of fairer assessment by paying greater attention to purposes, nature, administration, and scoring. Moreover, Willingham and Cole (1997) provided clear prescriptions for choosing different test design options. More specifically, both Chilisia (2000) and Stobart considered the issue of fairness within multicultural societies. Stobart, for example, emphasized that fairness and equity are issues of judgement and not just a matter of equal numbers. Chilisia noted that in multicultural systems the content and format of an examination raises the issue of whose knowledge is tested, and whether that knowledge is really of such great value. Thus, implicit in this issue is whether test designers are privileging certain types of knowledge in order to maintain a dominant culture. It is clear that large-scale assessment systems in multicultural societies such as Trinidad and Tobago must work towards the fairest possible test design fairness for gender and other grouping factors.

It is interesting to note that principles of test fairness have been published, including the *Code for Fair Testing Practices in Education*, which complements similar guidelines in the fields of licensing and credentialing (Council on Licensure, Enforcement, and Regulation &

National Organization for Competency Assurance, 1993; Joint Committee on Testing Practices, 2004; Zieky, 2002). These codes emphasize the different elements that contribute to fairness at various stages of the assessment cycle. The question is: *How does the SEA stand up to the principles in these codes*? Analysing the SEA on the 22 criteria proposed by the Joint Committee on Test Practices, most of the current SEA weaknesses centre on the failure to provide evidence for the current interpretation of test scores. If an examination is so important, then surely there should be greater efforts at collecting validation evidence, including studies of impact and fairness (Benson, 1998).

In Search of Evidence

Thus, the key gender fairness issues in Eleven-Plus testing are test design, variations in gendered achievement patterns, magnitude of gender differences, and differences in placement opportunities. This study will primarily consider the gendered impact of different Eleven-Plus test designs, including proposed changes to weighting the SEA composite score and use of the 30% remediation cutscore. Variations in gendered achievement patterns and differences in placement opportunities across regions are also analysed. Since there is no standard for determining the magnitude or impact of a gender difference, one approach might be to compare gender differences across the different design options and regions. In quantifying these differences, an effect size measure is essential to judging practical significance (Cohen, 1988; Schagen & Elliot, 2004; Thompson, 2002; Willingham and Cole, 1997). The magnitude (effect size) of the gender gap would allow us to judge the impact of the SEA as a gendered sieve. Following Willingham and Cole and Willingham (1999), standard deviation ratios and male to female outcome ratios were also used to measure adverse impact.

The five research questions guiding the study are:

1. How do gender differentials compare across the 1998–2000 CEE and 2001–2004 SEA administrations?
2. What was the pattern of gendered achievement and placement opportunities in the 2003 SEA?
3. How do gendered achievement patterns vary across construct, ability group, and regions in the 2003 SEA?

4. What is the gendered impact of changing the rules for computing the composite total score, based on the 2003 SEA data?

5. What is the accuracy and gendered impact of using an arbitrary 30% cutscore to select students for remediation, based on the 2003 SEA data?

Methods and Procedures

The main database was the 2003 SEA results for all 20,671 students taking the examination in that year. This database included raw and standard scores for each component and the composite score. The database was coded only for district, religion, class group, school, gender, age, number of attempts, and six parental choices. The Division of Educational Research and Evaluation (DERE) also provided selected data from the 1998–2000 CEE and the 2001 and 2002 SEA. In Microsoft EXCEL, new composite scores and rankings were created using the proposals put forward for the weighting and scoring scheme in the examination (T&T. DERE, 2004). To obtain the new weighting scheme of 5:3:2 for the Mathematics, Language Arts, and Creative Writing components, weights were applied directly to the standard score. This new standard score was then prorated to ensure equivalency, assuming an original standard score of 600±30. Alternative methods of calculating the standard score were also attempted, including (1) applying the weights directly to the raw scores, and (2) combining the raw scores of English and Creative Writing in a ratio of 3:2 and then obtaining a combined standard score.

In this study, both *r* (measures of association) and *d* (standardized mean differences) families of effect sizes were used to measure gender impact (Kline, 2004; Walker, 2004; Willingham & Cole, 1997). The *d* group effect size was the Cohen's *d* and the *r*-group effect size was eta-squared. The formula used to calculate the Cohen's *d* was $d = M_1 - M_2 / \square_{pooled}$, where M_1 is the mean for males, M_2 the mean for females and $\square_{pooled}$ the pooled standard deviation. Eta-squared is the proportion of the total variance attributed to an effect and is generated directly from the output of SPSS, version 12.0. It is formally defined as the ratio of the effect variance (SS_{effect}) to the total variance (SS_{total}). Two rubrics were used for

making a qualitative decision on the size of the difference. For Cohen's *d*, the benchmarks used for assessing the size of the impact are 0.2, 0.5, and 0.8 and for eta-squared 0.1, 0.3, and 0.5. Odds and risk ratios were also used to measure relative impact in group distribution. The odds ratio was defined as the ratio of odds for an event in one group divided by the odds in another group. The odds ratio may be viewed as an effect size measure for categorical variables (Fleiss, 1994; Kline 2004). While there were no available benchmarks for making a qualitative inference on magnitude, both the odds and risk ratios allow assessment of the relative proportion assigned to each group (Grissom & Kim, 2005).

Impact data were obtained from a specially constructed Microsoft EXCEL database, with formulae inputted for calculating the standardized mean difference, confidence intervals, and female to male and standard deviation ratios (Thompson, 2002). Differences in effect size were calculated across educational districts and five ability groups based on percentile data. These groupings were (1) below the 25^{th} percentile ($\leq 25^{th}$), (2) below the 50^{th} percentile ($\leq 50^{th}$), (3) above the 50^{th} percentile ($\geq 50^{th}$), (4) above the 75^{th} percentile ($\geq 75^{th}$), and (5) above the 90^{th} percentile ($\geq 90^{th}$). These percentile ranges corresponded to the lower and upper quartile, the lower and upper two quartiles, and the first decile. Based on the different raw scores, students were ranked in the EXCEL database using the rank function and the different rankings were compared. Changes of ranks across different percentiles were also calculated as a measure of impact. For the CEE and SEA data that covered the periods 1999–2002, mean scores and standard deviations provided by the Ministry of Education were plugged into the effect size spreadsheets.

In judging the utility and impact of the current 30% cutscore, classification accuracy was considered equivalent to consistency (Young & Yoon, 1998). Rates of misclassifications were therefore compared to (1) the criterion-referenced cutscore and (2) a hypothetical cutscore that assumes equal performance of males and females. To obtain the criterion-referenced cutscore, a standard setting panel of 12 teachers was convened (Hurtz & Hertz, 1999). From the two-week process, nine usable forms were obtained. The panellists were all first-year undergraduate students pursuing the Bachelor's in Education or Bachelor's degree in Mathematics and Education. All students were

scholarship recipients and had taught for 2 to 10 years prior to attendance at the training colleges. Five had substantial experience at the fifth standard level, four were male, two were from Tobago, and one successfully completed the Measurement and Evaluation elective at training college. The group underwent extensive training in standard setting and received materials related to the 2003 SEA, including question papers and table of specifications.

Panellists were provided with rating forms, guidelines, and original papers related to standard setting. The original design of the standard setting protocol had to be altered because item scoring schemes or rubrics were not available. Therefore, instead of assigning specific item weights, all questions were weighted equally, with task demands incorporated in the probability judgements. Both probability and extended Angoff[2] procedures were used, which required panellists to either estimate the probability of the minimally competent student getting each answer right or select the most likely score (Cizek, Bunch, & Koons, 2004; Hambleton & Plake, 1995). The extended Angoff procedure was used for setting standards on the essay paper. The procedure involved analysis and discussion of the test and test blueprint, development of specific descriptors for each performance level, individual rating, and discussion of group results (Brandon, 2004). The intense discussion during the process was necessary to ensure consensus and accuracy of judgement (Hurtz & Auerbach, 2003).

Four levels of performance were considered, which demanded three separate decisions for each subject area (Cizek, Bunch, & Koons, 2004). The performance levels corresponded to the definitions constructed for the 2005 Trinidad and Tobago national tests. The focus of this study was on the last two categories, at which the pass/fail boundary is assumed to exist. Level 2 was defined as *"marginal academic performance, with work approaching but not yet reaching satisfactory performance, indicating only a partial understanding and limited display of skills required."* This level is sometimes labelled as "Partially Proficient," "Borderline," or "Nearly Meets the Standard of Work Required at this Level." Level 1 was defined as "*inadequate academic performance indicating little understanding and minimal display of required skills, with a major need for additional instructional opportunities, remedial assistance, and/or increased student academic commitment to achieve at*

the Proficient Level." This level is sometimes labelled as "Below Basic," "Well Below Proficient," or "Well below the Standard of Work Required at this Level." Students at this performance level would normally be assessed for remediation.

Results

1. How do gender differentials compare across the 1998–2000 CEE and 2001–2004 SEA administrations?

Table 2 shows the changes in the standardized mean difference (Cohen's *d*) for each subject component across three CEE and four SEA administrations for the period 1998–2003.As shown, gender differences were negligible to small for Mathematics across all seven cohorts regardless of design (Range of Cohen's d = 0.169–0.274). The gender gap was also consistently negligible for Science (Range of Cohen's d = 0.127–0.176) and negligible to small for Social Studies (Range of Cohen's d = 0.171–0.233). In contrast, the gender gap for performance on the Creative Writing component was medium-sized for five of eight cohorts (Range of Cohen's d = 0.466–0.585). The greatest change in differentials across the CEE and SEA occurred in the Language Arts component where the standardized mean difference varied from 0.208 to 0.336 for the CEE, but was consistently over 0.4 for the SEA (Range of Cohen's d = 0.417–0.436). This difference was likely a consequence of changes in format and content across the two examinations, including the removal of the more "gender-neutral" components of Science and Social Studies (Cohen's d = 0.290 [1998] to 0.340 [2000]) compared with the CEE (Cohen's d = 0.391 [2004] to 0.421 [2001]). While this change is small, it may be notable because the composite score is used to determine placement (Olejnik & Algina, 2000).

Table 2. Changes in Component and Composite Scores for the CEE & SEA Eleven-Plus (1998—2004)

Year	Exam	Number		Cohen's *d*						*SDR*					
		Male	Female	Maths	Science	Lang Arts	Social Studies	Essay	Comp. Score	Maths	Science	Lang. Arts	Social Studies	Essay	Comp. Score
1998	**CEE**	13,961	14,665	0.17	0.13	0.31	0.21	0.51	0.29	0.94	0.94	0.90	0.91	0.88	0.90
1999	**CEE**	7,976	8,408	0.19	0.15	0.28	0.17	0.55	0.29	0.95	0.94	1.00	0.96	0.88	0.94
2000	**CEE**	7,914	8,470	0.22	0.18	0.34	0.23	0.59	0.34	0.97	0.94	0.95	0.91	0.89	0.92
2001	**SEA**	8,168	8,216	0.27		0.44		0.47	0.42	0.92		0.91		0.90	0.90
2002	**SEA**	8,095	8,289	0.24		0.43		0.51	0.42	0.93		0.88		0.88	0.89
2003	**SEA**	8,188	8,196	0.22		0.42		0.51	0.41	0.89		0.80		0.84	0.84
2004	**SEA**	10,338	10,476	0.20		0.43		0.48	0.39	0.91		0.86		0.86	0.88

Table 3. Patterns of Gendered Achievement on 2003 SEA Subject Components Across Five Ability Groups

SEA Construct	Total Population				*Indices of gendered achievement across different ability groups*				
	Male		Female		≤25th	≤50th	≥50th	≥75th	≥90th
	Mean	*SD*	Mean	*SD*	---------- Cohen's *d* ----------				
Mathematics (100)	56.62	26.61	62.31	23.78	0.15	0.19	-0.13	-0.20	-0.23
Language Arts (100)	59.02	26.11	68.83	20.85	0.56	0.49	0.20	0.19	0.11
Creative Writing (12)	7.17	2.85	8.50	2.42	0.58	0.57	0.34	0.33	0.20
Composite Total (212)	122.71	53.56	139.64	45.05	0.46	0.40	0.05	0.03	-0.04
	P	Cohen's *d*	*SDR*						
Mathematics	0.00	0.23	0.89		0.86	0.87	1.05	1.07	1.11
Language Arts	0.00	0.42	0.80		0.89	0.83	0.92	0.94	0.96
Creative Writing	0.00	0.50	0.85		0.88	0.86	0.92	0.87	0.87
Composite Total	0.00	0.34	0.84		0.85	0.83	1.00	1.00	0.99

Changes in the standard deviation ratios were small in both subject component and composite scores for the CEE and SEA. However, again the most notable differences across the seven administrations and two test designs were in Language Arts (Range of SDR_{CEE} = 0.904–0.948; Range of SDR_{SEA} = 0.855–0.911) and total composite score (Range of SDR_{CEE} = 0.902–0.924; Range of SDR_{SEA} = 0.836–0.904), with smaller values indicating that the distribution of males is increasingly more variable. This suggests that more males were likely to be found in the tails of the distribution for the SEA composite score. Again, while this is a relatively small change, the overall impact of these differences will be notable when the 30% cutscore is applied.

2. What was the pattern of gendered achievement and placement opportunities in the 2003 SEA?

Table 3 provides data on the pattern of gendered achievement in each component of the 2003 SEA. Apart from the mean raw scores for males and females, indices of statistical and practical significance are shown along with standard deviation ratios. The table shows that while all differences were statistically significant, only differences on Creative Writing were medium-sized or practically significant (Cohen's *d* [*Creative Writing*] = 0.503). This effect size indicates a notable difference in performance between males and females. However, the standard deviation ratio for Language Arts showed that a greater number of males were located in the tails of the distribution (SDR [*Creative Writing*] = 0.799).

Table 4 captures the gendered impact of placement by providing data on the mean, maximum, and minimum score of males and females obtaining their six choices. In this instance, eta-squared was the measure of practical significance used for continuous data and both risk and odds ratios for categorical outcomes (placement received).

Table 4. Gender Differences in Composite Total Score and Percentage of Males and Females Obtaining Choices 1 to 6

Parental Choices	Gender	Mean	*SD*	Max	Min	Median	Eta ^{2*}	*% Male and Female Obtaining Choice* %	Chi ^{2}	P-value	Risk Ratio	Odds	Odds Ratio
First	F	673.75	70.92	738	367	703	0.03	14.9	4.96	0.03	0.93	0.18	0.92
	M	644.05	97.45	739	368	685		16.0				0.19	
Second	F	650.51	73.96	723	369	674	0.05	11.6	0.00	1.00	1.00	0.13	1.00
	M	611.24	93.58	722	367	636		11.6				0.13	
Third	F	629.08	73.10	720	375	653	0.03	12.0	0.17	0.68	0.99	0.14	0.98
	M	603.88	82.87	715	365	631		12.2				0.14	
Fourth	F	617.21	69.58	713	370	638	0.03	11.9	11.11	0.00	0.89	0.14	0.87
	M	589.57	77.91	715	370	612		13.4				0.16	
Fifth	F	606.40	64.31	710	381	615	0.06	13.1	0.70	0.40	1.03	0.15	1.04
	M	570.35	77.05	693	366	578		12.7				0.15	
Sixth	F	592.69	58.45	702	376	598	0.09	15.7	5.39	0.02	1.08	0.19	1.09
	M	553.29	71.00	689	367	563		14.5				0.17	
None	F	575.71	64.74	698	367	589	0.10	20.9	5.17	0.02	0.98	3.80	0.92
	M	530.87	70.75	695	365	534		19.6				4.11	

*****All p < .001

Table 5. Changes in Ranks Based on New Weighting of Composite Score as Applied to the 2003 SEA Data

% Change in Rank*	Magnitude of Change by Rank	No. of Students Changing Rank						No. of Students Changing Rank by Direction of Change				Overall Change in Rank		
		Top 100	Quartiles				Lower 30%	Positive		Negative		F	M	F/M ratio
			1st	2nd	3rd	4th		M	F	M	F			
< 1	**1-199**	100	1,774	896	1,062	2,787	3,051	1,786	1,779	1,667	1,282	3,453	3,061	1.13
1-4.9	**200-999**		2,808	2,952	3,350	2,271	3,010	2,289	3,582	3,219	2,295	5,508	5,877	0.94
5-9.9	**1,000-1,999**		556	1,132	784	110	195	412	924	863	384	1,275	1,308	0.97
10-19.99	**2,000-3,999**		24	121	40	2	7	40	99	37	11	77	110	0.70
Overall % Change		82	99.65	100.00	99.94	99.65	99.81							
F/M Ratio	**Old Score**	1.17	0.55	1.01	1.20	1.51	0.61							
	New Score	1.17	0.58	1.04	1.21	1.38	0.64							

*No recorded changes greater than 20%

As shown, differences between achievement scores for females and males were statistically significant for all six choices and for students assigned to a school by the Ministry of Education. However, as judged from the values of eta squared (Range of eta^2 = 0.025–0.099), these differences were all negligible in terms of practical significance. Likewise, the value of chi-square was statistically significant for the fourth and sixth choices, indicating that relatively more males received their fourth choice and relatively more females received their sixth choice. Nonetheless, the magnitude of this impact was relatively small as measured by the odds and risk ratios (Range of odds ratios = 0.870–1.094; Range of risk ratios = 0.855–1.080).

3. How do gendered achievement patterns vary across construct, ability group, and regions in the 2003 SEA?

Table 2 also provides data on the standardized mean difference and standard deviation ratios across five ability groups: (1) below the 25th percentile, (2) below the 50th percentile, (3) above the 50th percentile, (4) above the 75th percentile, and (5) above the 90th percentile. As shown, gender performances in the top quartile were different to that in the lower quartile. In the top quartile, the difference in Mathematics (Cohen's d = -0.199) and Language Arts (Cohen's d = 0.191) was small. The difference between males and female scores on Creative Writing was also small but relatively larger (Cohen's d = 0.325). The consequent difference in the composite score was also negligible (Cohen's d = -0.34). For the lower quartile, only the difference in Mathematics achievement scores was small (Cohen's d = 0.150); whereas gender differences on Language Arts (Cohen's d = 0.560), Creative Writing (Cohen's d = 0.569), and the composite score (0.457) were all close to or above the benchmark of a medium-sized effect.

Thus, males in the lower quartile were primarily the ones who were underachieving compared with females in the same ability grouping. Moreover, this underachievement was primarily on measures of literacy. Males in the upper quartile and top 10% clearly do equally well in the SEA, and in the case of the top 10%, the gender gap on Mathematics actually showed a small advantage to males (Cohen's d = -0.230). The standard deviation ratios for the top and lower quartiles also indicated significant differences, with male scores more variable on all three

components in the lower quartile, but in the upper quartile, female scores were more variable (SDR = 1.109). However, even in this ability group, male scores were still more variable in Creative Writing (SDR = 0.867).

Table 4 provided data on the pattern of gendered achievement across districts and changes in the female to male ratio across the six choices and assigned group. The data suggested that both the gender gap and placement opportunities varied substantially across districts. In Tobago, for example, females had a small- to medium-sized advantage on the three SEA components, whereas in Victoria the differences on the components were negligible to small. Thus, the gender gap was relatively larger in low-performing districts such as Tobago and in the rural South Eastern and North Eastern regions. Overall, males were more likely to receive their first and fourth choice; however, there were also significant differences across educational districts. In Victoria, for example, males were more likely to receive their first choice (F/M ratio = 0.892), whereas in Port of Spain and Environs the chances are about equal (F/M ratio = 1.045). Males were severely disadvantaged on choices 1 to 4 in the Caroni region (F/M range = 0.707 [3rd choice] to 0.922 [1st choice]). Females also appeared especially disadvantaged on choices 4 to 6 in Port of Spain and Environs and St. George East. In both Victoria and Tobago, females were close to 1.5 times more likely to receive their second choice.

4. What is the gendered impact of changing the rules for computing the composite total score, based on the 2003 SEA data?

Table 5 provides data on the numbers changing rank in different parts of the score distribution when the new weighting scheme is applied. The changes resulting from the new scoring regime were relatively small, with 121 students in the second quartile changing rank, amounting to a 10–20% change. As shown, in the top 100 students, most changes were relatively small with no overall change in the female to male ratios. However, in the lower 30%, relatively more females were found under the new scoring regime, although again the magnitude of the change is relatively small, of the order of 0–5%. The greatest change occurs in the fourth quartile where relatively more males were positioned. Table 6 also provides details of the changes in ranks by gender. Although changes in

the weighting regime were designed to reduce the gender differences, males were still more disadvantaged, with most of the larger negative changes (decrease in rank) occurring for this group. This is a notable effect considering that the numbers of male and female students with changes are overall quite similar. It is also evident, however, that the new scoring regime will improve the position of males in the first (F/M [old scoring] = 0.55 & F/M [new scoring] = 0.58) and fourth quartiles (F/M [old scoring] = 1.51 & F/M [new scoring] = 1.38). At the same time, more males will be placed in the remedial class assuming application of the 30% cutscore (F/M [old scoring] = 0.61 & F/M [new scoring] = 0.64).

Table 6. Misclassification Rates for the Current Norm-Referenced Remediation Cutscore

Cut-score	Numbers of Candidates			F/M Ratio	Misclassification Rates for Equal Percentages		Misclassification Rates for a CRT Cutscore of 40	
	Total	F	M		F	M	F	M
5	164	145	19	0.13	-38.41	+38.41	-23.37	+23.37
10	522	432	90	0.21	-32.76	+32.76	-17.71	+17.71
15	913	732	181	0.25	-30.18	+30.18	-15.13	+15.13
20	1,408	1,079	329	0.30	-26.63	+26.63	-11.59	+11.59
25	1,986	1,466	520	0.35	-23.82	+23.82	-8.77	+8.77
30*	2,600	1,839	761	0.41	-20.73	+20.73	-5.68	+5.68
35	3,418	2,311	1,107	0.48	-17.61	+17.61	-2.57	+2.57
40 º	4,220	2,745	1,475	0.54	-15.05	+15.05	0.00	+0.00
45	5,220	3,256	1,964	0.60	-12.38	+12.38	+2.67	-2.67
50	6,375	3,852	2,523	0.65	-10.42	+10.42	+4.62	-4.62
55	7,481	4,388	3,093	0.70	-8.66	+8.66	+6.39	-6.39
60	8,826	5,062	3,764	0.74	-7.35	+7.35	+7.69	-7.69

º Levels 1/2 cutscore of 85/212 (Combined raw score) equivalent to 40.09%

5. What is the accuracy and gendered impact of using an arbitrary 30% cutscore to select students for remediation, based on the 2003 SEA data?

Table 6 shows the misclassification rates for the 30% cutscore in two situations: (1) a hypothetical situation in which the numbers of males and females are always equal and (2) the obtained criterion-referenced cutscore of 40. As shown, assuming hypothetical equality of male and female performance at each score point, misclassification rates decrease significantly at higher percentiles. This pattern occurs because of

differences in both score distribution and variability for male and female sub-populations, with male scores having a lower mean and higher variability at lower percentiles. As a result, there were many more males at lower percentiles. The norm-referenced 30% cutscore chosen by the Ministry of Education was well below the standards-referenced cutscore of 40%, which resulted in as many as 1,620 students being misclassified or categorized as not requiring remediation. This number included 906 females and 714 males. Thus, females were at a greater relative disadvantage because of the inaccurate determination of the cutscore (Relative misclassification rate = -5.68 %).

Discussion

This study was designed to evaluate the gender fairness of the Eleven-Plus. Using an evaluation protocol developed by Willingham (1999), gender fairness issues were first identified and analysed using an argument-based approach to test validation. This process led to the development of semantic and policy inferences on gender fairness of the SEA. This first stage was followed by an evaluation of the relative impact of the different CEE and SEA designs and variations in gender achievement patterns across educational districts. The findings of this stage confirmed that there were notable variations in gender differences for the Eleven-Plus composite score across different test designs, possibly because of (1) changes in the item format and (2) exclusion of gender-neutral subject components. There were also notable variations in the gender gap across educational districts. Based on the evaluation protocol, it is now left to resolve some of the fairness issues identified and evaluated.

The evidence suggested that tinkering with the system to alter the weighting of the three SEA components from 1:1:1 to 5:3:2 will only slightly improve the ranking of some high- and low-scoring males, but will not reduce the female advantage for the great majority of students. In fact, females in the middle parts of the distribution will gain a further advantage. These unexpected results may be explained by substantial differences in the shape and spread of score distributions for males and females. The data showed that the distribution of male scores were more platykurtic (flatter at the top) but less negatively skewed (fewer higher scores) (Male Distribution: Mean = 583.3, *SD* = 89.86, Kurtosis = -0.659,

Skewness = -0.425; Female Distribution: Mean = 617.0, *SD* = 75.9, Kurtosis = -0.139; Skewness = -0.638). Thus, male and female distributions on the SEA composite score were very different in terms of both score location and variability, with more females located in the centre (the female distribution is taller at the centre), and increasingly more males found in the lower tail (Halpern, 2000). Thus, the larger number of females in the centre of the distribution reduced the utility of the new changes in component weighting.

The key questions in this resolution phase is whether the SEA should be modified or is currently acceptable for fair use and, if used, what constraints should be put in place. There is no single answer to these questions because the goals and purposes of a selection system cannot be resolved through using or modifying a single examination. Indeed, rewarding accomplishment, improving efficiency, and ensuring equal opportunity are competing objectives that will always result in contradictory choices. Three general approaches to resolving fairness issues are to: (1) remove existing barriers, (2) implement a compensatory approach to mitigate disadvantages, and (3) adopt a more democratic approach emphasizing formative testing and individual development. While the former two approaches have been tried in the past, the continued fallibility of Eleven-Plus testing suggests a need for considering the last approach.

Perhaps, the weakness of the SEA as a gendered sieve was most evident in the sharp variations in gendered achievement patterns and placement opportunities across the eight educational districts. If SEA scores were a fair and equivalent separator for males and females, then large differences in achievement scores and placement opportunities across regions should not be so readily apparent. However, this is not the case and there were very large differences in favour of females in the rural South Eastern and North Eastern regions, and on the island of Tobago. Of course, one might explain this finding by arguing that most males in this region were simply of lower ability. However, it is more likely that limitations in the instrument coupled with the variations in the quality of instruction and opportunities to learn were the main factors contributing to lowered performance among some males. It is also possible that differing definitions of male achievement and success are part of the complex of factors that work to create larger differentials in these rural communities.

These impacts may be more notable on some constructs, such as literacy (Chevannes, 2001; Connell, 1995, 2000; Rowan, Knobel, Bigum, & Lankshear, 2002).

As suggested earlier by Jules (1994), the placement system operated quite differently to the achievement component and was a separate source of inequity, especially for some regions and placement choices. Females certainly appeared slightly disadvantaged on the highly desired first choice placement, most likely because of the very high prestige (desirability) of one or two all-girl secondary schools and the performance of the females in the top ability group. Males were more likely to receive their lower choices in Port of Spain and Environs, probably because of the greater number of school places in this district. However, in Caroni educational district, males were severely disadvantaged on the first four choices, as there were fewer high-achieving, all-male schools in this region and spaces in the co-educational schools were also limited.

Resolving fairness issues also requires greater attention to the processes of test design and test evaluation. For test design, one solution might be to decide what type of gatekeeping or sieve role is required, what type of selection system will achieve the desired purpose, and then collect evidence that proves a particular design is comparatively fairer and more valid than alternatives. Indeed, with full universal secondary education implemented, it just seems easier to focus on improving the quality of secondary schools currently considered less prestigious. This will certainly reduce the need for a selection system to allocate students to different school types. Indeed, even if a fairer examination could be designed, just how useful would it be to track students towards differential opportunities at such an early age? Nevertheless, a fairer Eleven-Plus examination is possible. For example, a meaningful composite score can be achieved through multiple measures spaced over time, format, and construct (Goldberg & Roswell, 2001; Schafer, 2003). This is by no means a plug for the inclusion of high-stakes school-based assessment because this will raise the additional question of, *how exactly do we maintain fairness and integrity for teacher-administered school-based assessments when such high-stakes are involved?* (Murphy, 2000) The answer to this new question cannot lie in trusting teachers more, because the issue at hand is about conflict for highly valued, sparsely distributed resources in a society that has never come to grips with fairness and

meritocracy. Therefore, an externally administered and managed examination might be the only legitimate option at this time.

Assessment experts must pay greater attention to the principles of evidence-based assessment design, which is based on three premises:

1. Assessments should be built upon important knowledge, skills, and attitudes and focus on how these competencies are acquired and put to use.
2. The line of reasoning from what test-takers say and do, and inferences of what they know, can do, or should do next must be based on the principles of evidentiary reasoning.
3. The test purpose should be the driving force behind all design decisions, reflecting constraints, resources, and conditions of use.

According to Mislevly, Almond, & Lukas (2004), the three core design questions are:

1. What complex of knowledge, skills, or attitudes should be assessed?
2. What behaviours or performances reveal these constructs?
3. What tasks or situations might elicit these behaviours?

One benefit of such an evidenced-based approach is that the design process becomes systematic and tasks are matched to key outcomes. Additionally, test developers should be required to evaluate the impact of newly proposed designs prior to full implementation. Clearly, more evaluation information is required to support current test use and interpretation associated with the Eleven-Plus. Certainly, there is little evidence in this study to support the continued use of a 30% cutscore. Moreover, there is no theoretical rationale or educational criteria for applying such a cutscore, especially when the evidence indicates that SEA operates differently for lower-ability males and females. It is also questionable whether useful diagnostic information is available from the current SEA design (Green & Weir, 2004). As such, even with a criterion-referenced "remediation" cutscore, a properly designed and developed diagnostic test must be applied afterward. Indeed, the act of misclassifying potentially remedial students has grave implications for constructing a fairer selection system.

(Re)organizing to Find the Evidence

Relevant departments in the Ministry of Education must be retooled and better resourced to ensure that, as a public agency, evidence is collected for current interpretations and test score use. This would require a fuller range of assessment and psychometric services, including procedures and personnel to design, develop, and evaluate assessments. The evaluation role is critical, as there are currently insufficient validity studies of high-stakes, large-scale assessments used for placement or accountability purposes. The role of the higher education institutions is also important in supporting such an agency. This requires improvement and updating of programmes in the area of measurement theory. Cooperation between different institutions is essential if national goals are to be achieved and if evidence-based decisions are to be used to guide further development of the SEA.

Considering the current role of high-stakes examinations in the Caribbean, stakeholders must receive more information about important issues associated with the use of test scores. Blanket acceptance of test scores as prescriptive of a child's future is outdated. The role of illuminators, then, becomes central to good measurement practice and greater consideration of these issues is required. Academic illuminators must obtain the tools to study large-scale assessments and prescribe ways in which testing can become fairer. In hindsight, then, there seems little advantage in moving from the CEE to the SEA. There is little available evidence that the SEA has greater predictive validity or can adequately measure or promote critical thinking and problem solving (T&T. Task Force, 1998). Instead, the evidence from this study suggests that the SEA is considerably less gender fair than the CEE. It seems incredible, then, that such a fallible examination might continue to have such a critical role sieving males and females towards future opportunities at such an early age.

In Memory of

Omesh Matura, 2004 Winner of the President's Medal for Teachers and 2005/2006 Year 1 B.Ed UWI, St. Augustine student; a potentially great measurement student who was involved in this study's standard setting exercise.

Notes

[1] The Concordat is a formal agreement between Government and the Denominational School Boards. One element of this agreement is that 20% of the Eleven-Plus places assigned to a denominational school are reserved for nominees of the Board, subject to criteria laid out by the Ministry of Education.

[2] Standard setting is the process used to obtain a passing score for an examination. The Angoff procedure is one of the more widely used test-based methods.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: AERA.

Archer, L., & Yamashita, H. (2003). Theorising inner-city masculinities: 'Race', class, gender and education. *Gender and Education, 15*(2), 115–132.

Bagley, C. A., Woods, P. A., & Glatter, R. (2001). Rejecting schools: Towards a fuller understanding of the process of parental choice. *School Leadership and Management, 21*(3), 309–325.

Bailey, B. (2000). School failure and success: A gender analysis of the 1997 General Proficiency Caribbean Examinations Council examinations for Jamaica. *Journal of Education and Development in the Caribbean, 4*(1), 1–18

Ball, S., Bowe, R., & Gewirtz, S. (1995). Circuits of schooling: A sociological exploration of parental choice in social class contexts. *Sociological Review, 43*(1), 52–78.

Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. *Educational Measurement: Issues and Practices, 17*(1), 10–17.

Brandon, P. R. (2004). Conclusions about frequently studied modified Angoff standard-setting topics. *Applied Measurement in Education, 17,* 59–88.

Camara, W. J., & Kimmel, E. W. (Eds.). (2005). *Choosing students: Higher education admissions tools for the 21st century.* Mahwah, NJ: Lawrence Erlbaum.

Chevannes, B. (2001). *Learning to be a man: Culture, socialization, and gender identity in five Caribbean communities.* Mona, Jamaica: UWI Press.

Chilisa, B. (2000). Towards equity in assessment: Crafting gender-fair assessment. *Assessment in Education: Principles, Policy, and Practice, 7*(1), 61–81.

Cizek, G. J., Bunch, M. B., & Koons, H. (2004). Setting performance standards: Contemporary methods. *Educational Measurement: Issues and Practice, 23*(4), 31–50.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Connell, R. W. (1995). *Masculinities.* Sydney: Allen & Unwin.

Connell, R. W. (2000). *The men and the boys*. Berkeley, CA: University of California Press.

Council on Licensure, Enforcement, and Regulation, & National Organization for Competency Assurance. (1993). *Principles of fairness: An examining guide for certification boards*. Lexington, KY: Author.

Cronbach, L. J. (1988). Five perspectives on validity argument. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 3–17). Hillsdale, NJ: Lawrence Erlbaum.

DeBlase, G. L. (2003). Missing stories, missing lives: Urban girls (re)constructing race and gender in the literacy classroom. *Urban Education, 38*(3), 279–329.

De Lisle, J., & Smith, P. (2004). Reconsidering the consequences: Gender differentials in performance and placement in the 2001 SEA. *Caribbean Curriculum, 11*, 23–55.

De Lisle, J., Smith, P., & Jules, V. (2005). Which males or females are most at risk and on what? An analysis of gender differentials within the primary school system of Trinidad and Tobago. *Educational Studies, 31*(3), 393–418.

Drenth, P. J. D., van de Flier, H., & Omari, I. M. (1983). Educational selection in Tanzania. *Evaluation in Education, 7*(2), 93–217.

Durbrow, E. H., Schaefer, B. S., & Jimerson, S. (2002). Diverging academic paths in rural Caribbean village children: Predicting secondary school entrance from the St. Vincent Child Study. *School Psychology International, 23*(2), 155–168.

Edwards, T., & Tomlinson, S. (2002). *Selection isn't working. Diversity, standards and inequality in secondary education* (Catalyst working paper). London: Catalyst Forum.

Elwood, J. (1999). Equity issues in performance assessment: The contribution of teacher-assessed coursework to gender-related differences in examination performance. *Educational Research and Evaluation, 5*(4), 321–344.

Elwood, J. (2005). Gender and achievement: What have exams got to do with it? *Oxford Review of Education, 31*(3), 373–393.

Fleiss, J. L. (1994). Measures of effect size for categorical data. In H. Cooper & L. V. Hedges (Eds.), *Handbook of research synthesis* (pp. 245–260). New York: Russell Sage Foundation.

Gardner, J., & Cowan, P. (2005). The fallibility of high-stakes '11-plus' testing in Northern Ireland. *Assessment in Education 12*(2), 145–165.

Gersten, R., & Baker, S. (2002). The relevance of Messick's four faces for understanding the validity of high-stakes assessments. In G. A. Tindal & T. M. Haladyna (Eds.), *Large-scale assessment programs for all students: Validity, technical adequacy, and implementation* (pp. 49–66). Mahwah, NJ: Lawrence Erlbaum.

Gipps, C., & Murphy, P. (1994). *A fair test?: Assessment, achievement and equity.* Philadelphia, PA: Open University Press.

Goldberg, G. L., & Roswell, B. S. (2001). Are multiple measures meaningful?: Lessons from a statewide performance assessment. *Applied Measurement in Education, 14*(2), 125–150.

Gould, S. J. (1996). *The mismeasure of man* (revised & expanded). New York: Norton & Norton.

Greaney, V., & Kellaghan, T. (1995). *Equity issues in public examinations in developing countries* (World Bank Technical Paper No 27). Washington, DC: World Bank.

Green, A. B., & Weir, C. J. (2004). Can placement tests inform instructional decisions? *Language Testing*, *21*(4), 467–494.

Grissom, R. J., & Kim, J. J. (2005*). Effect sizes for research: A broad practical approach*. Mahwah, NJ: Lawrence Erlbaum.

Haertel, E. H. (1999). Validity arguments for high-stakes testing: In search of the evidence. *Educational Measurement: Issues and Practice, 18*(4), 5–9.

Haertel, E. H., & Herman, J. L. (2005). A historical perspective on validity arguments for accountability testing. *Yearbook of the National Society for the Study of Education, 104*(2), 1–34.

Haladyna, T. M., & Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice, 23*(1), 17–27.

Halpern, D. F. (2000). *Sex differences in cognitive abilities (3rd ed.).* Mahwah, NJ: Lawrence Erlbaum.

Hambleton, R. K., & Plake, B. S. (1995). Using an extended Angoff procedure to set standards on complex performance assessments. *Applied Measurement in Education, 8*(1), 41–55.

Handy, C. (1989). *The age of unreason.* Boston, MA: Harvard Business School Press.

Heyneman, S. P. (1987). Uses of examinations in developing countries: Selection, research, and education sector management. *International Journal of Educational Development, 7*(4), 251–263.

Hickling-Hudson, A. (2002). Re-visioning from the inside: Getting under the skin of the World Bank's Education Sector Strategy. *International Journal of Educational Development, 22*(6), 565–577.

Hildebrand, G. M. (1996). Redefining achievement. In P. F. Murphy & C. V. Gipps (Eds.), *Equity in the classroom* (pp. 149–171). London: Falmer Press.

Hospedales, C. E. (1981). Prestige and success in schools. *Trinidad and Tobago Education Forum, 2*(3), 1–18.

Hurtz, G. M., & Auerbach, M. A. (2003). A meta-analysis of the effects of modifications to the Angoff method on cutoff scores and judgment consensus. *Educational and Psychological Measurement, 63*(4), 584–601.

Hurtz, G. M., & Hertz, N. R. (1999). How many raters should be used for establishing cutoff scores with the Angoff method? A generalizability theory study. *Educational and Psychological Measurement, 59*(6), 885–897.

Jackson, C. (2003). Motives for 'laddishness' at school: Fear of failure and fear of the 'feminine'. *British Educational Research Journal, 29*(4), 583–598.

Jennings, Z. (2001). Teacher education in selected countries in the Commonwealth Caribbean: The ideal of policy versus the reality of practice. *Comparative Education, 37*(1), 107–134.

Joint Committee on Testing Practices. (2004). C*ode of fair testing practices in education*. Washington, DC: Author. Retrieved, September 6, 2005, from http://www.apa.org/science/FinalCode.pdf.

Jules, V. (1994). A study of the secondary school population in Trinidad and Tobago: Placement patterns and practices – A research report. St. Augustine, Trinidad: Centre for Ethnic Studies, UWI.

Kane, M. (2001). Current concerns in validity theory. *Journal of Educational Measurement, 38*(4), 319–342.

Kane, M. (2002). Validating high-stakes testing programs. *Educational Measurement: Issues and Practice, 21*(1), 31–41.

Kane, M. (2004a). Certification testing as an illustration of argument-based validation. *Measurement: Interdisciplinary Research and Perspectives, 2(3), 135–170.*

Kane, M. (2004b). The analysis of interpretive arguments: Some observations inspired by the comments [Rejoinder]. *Measurement: Interdisciplinary Research and Perspectives, 2*(3), 192–200

Kellaghan T., & Greaney, V. (1992). *Using examinations to improve education: A study in fourteen African countries.* Washington DC: World Bank.

Kline, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioural research.* Washington: APA.

Kutnick, P. (2000). Boys, girls 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 Education and Development, 20*(1), 65–84.

Layne, A., & Kutnick, P. (2001). Secondary school stratification, gender, and other determinants of academic achievement in Barbados. *Journal of Education and Development in the Caribbean, 5*(2–3), 81–101.

Linde, G. (2003). A journey with Durkheim through an examination driven school system. *Educational Studies, 29*(4), 323–335.

London, N. A. (1989). Selecting students for secondary education in a developing society: The case of Trinidad and Tobago. *McGill Journal of Education, 24(3),* 281–291.

MacKenzie, C. G. (1989). The eleven-plus examination in developing countries: A case study. *Educational Studies, 15*(3), 281–300.

Messick, S. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 33–45). Hillsdale, NJ: Lawrence Erlbaum.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: Macmillan.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2004). *A brief introduction to evidence-centered design* (CSE Report 632). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing, UCLA. Retrieved, January 14, 2006, from http://www.cse.ucla.edu /reports/r632.pdf.

Mislevy, R. J., Steinberg, L. S., & Almond, R. A. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1*(1), 3–62.

Murphy, P. (2000). Equity, assessment, and gender. In J. Salisbury & S. Riddell (Eds.), *Gender, policy and educational change: Shifting agenda in the UK and Europe* (pp. 134–152). London: Routledge.

Myhill, D. A. (2005). Testing times: The impact of prior knowledge on written genres produced in examination settings. *Assessment in Education, 12*(3), 289–300.

Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology, 25*(3), 241–286.

Overdorf, J. (2004). Failing the grade. *Far Eastern Economic Review, 167*(26), 54–56.

Parsons, E., Chalkley, B., & Jones, A. (2000). School catchments and pupil movements: A case study in parental choice. *Educational Studies, 26*(1), 33–48.

Payne, M. A., & Barker, D. (1986). Still preparing children for the 11+: Perceptions of parental behaviour in Barbados. *Educational Studies, 12*(3), 313–325.

Rampersad, J. (1999). Patterns of achievement by gender in school science. *Caribbean Curriculum, 7*(1), 37–50.

Rowan, L., Knobel, M., Bigum, C., & Lankshear, C. (2002). *Boys, literacies and schooling: The dangerous territories of gender-based literacy reform.* Buckingham, UK: Open University Press.

Ryan, J. M., & DeMark, S. (2002). Variation in achievement scores related to gender, item format, and content area tested. In G. A. Tindal & T. M. Haladyna (Eds.), *Large-scale assessment programs for all students: Validity, technical adequacy, and implementation* (pp. 67–88). Mahwah, NJ: Lawrence Erlbaum.

Sacks, P. (2001a, June 8). How admissions tests hinder access to graduate and professional schools. *Chronicle of Higher Education, 47*(39), 11.

Sacks, P. (2001b, December/January). Pseudo-meritocracy: A response to "The future of affirmative action" by Susan Sturm & Lani Guinier. *Boston Review*. Retrieved December 26, 2005, from http://bostonreview.net/BR25.6/sacks.html.

Schafer, W. D. (2003). A state perspective on multiple measures in school accountability. *Educational Measurement: Issues and Practice, 22*(2), 27–31.

Schagen, I., & Elliot K. (Eds.). (2004). *But what does it mean? The use of effect sizes in educational research.* Slough, UK: NFER.

Schrag, F. K. (2004). High stakes testing and distributive justice. *Theory and Research in Education, 2*(3), 255–262.

Siegel, H. (2004). High-stakes testing, educational aims and ideals, and responsible assessment. *Theory and Research in Education, 2*(3), 219–233.

Stanley-Marcano, J., & Alexander, M. C. (1998). Trinidad and Tobago [Country report]. In M. Bray & L. Steward (Eds.), *Examination systems in small states: Comparative perspectives on policies, models and operations* (pp. 113–119). London: Commonwealth Secretariat.

Stobart, G. (2005). Fairness in multicultural assessment systems. *Assessment in Education: Principles, Policy, and Practice, 12*(3), 275–287.

Thompson, B. (2002). What future quantitative social science research could look like? Confidence intervals for effect sizes. *Educational Researcher, 31*(3), 25–32.

Trinidad and Tobago. Division of Educational Research & Evaluation. (2002). *Secondary entrance assessment guidelines 2003–2004*. Port of Spain, Trinidad: Author.

Trinidad and Tobago. Division of Educational Research & Evaluation. (2004). *A proposal for a change in weight and modification of the table of specifications of the Secondary Entrance Assessment (SEA) papers* (2005–6). Port of Spain, Trinidad: Author.

Trinidad and Tobago. Task Force for the Removal of the Common Entrance Examination. (1998). *Report.* Port of Spain, Trinidad: Ministry of Education.

Walker, D. A. (2004). The importance of drawing meaningful conclusions from data: A review of the literature with meta-analytic inquiry. *NASPA Journal, 41*(3), 452–469.

Willingham, W. W. (1999). A systemic view of test fairness. In S. J. Messick (Ed.), *Assessment in higher education: Issues of access, quality, student development, and public policy* (pp. 213–242)**.** Mahwah, NJ: Lawrence Erlbaum.

Willingham, W. W. (2002). Seeking fair alternatives in construct design. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), The *role of constructs in psychological and educational measurement* (pp. 193–206). Mahwah, NJ: Lawrence Erlbaum.

Willingham, W. W., & Cole, N. (1997). *Gender and fair assessment*. Mahwah, NJ: Lawrence Erlbaum.

Young, M. J., & Yoon, B. (1998). *Estimating the consistency and accuracy of classifications in a standards-referenced assessment (CSE Technical Report 475)***.** Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing, UCLA. Retrieved, January 26, 2006, from http://www.cse.ucla.edu/Reports/TECH475.pdf.

Zieky, M. (2002). Ensuring the fairness of licensing tests. *CLEAR Exam Review, 12*(1), 20–26.

Zwick, R. (2002). *Fair game?: The use of standardized admissions tests in higher education.* New York: RoutledgeFalmer.

Zwick, R., & Schlemer, L. (2004). SAT validity for linguistic minorities at the University of California, Santa Barbara. *Educational Measurement: Issues and Practice, 23*(1), 6–16.

Notes on Contributors

Iris Hewitt-Bradshaw is Ph.D. student at the School of Education, UWI, St. Augustine and a teacher educator at the Corinth Teachers' College. She has been involved in language education since 1981, first as a secondary school teacher and then as a teacher educator. Her research interests include the development of pupils' oral language skills and language use across content areas.

Margaret Cain is a Ph.D. student at the School of Education, UWI, St. Augustine. She has been involved in science education since 1972, first as a secondary school teacher and then as a teacher educator at both the Corinth Teachers' College and the School of Education, UWI, St. Augustine. Her research interests include teachers' beliefs and alternative assessment approaches.

Jerome De Lisle is a Lecturer in Educational Administration at the School of Education, UWI, St. Augustine. For the period 1998–2006, he worked as a Lecturer in Measurement and Evaluation at the Faculty of Medical Sciences, UWI, St. Augustine. In 2005, he developed and implemented the current standard setting plan for national tests in the primary school system. His current research interests include test fairness, gendered achievement, and standard setting.

Janet Fullerton-Rawlins is the Librarian/Documentalist at the School of Education Library, UWI, St. Augustine. She has been a librarian for over 20 years, and has worked in the public service of Trinidad and Tobago as a School Librarian, and at the Seismic Research Unit, UWI, St. Augustine.

Frank C. Worrell is Director of the School Psychology program and Faculty Director of the Academic Talent Development Program at the University of California, Berkeley, where he also serves as co-Director for Research and Development for the California College Preparatory Academy, a charter school involving UC Berkeley and Aspire Public Schools. His research interests include academic talent development, at-risk youth, scale development and validation, and teacher effectiveness. More specifically, he examines psychosocial variables related academic achievement in adolescent populations, and the reliability and validity of instruments used to measure psychosocial constructs. He is a member of the editorial boards of *Assessment*, *Gifted Child Quarterly*, the *Journal of Secondary Gifted Education*, *Roeper Review,* and *School Psychology Quarterly*, and is President-Elect of the Division 16 (School Psychology) of the American Psychological Association.

Jennifer Yamin-Ali is a lecturer in Foreign Language Education at the School of Education, UWI, St. Augustine. She taught Spanish at the secondary school and tertiary levels for over 20 years. She has served as a CXC Examiner and a CAPE resource person and is currently on the National Textbook Evaluation Committee for Spanish. Dr. Yamin-Ali pursued her post-graduate studies at UWI and at the University of Sheffield. England.

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. l, No. l, 1985	5.00
	Vol. 2, No. l, l987	5.00
	Vol. 3, No. 1, 1993	10.00
	Vol. 3, No. 2, 1993	10.00
	Vol. 4, No. 1, 1994	15.00
	Vol. 5, No. 1, 1995	15.00
	Vol. 5, No. 2, 1995	15.00
	Vol. 6, No. 1, 1996	15.00
	Vol. 6, No. 2, 1996	15.00
	Vol. 7, No. 1, 1999	15.00
	Vol. 7, No. 2, 2000	15.00
	Vol. 8, 2001	20.00
	Vol. 9, 2002	20.00
	Vol. 10, 2003	20.00
	Vol. 11, 2004	20.00

	Vol. 12, 2005	20.00
Kelvin Jarvis & Sonia Manjoo	A bibliography of staff publications l965-l985. Faculty of Education Library, l986, 48 pp.	15.00
E.P. Brandon & P.N. Nissen (Comps.)	Proceedings of the l990 Cross-Campus Conference on Education, 3-6 April, l990. Kingston, Jamaica. Jamaica: Faculty of Education, l99l, 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. (*Back issues available at reduced rates*).	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. l.	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*.

ORDER FORM

Enclosed is my cheque or money order for:

Title ordered:

.....

.....

.....

No. of copies: .....

Total sum enclosed $ .....

Name:.....

Address:.....

.....

.....

.....

E-mail:……Tel:…….Fax:……...

Information for Contributors

(1) Manuscripts for publication that have not been published elsewhere and all correspondence regarding articles should be submitted to *The Librarian/Documentalist, School of Education, The University of the West Indies, St. Augustine, Trinidad, West Indies.*

(2) Articles from all relevant subject areas will be considered by the editorial committee and priority will be given to those that, in addition to showing acceptable scholarship, fulfil one or more of the following criteria:

(a) describe or evaluate experiments in curriculum intervention that have relevance for Caribbean school system;
(b) discuss theoretical matters that can favourably affect thought on education in the Caribbean;
(c) promote continuous curriculum renewal in the Caribbean region.

(3) Manuscripts should be typed double-spaced (including quotations, footnotes, and references) on standard quarto paper (8 1/2“ x 11”), with ample margins. The author's name and affiliation should appear on a separate cover page, and only on this page, to ensure anonymity in the reviewing process. Include, also on separate pages, an abstract of l00-200 words, and a short biographical profile of the author (and all co-authors, if any). The APA Style Manual (5th edition) should be followed.

(4) The list of references should be arranged alphabetically by authors’ last names, and placed at the end of the manuscript. Reference citations are not to be used as footnotes.

(5) Endnotes for the purpose of clarifying the text are permitted and should be denoted by superscript numbers in the text. The numbered endnotes should be typed double-spaced and placed before the list of references.

(6) Tables and figures should be numbered, captioned, and inserted at appropriate points in the text.

(7) Kindly send three copies of each manuscript for editorial review, excluding originals. All manuscripts will be refereed.

(8) Manuscripts that do not conform to all the above specifications will be returned immediately to the author for conformity.

(9) Authors will receive five (5) copies of their articles without charge.