Comparing the Effects of Reading and Writing on Writing Performance

WAI-KING TSANG

City University of Hong Kong

The study compares the effects of an enriched syllabus which included extensive reading and frequent writing assignments on English descriptive writing performance at different form levels It examines a group of Cantonese-speaking students at four form levels in Hong Kong who participated in three English programs (A) regular plus unrelated (mathematics) enrichment program, (B) regular plus extensive reading, and (C) regular plus frequent writing practice Results demonstrated significant main effects due to the nature of program and form level with no significant interaction of these factors The regular plus extensive reading program was overall significantly effective, while both the regular plus mathematics program and the regular plus frequent writing practice were not In the area of content, the reading program was the only one which showed a significant positive effect Similarly, in the area of language use, the reading program was the only one of the three shown significantly effective

1 INTRODUCTION

This paper compares the effectiveness of an extensive reading program and a frequent writing program on the acquisition of descriptive writing skills in English by secondary students The study is motivated by (1) the extensive reading and frequent writing extension programs widely practiced in Hong Kong, (2) Krashen's Input Hypothesis (1980, 1982, 1985, 1989) and Swain's Output Hypothesis (1985, 1993) about second language acquisition, and (3) the findings of previous reading–writing research, which suggest additional efforts given to reading and/or writing promote the acquisition of the two literacy skills, and language proficiency in general

The English syllabus for secondary schools in Hong Kong encourages the use of graded readers to help language learning The various objectives include 'extending students' writing skills' (Hong Kong Education Department 1983 29) Such a position is supported by the literature Nuttall states that 'the best way to improve your knowledge of a foreign language is to go and live among its speakers The next best way is to read extensively in it' (1982 168) In a similar vein, Krashen theorizes that 'if second language acquisition and the development of writing ability occur in the same way, writing ability is not learned but is acquired via extensive reading in which the focus of the reader is on the message, i e reading for genuine interest and/or pleasure' (1984 23) An evaluation of an extensive reading program and a frequent writing program (which many local

Applied Linguistics, Vol 17, No 2 © Oxford University Press 1996

teachers intuitively believe to be beneficial) will certainly contribute towards the improvement of education

The present study is an evaluation of education in that it makes 'judgments about the merit, value, or worth of educational programs' (Borg and Gall 1989 742) Its practical orientation is to decisions and utilization (Krathwohl 1993 524) and is consistent with that of 'the analyzer', seeking 'to confirm rules, principles, or propositions that relate variables or events' (Krathwohl 1985 163) in arriving at pedagogic decisions and making use of existing resources

The Hong Kong context

In local secondary schools, where English is a non-native language, extensive reading and frequent writing assignments, in addition to the regular syllabus, are often given to help students improve their writing and general language proficiency While extra reading assignments usually require students to write reading reports with little instruction or feedback, similar writing assignments are usually given out during the longer stretches of holidays These writing assignments are simply graded generally, given summary comments without highlighting the mistakes, and marked partially for spelling or certain mechanical aspects rather than marked in detail Class discussion of high-frequency errors is rarely held, and essays are only occasionally returned Post-writing one-to-one conferences are even rarer Few student corrections or revisions are assigned

The present pedagogic practice in Hong Kong schools reflects the teacher belief that input-based reading schemes and output-based writing practice are both means of helping students to improve their language proficiency in general and their writing skills in particular Although there may be some evidence in the literature in support of such a belief (cf the section on 'Previous studies on the role of reading and writing'), the differential effectiveness of the input-based and output-based approaches has not been measured To focus solely on the effects of these two approaches to writing, it would certainly inform the teaching field if research could demonstrate the comparative effectiveness of the input-based and output-based approaches Comparative studies of the two approaches would be revealing if an input-based extensive program is more effective, or as effective as output-based frequent writing assignments, the reading program may be preferred because it requires no teacher editorial time, minimal supervision and feedback, and can be a learner activity carried out independently of the teacher If a significant difference between the two approaches is found only at some form levels but not others, input-based and output-based programs may play different roles and therefore may be applied at different form levels to meet the respective needs of each If learners in different programs benefit in different areas of writing—content, organization, vocabulary, language use, and mechanics—these programs may be incorporated into the syllabus concurrently or consecutively Other new programs may have to be developed to supplement the existing ones in areas where learners show few gains

Theoretical bases

Krashen's Input Hypothesis (1980, 1982, 1985, 1989) is one of five hypotheses constituting his Monitor Theory of language acquisition. According to the Input Hypothesis, the learner acquires language by comprehending linguistic input that is slightly beyond his or her level of competence (i + 1), and

> a necessary (but not sufficient) condition to move from stage i to stage i + 1 is that the acquirer understands input that contains i + 1, where 'understand' means that the acquirer is focussed on the meaning and not the form of the message (Krashen 1982 21)

The strongest criticism raised against Krashen's Input Hypothesis is the untestability of 'i + 1'. Despite this, Larsen-Freeman (1983) supported its formulation and viewed comprehensible input as a variable, differing between learners. Research has also been conducted in order to test the Input Hypothesis, which states that the key factor determining acquisition of competence in an L2 is exposure to large amounts of L2 input material which is meaningful, interesting, or relevant. With similar aims, Krashen and Terrell (1983), Faerch and Kasper (1985), and Pica and Doughty (1985) investigated the nature of oral interaction in the language classroom and the means of generating productive interactive patterns between participants.

Anderson, Wilson, and Fielding (1988) investigated the relation between amounts of reading and reading achievement of 155 fifth graders. The students filled out daily activity forms for periods ranging from eight to twenty-six weeks, which were analyzed as activity data. They took a battery of reading tests—a reading comprehension test, a checklist vocabulary measure, and a measure of reading speed—twice, once at the beginning of the form-filling period and again following this period. These were analyzed as reading proficiency data together with a measure of amount of reading. Results demonstrate that the number of books read is the best predictor of several measures of reading achievement, including gains in reading comprehension. Growth in reading proficiency may probably be attributable to book reading, suggesting the role of large-scale intralingual input in the acquisition of literacy skills.

Swain (1985) studied Grade 6 learners of French as a second language in a complete immersion program. Results showed that having achieved native-like discourse and sociolinguistic competence, students failed to achieve native-like grammatical competence, despite seven years' comprehensible input. Comprehensible output seemed to be lacking in the program.

Swain (1993: 159–61) proposes the following ways in which output plays a necessary role in second language learning: (1) it provides opportunities for practice of linguistic resources permitting the development of automaticity in their use, (2) it forces the learner to move from semantic to syntactic processing, (3) it provides the opportunity to test hypotheses, and (4) it generates feedback from speakers of the language which can lead learners to make modifications.

Although there is a prima-facie case for the contribution of input and output

in second language acquisition, there is still an on-going debate whether academic language abilities can be best supported and enhanced by students engaging in additional writing (output-based) or wide reading (input-based) Krashen's Input Hypothesis, which partly motivated the present study (especially the design of the input-based extensive reading enrichment program, i e , one of two experimental treatments), does not acknowledge possible differences between L1 and L2 in the discussion of the reading process The present study compares the effect of output-based activities with that of an input-based approach to acquisition of second language writing skills

Previous studies on the role of reading and writing

Smith (1984) examined the effects of frequent writing without teacher correction on the writing of native speakers of English in Grades 1 and 2 His results showed significant improvement in both grade levels in all categories measured general impression, length, vocabulary, spelling, and syntactic development In Hong Kong, however, where English is learned as a non-native language, it is not at all clear whether students' writing can improve with activities that exclusively focus on output

Krashen's Input Hypothesis (1982, 1985) states that the key factor determining acquisition of competence in an L2 is exposure to large amounts of meaningful, interesting, or relevant L2 input material In his (1989) review, studies indicated learners acquired vocabulary and spelling through reading In the few studies which examined the explicitly input-oriented component in L2 learning, both aural and written, Davies (1983) compared a French program geared around oral production practice and an input-based program on L2 skill development of Swedish high school pupils aged between sixteen and nineteen The results showed greater gains on both receptive and productive tests for the participants receiving the input-based instruction

In a study by Schleppegrell (1984), adult Egyptian economists enrolled in an ESP writing course focused on extensive reading and listening produced greater gains in writing ability than those with grammar instruction, extensive practice in conversation, and out-of-class writing As reported at the end of Schleppegrell's paper, the study was an outgrowth of an ESP program and not a controlled experiment Any attempt to replicate the study must take into account the selection of the participants and the other variables which may also affect the outcome of the study Nevertheless, two conditions of the study are similar to those in Hong Kong secondary schools First, in the Egyptian language environment (as well as in the Hong Kong language environment), students do not find it easy to see how they may apply grammar learned in writing and speaking Second, very few Egyptian students (and similarly Hong Kong students) have direct exposure to English through travel or close contact with native speakers Even though some schools in Hong Kong employ two native speakers to teach English or other subjects, the input may be restricted to an average of a thirty-minute lesson per week Thus, in these respects, the Hong Kong case is parallel to the Egyptian case

Examining reading materials alone, Fader and McNeil's (1968) and Thorndike's (1973) studies yielded results showing a strong correlation between amount of reading for pleasure and writing proficiency in L1 However, as Krashen (1984) emphasizes, there is little evidence for a similar relationship in L2 Elley (1991) reviewed nine studies—eight in L1 and one in L2—which involved learners between six and twelve years of age in book-based programs lasting three to thirty-six months All the studies show learners' rapid growth in language development compared with learners in regular language programs The most striking finding among these studies is 'the spread of effect from reading competence to other language skills—writing, speaking, and control over syntax' (Elley 1991 404) Four other sources of data in L2 came from Janopoulos (1986), Flahive and Bailey (1993), Elley and Mangubhai (1983, see also Elley 1991), and Hafiz and Tudor (1989, 1990, Tudor and Hafiz 1989, see also Elley 1991)

In Janopoulos's (1986) study, seventy-nine ESL graduate students of seven language groups produced a one-hour writing sample and completed a questionnaire to estimate the amount of time they spent each week reading for pleasure in their native language and in English Writing samples were evaluated holistically on a four-point scale by two trained raters Results established that proficiency levels for L2 reading and writing are closely associated The results, however, were correlational rather than causal Flahive and Bailey (1993) also used a questionnaire to determine the amount of pleasure reading learners engaged in In that study, the amount of pleasure reading correlated with reading proficiency but not with writing proficiency, as measured by an argument style placement essay Pleasure reading also correlated with grammar proficiency as measured by the CELT examination These contradictory findings beg the question whether pleasure reading improves only grammar, reading proficiency, and writing proficiency in general but fails to contribute towards writing-specific—for example, argumentative—types of discourse

Elley and Mangubhai (1983) employed simplified readers with Fijian learners of English at primary levels 4 and 5 They compared 380 learners who participated in a book flood program in addition to the regular English curriculum, with 234 learners who took the regular English language program with little emphasis on reading Learners in the book flood program read more books and demonstrated significantly greater gains in both receptive and productive skills as measured by a reading test, a listening test, and an open-ended English structures test

Hafiz and Tudor (1989, Tudor and Hafiz 1989) studied ESL learners aged ten and eleven in Leeds, UK Those who read simplified readers (in the Heinemann, Longman, and Macmillan series) exhibited significant gains in both reading and writing in NFER Tests of Proficiency in English A later study by Hafiz and Tudor (1990) involved secondary school pupils in Pakistan aged fifteen and sixteen Students who read graded readers (in the Collins, Heinemann, and Longman series) demonstrated significant gains in an essay writing test compared to those receiving no such treatment Unfortunately,

there were variables in Hafiz and Tudor's (1990) study that may have distorted the results different school background, different kinds of regular English classroom instruction, and diverse social background It is significant that the measure in the study did not involve content and organization, two important aspects of assessment in writing In addition, the test topics—'my family', 'the fire', and 'on coming to England'—which were given without specific instructions, might have been interpreted as descriptive or narrative topics Thus these may have been somewhat biased in favor of the treatment group, who would have had more exposure to these types of discourse

The previous studies have involved learners of English as a first language (as in the United States) or a second language (as in the United Kingdom)—non-equivalent conditions This study is concerned with Hong Kong learners of English as an auxiliary language [1] The participants are Cantonese speaking secondary students Their language environment is poor in that contact with English outside the classroom is severely limited Graded materials (simplified classics, original readers, and information-based books in various areas of interest) were used because they were most accessible to the learners, who were none the less encouraged to read extensively any ungraded books written in English In the present study, reading was done outside the classroom in the student's own free time to ensure a tension-free environment The study compared the effects of (1) extensive reading (which included, but was not limited to, description and which was synonymous with McCracken's 1971 guidelines for SSR in the sense that no restrictions were imposed as to type of materials read), and (2) frequent writing (of various genres, including description) on the acquisition of descriptive writing skills The descriptive genre was selected as the matter for investigation because (1) the content of descriptive writing is of general interest, (2) it is relatively neutral as regards the background knowledge required for its interpretation, (3) its organization is relatively unmarked, and (4) there are no rigid formal schemata for descriptive writing An added advantage of selecting the descriptive genre was that it was not unfamiliar to all participants across form levels, since it was included in their regular syllabus In addition, the possibility that the form levels (equivalent to grade levels in the North American school system) are a factor in writing was also examined

2 HYPOTHESES

The present study aimed to determine whether nature of program affects writing by comparing the effectiveness of different enrichment programs on writing performance, whether form level affects writing by comparing writing performance at different form levels, and whether various programs have differential effects on writing performance at various form levels by checking on the possibility that a program is more effective at one level than at the other levels Five null hypotheses were thus formulated and tested in this study

1 There is no significant main effect for nature of program (see Materials section below for description of programs) as a factor in writing performance of secondary students

2 There is no significant difference between the regular plus mathematics program and the regular plus English enrichment programs
3 There is no significant difference among the English enrichment programs
4 There is no significant main effect for form level as a factor in descriptive writing performance
5 There is no significant interaction effect between nature of program and form level as factors in descriptive writing performance

The alpha level for all statistical decisions is set at $p < 01$, as this level of significance is seen as neither overly liberal nor overly conservative as a measure of results

3 METHOD

Participants

The one hundred and forty-four participants in this study are all Cantonese speakers and learners of English as an auxiliary language Their English proficiency can be generally described as high elementary to low intermediate [2] They were pre-selected on the basis of their overall academic performance in the previous year They were from the best classes in Forms 1 to 4[3] and were thus expected to be easily motivated and co-operative. The student population employed was from the local Chinese majority, enrolled in a Government-aided Anglo-Chinese grammar school It is an English medium, as opposed to a Chinese medium, secondary school As in most secondary schools, students of Forms 1 to 3 in this school study both arts and sciences, students of Forms 4 to 7 are channeled into either arts or sciences (See Table 1 for distribution of participants by major, age, and gender.)

Within each level, students were randomly placed in three equal-sized groups by employing a random number list generated by Lotus 1-2-3 (Lotus Development Corporation 1990) There were twelve groups in total The participants were instructed at the beginning of the program that there would be an after-school work achievement contest (described in detail in the next section) They

Table 1 Distribution of participants by major, age, and gender

		Major (%)			Age		Gender (%)	
Form level	*n*	Arts	Sciences	Undeclared	Range	Mean	Female	Mal
4	36		100		14–17	15	44	56
3	36			100	14–17	14	53	47
2	36			100	13–15	13	50	50
1	36			100	12–14	12	50	50

were mostly from East Kowloon, a city area The majority of their parents were laborers in manufacturing and service sectors

Materials

All students in this study received regular English instruction of forty-minute sessions seven to nine times over a six-day cycle (students of Forms 1, 2, 3, and 4 had nine, eight, seven, and nine English sessions, respectively, per cycle) The random assignment of students into the three treatment groups (leaving each class intact) was meant to factor out possible variations due to different processes of teaching, learning, and interaction that took place in respective classrooms As there was no classroom observation during the period of investigation, a preliminary survey on teachers' approaches to teaching reading and writing was conducted before the start of the study in order to obtain some information on the actual instructional background, instructional conditions, and quality of interaction in each classroom concerned (see Appendix 1 for a summary of the survey)

The study spread over twenty cycles (equivalent to twenty-four weeks of instruction) during the regular school year The pre-test was held in the first cycle of the study, and the post-test in the twentieth

The treatments applied in this study were an input-based reading program and an output-based writing program In the reading program, students were referred to a list of graded materials (simplified classics, original readers, and information-based books in various areas of interest) Both the list and the books were available in the school library In twenty-four weeks the students were required to read eight books and complete eight review forms (requiring minimal writing) during their own free time The form completions made the activity purposeful while at the same time adding some degree of control of students' independent work At the end of every third week, the reviews were graded (using letter grades) on details and persuasiveness The best ones were selected by the researcher and then posted on the class display board The best reviewers were commended in front of the other students to provide positive reinforcement for all The students were encouraged to read books of any genre written in English either on or off the reading list The free choice of books was intended to help maintain interest

Students in the writing program were given eight essay-writing tasks to complete in twenty-four weeks The eight topics covered a variety of content and different genres targeted in the syllabus The essays were impressionistically graded, given brief positive comments by the researcher, and returned to the students Every third week, the best pieces of writing were selected, and the best writers were commended before all of the students as motivation to continue in the study and exert their best efforts The researcher has to acknowledge the apparent contrast between the precise criteria (details and persuasiveness) for the assessment of the reading group's reviews and the impressionistic grading with positive comments for that of the writing group's essays These assessment criteria were selected in the present study to approximate the current practice of

the co-operating school, so that whatever results the study achieved could easily be compared against the local context, and the implications interpreted could be applied to that context

Students in the regular English program plus unrelated (mathematics) enrichment acted as a control group They did some after-school exercises which provided no significant English input, in addition to the current syllabus Eight mathematics assignments, each of which included ten to fifteen multiple-choice problems, were given in twenty-four weeks as the control group treatment The researcher marked and returned the problems to the students Every third week, the paper with the highest score was posted on the class display board, and the best mathematician was praised before the entire class as a form of motivation

Procedure

To introduce the program, the researcher[4] told each class that there would be an after-school work achievement contest in which they were placed into three groups doing different activities Group A Mathematics, Group B Reading, and Group C Writing Each group was told that its best achievers would be named and commended The names of the best achievers later appeared on enlarged name lists serving as Achievement Charts This was supposed to be a form of reinforcement that also created peer pressure to work hard Students were informed of the procedure at the beginning of the first cycle

Since the three groups were assigned the same topic, topic across programs was also controlled The same test was administered to all twelve groups Hence there was no test difference to control across groups In addition, the students were all equally motivated by the suggestion of an 'achievement contest' They were also informed that there would be two writing tests, the first one unmarked until the second one was written, to compare their improvement throughout the year The researcher did not tell them the same topic would be used and tried to avoid sensitizing them to the true purpose of the study Novelty effects, if present, were assumed to be equalized across the twelve groups because of random assignment and because these possible effects existed for all groups The use of MANCOVA[5] (Multiple Analysis of Covariance) with pre-test scores as covariates was another measure to take account of the variation between students before they took part in the treatment programs

For the first two months, the researcher visited the school every two weeks to verify that the teachers were doing what they were supposed to do, according to the design of the experiment From the third month onwards, when the teachers generated more feedback on the contest, the researcher paid weekly visits to the school and kept a record of work, including conversations with the students, the co-operating teachers, and other staff After the completion of the programs, the teachers wrote their evaluation freely (see Appendix 2 for a summary of teachers' program evaluation)

In the first of the twenty cycles, all students were given the same essay topic 'My favourite person' In the twentieth cycle, they were assigned the same essay

topic as the one in the pre-test The students were encouraged to write as much as, and as well as, they could They were allowed thirty minutes to finish the test during an English lesson in their classrooms There were standardized test instructions and exact timing

The essays were randomized and graded impressionistically on five analytic scales—content, organization, vocabulary, language use, and mechanics—by two independent raters[6] to ensure blind rating and inter-rater reliability The raters had no information on which paper was written in the pre- or post-test The grades (sub-totals and totals) were recorded on name lists

The paradigm used was the Jacobs, Zinkgraf, Wormuth, Hartfiel, and Hughey (1981) ESL Composition Profile with a total score ranging between 34 and 100 points The marking protocols were explained to the raters, who then underwent training by going through the exercise on Samples 1 to 3 taken from Jacobs *et al* (1981) The raters later graded five sample test papers randomly taken from the pre- and post-tests Reliability of training has been defined by Jacobs *et al* (1981) as a divergence of ten points or less on the total scores (ranging from 34 to 100 points) between the raters Inter-rater reliability during the training was within the acceptable range of a divergence of less than ten points The exercise with the five papers was redone to check for individual rater reliability, which proved to be at an acceptable level (of a difference in total scores of no more than ten points) between two ratings of the same paper After the actual rating, approximately 7 per cent of the papers previously assigned total scores diverging by more than ten points were re-graded by the two raters The new scores were assigned as the final rating for that paper The scores of the two raters were averaged and entered for statistical analyses

4 ANALYSIS

The first independent variable in this study is the nature of program There are three levels of this factor (A) regular plus unrelated extra practice (the control), (B) regular plus extensive reading, and (C) regular plus frequent writing practice The second independent variable is form level The four levels of this factor include (1) Form 1, (2) Form 2, (3) Form 3, and (4) Form 4 The dependent variable investigated is the post-test score analyzed with its pre-test score as a covariate Twelve cells—three programs by four form levels—represent the design of the research SPSS/PC + (SPSS Incorporation 1990) was used to run all statistical procedures in the present study

Hypotheses were constructed and tested using a MANCOVA design to examine three factors (1) nature of program, (2) form level, and (3) interaction between nature of program and form level, with post-test scores as the dependent variables and pre-test scores as covariates The hypotheses were tested first by examining the overall F ratio Then main effects and interaction were investigated for significant F values MANCOVA was run on the six impression variables (impression total and sub-scales) entered simultaneously

Simple contrasts listing parameters and significance for all possible a posteriori mean comparisons were conducted in the investigation of simple

main effects Subsequent univariate tests were performed on each dependent variable adjusted for its paired covariate Ninety-nine per cent simultaneous confidence intervals (CI) were determined in the interpretation of significance at different levels within a factor for each variable

5 RESULTS

Descriptive statistics

Table 2 reports the means ($\bar{x}$), ranges, standard deviations (s), and number of participants (n) according to the nature of the program, and form level, handling only Impression Totals Tables 3 and 4 show descriptive statistics by program and form level respectively Interpretation of the observed differences depends upon the MANCOVA, determining whether any are statistically significant

Table 2 Descriptive statistics—pre-test and post-test impression

		Program							
		A		B		C		Total (34–100)	
Form level	Statistics	Pre	Post	Pre	Post	Pre	Post	Pre	Post
4	$\bar{x}$	68 33	70 42	68 58	72 71	70 63	71 00	69 18	71 38
	Range	18 00	22 00	16 50	14 00	13 50	12 00	19 50	22 00
	s	5 25	6 69	4 81	4 91	4 44	4 38	4 82	5 35
	n	12	12	12	12	12	12	36	36
3	$\bar{x}$	64 08	66 08	61 92	69 00	65 63	68 96	63 88	68 01
	Range	26 00	24 00	20 50	16 50	18 50	18 00	27 00	24 50
	s	8 85	7 27	6 76	5 05	4 43	4 81	6 89	5 81
	n	12	12	12	12	12	12	36	36
2	$\bar{x}$	64 42	65 00	63 54	70 75	60 83	66 33	62 93	67 36
	Range	11 00	15 00	24 00	18 00	24 00	11 00	29 00	23 00
	s	3 46	3 98	6 66	4 78	7 33	3 31	6 08	4 67
	n	12	12	12	12	12	12	36	36
1	$\bar{x}$	50 71	68 08	53 04	69 79	53 67	66 92	52 47	68 26
	Range	26 50	14 00	22 50	12 00	24 50	23 50	29 50	24 00
	s	8 52	3 64	7 25	3 30	8 60	6 43	8 01	4 69
	n	12	12	12	12	12	12	36	36
Total	$\bar{x}$	61 89	67 40	61 77	70 56	62 69	68 30	62 12	68 75
	Range	38 50	28 00	35 50	24 00	35 00	24 50	40 00	29 00
	s	9 48	5 83	8 42	4 63	8 89	5 06	8 89	5 33
	n	48	48	48	48	48	48	144	144

*$p < 01$
(34–100) = range of possible total scores
A = Mathematics
B = Reading
C = Writing

Table 3 Descriptive statistics—pre-test and post-test impression by program

		Program							
		A		B		C		Grand	
Scale	Statistics	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Cont	$\bar{x}$	19 67	21 80	19 50	23 25	19 65	21 82	19 60	22 29
(13–30)	Range	11 00	13 50	11 50	11 50	12 50	11 50	12 50	13 50
	s	3 91	2 41	3 59	1 73	3 92	2 53	3 78	2 34
	n	48	48	48	48	48	48	144	144
Orga	$\bar{x}$	12 49	13 52	12 42	13 79	12 62	13 76	12 51	13 69
(7–20)	Range	9 00	5 00	7 00	5 50	8 00	5 00	9 00	5 50
	s	2 00	1 20	1 73	1 22	1 93	1 19	1 88	1 20
	n	48	48	48	48	48	48	144	144
Voca	$\bar{x}$	12 54	13 43	12 40	13 66	12 77	13 54	12 57	13 54
(7–20)	Range	7 50	5 00	7 00	3 50	5 50	4 50	8 00	5 00
	s	1 56	1 14	1 42	0 87	1 38	0 98	1 46	1 00
	n	48	48	48	48	48	48	144	144
Lang	$\bar{x}$	13 97	15 20	14 21	16 41	14 41	15 58	14 19	15 73
(5–25)	Range	10 50	8 00	9 50	6 00	8 50	5 50	11 50	9 00
	s	2 33	1 82	2 05	1 46	2 04	1 38	2 14	1 63
	n	48	48	48	48	48	48	144	144
Mech	$\bar{x}$	3 22	3 45	3 25	3 48	3 25	3 59	3 24	3 51
(2–5)	Range	2 00	1 50	2 00	1 00	2 00	1 00	2 00	1 50
	s	0 50	0 42	0 47	0 37	0 44	0 37	0 47	0 39
	n	48	48	48	48	48	48	144	144
ImpT	$\bar{x}$	61 89	67 40	61 77	70 56	62 69	68 30	62 12	68 75
(34–100)	Range	38 50	28 00	35 50	24 00	35 50	24 50	40 00	29 00
	s	9 48	5 83	8 42	4 63	8 89	5 06	8 89	5 33
	n	48	48	48	48	48	48	144	144

*$p < 01$

Cont = Content, Orga = Organization, Voca = Vocabulary, Lang = Language Use, Mech = Mechanics, ImpT = Impression Total, ( ) = Range of possible scores, A = Mathematics, B = Reading, C = Writing

Reliability estimates

Inter-rater reliability was calculated for the final impressionistic scores submitted by the independent raters at the end of the actual rating (or subsequent re-rating) The best possible control over test time, coupled with standardized, detailed test protocols across both pre- and post-tests, was administered to ensure test reliability Reliability estimates were calculated separately for the pre- (09*) and post-test (07*) in terms of Pearson product-moment correlation coefficient The pre-test reliability is higher than the post-test reliability by a

Table 4 Descriptive statistics—pre-test and post-test impression by form level

		Form Level									
		1		2		3		4		Grand	
Scale	Statistics	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Cont	$\bar{x}$	15 74	22 39	20 14	22 26	20 42	21 83	22 13	22 68	19 60	22 29
(13–30)	Range	9 50	10 50	11 50	13 00	11 00	11 50	11 50	13 50	12 50	13 50
	s	3 41	1 99	2 99	2 01	3 07	2 57	2 37	2 70	3 78	2 34
	n	36	36	36	36	36	36	36	36	144	144
Orga	$\bar{x}$	10 74	13 78	12 58	13 39	12 89	13 58	13 82	14 01	12 51	13 69
(7–20)	Range	7 00	4 50	6 00	5 50	8 00	5 00	5 00	4 50	9 00	5 50
	s	1 80	1 14	1 26	1 15	1 74	1 27	1 20	1 19	1 88	1 20
	n	36	36	36	36	36	36	36	36	144	144
Voca	$\bar{x}$	11 06	13 28	12 58	13 14	12 83	13 49	13 81	14 26	12 57	13 54
(7–20)	Range	6 00	3 50	5 00	4 50	3 50	4 00	3 50	3 50	8 00	5 00
	s	1 41	0 87	1 00	0 83	0 90	0 97	0 93	0 98	1 46	1 00
	n	36	36	36	36	36	36	36	36	144	144
Lang	$\bar{x}$	12 08	15 25	14 36	15 22	14 43	15 64	15 90	16 81	14 19	15 73
(5–25)	Range	8 50	5 00	6 50	7 50	8 00	7 00	5 50	6 50	11 50	9 00
	s	2 05	1 22	1 64	1 61	1 59	1 77	1 25	1 42	2 14	1 63
	n	36	36	36	36	36	36	36	36	144	144
Mech	$\bar{x}$	3 57	3 57	3 26	3 38	3 31	3 47	3 53	3 61	3 24	3 51
(2–5)	Range	1 00	1 00	1 00	1 00	2 00	1 50	1 00	1 00	2 00	1 50
	s	0 32	0 32	0 30	0 40	0 45	0 36	0 40	0 43	0 47	0 39
	n	36	36	36	36	36	36	36	36	144	144
ImpT	$\bar{x}$	52 47	68 26	62 93	67 36	63 88	68 01	69 18	71 38	62 12	68 75
(34–100)	Range	29 50	24 00	29 00	23 00	27 00	24 50	19 50	22 00	40 00	29 00
	s	80 12	4 69	6 08	4 67	6 89	5 81	4 82	5 35	8 89	5 33
	n	36	36	36	36	36	36	36	36	144	144

*$p < 01$

Cont = Content, Orga = Organization, Voca = Vocabulary, Lang = Language Use, Mech = Mechanics, ImpT = Impression Total, ( ) = Range of possible scores

margin of 0 02 This fairly big difference might be attributable to the fact that the first essays were on average shorter, more modest, and generally of 'poor' quality, so the raters did not have much to disagree upon, but the second essays were longer and more varied, with some showing obvious improvement, thus inviting more diverse interpretation and judgment on the part of the raters As a result, the correlation of the ratings between the raters in the post-test is not as strong as that in the pre-test

MANCOVA

As shown in Table 5, giving the results of the MANCOVA, nature of program and form level yield significant main effects, while the interaction of the two does not show significant main effects

Table 5 Results of MANCOVA on impression scales—main effects (test Pillais)

Source of variation	Value	Approx *F*	Hypoth *DF*	Error *DF*	Sig of *F*
Within cells regression	0 53	2 49	30 00	630 00	0 00
Form level by nature of program	0 25	0 94	36 00	762 00	0 58
Nature of program	0 32	3 86	12 00	246 00	0 00
Form level	0 34	2 67	18 00	372 00	0 00
Constant	0 48	18 89	6 00	122 00	0 00

*$p < 01$

Gains from pre-test to post-test for each group separately

Univariate tests for effects of nature of program, subsequent to the MANCOVA, were performed on each dependent variable adjusted for its paired covariate Significant changes brought about by the nature of the program are found in content, language use, and impression total scores, as shown in Table 6

Simple contrasts were conducted to determine exactly where the statistical

Table 6 Results of univariate tests for program effects subsequent to MANCOVA on each dependent variable (impression total and sub-scales)

Variable	Hypoth SS	Error SS	Hypoth MS	Error MS	*F*	Sig of *F*
Cont	73 50	535 58	36 75	4 21	8 71	0 00
Orga	1 78	157 04	0 89	1 24	0 72	0 49
Voca	1 44	83 88	0 72	0 66	1 09	0 34
Lang	36 78	190 94	18 39	1 50	12 23	0 00
Mech	0 33	15 29	0 16	0 12	1 36	0 26
ImpT	278 33	2364 85	139 16	18 62	7 47	0 00

*$p < 01$

Cont = Content, Orga = Organization, Voca = Vocabulary, Lang = Language Use, Mech = Mechanics, ImpT = Impression total

changes lie Ninety-nine per cent confidence intervals are reported in Table 7 If the numbers (Lower and Upper) on the ends of a confidence interval do not have zero between them, that is, if both numbers are positive (greater than the group mean and hence a significant rise) or both negative (smaller than the group mean and hence a significant drop), then the null hypothesis can be rejected for that particular dependent variable (Huck, Cormier, and Bounds 1974)

It can be noted from Table 7 that only Program B (Reading) showed significant overall performance in terms of impressionistic total scores It also produced a statistical improvement in content (referring to knowledge, substance, development of thesis, and relevance to assigned topic) and language use (referring to constructions, agreement, tense, number, and word order/function) Program A (Mathematics) showed a significant drop in language use, while Program C (Writing) brought about a significant change in neither the overall performance nor any specific aspect of writing

Table 7 Gains from pre-test to post-test for each group separately—CI for significant impression variables

Contrast	CI	Impression scales		
		Cont	Lang	ImpT
A	Lower	−1 20	−0 83*	−2 70
	Upper	0 19	−0 03*	0 16
B	Lower	0 29*	0 27*	0 50*
	Upper	1 68*	1 07*	3 37*
C	Lower	−1 47	−0 81	−2 69
	Upper	0 51	0 33	1 36

*$p < 01$

Contrasts A, B, C = Mathematics, Reading, Writing Groups respectively, Cont = Content, Lang = Language Use, ImpT = Impression total

Comparing the three groups

Since the F values for this study were significant, the main effects were examined and found to be significant for both nature of program and form level However, the effect due to the interaction of these main effects was not significant Since there was a significant main effect for nature of program and follow-up univariate tests on each dependent scale identified statistically meaningful program effects in content, language use, and impression total (see Table 6 for results), simple contrasts were conducted The resulting CI values are reported in Table 8

Table 8 Comparing the three groups—CI for significant impression variables

Contrast	CI	Impression scales		
		Cont	Lang	ImpT
A–C	Lower	−1 24	−0 88	−3 09
	Upper	1 18	−0 51	1 88
B–C	Lower	0 26*	0 22*	0 12*
	Upper	2 67*	1 61*	5 09*
A–B	Lower	−2 48*	−1 67*	−5 24*
	Upper	−0 50*	−0 53*	−1 18*

*$p < 01$

Contrasts A, B, C = Mathematics, Reading, Writing Groups respectively, Contrast A–C = A versus C (mean for A minus mean for C) Contrast B–C = B versus C (mean for B minus mean for C), Contrast A–B = A versus B (mean for A minus mean for B), Cont = Content, Lang = Language Use, ImpT = Impression total

Table 8 shows that for impression total scores, the mean for Program B (Reading) was significantly higher than those for Program A (Mathematics) and Program C (Writing) The mean for Program A (Mathematics) was not statistically different from that for Program C (Writing)

With reference to the various aspects of writing, both Programs A (Mathematics) and B (Reading) brought about a significant effect in language use, with Program B (Reading) showing a statistically greater effect than Program A (Mathematics), the latter in fact performed statistically worse in the post-test Program B (Reading) also engendered a significantly positive effect in content, which was lacking in the other programs

6 DISCUSSION

The first four of the five hypotheses originally proposed were rejected They are restated below according to the findings of the study

1 There was a significant main effect for nature of program as a factor in writing
2 There was a significant difference between the regular plus mathematics program and the regular plus English enrichment programs the mathematics program was significantly less effective than the reading program on the scale of language use, nevertheless, the mathematics program was not significantly different from the writing program
3 There was a significant difference among the English enrichment programs the reading program was significantly more effective than the writing program

4 There was a significant main effect for form level as a factor in descriptive writing

There was no significant interaction effect between nature of program and form level as predicted in Hypothesis 5

Gains from pre-test to post-test for each group separately

The mathematics group, who received unrelated enrichment and acted as the control in this study, did not show significant overall gain in terms of impressionistic total scores Additionally, they showed a statistical drop in their language use scores Their 'poor' performance in the post-test overall and in various aspects of writing was not significantly different from that of the Form 4 students, both might be caused by a lack of input

The writing program with minimal feedback, which was, by virtue of the design, an output-based approach, may lack the necessary input for descriptive writing It did not show any significant gain

The reading group exhibited significant gains in content and language use, and in overall improvement of the quality of writing. They illustrated no statistical gain in organization, vocabulary, or mechanics One of the reasons for the lack of effectiveness in the area of organization may be that this aspect is more abstract than specific language points Mastery of organizational skills surely involves cognitive development and perhaps explicit teaching

Where the lack of improvement in vocabulary in the reading group (contrary to results reported by Krashen 1989) is concerned, there would seem to be four possible explanations Firstly, the lack of effect found for the vocabulary variable might be attributable to book selection Since the students were provided with a menu of possible items to read and, in addition, were given the option of reading things not included in the menu, it seems possible that what the students in the reading group read could have affected the outcome of the study That is, simplified classics, original readers, and information-based books in various areas of interest were included in the menu, and students were free to choose the types of books they read Consequently, the materials actually read might not expose them to vocabulary they could use in the writing test Secondly, some vocabulary may require elicitation by more focused topics than the one in the pre- and post-tests Thirdly, the reading program may have enriched the reading comprehension of the students None the less, it may take a longer time to internalize reading knowledge/skills and transfer such knowledge/skills to writing, as learners may need multiple exposures to new words before they achieve productive control If the study had investigated receptive vocabulary size and knowledge, some effect, or the lack of it, for the reading program might have been found A fourth possibility is that there may be a ceiling effect on the acquisition of vocabulary in a second language Cumming (1989. 121) argues that high L2 proficiency has an additive effect on writing performance Since the English proficiency of the participants in the present study can generally be described as high elementary to low intermediate, such limited L2 proficiency may impose a ceiling upon the attainments in writing

proficiency in general, and possibly in productive control of vocabulary in particular Therefore, the lack of improvement in vocabulary does not necessarily discount entirely the effectiveness of the reading program in this respect On the research front, future efforts are needed to uncover the source of lack of vocabulary effects in reading programs of the present type

The lack of gains in mechanics may be caused by insufficient input Improvement in spelling may be brought about with more leisure reading as observed by Krashen (1989), and Polak and Krashen (1988) Punctuation and paragraphing may require instruction for improvement to take place

Comparing the three groups

The mathematics group was not significantly different from the writing group in language use or in other aspects of writing Figures in Table 3 suggest some numeric, though statistically non-significant, gains for both groups in the post-test, indicating some positive trends similar to those for the reading group, which achieved statistically significant improvement Therefore caution is needed in formulating pedagogic implications

As indicated by the non-significant interaction between nature of program and form level among all enrichment programs, in terms of its overall performance, the reading program was found to be the most effective in descriptive writing at all form levels dealt with in this study The reading program appeared to be significantly effective in two ways the reading group demonstrated a significant gain in content, which was absent from the mathematics and writing groups, and the reading group showed a statistically greater gain in language use than the other two groups

The present study draws attention to the fact that frequent writing practice without teacher feedback brought about little improvement in writing of non-native speakers Though writing for its own sake has been shown to be valuable in L1 contexts as in Smith's (1984) study, conventional writing extension programs in L2 with minimal teacher support may not work (see Swain 1993 for a discussion of the role of feedback on output) The frequent writing group demonstrated no gains over the mathematics control group, suggesting L2 writing needs instruction It seems likely from the present findings that the value of uninstructed writing in L2 increases as proficiency in L2 increases It is also possible that the students in the writing group concerned did not achieve an L2 level where writing practice with little feedback or consultation might be of value

The advantage for the reading group in the present investigation lends support to Krashen's Input Hypothesis (1980, 1982, 1985, 1989) and strikes a resonant note with Janopoulos (1986), Elley and Mangubhai (1983), and Hafiz and Tudor (1989, 1990, Tudor and Hafiz 1989) in that it demonstrates the contribution of extensive L2 reading to L2 writing Two observations can be made

1 The reading program described here helps improve writing and is thus recommended in the English syllabus

2 The reading program had differential effects on various aspects of writing—more effective in some areas, but less effective in others—therefore, it is inadequate (in its present form) in effecting all-round improvement in writing

The reading program exposed students to an appropriate model of the target language at an appropriate level (see Krashen 1982), it improved general knowledge and thus helped develop content in writing It also exposed students to appropriate models of construction, agreement, tense, number, and word order/function, which strengthened their use of the language Among the mathematics, reading, and writing enrichment programs, reading is thus recommended for improvement in content and language use, as well as overall quality of descriptive writing performance

The lack of improvement in organization, vocabulary, and mechanics may characterize a relatively short reading program like the present one None the less, these aspects of writing may be influenced in the long term through continuous exposure to reading materials Other programs specifically addressing organization, vocabulary, and mechanics may also be needed to bring about improvements in these respects In terms of substance, first, such a program may need to include as reading materials more graded readers (as opposed to unsimplified fiction and information-based material) These are simplified or written in simple English to allow students to carry out large quantities of reading within their limited vocabulary, and this may ultimately widen their vocabulary base As Wodinsky and Nation (1988) put it, simplified texts provide very favourable repetition for vocabulary learning Second, improvement in spelling may be brought about with more leisure reading, whereas punctuation and paragraphing may require instruction for improvement to take place Third, as far as organization is concerned, the program may have to incorporate explicit teaching of material at levels appropriate to the students' cognitive development Exposing learners to reading materials related to specific types of discourse they need to produce may bring additional benefits (cf the results of Flahive and Bailey 1993)

A major constraint of the present study is its focus on intralingual input (influence of L2 reading on L2 writing) without investigating interlingual transfer (influence of L1 on L2 literacy), both of which are sources of L2 literacy skills (Eisterhold 1990) Research in the direction of interlingual transfer and its interaction with intralingual input will certainly help clarify issues of acquisition of descriptive writing skills in second language Indeed more research is needed to find out exactly the best way to modify the reading program described in this study as a means to improve the acquisition of descriptive writing skills Research related to the acquisition of writing skills of other genres, for example, exposition, will also be useful to secondary students Finally, projects involving speakers other than Cantonese and those of other levels of proficiency are needed to further test the generalizability of the present findings

(Revised version received August 1995)

NOTES

[1] In Hong Kong, English is co-official with the Chinese language, English is a medium of instruction in local secondary schools and tertiary institutions Therefore, it cannot be defined as a foreign language Nevertheless, English is rarely functional outside the domains of government, education, trade and commerce, and some professions Its use is severely limited in daily activities, the language cannot be adequately considered a full second language, i e as in a bilingual situation In view of its compartmentalized functions and its status in the territory, English is appropriately described as 'auxiliary' as defined by Luke and Richards (1982)

[2] Prior to the start of the present study, Forms 1, 2, 3, and 4 students took Attainment Tests JS 1A, JS 1B, JS 2, and J 3, respectively These tests were administered by the Hong Kong Education Department as measures of the general English proficiency level of junior secondary students across the territory The raw scores were standardized with a mean of 100 and a standard deviation of 15 ANOVA was performed on the standardized scores of different treatment groups Results showed that there was no significant difference between the treatment groups as entities, or between treatment groups at different form levels This confirms the assumption that the general English proficiency level prior to the start of the study was controlled for by the random assignment of participants into different treatment groups

[3] Form levels in Hong Kong are equivalent to grade levels in the North American school system Forms 1 to 4 are equivalent to what would have been Grades 7 to 10 in the North American system

[4] The researcher of the present study had taught in the school for three years and had resigned by the time the research took place Most of the student participants, except those from Form 1, knew her Thus the novelty effect might not have been as great as that brought about by a researcher unknown to the students

[5] There are two parts to the decision on employing the MANCOVA First, analysis of covariance is preferred over analysis of variance since the groups did not start the academic term with the same pre-test scores, a comparison of the post-test means or a comparison of gain scores alone does not represent a legitimate way to answer the question The post-test means needed to be adjusted to account for the initial differences between the pre-test means (Huck, Cormier, and Bounds 1974, Huitema 1980, Bock 1985, Bryman and Cramer 1990) Second, multiple analysis of variance is more appropriate than multiple runs of one-way analyses of variance Multiple runs of ANCOVA will raise the family error rate above the level of significance and the univariate tests ignore intercorrelations between the dependent variables (Huitema 1980) In view of these, MANCOVA rather than multiple ANCOVAs is selected for the present analysis

[6] One of the raters is a holder of a Certificate in Education, a BA in English, and an MA in Teaching English as a Second Language She had been teaching English in local secondary schools for five years by the time of rating The other rater is a holder of a BA in English Studies and Comparative Literature, and a Diploma in Education She had also been teaching English in local secondary schools for eight years by the time of rating

REFERENCES

Anderson, R. C., P. T Wilson, and L. G. Fielding. 1988 'Growth in reading and how children spend their time outside of school' *Reading Research Quarterly* XXIII/3 285–303

Bock, R. D. 1985 *Multivariate Statistical Methods in Behavioral Research* Scientific Software Inc

Borg, W. R. and **M. D. Gall.** 1989 *Educational Research An Introduction* (5th edn), New York Longman

Bryman, A. and **D. Cramer.** 1990 *Quantitative Data Analyses for Social Scientists* New York Routledge

Cumming, A. 1989 'Writing expertise and second language proficiency' *Language Learning* 39/1 81–141

Davies, N. F. 1983 'The receptive way—to activate learning' *System* 3 245–8

Eisterhold, J. C 1990 'Reading–writing connections Toward a description for second language learners' in B Kroll (ed) *Second Language Writing Research Insights for the Classroom* Cambridge University Press

Elley, W. B. 1991 'Acquiring literacy in a second language The effect of book-based programs' *Language Learning* 41/3 375–411

Elley, W. B. and **F. Mangubhai** 1983 'The impact of reading on second language learning' *Reading Research Quarterly* 19 53–67

Fader, D. and **E. McNeil.** 1968 *Hooked on Books Program and Proof* New York Putnam

Faerch, C. and **G. Kasper.** (eds) 1985 'Language learning under classroom conditions' A special issue of *Studies in Second Language Acquisition* 7/2

Flahive, D. and **N. Bailey.** 1993 'Exploring reading/writing relationships in adult second language learners' in J Carson and I Leki (eds) *Reading in the Composition Classroom Second Language Perspectives* Boston, MA Heinle & Heinle Publishers

Hafiz, F. M and **I. Tudor.** 1989 'Extensive reading and the development of language skills in an L2' *ELT Journal* 43 4–13

Hafiz, F. M and **I. Tudor.** 1990 'Graded readers as an input medium in L2 learning' *System* 18 31–42

Hong Kong Education Department. 1983 *Syllabus for English (Forms I–V) One of a Series of Syllabuses Recommended by the Curriculum Development Committee, Hong Kong* Hong Kong Government

Huck, S. W., W. H. Cormier, and **W. G. Bounds Jr.** 1974 *Reading Statistics and Research* New York Harper Collins Publishers

Huitema, B. E. 1980 *The Analysis of Covariance and Alternatives* Chichester John Wiley & Sons

Jacobs, H L, S. A. Zinkgraf, D. R. Wormuth, V. F. Hartfiel, and **J. B Hughey** 1981 *Testing ESL Composition* Rowley, MA Newbury House

Janopoulos, M. 1986 'The relationship of pleasure reading and second language writing proficiency' *TESOL Quarterly* 20/4 763–8

Krashen, S. D. 1980 'The input hypothesis' in J Alatis (ed) *Current Issues in Bilingual Education* Washington, DC Georgetown University Round Table on Languages and Linguistics

Krashen, S. D. 1982 *Principles and Practice in Second Language Acquisition* Oxford Pergamon

Krashen, S. D. 1984 *Writing* Oxford Pergamon

Krashen, S. D. 1985 *The Input Hypothesis Issues and Implications* Harlow Longman

Krashen. S. D 1989 'We acquire vocabulary and spelling by reading Additional evidence for the input hypothesis' *Modern Language Journal* 73 440–64

Krashen, S. D. and **T. D Terrell.** 1983 *The Natural Approach* Oxford Pergamon/ Alemany

Krathwohl, D R. 1985 *Social and Behavioral Science Research A New Framework for Conceptualizing, Implementing, and Evaluating Research Studies* San Francisco, CA Jossey-Bass Publishers

Krathwohl, D. R. 1993 *Methods of Educational and Social Science Research An Integrated Approach* New York Longman

Larsen-Freeman, D. 1983 'Second language acquisition Getting the whole picture' in K Bailey, M Long, and S Peck (eds) 1983 *Second Language Acquisition Studies* Rowley, MA Newbury House

Lotus Development Corporation. 1990 *Lotus 1-2-3 Release 2 4* [Computer program] Cambridge, MA Author

Luke, K. K. and **J C. Richards** 1982 'English in Hong Kong Functions and status' *English World-wide* 3/1 47–64

McCracken, R. 1971 'Initiating sustained silent reading' *Journal of Reading* 14 521–4

Nuttall, C. 1982 *Teaching Reading Skills in a Foreign Language* London Heinemann

Pica, T and **C. Doughty** 1985 'Input and interaction in the communicative language classroom A comparison of teacher-fronted and group activities' in S M Gass and C G Madden (eds) 1985 *Input in Second Language Acquisition* Rowley, MA Newbury House

Polak, J. and **S Krashen.** 1988 'Do we need to teach spelling? The relationship between spelling and vocabulary reading among community college ESL students' *TESOL Quarterly* 22 141–6

Schleppegrell, M 1984 'Using input methods to improve writing skills' *System* 12 287–92

Smith, V. H. 1984 'Teaching beginning writing vs nurturing beginning writers Paper presented at the Fourth Annual Meeting of the Indiana Teachers of Writing Conference (Indianapolis, IN, September 28–29, 1984)

SPSS Incorporation. 1990 *SPSS/PC+ 4 0* [Computer program] Chicago, IL Author

Stewart, M. and **Cheung, M. Y M** 1989 'Introducing a process approach in the teaching of writing in Hong Kong' *Institute of Language in Education Journal* 6 41–8

Swain, M 1985 'Communicative competence Some roles of comprehensible input and comprehensible output in its development' in S M Gass and C G Madden (eds) *Input in Second Language Acquisition* Rowley, MA Newbury House

Swain, M 1993 'The Output Hypothesis Just speaking and writing aren't enough' *Canadian Modern Language Review* 50/1 158–64

Thorndike, R. L. 1973 *Reading comprehension education in 15 countries (International Studies in Evaluation III)* New York Halsted Press

Tudor, I. and **F. M Hafiz.** 1989 'Extensive reading as a means of input in L2 learning' *Journal of Research in Reading* 12 164–78

Wodinsky, M. and **P. Nation.** 1988 'Learning from graded readers' *Reading in a Foreign Language* 5 155–61

APPENDIX 1

Summary of preliminary survey on teachers' approach to teaching reading and writing

Open-ended questionnaire

The five co-operating teachers of English (one chairing the English Subject Panel in the school concerned) in the present study completed an open-ended questionnaire before the start of the programs They reported on (1) their methodology of teaching reading

and writing, (2) the type of instruction they gave while teaching class readers (books specified on the school booklist or recommended by the teacher to be read at home), and (3) the instruction they gave while teaching in-class composition writing and giving take-home writing assignments The main points of their practices, which the teachers characterized anecdotally as 'not uncommon in Hong Kong', are summarized below as background information on the setting of L2 literacy education in which the research reported here was conducted

According to the teachers' self-reports, teaching reading might start with the teacher summarizing the story in class, having assigned home reading before coming to class In-class story telling, dramatization, and questioning might be used to ensure comprehension Consolidation exercises might take the form of worksheets for upper form students As the teachers assign home readers to lower form students, they would make suggestions as to what kinds of book to borrow and how to complete book report forms *Students in the upper forms are encouraged to choose whatever suits their level and to* guess the meaning of unknown words

In the teaching of writing, teacher input apparently varies greatly across form levels and moves from very controlled to less controlled situations At Form 1, the teacher reported providing lots of language support using guiding questions and answers, at Form 2, grammar basics and vocabulary would be provided, at Form 3, discussion of organization is involved, and at Form 4, pre-writing activities including an analysis of the topic are the routine before the actual writing task While the Form 2 English teacher does not give take-home writing assignments, she commented that such assignments are prescribed and thus given out to her students under the general agreement with the English Subject Panel Chair

The co-operating English teachers of the school where the present study is based indicated that some students tend to read more than to write, as their receptive skills are superior to their productive skills Another observation made by the English teachers is that it usually takes students longer to finish a book (two to three weeks, because they like to read only a few pages every day in a leisurely manner) than to write a composition, because they are inclined to plan, draft, and edit their work intensively (sometimes, in a week)

Post-questionnaire interview

It seems that teachers in the co-operating school generally take an eclectic approach to teaching reading and writing for two mains reasons (1) teachers' lack of training in teaching these skills, (2) students' lack of training in them (a possible consequence of the first reason) and their low proficiency in English These reasons may also become the causes of students' difficulties in ESL reading and writing Such observations are not uncommon in other schools over the territory (see Stewart and Cheung 1989)

APPENDIX 2
Summary of teachers' program evaluation

Teachers generally had a positive evaluation of the contest They found the structure of the contest and the researcher's regular school visits effective

At Form 1 level, some students were reported to have formed a reading habit, some were said to have improved their writing This teacher observation was consistent with the significant overall impressionistic gain Teachers at all levels (except Form 2) reported that the enrichment programs brought positive effects on learning attitudes in general

Students at Form 3 were reported to have obtained 'a sense of doing something which was done well and checked thoroughly' Students at Form 4 had, in a teacher's words, 'more participation in academic activities' At Form 1 level, to quote from teachers' evaluation, 'academic standards in general showed signs of improvement'

Teachers at Form 3 noted that students were losing enthusiasm towards the end of the year, possibly because of heavy school work Teachers at other levels suggested control over the reading group for better effects Some proposed more severe punishment for a few lazy students (Form 1) Others suggested setting lower limits to pages to be read and requiring periodic writing out of the story (Form 2), or handing out copies of reading lists and giving more selection guidelines (Form 4) These numerous suggestions indicate that for extensive reading to be feasible in secondary schools at Band 3 or 4, more control is needed, at least at the preliminary stage when an entirely favorable environment is not available and when students have not yet formed a reading habit

To summarize, remarks made by the co-operating school in the present study suggest effects of the contest on general academic performance There were also remarks which indicate the possible effects of the enrichment programs on improving English reading and writing The new phenomenon that many students borrowed English books from the school library might be a consequence of the implementation of the reading enrichment program in addition to the school's regular reading scheme This might indicate that the enthusiasm and goodwill factors were also effects of this type of enrichment program