
Title	Testing students' bilingual ability in a bilingual manner
Author	Kay Cheng Soh
Source	<i>Journal of Linguistics and Language Teaching</i> , Volume 2 (2010), Issue 2, pages 253-266

This document may be used for private study or research purpose only. This document or any part of it may not be duplicated and/or distributed without permission of the copyright owner.

The Singapore Copyright Act applies to the use of this document.

Testing Students' Bilingual Ability in a Bilingual Manner

Kay Cheng Soh (Singapore)

Abstract

Bilingual ability of students has traditionally been tested by using separate monolingual tests. Such tests more often than not differ in both substantive and linguistic content and task format, thus testing something different from one another. This mode of assessing bilingual ability introduces more error variance due to the extraneous factors and thus under-estimates the students' ability to function with two languages. Bilingual tests can be constructed simply by bringing two languages into one testing task using one language for the questions and the other for the answers (options in the case of multiple-choice items). The results of testing a group of Grade 3-5 students in Singapore support this contention. Implications for curriculum and instruction for developing bilingual ability are discussed.

Key words: tests, bilingual ability, bilingual testing, English, Chinese

1 Introduction

“The concept of bilingualism seems at first sight to be non-problematical.” So Hamers & Blanc (2000, p.1) begin their discussion on the definition of bilingualism. They soon, after canvassing some definitions, conclude that

All these definitions, which range from a native-like competence in two languages to a minimal proficiency in a second language, raise a number of theoretical and methodological difficulties.

Indeed, trying to define bilingualism is an unrewarding task when there can be as many definitions as language researchers who have to deal with a seemingly simple concept used in a great variety of contexts.

2 Additive versus Interactive Testing

In the school context, bilingual ability is, more often than not, taken to mean the ability to understand, speak and possibly write in two languages at an acceptable level. The two languages of bilingual learners are kept apart in the learning process as if they were learning two entirely unrelated sets of linguistic knowledge and skills. Teachers may go to the extent of preventing, discouraging, and even punishing the *interactive* use of the two languages, tacitly assuming that such interaction is harmful and leads to interference or errors. In this sense, *bilingualism* means exactly what its literal meaning suggests: the two languages are regarded as additive, i.e. L1 + L2.

This view has not only influenced the curriculum and pedagogy embracing two languages. It also influences the way bilingual ability is assessed. Normally, bilingual learners are tested for their bilingual ability through the use of two *monolingual* tests. Each test only uses one of the two languages being learnt and is usually set by different teachers on different substantive content and measures different parts of linguistic knowledge and different skills. Thus, a student taking such a test only needs to use one of his two languages at a time. At such testing, he is supposed to (and should be able to) switch off the language that is not being tested. A student's bilingual ability is then determined by his scores for the two independent monolingual tests. This being the case, for a group of students, if their scores for the two monolingual tests are correlated, the correlation is seen as incidental and not of much pedagogical interest.

An alternative view of bilingual ability is the *interactive* perspective in which the two languages of bilingual learners are supposed to be interacting. Cross-language reference is not seen as a source of problems and is, on the contrary, encouraged; code-switching or code-mixing between languages is not condemned as a bane but welcome as a boon. Reference across languages is considered as a strategy to facilitate and reinforce the learning of both languages and, hence, enriches them. In

recent years, this view - i.e. turning a supposed liability (of inter-language interference) into an asset (of positive transfer of training) - has become increasingly accepted. Theoretically, this is congruent to Paivio-Resrocherss' model of bilingual dual-coding (1980) which is an extension of Paivio's (1971) model of dual-coding. The validity of bilingual dual-coding has received much support of empirical evidence in studies involving European and Asian languages (Soh, 2010).

3 Correlations between L1 and L2

As mentioned above, the correlation between scores for two monolingual tests is of little interest to language educators who consider the bilingual learners' two languages as separate and discrete. If any, it is the probable negative correlation which is usually highlighted with instances of some relatively minor interference or errors, e.g. Chinese-speaking learners tend to leave out plural indicators when speaking or writing English. Nevertheless, it is well-documented that intra-language errors by far outnumber inter-language ones (Soh 1985), indicating that difficulties in learning a second language mainly come from *this very* language itself rather than from *the other* language being learnt concurrently. Conversely, inter-language correlation implies that there could be a facilitating effect from the one language to the other, and interaction across languages can be capitalized to benefit bilingual learners' efforts to develop bilingual ability. This is illustrated in the recent studies summarized below.

For a group of undergraduates (N = 119), Mostafa (2002) reported a correlation of 0.34 between their scores for Malay (L1) and English (L2). These scores were obtained by using two one-word vocabulary tests in the two languages. Kohro (2009) assessed the compositions of a group of 19 English majors. They were asked to write two narrative compositions on personal experiences, one in Japanese and the other in English. The correlation was 0.30. Ito's study (2009) involved 317 Japanese undergraduates learning English as a second language. The subjects were classified

into five level of L2 proficiency. It is interesting that the magnitude of correlations between L1 and L2 writing increased from a low and non-significant 0.13 to a moderate and statistically significant 0.50, confirming that there was an L2 threshold below which transfer could not be found. Alsamadani (2010) found a correlation of 0.64 between writing in Arabic and English in response to a scenario about going overseas on scholarship. The study involved 35 male undergraduates in an English Language department. The essays were rated on content, organization, vocabulary, language use, and mechanics.

At the lower end of the educational scale, Bialystok, Luk & Kwan (2005) tested three groups of Grade-One children (N = 132) with the Forward Digit Span subtest of the Wechsler Intelligence Scale for Children (Revised) in their two languages (Chinese-English, Hebrew-English, and Spanish-English). On this measure of short-term verbal memory, the correlations are 0.48, 0.54, and 0.38 for the language pairs. These correlations indicate that children could express in *another* language what they could express in one language. Hayashi (2005) reported a study conducted with 4th and 5th Graders in three bilingual education programs: an English Immersion Program in Japan, a Japanese Bilingual Program in California, and a Transitional Bilingual Education Program in Massachusetts. Children were tested for their holistic oral skills with items taken from the Bilingual Language Assessment Record. Written skills were assessed through the writing of one-page English essays on the topics “The Most Exciting Experience in My Life” in Japanese and “What I Enjoyed Recently” in English. Contrary to expectation, there were negative correlations between the scores for responses in Japanese and English, -0.72 for oral skills and -0.39 for writing skills ($p < 0.1$). It is not known whether it was differences in the programs, children’s pre-program English proficiency, or instruments (essay topics) that contributed to these findings. Although the direction of correlation comes as a surprise, the interaction found between languages is real and sizeable.

In a review of some product-oriented studies on reading comprehension, Yamashita (n.d.) found only low to moderate correlations between L1 and L2 abilities, with correlation coefficients varying from 0.20 to 0.64 for six studies published between

1992 and 1999. This leads to the conclusion that “the importance of [the relationship between – Soh] reading ability in two languages has been less acknowledged” (Yamashita n.d.: 281).”

First languages studied in the study programs cited above include Arabic, Malay, Hebrew, and Japanese, with English being the L2. Both university undergraduates and elementary pupils were involved, while high school students were obviously absent for unknown reasons. Language tasks used to test bilingual ability mainly included writing skills though vocabulary and reading were also tested. It is difficult to come to any definitive conclusion with such a limited number of studies; the correlations reported, however, are invariably low with occasional moderate ones.

Against the background of these low to moderate L1-L2 correlations, a question can be asked whether the relationship between the two languages of bilingual learners has been under-estimated because of the way it has been tested.

4 Bilingual Testing

In the school context, the pupils’ bilingual ability is usually tested by means of two monolingual tests which are based on different substantive content and assess different parts of linguistic knowledge and different skills and by means of different tasks, i.e. items and question formats. In this sense, the assessment of bilingual ability is, in fact, an assessment of the ability to function in *in* two languages (comparable to the ability to respond to tests of two different subjects in the curriculum) and not an assessment of the ability to function *with two languages*. From the measurement perspective, differences in substantive and linguistic content and task formats (i.e. item types and scoring procedure) introduce into the test results error variance specific to the testing procedures leading to a deflation of the inter-language correlation as an indicator of bilingual ability.

Assessment basically involves three steps: (1) presenting a task to the student to elicit (2) his response for (3) the judgment of its acceptability. While the last

(judgment) is done ‘behind the scene’, unseen by the student, the first two involve him explicitly. With two languages being involved in bilingual ability, there are four combinations as shown in figure 1 below (adapted from Soh, 1985).

As shown in Figure 1, when both the task (question or item stem) and the response (option) are in the *same* language, the test is an L1-L1 or an L2-L2 monolingual test, usually used for assessing bilingual learners’ abilities in two languages considered separately in the school context. This is also the kind of test used in those studies cited above. When taking a monolingual test, the student only needs one language and is supposed to switch off the other language in his mind. The ability of the other language is assumed to have nothing to do with the student’s performance. The results for two monolingual tests are treated separately without cross-language reference

On the other hand, and in contrast, when the task and the response involve two different languages, the test is bilingual (L1-L2 or L2-L1) in nature, since the student needs both languages in the same process when attempting the task. When taking a bilingual test, the student is assessed for his ability to make use of information given in one language and responds in the other language. In short, his ability to function interactively with two languages is assessed. One presumed advantage of this form of testing is that substantive content and linguistic content are kept constant across languages, by which the effect of extraneous factors is reduced and, hence, error variance is minimized. This approach obviously is in contrast with the usual method of assessing bilingual ability in which two separate monolingual tests are employed.

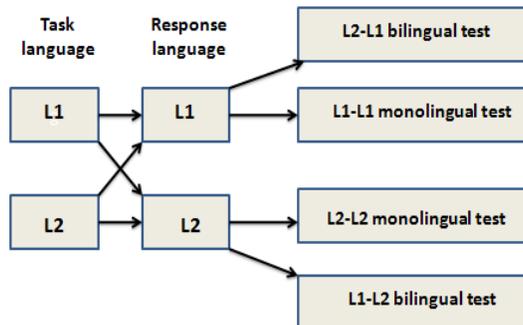


Fig.1: Combinations of languages and testing

Shown below are two parallel items of the English-Chinese and the Chinese-English bilingual tests. For the first sample item, the stem poses a question in English for which options in Chinese are provided. Thus, within the same task, both languages are involved and the student needs to code-switch at the word level in the context of the item-stem to answer the question. This is different from the usual monolingual test for which no bilingual ability is called for:

Kong Wah and Ali are _____. They always play together.

1. 兄弟
2. 姐妹
3. 朋友
4. 亲戚

The second sample item below has been taken from the Chinese-English bilingual test. The thinking process in answering this question is just the opposite of that for the sample item above:

光华和阿里是_____. 他们常在一起玩耍。

1. brothers
2. sisters
3. friends
4. relatives

5 An Exploratory Study

The idea of interactive testing was trailed with 211 Grades 3 to 5 boys and girls in Singapore. There were ethnic Chinese children who mainly spoke Chinese (or one of its variants) at home but went through an English-based curriculum in which all subjects were learnt through using English except the Chinese Language. The

proportion of curriculum time was about 85% of English and 15% of Chinese.

They were tested by means of two monolingual and two bilingual vocabulary tests in English and Chinese. Following the model as shown in Figure 1, the monolingual English test was designed first, and it consisted of 65 four-option multiple-choice items. Next, it was translated to form the monolingual Chinese test. Then, the items in one language were matched with the respective options in the other language, which resulted in an English-Chinese and a Chinese-English bilingual test. Thus, the substantive and linguistic content and task format were controlled among the four resultant tests. With both the substantive content and linguistic content equalized, correlations greater than those reported in the literature were expected, for the reduction of error variance.

Table 1 shows the means (and standard deviations) for the four tests, by sex, by level, and for the sample as a whole. Certain general patterns of the performance are of interest. Firstly, as would be expected, the means tend to increase with grade level. Secondly, girls tended to score higher than boys did. Thirdly, there were not many differences in the means of the four tests at grade level. This suggests that, they were generally balanced in the two languages:

Level	Sex	N	E-E	C-C	E-C	C-E
Primary 5	Boys	21	54.6	54.5	54.2	57.5
Primary 5	Girls	40	55.6	55.5	54.9	55.2
Primary 4	Boys	40	40.7	46.8	46.4	46.5
Primary 4	Girls	40	53.7	53.7	49.3	53.4
Primary 3	Boys	40	35.3	34.8	34.9	34.8
Primary 3	Girls	40	42.8	42.9	36.2	36.3
All		221	46.5	47.5	45.3	46.4

Note: E-E = English-English monolingual test; C-C=Chinese-Chinese monolingual test;
E-C = English-Chinese bilingual test; C-E=Chinese-English bilingual test.

Tab. 1: Means of the Monolingual and Bilingual Tests

Table 2 shows the correlations between the two English-English and Chinese-Chinese monolingual tests. As shown in the table, correlations are high, varying from

0.85 to 0.94, with an average of 0.90. These data indicate that the two monolingual tests shared variances from 72% up to 89%, with an average of 86%. Compared with the L1-L2 correlations reported by the studies cited above, these correlations are obviously much higher, confirming the contention that inter-language correlations have been under-estimated due to error variance caused by differences in substantive and linguistic content as well as task formats as found in the conventional method of estimating bilingual ability:

Level	Sex	N	r	r-squared
Grade 5	Boys	21	0.85	72%
Grade 5	Girls	40	0.88	77%
Grade 4	Boys	40	0.90	81%
Grade 4	Girls	40	0.94	89%
Grade 3	Boys	40	0.88	78%
Grade 3	Girls	40	0.93	86%
All		221	0.90	81%

Tab. 2: Correlations between Bilingual Tests

It is interesting, then, to find out the extent with which performance on a monolingual test predicts performance on a bilingual test and, further, how much performance on another monolingual test enhances the prediction. This was studied through regression analyses, first with one monolingual test as a predictor and then by adding another monolingual test as a second predictor.

Table 3 shows the result for the English-Chinese bilingual test. The correlations between English-English monolingual test scores and English-Chinese bilingual test scores vary from 0.76 to 0.88, with an average of 0.83. These indicate that 63% to 77% of the bilingual test variances were predicted by the monolingual test, with an average of 68%. When the Chinese-Chinese monolingual test scores were added to the prediction, there was a general increase in the variances predicted. The additional variances predicted varied from 0% to 14%, with an average of 6%:

Level	Sex	N	Predicting E-C scores with E-E scores		Predicting E-C scores with E-E and C-C scores		Increase in variance predicted
Grade 5	Boys	21	0.76	57.8%	0.85	72.1%	14.3%
Grade 5	Girls	40	0.79	62.6%	0.81	65.3%	2.7%
Grade 4	Boys	40	0.82	67.7%	0.82	67.6%	-0.2%
Grade 4	Girls	40	0.82	66.7%	0.89	78.5%	11.8%
Grade 3	Boys	40	0.86	73.4%	0.88	77.6%	4.2%
Grade 3	Girls	40	0.88	76.6%	0.91	82.3%	5.7%
All		221	0.83	68.3%	0.86	74.0%	5.7%

Tab. 3: Predicting English-Chinese Bilingual Ability from Monolingual Tests

Table 4 shows the result for the Chinese-English bilingual test. As shown, the unstandardized regression coefficients for English-English monolingual test scores, predicting Chinese-English bilingual test scores, vary from 0.61 to 0.88, with an average of 0.83, indicating predicted variances from 37% to 77%, with an average of 65%. When the Chinese-Chinese monolingual test scores were added to the regression, there was a general increase in the magnitudes in the multiple regression coefficients. In contrast, the magnitudes of the predicted variance increased, varying from 0% to 17%, with an average of 5%:

Level	Sex	N	Predicting C-E scores with E-E scores		Predicting C-E scores with E-E and C-C scores		Increase in variance predicted
Grade 5	Boys	21	0.83	68.2%	0.89	78.5%	10.3%
Grade 5	Girls	40	0.61	37.2%	0.74	54.5%	17.3%
Grade 4	Boys	40	0.87	75.9%	0.90	80.5%	4.6%
Grade 4	Girls	40	0.67	44.4%	0.67	44.4%	0.0%
Grade 3	Boys	40	0.88	76.9%	0.88	76.9%	0.0%
Grade 3	Girls	40	0.93	85.9%	0.93	85.9%	0.0%
All		221	0.80	64.5%	0.83	69.4%	5.0%

Table 4. Predicting Chinese-English Bilingual ability from Monolingual Tests

By way of summary, the following findings can be accentuated:

- When young bilingual learners were tested for their bilingual ability through the use of bilingual tests, they were able to respond with both languages with almost equal effectiveness.
- When substantive content, linguistic content, and task format were controlled, the L1-L2 correlations were definitely higher than those reported in studies in which these factors were not equalized in the two language tests.
- The students' performance in the bilingual tests (which required using information in one language to solve a problem posed in the other language) can largely be predicted by their ability in either language.

6 Further Implications

It is a cliché that the total is more than the sum of its parts. Putting together the scores for two monolingual tests constructed and administered independently may not really assess a student's bilingual ability. There may be something not manifested in the scores – the ability to use two languages interactively or bilingually. This may be the key of bilingualism. The findings of this study support the contention that the correlations between scores for two monolingual tests under-estimate the bilingual ability, that is, the ability to use two languages in an interactive manner.

In the school context, languages concurrently learnt are kept apart (just like the various subjects are) both in instruction and assessment. In contrast, in real life, bilinguals often find themselves having to re-code information or code-switch from one to the other of their two languages. This process may well reflect the L1-L2 linkages in the Paivio-Resrochers model of bilingual dual-coding (Soh 2010). The ability to do so effectively and efficiently is a sign of true bilingualism.

This study is admittedly exploratory in nature and relatively narrow in scope. As is true for many similar studies, the conclusion of the present study is not a final one,

and more research is indicated. It will therefore be beneficial, both for theoretical and practical purposes, to extend the study to the higher linguistic levels of *phrase*, *sentence*, and *discourse*.

However, as building up a sizeable vocabulary is, at least in the early stages of learning, always the first and main target of L2 teaching, the findings of this study have considerable implications for curriculum and instruction. Firstly, to maximize the L1-L2 linkage for a more effective development of bilingual ability, the syllabi for the two languages need to be coordinated or correlated so as to cover similar substantive and linguistic content where feasible. This will influence the textbooks based on the syllabi, and, as a result, students' learning will get oriented towards cross-language referencing, leading to the development of the ability to use their languages interactively or bilingually.

An extension of this bilingual vocabulary development is that language educators need to consciously - and consistently – promote cross-language referencing in their lessons and assignments in order to maximize the opportunity afforded by the coordinated syllabuses and textbooks. Thus, instead of avoiding cross-language referencing, as has been traditionally done with the assumption of it causing inter-language errors, language educators will capitalize on the opportunity of mutual reinforcement and facilitation of the languages. Assessing students with two monolingual tests shows the extent to which they are able to use their languages separately, but it does not show to what extent they can interactively use the languages. To truly assess students' bilingual ability, they need to be tested *bilingually* with bilingual tests.

References

- Alsamadani H. A. (2010). The relationship between Saudi EFL students' writing competence, L1 writing proficiency, and self-regulation. *European Journal of Social Sciences*, 16 (1), pp. 53-63.
- Bialystok E., Luk G. & Kwan E. (2005). Bilingualism, Biliteracy, and Learning to Read: Interactions Among Languages and Writing Systems. *Scientific Studies of Reading*, 9(1), pp. 43-61.
- Hamers J. F. & Blanc M. H. (2000). *Bilinguality and bilingualism, 2nd Edition*. Cambridge: Cambridge University Press.
- Hayashi A. (2005). Japanese English bilingual children in three different educational environments. In James Cohen, Kara T. McAlister, Kellie Rolstad, and Jeff MacSwan (eds.). *Proceedings of the Fourth International Symposium on Bilingualism*. MA: Cascadilla Press Somerville, pp. 1010-1033.
- Ito F. (2009). Threshold to Transfer of Writing Skills from L1 to L2. Paper presented at the Symposium on Second Language Writing. Nagoya, Japan, 2007. ERIC on-line submission.
- Kohro Y. (2009). A contrastive study between L1 and L2 compositions: Focusing on global text structure, composition quality, and variables in L2 writing. *Dialogue* 8, pp. 1-19.
- Mostafa N. A. B. (2002). Bilingualism and the academic achievement of Malay students. Paper presented to the International Conference IPBA 24-26 September 2002.
- Paivio A. (1971). *Imagery and Verbal Processes*. New York: Holt, Rinehart & Winston. (Reprinted by Lawrence Erlbaum Associates, Hillsdale, NJ, 1979).
- Paivio A. and Resrocherss A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of psychology, Review of Canadian Psychology*, 34(4), pp. 388-399.
- Soh K. C. (1985). *Code Switching of Primary Children*. Unpublished doctoral thesis. Singapore: National University of Singapore.
- Soh K. C. (2010). Bilingual dual-coding and code-switching: implications for L1 in L2 Learning. *Journal of Linguistics and Language Teaching*, 1(2), pp. 271-295.
- Yamashita J. (n.d.). Influence of L1 reading on L2 reading: Different perspectives for the process and product of reading. *Essay on Language and Culture (言语文化论文集)*, 23 (2), pp. 271-283.

Author:

Dr. Kay Cheng Soh

50 Lorong 40 Geylang #07-29

The Sunny Spring

Singapore 398074

E-Mail: sohkc@singnet.com.sg