

A Short Test of English Silent Word Reading for English Language Learners

Sylvia C. Kalindi **Catherine McBride*** **Shingfong Chan**

The Chinese University of Hong Kong, Hong Kong

Kien Hoa Kevin Chung **Chia-Ying Lee**

The Hong Kong Institute of Education, Hong Kong Academia Sinica, Taiwan

Urs Maurer **Xiuhong Tong**

The Chinese University of Hong Kong, Hong Kong

We developed a test of English silent word reading, following work by Mather, Hammill, Allen and Roberts (2004) and Bell, McCallum, Krik, Fuller, and McCane-Bowling (2007), in order to tap Hong Kong Chinese children's reading of English as a foreign language. We created one subtest of individual word reading and another of word reading contextualized within sentences; together, these tests require no more than 10 minutes for administration. In Study 1, we administered the entire test to 552 second grade Hong Kong Chinese children between the ages of 70 and 121 months old, from five different primary schools. The association between the subtests of English silent word reading and contextual reading was positively correlated (.78). In Study 2, 77 Hong Kong Chinese second graders were tested on our newly developed English silent word reading test, together with non-verbal IQ, an English word reading and a Chinese character recognition test (both read aloud). With age and non-verbal IQ statistically controlled, there was a significant correlation between English silent word reading and the more standard English word reading, read aloud, (.78); the association between English silent word reading and Chinese character recognition was also positively correlated (.49). This newly created test is a quick and reliable measure, suitable for both educators and researchers to use to identify poor readers who learn English as a foreign or second language.

Keywords: English Language Learners, reading ability, short test, silent reading test

In many regions in the world, including Hong Kong, Zambia, and India, governments have adopted English as their official language, making it the medium of instruction in classrooms, although the majority of citizens rarely use English. Typically, in such regions, people use their native language for most communication outside of school or formal

contexts. For example in Hong Kong, which is largely monolingual (Cheung et al., 2010; Chung, McBride-Chang, Cheung, & Wong, 2011), both children and adults conduct most daily communication in Chinese. In this case, the manner in which children learn English as a foreign language is considerably different from those discussed in Western contexts (e.g. McBride-Chang, 2004) where the school and peer language is typically English (Abella, Urrutia, & Shneyderman, 2005; Chiappe &

* **Corresponding Author:** Catherine McBride, Professor, Psychology Department, The Chinese University of Hong Kong, Shatin, Hong Kong. Email: cmcbride@psy.cuhk.edu.hk

Siegel, 1999). In Hong Kong, English is truly a foreign language as its use is limited to formal contexts (Cheung et al., 2010; Lin, Wong, & McBride-Chang, 2012). In spite of this characteristic, English literacy instruction (including Chinese) begins early, at 3.5 years in kindergarten, and society places a strong emphasis on English learning (Lin et al., 2012; McBride-Chang, 2004).

Over the years, the Hong Kong government has invested considerable resources, with numerous government initiatives, aimed at enhancing students' reading performance. In this regard, Chinese and English are part of the three 'core' subjects (including math) that are widely tested and emphasized across primary and secondary education (Lin et al., 2012). Many schools in Hong Kong and elsewhere use English as a medium of instruction for most or all other subjects from early on in primary school. Given the critical importance of English as a foreign language in such societies, identification of those children with specific difficulties in learning English may be important in order to ensure that they get help before they fall too far behind (e.g., McBride-Chang et al., 2013; McBride-Chang, Liu, Wong, Wong, & Shu, 2012). A full assessment method is usually the best way of evaluating the reading performance of children. However this method has been observed to be very expensive and time consuming and may require a professional to be administered (Sattler, 2001). In Hong Kong, elementary teachers use a various assessment methods including dictation, reading comprehension, writing composition, listening and oral language tests to evaluate English reading among pupils. In as much as these evaluation measures touch on different aspects of language skill development, they can also be time consuming as they may require one-to-one testing, which can be cumbersome for the teachers. On the other hand researchers mainly use comprehension as well as word recognition tasks to evaluate elementary children's English reading skills. In this regard, given that some of these tasks may also be relatively fast to administer, we thought our short test of word

reading for English language learners could potentially not only supplement the already available means of assessment but also provide quick feedback and be used to research the nature of possible reading difficulties in the participants.

At this point it is important to give a theoretical background on language proficiency among bilingual students before we look at assessment issues in bilingual programs. There are several theories that have explained language proficiency among bilingual students. For example, Bialystok (1978) proposed a model of second language learning which accounts for discrepancies both in individual achievement and achievement in different aspects of second language learning. The model is both descriptive and explanatory in that it incorporates the factors suggested by the literature to be relevant for second language learning, and it also postulates processes describing the nature of the effect of these factors and the interactions among them. The model is organized on three levels, i.e. input (language exposure), knowledge (explicit and implicit linguistic knowledge as well as other knowledge) and output (inferencing and monitoring), each of which provides a unique stage in the learning and use of a second language. It is worth noting that to the extent that this model is explanatory, it is also predictive—language learning outcomes may be predicted by considering the state of the conditions presented in the model. A critical feature of Bialystok's model is its generality, in that it describes the way in which humans, given biological, social, and other restrictions, learn a second language.

Another interesting model of second language learning is Cummins' 'common underlying proficiency' (CUP) model of bilingual proficiency in which the cognitive academic language proficiency in both languages, e.g. Chinese and English, are regarded as manifestations of one underlying dimension (Cummins, 1980; 1981). This model is based on the interdependence hypothesis which suggests that first language (L1) and second language (L2) academic proficiency, such as

reading ability, tend to be developmentally interdependent. Thus, for example, the development of L2 proficiency is partially dependent upon the development of L1 proficiency in educational contexts. According to Cummins (1984) the CUP is theoretically capable of being developed through instruction in either language. However, whether or not instruction in a particular language will successfully develop cognitive academic language proficiency depends on sociocultural factors as much as on pedagogical factors. Assessment of language proficiency in bilingual children reveals an enormous variation in what the different tests purport to measure (Dieterich, Freeman, & Crandall, 1979). Some tests focus on measuring phoneme production, various levels of lexical ability, oral syntax comprehension, or pragmatic aspects of language (DeAvila & Duncan, 1978). Initially, assessment of language proficiency skills were often targeted at minority and immigrant students to determine their skills for the purpose of entry into or exit from bilingual programs. Thus, it is important to determine whether the measures of language proficiency are related to constructs of intelligence and academic achievement or not. Some assessment batteries such as the Language Assessment Scale (LAS) (DeAvila & Duncan, 1978) consistently reported moderate correlations with overall achievement, while the Bilingual Syntax Measure (BSM) (Burt, Dulay, & Hernandez-Chavez, 1976) reported much lower correlations with achievement. In any case, showing how the validity of language proficiency measures in a specific context of bilingual education is determined is very important.

The present study was a first attempt to design a standardized short test of reading ability that might be adequate for evaluating learners of English as a foreign language. There are a number of studies that have designed brief assessments because of the ability of such tests to give quick feedback and allow for prompt intervention measures. Examples include the French Kit (French, Ekstrom, & Price, 1963) and the brief assessments of math achievement

(Steiner & Ashcraft, 2012).

A substantial amount of work has been devoted to optimally tapping children's reading skills. This has basically involved coming up with assessment instruments that vary in scope, content, tasks, scoring procedure, and interpretation. Most assessments pose tasks that involve comprehension and production of oral language, while others also include written tasks in a bid to roughly answer the question of how well a student knows a particular language such as English (Dieterich, Freeman, & Crandall, 1979). Examples of such English language proficiency tests may include (apart from others already highlighted) Silverman and colleagues' *Oral Language Tests for Bilingual Students* (1976), the *Short Test of Linguistic Skills, Comprehension of Oral Language* as well as the *Oral Language Evaluation* (refer to Dieterich, Freeman, & Crandall, 1979, for details). Fundamentally, Torgesen, Rashotte, and Alexander (2001), among others, observed that good and poor readers differ significantly in the number of words they can recognize quickly. Building on this, Fuchs, Fuchs and Compton (2004) further highlighted the importance of reading fluency as a powerful indicator of overall reading competence. Several studies have shown a moderate to strong correlation between reading fluency (both word and text level reading) and reading comprehension for students in the elementary grades (e.g., Hosp & Fuchs, 2005; Marston, 1989; Speece & Ritchey, 2005). Although most of these tests have been in the area of Oral Reading Fluency (ORF), a recent study by Wagner, Torgesen, Rashotte, and Pearson (2010) also revealed relations between silent reading fluency and comprehension for the elementary grades, noting a decline in the relation as the grades increased from 4 to 10. Importantly, all of this research has focused on reading in a first language. Our primary target group in the present study was Hong Kong Chinese children learning English as a foreign language.

Measures of Silent reading fluency

Given that a lot of time is required to individually administer tests of ORF, assessments of silent reading fluency may be particularly useful for identifying students with reading difficulties (Denton et al., 2011). For instance, some standardized measures of silent reading fluency take only three minutes and can either be individually or group administered. According to Mather et al. (2004) the Test of Silent Word Reading Fluency (TOSWRF) measures word identification, accuracy and speed as well as word comprehension. The authors opted to use word chains as opposed to letter chains based on the literature review and field testing which both indicated that for the TOSWRF, word chains had a strong correlation with reading, and provided better prediction and identification of poor readers. In this test, students are presented with a string of unrelated words without spaces which they have to separate using lines (TOSWRF: Mather et al., 2004). Timed word-strings, presented in a “without spaces” format, are successful in measuring word identification speed, e.g., in assessments such as Guilford's (1959) *Structure of Intellect Model* and Meeker, Meeker and Roid's (1985) *Structure of Intellect Learning Abilities Test*. According to Williams and Bell (2005) the TOSWRF has solid reliability and validity as an assessment tool (refer to Williams & Bell, 2005, for details on how the psychometric attributes were demonstrated).

The other standardized test of oral reading fluency is the Test of Silent Contextual Reading Fluency (TOSCRF) (Bell et al., 2007). This test is similar to the TOSWRF, except that the students are now presented with text passages that are all printed in uppercase letters without spaces or punctuation (Bell et al., 2007). Like the TOSWRF, students have to draw lines between each word of the text as they read the passages. According to Hammill, Wiederholt, and Allen (2006) the TOSCRF measures many aspects of silent reading such as syntax and morphology, thereby providing ‘extraordinarily accurate measures of silent reading competence’ (p.2). In

a study to establish the psychometric integrity of the TOSCRF, Bell and colleagues (2007) demonstrated that as an assessment tool, the TOSCRF has both excellent reliability and validity.

In the present study, we included both word and sentence reading, following Mather et al. (2004) and Bell et al. (2007) in order to tap Hong Kong Chinese children's reading of English as a foreign language. In Study 1, we pilot-tested items and eventually normed the test with second grade Hong Kong Chinese children. In Study 2, we compared the results of our newly created test of English word reading as a foreign language with another previously used test of oral word recognition in English. The present study was divided into two studies because we first needed to pilot the test items of our research tool, which not only involved establishing the reliability of the measures but also establishing a basic norm for second graders. In addition, we needed to illustrate the independent performance of children on a more traditional individually administered oral word recognition tests done in both English and Chinese. In comparing the newly developed tests with previously used oral word recognition tests in English and Chinese, we wanted to establish concurrent validity.

Materials for this short test of English reading as a foreign language were created with reference to major textbooks designed for teaching the local English curriculum in Hong Kong. We chose the kind of words that are frequently used in English textbooks, and the words included parts of speech such as nouns (e.g. tiger), connectives (e.g. but), verbs (e.g. run), adjectives (e.g. bad), verbs (laugh), and sight words (e.g. girl).

Study 1

In this pilot study, we hypothesized that the two tests of silent word reading would be correlated in that they measured similar aspects of silent word reading, i.e. speed and accuracy.

Method

Participants

In Study 1, there were two phases, a pilot study and a norming phase. Fourteen second graders participated in the pilot study. There were 8 boys and 6 girls and they were between the ages of 86 to 93 months old (Mean = 88.93, $SD = 2.27$). They were from 4 primary schools in the New Territories in Hong Kong. In the norming phase, there were 552 second graders, of whom 290 were boys and 262 were girls. They were between the ages of 70 to 121 months old (Mean = 89.85, $SD = 5.26$). They were from 5 primary schools in the New Territories and Kowloon in Hong Kong. All children were born in Hong Kong, and Cantonese was their native language. They had all learned English as a foreign or second language from the age of about 3.5 to 4 years old.

Procedure

In the pilot study, informed consent was obtained from all parents. Both the individual word reading and contextual word reading subtests were administered by a trained undergraduate psychology major student in a room at the child's school. In the norming phase, invitation letters were first sent to schools and informed consent was again obtained from parents of all participating children. The test was group administered by trained undergraduate psychology majors. One tester was responsible for one class of students and a teacher helped to maintain the class discipline. Also, half of the children were given the individual word reading test and then the silent contextual reading test, and the other half of the children were given the silent contextual reading test first and then the silent word reading test.

Materials

Nonverbal IQ: Sets A and B of Raven's Standard Progressive Matrices (Raven, 1976) were used to measure children's general

reasoning ability, as a proxy for general IQ. This task is a non-verbal multiple choice test. There were 24 items, and the maximum score was 24 points. For each test item, children were asked to identify the missing element that completed a pattern from among six choices.

English silent word reading—subtests 1 (individual word reading) and 2 (contextualized word reading): The test and instructions for this task were the same as those used in the norming phase in Study 1.

English word reading (read aloud): This test was successfully used in previous studies (e.g. Chung et al., 2011; Li, McBride-Chang, Wong, & Shu, 2012). There were 40 words in this test. They were arranged in four levels with increasing difficulty, and there were 10 words in each level. Children were asked to read aloud each word, individually to the experimenter. When children answered 4 consecutive items incorrectly in the same level, the testing stopped. Participants were awarded one point for each word they correctly identified for this task.

Chinese character reading: This task was adopted from the Hong Kong Test of Specific Learning Difficulties in Reading and Writing (Ho, Chan, Tsang, & Lee, 2000). There were 150 two-character words in this test. Children were required to read the words one by one aloud, individually to the experimenter. The children were awarded one point when they read both characters correctly in a word. Testing stopped when 15 consecutive words were incorrectly identified by the children.

Results

Descriptive statistics for Study 1 are shown, respectively, in Tables 1 and 2. Across each part of the test, internal consistency reliabilities were .96. In addition, the correlation between silent word reading and silent contextual reading was .78 ($p < .001$), confirming our hypothesis for

Table 1
Descriptive Statistics of Word Reading Measures and Age in Norming Phase in Study 1.

	Means	SD	Range	Reliability
Age in months	89.85	5.26	70-121	N/A
English silent word reading	31.86	19.29	0-150	.96
English silent contextual reading	48.47	19.46	0-152	.96
English silent reading (word + contextual)	80.33	36.55	0-302	.97

Note: $N=552$.

Table 2
Correlations among Test Measures and Age in Norming Phase in Study 1.

	1	2	3
1. Age in months	---		
2. English silent word reading test	.14**	---	
3. English contextual reading test	.15***	.78***	---

Note: $N=552$; ** $p < .01$, *** $p < .001$.

this study. Given that both individual word recognition and reading comprehension are important skills in foreign language learning (August & Shanahan, 2008), and because the correlation between the two sections of this test was relatively high, these were combined for Study 2.

Study 2

In this study, we hypothesized that the newly developed test would significantly correlate with the previously developed, and more traditional oral word recognition tasks in English and Chinese.

Method

Participants

There were 77 second graders, 46 boys and 31 girls, included in Study 2. They ranged in age from 90 to 124 months old (Mean = 97.94, $SD =$

7.30). They were from 3 primary schools in the New Territories and Kowloon in Hong Kong. All children were born in Hong Kong, and Cantonese was their native language. They had all learned English as a second language starting at about the age of 4. Among the 77 children, 9 were dyslexics as identified by school psychologists.

Procedure

Invitation letters were sent to schools and informed consent was obtained from parents for this study, the purpose of which was to examine the association between a more traditional individually administered oral word recognition test of English and the newly created silent word reading test. We also included a traditional task of Chinese word reading in order to examine the association between Chinese and English word reading in these children. Each task was administered by trained undergraduate psychology majors. English word reading (read aloud) and Chinese character reading were individually administered while the English silent reading tests and the Raven's test were

group administered. Thus we hypothesized that our English silent reading tests would be associated with the more traditional individually administered oral word recognition tests in both English and Chinese.

Materials

Nonverbal IQ. Sets A and B of Raven's Standard Progressive Matrices (Raven, 1976) were used to measure children's general reasoning ability, as a proxy for general IQ. This task is a non-verbal multiple choice test. There were 24 items, and the maximum score was 24 points. For each test item, children were asked to identify the missing element that completed a pattern from among six choices.

English silent word reading-subtests 1 (individual word reading) and 2 (contextualized word reading). The test and instructions for this task were the same as those used in the norming phase in Study 1.

English word reading (read aloud). This test was successfully used in previous studies (e.g. Chung et al., 2011; Li, McBride-Chang, Wong, & Shu, 2012). There were 40 words in this test. They were arranged in four levels with increasing difficulty, and there were 10 words in each level. Children were asked to read aloud each word, individually to the experimenter. When children answered 4 consecutive items incorrectly in the same level, the testing stopped. Participants were awarded one point for each word they correctly identified for this task.

Chinese character reading. This task was adopted from the Hong Kong Test of Specific Learning Difficulties in Reading and Writing (Ho, Chan, Tsang, & Lee, 2000). There were 150 two-character words in this test. Children were required to read the words one by one aloud, individually to the experimenter. The children were awarded one point when they read both characters correctly in a word. Testing stopped when 15 consecutive words were incorrectly identified by the children.

Results

The means, *SDs*, ranges, and reliabilities of each task are shown in Table 3, and results of correlations are shown in Table 4. In the correlation analyses, age and Raven's test were statistically controlled. They were sometimes significantly associated with reading. In this study, correlations of age with Chinese character reading, English silent word reading, and English word reading (read aloud) were minimal, i.e., .07, -.18, and -.23 respectively, while the correlations of the Raven's test with Chinese character reading, English silent word reading and English word reading (read aloud) were .43 ($p < .001$), .30 ($p < .01$), and .30 ($p < .01$), respectively. The correlation between the English silent reading test and the English word reading test (read aloud) was .78 ($p < .001$). This correlation was larger than that between English silent reading and Chinese character recognition which was .49 ($p < .001$). Our English silent reading tests were relatively strongly associated with the more traditional individually administered oral word recognition tests in both English and Chinese, thereby confirming our hypothesis for study 2. Because, among the 77 children in Study 2, there were 9 dyslexics, we examined these associations with and without them included. The correlation between English silent reading and English word reading test (read aloud) remained .78 ($p < .001$) when these 9 dyslexics were excluded.

Discussion

In the present study, we created a test of English silent reading for second grade Hong Kong Chinese children learning English as a second or foreign language. The two parts of the test, focused, respectively, on isolated word reading and contextualized (within sentences) word reading, were fairly highly associated in a sample of nearly 600 second graders. When these two parts of the test were combined and compared to an individually administered oral word recognition task, the association was also

Table 3
Descriptive Statistics on All Measures Included in Study 2.

	Means	SD	Range	Reliability
Age in months	97.94	7.30	90-124	N/A
Raven's test	17.96	3.37	10-23	.75
English silent word reading test	27.54	16.73	0-80	.96
English silent contextual reading test	44.78	19.66	0-100	.95
English silent reading test (word + contextual)	72.32	34.50	0-180	.97
English word reading test (read aloud)	13.51	11.87	0-40	.97
Chinese character recognition	87.39	34.69	0-133	.98

Note: $N=77$.

Table 4
Correlations Among Word Reading Measures with Age and Raven's Statistically Controlled in Study 2.

	1	2	3
1. English silent reading test (word + contextual)	---		
2. English word reading test (read aloud)	.78***	---	
3. Chinese character recognition	.49***	.54***	---

Note: $N=77$; *** $p < .001$.

relatively high. Moreover, the correlation between our English silent reading test and Chinese word reading was lower as compared to the one between the two English word reading tests, suggesting that our newly created test has some validity as a task of English word reading. It is worth noting that other studies have also found similar correlations (ranging from .51 to .61) between English and Chinese reading among Hong Kong Chinese children (Chung et al., 2011; McBride-Chang, Cheung, Chow, Chow, & Choi, 2006). This may be due in part to the fact that children in Hong Kong primarily learn using the look/say method for literacy instruction.

This new test of English word reading is a short and efficient way of tapping Hong Kong Chinese children's reading of English as a foreign language. Both sections of the test together require no more than 10 minutes for administration. This type of quick and reliable measure is of potential importance both for educators and researchers in efficiently

identifying poor readers (Bell et al., 2007; Williams & Bell, 2005). Furthermore, given the fact that many children worldwide learn using the medium of English for school learning even when it is not their native language, the measure of reading skill suggested in this study could be ideal not only for Hong Kong but also for many countries or regions around the world. This assessment tool can be particularly useful for busy teachers (Williams & Bell, 2005), perhaps, in the future, especially in Sub-Saharan Africa where teachers generally have to handle big multiple classes (Policy and Operations Evaluation Department, 2011; Sinyolo, 2007; UNESCO, 2009). We hope that this test can serve as a basis for others to create a similar test (or modify this one for English standards in a given region) in many different areas in which English is learned as a second or foreign language. The test created in the present study can probably be used for students as young as kindergarten age and as old as sixth grade. Local modifications, including cultural or local

references (e.g., we talked about Hong Kong in a contextualized sentence) or language (e.g., there are some British-English specific terms in Hong Kong English textbooks that are learned before other vocabulary; in other regions of the world there might be other vocabulary-specific changes to be made), would be necessary but hopefully minor in most cases.

We also view this task as having strong potential for researchers (Bell et al., 2007; Williams & Bell, 2005). The number of research studies on children learning English as a second or foreign language has increased in the past few years. The issue of language-learning in one's non-native language is important to consider in the world today, where more people speak English as a second language than as a mother tongue, and the popularity of English for learning and communicating with others worldwide continues to increase.

The reading measure suggested here is appropriate for any setting where brief screening on reading achievement would be useful for reasons of instruction or research. More lengthy and careful follow up evaluations are available, but this particular test should be one of the most efficient and possibly important ones to use initially to understand children's relative standing in English learning. We hope that by publishing these data, we will encourage others in different regions of the world to create and norm such a test where they are living. In such instances it would be important to ensure that issues of reliability and validity are attended to. For example it would be good to consider aspects of content-description validity, i.e. whether word chains would be familiar in the popular culture of some countries. Our measure is available to be considered and modified for different world regions from us upon request. We further hope that in the future, this test would be expanded, with regard to administration instructions, to include students with special needs such as the deaf and hard of hearing students.

Acknowledgements: We are grateful that this

research was supported by Research Grants Council of the Hong Kong Special Administrative Region (Grant No: 451811).

References

- Abella, R., Urrutia, J., & Shneyderman, A. (2005). An examination of the validity of English language achievement test scores in an English language learner population. *Bilingual Research Journal*, 29(1), 127-144.
- August, D., & Shanahan, T. (Eds.). (2008). *Developing reading and writing in second-language learners: Lessons from the report of the National Literacy Panel on language-minority children and youth*. London: Routledge
- Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning*, 28(1), 69-83.
- Bell, S. M., McCallum, R. S., Krik, E. R., Fuller, E. J., & McCane-Bowling, S. (2007). Investigation of the psychometric attributes of the test of silent contextual reading fluency. *Assessment for Effective Intervention*, 33(1), 39-46.
- Burt, M. K., Dulay, H. C., & Hernandez-Chavez, E. (1976). *Bilingual syntax measure: Technical handbook*. New York, NY: Harcourt Brace Jovanovich.
- Cheung, H., Chung, K. K. H., Wong, S. W. L., McBride-Chang, C., Penney, T. B., & Ho, C. S. -H. (2010). Speech perception, metalinguistic awareness, reading, and vocabulary in Chinese-English bilingual children. *Journal of Educational Psychology*, 102(2), 367-380.
- Chiappe, P., & Siegel, L. (1999). Phonological awareness and reading acquisition in English and Punjabi speaking Canadian children. *Journal of Educational Psychology*, 91(1), 20-28.
- Chung, K. K. H., McBride-Chang, C., Cheung, H., & Wong, S. W. L. (2011). General auditory processing, speech perception and phonological awareness skills in Chinese-

- English biliteracy. *Journal of Research in Reading*, 36(2), 202-222.
- Cummins, J. (1980). The entry and exit fallacy in bilingual education. *NABE Journal*, 4(3), 25-59.
- Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.). *Schooling and language minority students: A theoretical rationale* (pp. 3-49). Los Angeles, CA: California State University.
- Cummins, J. (1984). Wanted: A theoretical framework for relating language proficiency to academic achievement among bilingual students. In C. Rivera (Ed.), *Language proficiency and academic achievement: Multilingual matters 10* (pp.2-19). Clevedon, New Zealand: Multilingual Matters.
- DeAvila, E. A., & Duncan, S. E. (1978). *A few thoughts about language assessment: The LAU decision reconsidered*. Paper presented at the Conference on Research and Policy Implications of the Task Force Report of the United States Office for Civil Rights, Austin, TX. Abstract retrieved from <http://www.ncela.us/rcd/bibliography/BE001237>
- Denton, C. A., Barth, A. E., Fletcher, J. M., Wexler, J., Vaughn, S., Cirino, P. T., Romain, M., & Francis, D. J. (2011). The relations among oral and silent reading fluency and comprehension in middle school: Implications for identification and instruction of students with reading difficulties. *Scientific Studies of Reading*, 15(2), 109-135.
- Dieterich, T. G., Freeman, C., & Crandall, J. A. (1979). A linguistic analysis of some English proficiency tests. *TESOL Quarterly*, 13(4), 535-550.
- French, J. W., Ekstron, R. B., & Price, I. A. (1963). *Kit of reference tests for cognitive factors*. Princeton, NJ: Educational Testing Service.
- Fuchs, L. S., Fuchs, D., & Compton, D. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children*, 71(1), 7-21.
- Guilford, J. P. (1959). Three faces of intellect. *American Psychologist*, 14(8), 469-479.
- Hammill, D. D., Wiederholt, J. L., & Allen, E. A. (2006). *Test of silent word reading fluency*. Austin, TX: PRO-ED.
- Ho, C. S. -H., Chan, D. W., Tsang, S., & Lee, S. (2000). *The Hong Kong test of specific learning difficulties in reading and writing (HKT-SpLD) manual*. Hong Kong: Hong Kong Specific Learning Difficulties Research Team.
- Hosp, M. K., & Fuchs, L. S. (2005). Using CBM as an indicator of decoding, word reading, and comprehension: Do the relations change with grade? *School Psychology Review*, 34(1), 9-26.
- Li, T., McBride-Chang, C., Wong, A. M. -Y., & Shu, H. (2012). Longitudinal predictors of spelling and reading comprehension in Chinese as an L1 and English as an L2 in Hong Kong Chinese children. *Journal of Educational Psychology*, 104(2), 286-301.
- Lin, D., Wong, K. K., & McBride-Chang, C. (2012). Reading motivation and reading comprehension in Chinese and English among bilingual students. *Reading and Writing*, 25(3), 717-737.
- Marston, D. (1989). A curriculum-based measurement approach to assessing academic performance: What is it and why do it. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18-78). New York, NY: Guilford.
- Mather, N., Hammill, D. D., Allen, E.A., & Roberts, R. (2004). *TOSWRF: Test of silent word reading fluency: Examiner's manual*. Austin, TX: PRO-ED.
- McBride-Chang, C. (2004). *Children's literacy development*. London: Arnold.
- McBride-Chang, C., Cheung, H., Chow, B. W. -Y., Chow, C. S. -L., & Choi, L. (2006). Metalinguistic skills and vocabulary knowledge in Chinese (L1) and English (L2). *Reading and Writing*, 19, 695-716.
- McBride-Chang, C., Liu, P. D., Wong, T., Wong, A., & Shu, H. (2012). Specific reading difficulties in Chinese, English, or both: Longitudinal markers of phonological

- awareness, morphological awareness, and RAN in Hong Kong Chinese children. *Journal of Learning Disabilities*, 45(6), 503-514.
- McBride-Chang, C., Shu, H., Chan, W., Wong, T., Wong, A. M. -Y., Zhang, Y., Pan, J., & Chan, P. (2013). Poor readers of Chinese and English: Overlap, stability, and longitudinal correlates. *Scientific Studies of Reading*, 17(1), 57-70.
- Meeker, M. N., Meeker, R., & Roid, G. (1985). *Structure of intellect learning abilities test (SOI-LA)*. Los Angeles, CA: Western Psychological Services.
- Policy and Operations Evaluation Department (IOB), the Netherlands Ministry of Foreign Affairs. (2011). *Unfinished business: Making a difference in basic education: An evaluation into the impact of education policies in Zambia and the role of budget support*. The Hague: Ministry of Foreign Affairs of the Netherlands.
- Raven, J. C. (1976). *Standard Progressive Matrices: Sets A, B, C, D, and E*. Oxford: Oxford Psychologists Press.
- Sattler, J. M. (2001). *Assessment of Children: Cognitive applications* (4th ed.). San Diego, CA: Sattler Publishers.
- Silverman, R. J., Noa, J. K., & Russell, R. H. (1976). *Oral language tests for bilingual students: An evaluation of language dominance and proficiency instruments*. Portland, OR: Center for Bilingual Education, Northwest Regional Educational Laboratory.
- Sinyolo, D. (2007). *Educational international: Teacher supply, recruitment and retention in six anglophone sub-Saharan African countries in Gambia, Kenya, Lesotho, Tanzania, Uganda and Zambia*. Brussels, Belgium: Education International.
- Speece, D. L., & Ritchey, K. D. (2005). A longitudinal study of the development of oral reading fluency in young children at risk for reading failure. *Journal of Learning Disabilities*, 38(5), 387-399.
- Steiner, E. T., & Ashcraft, M. H. (2012). Three brief assessments of math achievement. *Behavioral Research Methods*, 44(4), 1101-1107.
- Torgesen, J. K., Rashotte, C. A., & Alexander, A. W. (2001). Principles of efficiency instruction in reading: Relationships with established empirical outcomes. In M. Wolf (Ed.), *Dyslexia Efficiency and the Brain* (pp. 333-355). Timonium, MS: York Press.
- UNESCO. (2009). *Regional overview of sub-Saharan Africa - EFA global monitoring report: Overcoming inequality: Why governance matters*. Paris: UNESCO.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). *Test of silent reading efficiency and comprehension*. Austin, TX: Pro-Ed.
- Williams, A., & Bell, S. M. (2005). Test review: Mather, N., Hammill, D. D., Allen, E. A., & Roberts, R. (2004). "Test of silent word reading fluency." Austin, TX: PRO-ED. *Journal of Psychoeducational Assessment*, 23(2), 182-188.

Received March 17, 2015
 Revision Received July 3, 2015
 Accepted July 31, 2015