Investigating the Psychometric Properties of IDAPEL

(Indicateurs Dynamiques d'Habiletés Précoces en Lecture)

French-language Early Literacy Measures with

Students Learning to Read in French

Chantal Dufour-Martel

Dynamic Measurement Group

Alain Desrochers

University Ottawa

Technical Report 12

Suggested Citation

Authors' Note

Abstract

This technical report presents findings of a validation study examining the psychometric properties of French-language early literacy experimental measures known as IDAPEL with a group of French-speaking students learning to read in French as their first language. The purpose of the study is to verify the predictive validity, reliability and construct validity of IDAPEL measures as well as their ability to predict later reading outcomes.

Kindergarten, first grade and second grade French-speaking student participants were assessed during one school year (2008 – 2009). Each student was assessed at three time-points (fall, winter, and spring) with the IDAPEL battery of tests. A subset of students in two grades was assessed at two-week intervals between benchmark periods using alternate forms of kindergarten and first grade measures. In addition, at the end of the school year, another subset of student participants was assessed on a set of ÉCOLE criterion-related battery of tests.

In this report, information about measure development, along with descriptive and correlational data are presented. For all three grade levels, the IDAPEL measures display adequate reliability and validity for decision making in early literacy. Implications of these findings are discussed.

Keywords: early literacy assessment, French-language, curriculum-based.

Investigating the Psychometric Properties of IDAPEL (Indicateurs dynamiques d'habiletés précoces en lecture) French-Language Early Literacy Measures with Students Learning to Read in French.

While literacy levels among Francophone populations in Canada have improved over the last two decades, a 2006 Statistic Canada report examining literacy status among official language minority groups indicates that Canadian Anglophones have continued to do better on literacy tests than their French-Canadian counterparts. The report cites that, on a national scale, 39% of the English population aged 16 to 24 failed to reach an adequate or competent level of reading comprehension. Comparatively, 56% of the Francophone population of the same age-range failed to reach an adequate reading comprehension level in their native language. (Statistique Canada[a], 2006).

The consequences of low reading comprehension skills cannot be undermined, and indeed have been linked to delinquency, high levels of high-school dropout and completion rates, and, in the long term, high unemployment rates (McGill-Franzen, 1987; Vacca, 2008). Realistically, completing a high school diploma is, at best, a minimal educational requisite for access to well-paying employment in the labor market. It is estimated that 40% of Canadian jobs created between 1989 and 2000 required more than 16 years of education and training (Schargel & Smink, 2001).

Reading and reading comprehension skills are acquired in a relatively predictable way (National Research Council, 1998) and most often, reading difficulties expressed at the elementary grades can be avoided or resolved in the early childhood years (Snow, Burns & Griffin, 1998). In the Canadian school context, nationwide policy efforts to ensure early reading success dovetails with most current empirical evidence on what it takes for children to learn to

read. Increasingly, early literacy research compels educators to begin the process of learning to read early.

Given the broad demographic and socioeconomic profiles of children who enter public schools, some children are less 'ready to learn' than others. In point of fact, children begin their formal education with widely discrepant language and emergent literacy experiences (Hart & Risley, 1995). Children who have been exposed to literacy experiences from an early age are typically more prepared to learn to read at the beginning of formal schooling than those who have not been exposed to as frequent early literacy experiences. In circumspect, prior to any formal instruction, determining a child's initial level of language and literacy skill remains vital to guiding instructional practices, and indeed, the most common reason for early assessment is to screen children who may be 'at risk' for future reading difficulties (Desrochers, Simon & Thompson, 2011; Good Simmons & Kame'enui, 2001).

Early screening for French reading difficulty within the French language school context in Canada is a challenge because few standardized, validated French language early literacy evaluation tools exist (Fréchette & Desrochers, 2011). Identifying a struggling French language reader at the primary level continues to be a challenge for teachers and school psychologists. Without adequate evaluation tools to assess students' educational needs, appropriate instruction frequently continues to be delayed putting students at increased risk for reading difficulties. The increased attention to accountability for positive reading outcomes of all primary grade Francophone students continues to be strong motivation for developing and validating early literacy measurement tools in the French language, as very few exist (for an example, see Saint-Laurent & Giasson, 2010).

The process of learning to read in a first language is multi-faceted, complex, and takes considerable time and resources to develop. Often, teaching reading is a challenge for teachers. In the French language, validated principles for optimizing reading performance of students learning to read in French are lacking. Knowing more about optimizing the reading performance of French speaking students during the early elementary years remains important to the later reading performance of these students. The degree of success students experience later on with the more abstract, decontextualized, literate language of core content text relates, in large part, to early success establishing sound foundational literacy skills during the primary grades.

In the English language first (EL1) research literature, an evidence-based body of research on early literacy has informed our knowledge about how children learn to read. This knowledge base is reflected in the works of the National Research Council's (1998) report, *Preventing Reading Difficulties in Young Children*, and from the report of the National Reading Panel (2000), *Teaching Children to Read*. Converging and convincing evidence substantiates the conclusion that reading competence in English is causally influenced by proficiency on foundational skills in beginning reading (National Reading Panel, 2000; National Research Council, 1998). Research with monolingual English readers convincingly identifies phonological awareness and knowledge about the alphabetic principle as crucial components of word recognition and beginning reading achievement (Adams, 1990; Blachman, 2000).

Put into perspective, phonological awareness and the alphabetic principle are the skills that boot-strap young children into reading. For some learners, this is a rather complex task. In order for children to read the printed word, they must first be aware that spoken words are made up of individual sounds. Phonological awareness is the explicit awareness that spoken words are made up of individual sounds or phonemes, and involves the ability to attend to and manipulate

the sounds of speech in words (Carnnine, Silbert & Kameenui, 1997). Next, children must learn

that these sounds are represented by symbols, or graphemes (letters). The key to unlocking the

'code' of the written language begins when the child learns to make associations between letters

and sounds (grapheme-phonemes) (Perfetti & Zhang, 1996). To learn how the writing system

works is to understand the alphabetic principle.

Thus, to become fluent in recognizing English words in print, children need a firm

foundation in phonological awareness and the alphabetic principle. Although most studies on

phonological awareness have been conducted in English, phonological awareness has been

investigated in many other alphabetic languages. There is growing consensus that phonological

awareness skills are important for learning to read in several orthographies that have been

studied, including Danish, German, Norwegian, Spanish, Italian, Greek, Chinese and French

(Borzone de Manrique & Signorini, 1994; Comeau et al., 1999; Content, 1993; Cossu,

Shankweiler, Liberman, Katz, & Tola, 1988; Demont, 1994; Desrochers, Cormier & Thompson,

2005; Grandmaison, 1996; Huang & Hanley, 1995; Lecocq, 1991; Lundberg, 1994; Porpodas,

1993; Schneider, Kuspert, Roth, Vise, & Marx, 1997). Findings from these studies indicate a

strong relation between phonological awareness and word decoding. Hence, phonemic awareness

appears to be a critical component of word recognition in various orthographies, to the extent

that developing phonemic awareness in young children makes noticeable differences in later

reading outcomes (Ball & Blachman, 1991; Yopp, 1992).

Equally remarkable in the English language first (EL1) research literature is strong

evidence to support educator use of Curriculum-Based Measures (CBM) as viable means to

assess students' academic skills. Curriculum-based measurement is a set of standardized and

well-researched procedures for assessing and monitoring student progress in reading, math,

spelling, and writing (Fuchs & Deno, 1991; Shinn, 1989, 1998; Tindal & Marston, 1990). Curriculum-based measurement has an impressive track record of randomized control trials showing improved learning outcomes for special education students (Fuchs & Fuchs, 1998).

To address the need for literacy assessments for French speaking student populations, a battery of tests known as *Indicateurs dynamiques d'habiletés précoces en lecture* (IDAPEL) was developed in the French language. IDAPEL® is a researched set of experimental French language curriculum-based measures. By design, they function as formative French language assessments measuring student progress on *critical early literacy skills* of French phonological awareness, the understanding of the alphabetic principle, and a child's fluency with connected text. IDAPEL test materials are of equal difficulty and represent the general curriculum. Because the measures are sensitive to student reading skill growth over time, they are used to establish an initial skill base-line at the beginning of the school year, and to measure reading skill growth across the year at the winter and spring benchmark periods. In this way, the measures help determine students' overall skill change during the academic school year. IDAPEL measures are administered three times a year for universal screening to students learning to read in French from kindergarten through fifth grade. The measures can be administered more frequently for progress monitoring the reading growth of students in need, and to evaluate the effects of instructional intervention. As formative assessments, they are used to adapt instruction to meet student need and to support student learning.

The IDAPEL battery of tests assesses students' early phonological awareness and alphabetic principle knowledge, and includes the following subtests: *Facilité à reconnaître le premier son* (FPS) [1], *Facilité à dénommer des lettres* (FDL) [2], *Facilité à segmenter les phonèmes*

[1]First Sound Fluency, [2]Letter Naming Fluency

(FSP) [3], *Facilité à lire des non-mots* (FNM) [4], *Facilité en lecture orale* (FLO) [5], and *Rapport du récit* (ROR) [6]. Their practicality as one-minute measures allows classroom teachers to assess student reading skill quickly and efficiently.

To date, few curriculum-based measures exist in the French language, specifically measures targeting key skills of phonological awareness and the alphabetic principle, and few studies have explored the technical adequacy and utility of curriculum-based measures in Canada with Francophone speaking populations (Saint-Laurent & Giasson, 2010). Three IDAPEL measures have been previously validated with a group of French immersion students. Two of the measures, *Facilité à lire des non-mots* (FNM) and *Facilité en lecture orale* (FLO), provide evidence for being strong, robust measures useful in predicting French reading outcomes of second-grade French immersion students (Dufour-Martel & Good, 2009).

Overview of the study

The purpose of this technical report is to present an analysis of the psychometric properties of French early literacy measures validated with a French-speaking student population learning to read in French as their first language. The technical adequacy, specifically the reliability, sensitivity, and predictive-criterion and concurrent-criterion related validity of the following six IDAPEL® measures will be reported: *Facilité à reconnaître le premier son* (FPS), *Facilité à dénommer des lettres* (FDL), *Facilité à segmenter les phonèmes* (FSP), *Facilité à lire des non-mots* (FNM), Facilité en lecture *orale* (FLO), and *Rappel oral du récit* (ROR). Information about the development of the measures will also be reported. For further

[3] Phoneme Segmentation Fluency, [4] Nonsense Word Fluency, [5] Oral Reading Fluency, [6] Retell Fluency.

information regarding the development of English-language curriculum-based measures and complementary technical reports can be found at the following web address: www.dibels.org.

Method

Participants

The participants in this study were 298 students (47 % girls and 53 % boys) sampled from 19 general education classrooms in 5 elementary schools from 3 separate school districts in a Canadian province where English is majority language and French a minority language. Of this total, 138 participants were in Kindergarten (Mean age at the end of the school year = 71 months), 110 in Grade 1 (Mean age = 82.2 months) and 50 in Grade 2 (Mean age = 95 months). Participants included all students who were receiving French language reading instruction, including students with disabilities.

Participant selection. French speaking school districts were recruited with the assistance of the Learning Disabilities Association of Ontario (LDAO) who spearheaded the research and provided information about the nature and intent of the research as well as detailed information about participation requirements. Study participation was open to school districts located within the vicinity of a large metropolitan city, with a population of 812,000 people (Statistics Canada[b], 2006). The city has a main urban area with several urban, suburban and rural areas within the city's limits. The main suburban area extends over an area of 1,065 square miles to the east, west and south of the city centre. Based on site participation requirements, three school districts agreed to participate. Parents from the three participating school districts were informed about the study at the beginning of the school year by means of a letter sent home by participating classroom teachers. Parents were requested to return passive consent forms only if they did not wish their child/children to participate in the study. Any student participant not returning the

passive consent form was included in the study. All data were collected during the 2008-2009

school year. All schools remained in the study until its conclusion at the end of the academic

year.

International immigration has become the most dynamic population growth factor in the

area where the study took place. In 2006, members of visible minority groups (non-

white/European) constituted 20.2 percent of the population, while those of Aboriginal origin

numbered 1.5 percent of the total population. The largest visible minority groups reported in

Canada census are Black Canadians: 4.9%, Chinese Canadian: 3.8%, South Asian: 3.3% and

Arab: 3.0%, as well as smaller proportions of mixed race and other East Asian groups (Statistics

Canada[b], 2006). Given these percentages, the majority population remains Caucasian. Table 1

profiles French language school board and French language student population information for

the three school districts. Table 2 outlines demographic profiles for all students in all schools

compared with the general student population in this particular Canadian province. Ethnicity

demographics were not available for primary aged student populations.

Table 1.
2005-2006 French Language School Board Profiles

Research Site	Area Type	Number of Schools in the District	Number of Students in the District
A	Urban	21 elementary 14 secondary	11,000
B	Suburban	38 elementary 8 secondary	17,000
C	Rural	32 elementary 7 secondary	12,000

Note. School level data were taken from Ontario Ministry of Education School Board Profiles
website.

Table 2.

Student demographic profiles in percent for all Sites Compared with General Student Population in the Province.

		Site A	Site B		Site C	
Demographics	Province	School 1	School 2	School 3	School 4	School 5
L1 is not English	21.8	85.4	88.3	69.3	99.4	97.5
L1 is not French	96.4	28.3	29.1	45.4	0.9	2.5
Special education services	12.5	4.9	2.3	9.7	19.8	15.4
Low income	16.5	34	24	19	1	2

Note. School level data were taken from Ontario Ministry of Education School Board Profiles website. http://esip.edu.gov.on.ca/english/profiles/quick_info_Comp.   L1 = first language.

Procedure

Training and data collection. A total of 35 French teaching staff and school district special services personnel received an initial 2-day IDAPEL Essential training workshop in the administration and scoring of the IDAPEL measures. A four-hour webcast refresher training, which included practice with simulated activities, occurred before the winter benchmark collection data point. All data for this study was collected exclusively by trained school personnel from participating school districts. Trained school personnel entered all IDAPEL data in the Web-Based Teaching Tool (WBTT) database, which was subsequently exported to the DMG database. Two French speaking WBTT coordinators collected two-week alternative form data for all the measures across two grades at two benchmark points. Criterion measures were collected at the end of the school year, between May 1st and June 1st, by members of the Reading

Development Research Group at the University of Ottawa. This research group, trained to

fidelity on the administration and scoring of the ÉCOLE measures, was responsible for scoring,

entering, and verifying these measures, and for merging IDAPEL® benchmark data with criterion

data in order to create a master data file. Table 3 outlines, by grade-level, IDAPEL benchmark,

two-week reliability, and criterion ÉCOLE measure administration timelines.

Table 3

Data Collection Timeline by Grade, by Measures, and by Time of the Year

Grade	Fall	Winter	Spring
K	FPS/FDL	FPS/FDL/FSP/ FNM I	FDL/FSP/FNM I
		2 wk alternative form for FPS/FDL/ FSP/FNM I	ÉCOLE criterion
1	FDL/FSP/FNM I	FSP/FNM II/ FLO/ROR	FSP/FNM II/FLO/ ROR
	2 wk – alternative form for FDL/ FSP/FNM I		ÉCOLE criterion
2	FNM/FLO/ROR	FLO/ROR	FLO/ROR 2 wk alternative form for FNM II
			ÉCOLE criterion

Note : FDL = Facilité à dénommer des lettres; FPS = Facilité à reconnaître le premier son; FSP = Facilité à segmenter les phonèmes; FNM = Facilité à lire des non-mots; FNM I = simple graphemes, FNM II = complex graphemes, FLO = Facilité en lecture orale; ROR = Rapport oral du récit.

Measures

Two sets of measures were used in this validation study. The first set, IDAPEL®

(Dufour-Martel & Good, 2009), is a research-based experimental set of formative French

language assessments designed to assess the basic early literacy skills. The second set, ÉCOLE

(Desrochers, 2011), is a research-based normed set of literacy performance indicators developed for a broad range of assessment purposes.

General description of the IDAPEL measures. Measures of early literacy skills in alphabetical languages are important because they reflect the level of knowledge and skills beginning readers bring to the task of learning to read. Research in several alphabetic languages has evidenced that certain skills are important and play a substantial role in learning to read. Critical among these are the early skills of phonological awareness and the alphabetic principle (National Research Council, 1998; National Reading Panel, 2000). The IDAPEL measures are intended as indicators of these foundational skills and based on the same theoretical and evidence-based principles of how children learn to read in alphabetic languages. The measures are modeled after the English measures known as DIBELS *(Dynamic Indicators of Early Literacy Skills)* (Kaminski & Good, 1996). Comparatively, in construct and in use, both IDAPEL and DIBELS measures assess the early literacy skills. The IDAPEL measures, as *adapted* from the English measures, have been designed to encompass and reflect the phonology, orthography, and natural syntactic discourse of the French language. The IDAPEL® measures were designed to provide classroom teachers with information on students' developing skills in the core components of beginning reading including phonemic awareness, alphabetic principle understanding, the ability to accurately and fluently read connected text, and to do so with comprehension.

IDAPEL measures were designed as general outcome performance measures to be administered three times a year. As general outcome measures, they are intended to address broader questions about student skills in general such as: what is a given student's level of reading skill compared to other students? And, is the student progressing sufficiently in reading

skill? The IDAPEL measures are designed to be simple to administer and score, and to be sensitive to growth and skill improvement. The measures may be used to progress monitor student reading skill. All IDAPEL administrative directions were given in French. Brief information about the development of the measures and measure descriptions follow. Indicateurs Dynamiques d'Habiletés Précoces en Lecture (IDAPEL) and Curriculum-Based Measurement of Facilité en Lecture Orale (CBM FLO) comprised the IDAPEL assessment materials at the time of this study. IDAPEL K-5 benchmark materials and K-2 progress-monitoring materials are available for all measures except for FLO. IDAPEL measures are available for download at www.dibels.org.

Facilité à dénommer des lettres (FDL) is a timed, one minute measure designed to be an indicator of letter naming knowledge. Students are presented with a page of upper- and lower-case alphabetic letters arranged in a random order and are asked to name as many letters as they can. Students are also asked to name letters having accents [e accent aigu (é), e accent grave (è)]. If the student does not know a letter, the examiner tells the student the letter-name. The student is allowed one minute to produce as many letter names as he/she can, and the score is the number of correctly named letters in one minute. The measure takes approximately 1 and 1/2 minutes to administer.

Facilité à Reconnaître le Premier Son (FPS) is intended for most children from the last year of preschool through the middle of kindergarten. It may also be appropriate for monitoring the progress of older children with very low skills in phonemic awareness. Stimulus words for the measure were taken from MANULEX (Lété, Springer-Charolles & Colé, 2004) a Web-accessible database that provides grade-level word frequency lists of non-lemmatized and lemmatized words (48,886 and 23,812 entries, respectively) computed from 1.9 million words

taken from 54 French elementary school readers. From the non-lemmatized 48,886 word entries,
898 one syllable words were pulled from the first 6000 word entries. The one-syllable inclusion
rule was to select three to five phonemes words. From the 898 one-syllable word pool, exclusion
rules were applied, including removing words related to violence, hate, death, god or words
having an obscure meaning without context (i.e., galbe, fronce), words beginning with French
aspirated /h/ sound, two-sound words having more than three phonemes such as nain, âge, ache.
Next, three rules were applied for stimulus word pool inclusion: words beginning with either a
consonant blend (i.e., bl, cr, dr), continuous stop sound (i.e., d, hard c, ch, d) and continuous first
sounds (i.e., s, f, m). A total of 698 words (i.e., sin, loupe, prune) were included in the word pool
and randomized.

The measure *Facilité à Reconnaître le Premier Son (FPS)* is a timed one-minute measure
designed to be an indicator of phonological awareness. The measure assesses a student's fluency
in identifying the initial sound in orally presented words. The ability to isolate the first sound in a
word is an important phonemic awareness skill that is highly related to reading acquisition and
reading achievement (e.g., Yopp, 1988).

To administer this measure, the examiner says a word and asks the student to produce the
first sound in the word. The examiner begins the stopwatch immediately after saying the first test
item, and continues saying a series of words one at a time for one minute. On the scoring page,
the assessor circles the corresponding sound or group of sounds the student says. Students
receive two points for saying the initial phoneme of a word, and one point for saying correct
initial sounds of words, such as correct initial consonant blend, consonant plus vowel, or
consonant blend plus vowel. A response is scored as correct as long as the student provides any
of the correct first sound responses listed for the word. Differential scoring for student responses

allows young children to receive partial credit for demonstrating beginning skills in phonemic awareness. For example, a student who is not able to isolate the initial phoneme /f/ would receive partial credit for providing the first group of sounds /fr/ or /frê/ in the stimulus word *frêle*, showing an emerging understanding that sounds in words can be broken up into parts. The goal is for students to be able to correctly isolate the first phoneme in words. The total score is based on the number of correct 1- and 2-point responses the student says in one minute. No words are presented at the end of one minute. The measure has over 20 alternate forms for progress monitoring.

 Facilité à segmenter les phonèmes (FSP): Stimulus words for the phonemic segmentation fluency task were garnered from internet searches using a French translator. The MANULEX database was not accessible at the point in time when this measure was originally developed ten years ago as its development preceded that of *Facilité à reconnaître le premier son* (First Sound Fluency). Three consecutive Web searches for French children's stories written were initiated. On the first search, keyword descriptor 'French children's stories' yielded 667 hits. Nine short stories, both fiction and nonfiction, were downloaded and saved as text files. For the second search, keyword description 'littérature enfantine' yielded 241 hits. A total of nine French short stories, originating from diverse French speaking countries from around the world, were downloaded and saved as text files. On the third search, keywords 'contes classiques pour petits enfants' yielded 8,910 hits. Of these, ten classic tales were selected. In all, 28 French-language short stories were selected and put into text files. A computer program, designed specifically to count words, compiled a list of 3,657 French words as found in the text files.

 From this initial pool of words, 2,424 words were excluded on basis of occurring only once in any of the short stories. Next, 14 one phoneme words were excluded, as well as 24

inappropriate words. (e.g., en, ait, fessée, meurent). Four hundred and twenty four words having more than two syllables or more than five phonemes were excluded. Ninety one words having two syllables and five phonemes but used fewer than three times were excluded. Eight two-syllable words having five phonemes but used in the imperfect tense were excluded for a total of 2,985 excluded words altogether. In this manner, a word pool consisting of 696 words was created having the following features: (a) all high frequency words, (b) with one or two syllables, and (c) more than 1 but less than 6 phonemes in length. Twenty alternate forms of the phonemic segmentation task were developed from the 696 word high frequency word pool.

The measure *Facilité à Segmenter les Phonèmes (FSP)* is a timed one-minute direct measure of phonemic awareness. The measure assesses a student's fluency in segmenting a spoken word into its component parts or sound segments. Using standardized directions, the assessor orally presents a real word and asks the student to produce verbally the individual phonemes for each word. The assessor underlines each correct sound segment of the word that the student says. For example, the examiner says, *regard*, and the student says, / ʀ / / ə / / g / / a / / ʀ / to receive 5 possible points for the word (one point for each phoneme – no point for the sound /d/). After the student responds, the examiner presents the next word. For students not segmenting at the individual phoneme level, partial credit is given for partial or overlapping segmentation. For example, a student may segment the word *regard* as / ʀ ə / /g/ / aʀ / or as / ʀə / /əgaʀ / indicating a developing sense of phonemic awareness. Alternatively, a student may segment a word at the syllable level and say / ʀə / /gaʀ/. Allowing partial credit in the scoring increases the sensitivity of the measure, and makes it possible to measure skill growth from partial to complete segmentation with one measure. Although partial credit is given, the

preferred response is for students to completely segment words at the individual phoneme level. The number of correct sound segments (parties totales, PT) or syllable segments (syllabes totales, ST) produced in one minute determines the final score. The measure takes about 2 minutes to administer and has over 20 alternate forms for monitoring progress.

Facilité à lire des non-mots (FNM): Given the orthographic depth of the French language, two French pseudo-word measures were constructed based on word-level difficulty, one for across-year Kindergarten to beginning of first grade measure (Form 1), and one for middle of first grade to beginning of second grade (Form 2). To develop these, the most frequent letter sound units and most frequent letter sound patterns of the French language (i.e., CV, CVV, CCV, CVC, VCV, CVCV, CVVCV, CVCVV) were identified using a pronunciation guide that follows linguistic regularities for French vowels and French consonants.

We began with Form 2. Firstly, to create the nonsense words, we determined all possible simple graph, di-graph and tri-graph combinations of the language (e.g., a, e, é, oi, ou, au, eu, eau, oir, etc). Following the most frequent letter sound patterns of the language, we charted them in a matrix along with most frequent consonants of the language. From this matrix of pseudo-words, real words and inappropriate words were deleted from the corpus. For the Kindergarten level measure (Form 1), pseudo-words having uncommon French patterns were removed as well as any CVC word having soft and hard c and g occurring in the initial and medial letter positions. The resulting product was a large matrix of nonsense words having familiar, common French vowel sound combinations and typical consonant vowels patterns. Twenty alternate forms of the nonsense word task were developed for both Form 1 and Form 2.

The measure *Facilité à Lire des Non-Mots (FNM)* is a timed one-minute direct measure of the alphabetic principle, including knowledge of letter-sound correspondences, and the ability

to blend letter-sound combinations representing common basic sounds. In addition to assessing basic phonics knowledge, the measure assesses a student's ability to decode and to read pseudo words as whole words. For both Form 1 and Form 2, the student is presented an 8.5" x 11" sheet of paper with randomly ordered one and two-syllable nonsense words having common letter-sound patterns such as CVC, CVCV, and CVV for the Kindergarten beginning grade 1 level (e.g., upa, rur, bémé) probe, and common letter patterns such as CVCV, CVCVV, CVCVC, CVVCV, CVCVVC, CVCCVC, or CVCVCV (e.g., jonjin, tetou, doivi) for the middle of first grade probe. The student is asked to produce verbally either the individual letter sound or letter-sound combinations of each word or to read the whole nonsense word. For example, if the stimulus word is *nedou* the student could say /n/ /e/ /d/ /ou/ to obtain a total of four letter-sounds correct. The assessor underlines each correct letter sound produced either in isolation or blended together. For example, a student may decode the word *nedou* as a whole word. Whole words read without sounding out are underlined in their entirety. To calculate the final score, the assessor counts both the 'Nombre de sons corrects' (NSC) and Nombre de mots recodés' (NMR). 'Nombre de sons corrects' (NSC) is the number of letter sounds produced correctly in one minute. For example, if the student reads the *upa* as /u/ /p/ /a/, the score for NSC is 3. If the student reads the word *upa* as /u/ /pa/, the score is also 3. 'Nombre de mots recodés' (NMR) is the number of pseudo-words read correctly as a whole word without first being sounded out. For example, if the student reads the word *upa* as 'upa', the score is 3 points for NSC and 1 point for NMR.

The student is allowed one minute to produce as many 'Nombre de sons corrects' (NSC) and/or 'Nombre de mots recodés' (NMR) as he/she can in one minute. Ultimately, the goal is for students to read whole words. However, an advantage of the measure is that it allows for

monitoring the development of the alphabetic principle as early as the middle of kindergarten, when producing individual letter sounds is the more common response, through the middle of first grade when students should be reading whole words.

Facilité en lecture orale (FLO): Curriculum-based French language oral reading fluency passages were developed following strict oral reading fluency passage design specifications. All passages were written by a native French speaker in order to capture contemporary language use, and animate knowledge about culture and cultural traditions. A total of 80 passages were written following design specifications allowing for grade-level and age-level appropriate topic content and appropriate language-level parameters. A mixture of narrative, expository and/or informative genres was used. After the initial production, all passages went through several edits and revisions. Next, all passages were calibrated using SATO Calibrage (Daoust et al., 1996) a software product designed to analyze running text and determine text readability for French-language texts. Based on SATO passage readability outcomes, passages were assigned grade levels, one through five and several passages were revised again to fit within grade-level parameters. Passages that were selected for use in the study all retained similar ranges of grade-level difficulty. Grade-level ranges for first grades were set at 4.0 to 5.1, second grade 5.2 to 6.3, third grade 6.4 to 7.5, fourth grade 7.6 to 8.7, and for fifth grade, 8.8 to 9.9.

The measures *Facilité en lecture orale (FLO)* and *Rapport Oral du Récit (ROR)* are both timed one-minute measures. FLO assesses students' rate of accuracy and fluency when reading connected text (Fuchs, Fuchs, Hosp & Jenkins, 2001; Hasbrouck & Tindal, 2006), and ROR assesses reading comprehension. FLO is administered first, followed by ROR. The first component (FLO) asks the student to read aloud a reading passage, and the second component (ROR) asks the student to recall what they have read. ROR is intended to provide a

comprehension check for the FLO assessment, and provides an indication that the student is

reading for meaning. With a prompted passage retell, the student is instructed to read for

meaning as opposed to reading for speed.

FLO is a standardized set of passages and administration procedures designed to (a)

identify children who may need additional instructional support, and (b) monitor progress toward

instructional goals. The student is directed to read unfamiliar, grade-level passages aloud for one

minute. Words that are omitted and/or substituted are scored as errors unless self-corrected

within three seconds. Hesitations of more than three seconds are also scored as errors. After one

minute, the number of words read correctly determines the score. Students are directed to read

aloud three unfamiliar selected passages each for one minute. The score is the median number of

words read correctly and the median number of errors across the three passages. The oral reading

fluency component can be used from beginning of winter of first grade through spring of fifth

grade.

Case studies have documented students, otherwise unimpaired, who can read words but

who do not comprehend what they read (Dewitz & Dewitz, 2003; Hamilton & Shinn, 2003).

ROR provides an efficient procedure to identify students who have difficulties retelling what

they have read. Inclusion of ROR explicitly instructs students to be reading fluently for meaning.

The quality of a student's retell provides valuable information about overall reading proficiency

and oral language skills as during ROR, the student is asked to tell about what he/she has read.

The assessor indicates the number of words in the retell that are related to the story by drawing a

line through a box of numbers whereby each number represents a word. Following a hesitation of

3 seconds, students are prompted to tell as much as they can about the story. If there is a second

hesitation of 5 seconds or more, or if the student is clearly responding in a way that is not

relevant to the passage for 5 seconds, the task is discontinued. At the end of one minute, the

assessor makes a judgment about the quality of the retell based on the level of detail reported in

the retell using the quality of response rating. The qualitative rating is based on how well the

student retold the portion of the passage that he/she read. Retell can be used from the middle of

first grade through the spring of fifth grade. Table 4 outlines the core components of early

literacy with corresponding IDAPEL® measure name.

Table 4

Core Components of Early Literacy Skill with Corresponding IDAPEL Measure Name

Core Component of Early Literacy Skill	IDAPEL Measure
Phonemic Awareness	Facilité à reconnaître le premier son (FPS) Facilité à segmenter les phonèmes (FSP)
Alphabetic Principle	Facilité à dénommer des lettres (FDL) Facilité à lire des non-mots (FNM) Facilité en lecture orale (FLO)
Accuracy and Fluency reading Connected Text	Facilité en lecture orale (FLO)
Text Comprehension	Facilité en lecture orale (FLO) Rappel oral du Récit (ROR)

General description of ÉCOLE criterion measures. ÉCOLE (*Épreuves de COmpétence

en LEcture*) includes a broad set of reading-related measures initially designed for the diagnostic

assessment of reading difficulties. These measures, however, were found useful as criterion

measures in validity studies, in early screening studies and program evaluation studies

(Desrochers, 2011). The measures used in the present study were selected to match as closely as

possible the constructs underlying the IDAPEL measures. Table 5 provides a list of the measures

along with their Cronbach Alpha coefficient and their test-retest reliability coefficients. The

constructs measured included phonological awareness, letter knowledge, oral reading, and

reading comprehension. In most tasks, items were presented in increasing order of difficulty.

When four errors were made within a series of six trials, the stop rule applied, and the test was

discontinued. We refer to this procedure as the standard scoring procedure. Descriptions of the

different measures follow.

Table 5
ÉCOLE Measures Selected for the Study

Measure	Cronback Alpha Coefficient
Identification of Initial Phoneme	.96
Phonemic Segmentation	.89
Phonemic Elision	.93
Contextualized Grapheme Sounding	.95
Word Reading	.95
Nonword Reading	.90
Text Reading (Marie)	n/a
Picture-Word Matching	n/a
Picture-Sentence Matching	n/a

Identification of Initial Phoneme (IIP). In this task, participants were asked to produce

the first phoneme of each word presented orally by the examiner. Participants were guided

through 4 practice items followed by the experimental items (Max = 25). The standard scoring

procedure, as described above, applied. Items were scored correct if the participant produced the

correct response initially, or if they self-corrected within 30 seconds. Items were scored

incorrect if the response was incorrect or if the participant failed to respond.

Phonemic Segmentation (PHS). In this task, participants were asked to verbally produce

each phoneme of each word presented orally by the examiner. The task included four practice

items again followed by the experimental items (Max = 25). Standard scoring procedures were

followed. Items were scored correct if all phonemes were produced in the right order; otherwise, they were scored incorrect.

Phonemic Elision (PHE). In this task, participants were asked to remove a phoneme in one of several possible positions and to indicate what the word was after phoneme removal. After the guided four practice item procedure, experimental items (Max = 40) were presented until the stop rule applied. All items were real words produced orally by the examiner. The standard scoring procedure was applied.

Contextualized Grapheme Sounding (GRS). In this task, participants were shown 44 short nonwords with one grapheme underlined and asked to sound out the underlined grapheme. This task permitted us to verify the mastery of contextual grapheme-phoneme correspondences in French. The standard scoring procedure was applied.

Word Reading (WRead). In this task, word items were shown in four rows per page. Participants were asked to read them aloud starting from the top of each page. Participants were guided through four practice items, followed by the experimental items (Max = 44). The words were constructed using several types of graphemes: single-letter regular graphemes, multiple-letter regular graphemes, contextual graphemes (e.g. c, g), or irregular graphemes. Regardless of their graphemic structure, these words were presented in increasing order of difficulty. The standard scoring procedure was applied.

Nonword reading (NWRead). This task is identical to the previous task, with the exception that word items were pronounceable nonwords.

Text Reading (TXTRead). This test was designed by Simonart (2006). Participants are asked to read aloud a story about a character called Marie. As the story unfolds, word and sentence structures become increasingly more complex and difficult. The task was stopped after

one minute of reading. Each line of text includes ten words for a total of 150 words over the entire text. The score was established by counting the number of words read correctly within one minute.

Word-Picture Matching (WPM). In this task, participants were given two minutes to categorize as many words as they could. On a given probe, five icons representing different semantic categories such as body parts, articles of clothing, colors, birds, and fruits are presented. Word items are shown in separate rows, with five rows per page. Participants are guided through five practice items, then the 35 experimental items. The standard scoring procedure was applied.

Sentence-Picture Matching (SPM). In this task, participants were asked to read sentences and indicate which of the four associated scenes it described. Each item was shown on a separate page and included a line drawing depicting four different scenes (e.g., a young girl wearing a skirt and a hat) and a sentence (e.g. *The girl is wearing pants and a hat*).This task involved one practice item and a total of 22 experimental items. The task was timed for 2 minutes. The standard scoring procedure was applied.

<div align="center">Overview of the Statistical Analyses</div>

The statistical analyses reported here were intended to address four distinct issues. First, analyses of variance with the repeated measures over the school year (Fall, Winter, and Spring) were carried out to track the sensitivity of the IDAPEL measures to successive skill gains. Similar analyses were performed on the ÉCOLE measures across grades. Second, we examined the reliability of the IDAPEL measures by examining the relationships between alternate forms of the same measures taken two weeks apart. Third, we established the construct validity of the IDAPEL indicators by calculating their correlations with an independent set of conceptually

related measures (École), both taken at year end. Finally, we investigated the predictive validity

of the IDAPEL indicators by computing the correlations between all performance measures

taken at the beginning and the end of the school year.

<div align="center">Results</div>

Because datasets needed to be merged for data analysis, data management procedures

were followed and unique student identification numbers were used to assure correct merging

and analysis of each dataset. The total sample size from the DMG dataset was 294. During an

examination of this data set, we observed that complete raw score IDAPEL data was entered for

the fall benchmark period for the three grade levels, but less so for the winter and spring data

sets. Researchers were not alerted about the missing data and WBTT data entry was not within

our control. Table 6 provides sample size information by grade and by time period, as well as the

percentage for which raw data was entered for each time period. Complete IDAPEL across-year

data is available for 59% of Kindergarten students, for 77% of First grade students and for 75%

of Second grade students.

Table 6
Sample by Grade, Size and Time Period, and Percentage of Data Entered by Time Period

Grade	Fall	% entered	Winter	% entered	Spring	% entered
K	132	100	122	80	92	66
1	110	100	87	81	87	81
2	50	100	37	79	37	79

Note. Data exported from the Web-Based Teaching Tool (WBTT) database sponsored by the Learning Disabilities Association of Ontario (LDAO).

Growth Over the School Year

IDAPEL measures were expected to reflect gradual increases in performance over the

school year at each grade level. Table 7 provides descriptive statistics on the rates of growth for

each of the IDAPEL measures across the three grades, Kindergarten to 2nd grade. For each grade

level, and for each of the measures administered at that grade, the mean, standard deviation and

sample size number are reported for the three data collection points, beginning of (fall), middle

of (winter), and end-of-year (spring). Evidence of skill growth, as measured by IDAPEL, is

distinct across each grade.

Table 7

IDAPEL Measure Descriptive Statistics Within and Across Grades

Measure	Beginning of year			Middle of year			End of year		
	Mean	*SD*	*n*	Mean	*SD*	*n*	Mean	*SD*	*n*
Kindergarten									
FPS	15.92	10.49	132	23.07	11.75	132	—	—	—
FDL	20.94	12.64	132	27.74	14.79	122	32.08	14.13	91
FSP	—	—	—	15.1	11.13	122	28.85	18.95	91
FNM	—	—	—	18.06	11.02	121	25.49	15.73	91
NMR	—	—	—	1.26	2.87	120	3.31	5.89	90
First Grade									
FDL	36.36	14.55	108	—	—	—	—	—	—
FSP	30.73	17.53	108	35.17	13.85	86	43.62	14.51	86
FNM	33.92	25.35	108	53.93	30.14	86	77.42	35.03	87
NMR	5.26	8.23	108	8.1	13.31	86	13.99	20.95	86
FLO	—	—	—	40.04	26.23	77	61.74	33.14	85
ROR	—	—	—	13.74	8.23	48	17.68	8.19	57
Second Grade									
FNM	66.64	39.2	47	—	—	—	—	—	—
NMR	15.4	10.7	44	—	—	—	—	—	—
FLO	62.5	36.14	47	73.84	36.29	35	81.02	40.76	36
ROR	2.84	1.07	11	20.01	13.97	35	18.1	11.88	36

Note. FPS = Facilité à reconnaître le premier son; FDL = Facilité à dénommer des lettres; FSP = Facilité à segmenter les phonèmes; FNM = Facilité à lire des nonmots; NMR = Nombre de mots recodés; FLO = Facilité en lecture orale.

The growth patterns observed through visual inspection of the mean scores were verified through analyses of variance with a Bonferonni correction. Table 8 reports the number of valid cases for each test, the F ratio and associated degrees of freedom, and the significance level. All analyses on the Kindergarten and Grade 1 data were carried out with sufficient statistical power ($\geq$ .92); however, for the analyses on the Grade 2 data, statistical power was found insufficient ($\leq$ .16) for ROR. Omnibus tests on data collected in the Fall, Winter, and Spring were followed with multiple comparisons. It was found that whenever the Omnibus test was significant mean scores differed significantly from Fall to Winter, and from Winter to Spring ($p < .001$).

Table 8

ANOVA on IDAPEL Measures as a Function of Time of Testing

Variables	Valid Cases	F ratio	df	Significance p
Kindergarten				
FPS	120	64.23	1, 120	< .001
FDL	82	96.02	2, 80	< .001
FSP	83	69.93	1, 82	< .001
FNM	83	54.53	1, 82	< .001
NMR	83	17.07	1, 82	< .001
Grade 1				
FSP	84	48.41	2, 82	< .001
FNM	84	108.56	2, 82	< .001
NMR	84	7.15	2, 82	0.001
FLO	84	184.6	1, 83	< .001
ROR	48	23.31	1, 47	< .001
Grade 2				
FLO	35	41.77	2, 33	< .001
ROR	10	0.74	2, 8	0.51

Note: FPS: Facilité à reconnaître le premier son; FDL: Facilité à dénommer des lettres; FSP: Facilité à segmenter le phonèmes; FNM: Facilité à lire des non-mots; NMR: Nombre de mots recodés; FLO: Facilité en lecture orale; ROR: Rapport oral du récit.

The results of these analyses are clear and consistent: on most measures the mean scores increased significantly over the course of the school year. Even though the items used in each test at the beginning, middle and end of the year were different, the general pattern is consistent with the conclusion that all IDAPEL measures are sensitive to learning growth over the three time points in Kindergarten, Grade 1, and Grade 2. The only exceptions to this pattern are observed in Grade 2, with ROR. In this case the exceedingly low sample size resulted in a loss of statistical power and may have made the statistical tests insensitive to the observed score differences. As a consequence the results observed for Grade 2 must be interpreted with caution.

Growth across School Years

In this validation study of IDAPEL measures, the ÉCOLE measures were primarily intended as outcome measures. As in the preceding section, we verified that the ÉCOLE measures are sensitive to growth in knowledge and skills from one year end to the next. Table 9 reports the mean scores, standard deviations and number of valid cases across school grades. The presence of differences across and between grade levels was tested with an omnibus analysis of variance and a test of multiple comparisons (Least Significant Difference) among mean scores. The results of these tests are reported in Table 10.

These results confirm that all ÉCOLE measures, except Identification of initial phoneme (IIP), are sensitive to change across grades. Multiple comparisons confirm that significant mean score differences are observed between Kindergarten and Grade 1 (and between K and Grade 2) for all measures except IIP. Significant differences between Grade 1 and Grade 2 are detected only for WRead, NWRead, TXTRead, WPM, and SPM. The last set of results must be interpreted with caution, however, because the number valid cases in Grade 2 is restricted to 23 students in this study

31

Table 9

End of Year École Subtest Descriptive Statistics by Grade Level

Kindergarten

	IIP	GRS	PHE	PHS	NWread	WRead	CWP	CSP	TXTRead
Mean	17.18	29.05	10.81	3.84	4.23	3.68	2.03	4.49	7.13
SD	7.14	10.16	8.10	4.31	4.73	2.32	1.77	3.44	8.68
n	80	80	80	80	80	80	67	69	66

First Grade

	IIP	GRS	PHE	PHS	NWread	WRead	CWP	CSP	TXTRead
Mean	17.02	65.10	24	9.69	22.78	17.29	10.65	21.71	52.35
SD	7.44	7.38	8.44	6.43	8.42	7.74	3.12	5.42	9.47
n	51	51	51	51	51	51	51	51	51

Second Grade

	IIP	GRS	PHE	PHS	NWread	WRead	CWP	CSP	TXTRead
Mean	18.36	68.36	25.04	8.40	27.12	22.96	12.16	25.84	66.32
SD	5.87	7.36	10.29	6.40	9.45	9.80	2.98	5.51	21.26
n	25	25	25	25	25	25	25	25	25

Note: IIP = First Phoneme Identification; GRS = Grapheme Identification; PHE = Phonemic Elision; PHS = Phonemic Segmentation;NWRead = Non-Word Oral Reading, Wread = Oral Word Reading; CWP = Word Picture Match; CSP = Sentence Picture Match; TXTRead = Oral reading of text « Marie ».

Table 10

ANOVA and Post Hoc Tests on École Measures Across Grades

Variables	Valid cases	F ratio	Df	Significance P	Least significant differences
IIP	159	0.14	2,156	0.87	K = Gr1 = Gr2
PHS	159	20.89	2,156	< .001	K < Gr1 = Gr2
PHE	159	49.39	2,156	< .001	K < Gr1 = Gr2
GRS	159	344.34	2,156	< .001	K < Gr1 = Gr2
WRead	159	137.74	2,156	< .001	K < Gr1 < Gr2
NWRead	159	164.54	2,156	< .001	K < Gr1 < Gr2
TXTRead	145	327.55	2,142	< .001	K < Gr1 < Gr2
WPM	148	308.09	2,145	< .001	K < Gr1 < Gr2
SPM	146	220.16	2,143	< .001	K < Gr1 = Gr2

Note: IIP: Identification of Initial Phoneme; PHS: Phonemic Segmentation; PHE: Phonemic Elision; WRead: Word Reading; NWRead: Nonword Reading; TXTRead: Text Reading (Marie); WPM: Word-Picture Matching in 2 min; SPM: Sentence-Picture Matching in 2 min.

Overall, there is clear empirical evidence that most of the chosen outcome measures are sensitive to the growth of knowledge and skills between the end of Kindergarten and the end of Grade 1. The absence of significant differences for IIP may indicate that the identification of initial phonemes, as a basic phonological awareness skill, is practiced extensively in Kindergarten and no further gain is to be expected beyond the Kindergarten school year. Importantly for further analyses, these scores are not constrained by ceiling effects; a reasonable amount of variance is observed in all cells of Table 9.

Reliability of Alternate Forms of the Same Test

Educational tools are considered reliable if they are stable across time, when minimal learning has occurred, and across items (Henson, 2001). An appropriate procedure for estimating alternate-form reliability over a two-week interval is the Spearman-Brown Prophesy Formula (Nunnally & Bernstein, 1994). The resulting alternate-form reliability index indicates the extent to which test results generalize to different item samples, different times of testing, different conditions, and different testers. In the present study, two different (i.e., alternate) forms of the same test were administered two weeks apart. Alternate-form reliability of a single-form was estimated by the correlation between the score recorded at time point A and the score recorded at time point B. If we view these two forms as two halves of the same test then this coefficient represents an estimate analogous to a split-half reliability coefficient – a reliability estimate of each half of the test. The Spearman-Brown Prophesy Formula can then be applied to obtain a reliability estimate for the two halves of the test, as if the test included twice as many items. Salvia, Ysseldyke, & Bolt's (2007) standards for evaluating reliability are as follows: .60 is a required minimum for administrative purposes and reported group scores, .80 is a required minimum for screening decisions, and .90 is a required minimum for important educational

decisions concerning an individual student. These are the decision rules we adopted in the present study.

Table 11 reports the zero-order correlation coefficients between the scores of the two forms (measured at time point A and two weeks later at time point B) and the Spearman-Brown coefficients for the following measures taken in Kindergarten and Grade 1: FPS (Facilité à reconnaître le premier son), FDL (Facilité à dénommer des lettres), FSP (Facilité à segmenter les phonèmes), FNM (Facilité à lire des non-mots), and NMR (Nombre de mots recodés).

Table 11

Two-week Alternate Form Reliability Estimates for IDAPEL measures

| Measure | N | Two-week Alternative form Reliability | | Time of testing |
		Single-Form	Two-Form	
Kindergarten				
FPS	47	.77(B)	.87(B1)	Winter
FDL	47	.81	.90	Winter
FSP	47	.76	.86	Winter
FNM I	47	.73	.84	Winter
NMR	47	.72	.84	Winter
First Grade				
FDL	47	.72(A)	.84(A1)	Fall
FSP	47	.53	.69	Fall
FNM I	47	.78	.88	Fall
NMR	47	.74	.85	Fall
FNM II	46	.80(C)	.89(C1)	Spring
NMR	46	.33	.50	Spring

Note: Alternate form reliability coefficients are based on benchmark administration time point A, B and C with respective two-week alternate form data (A1, B1, C1). FDL = Facilité à dénommer des lettres; FPS = Facilité à reconnaître le premier son; FSP = Facilité à segmenter les phonèmes; FNM I= Facilité à lire des non-mots, simple graphemes; NMR = Nombre de mots recodés; FNM I = Facilité à lire des non-mots, complex graphemes; A = Beginning-of-year; B = Two weeks later beginning-of-year.

Although these coefficients are influenced by two sources of variability, alternate forms of the same test and time of testing, they provide an estimate of common variance between forms. All coefficients, except for 1, range between .53 and .81. Most of them approached Salvia et al.'s (2007) reliability criterion for screening decisions (i.e. .80). The phoneme segmentation

fluency (FSP) coefficient at beginning-of-year first grade is within the administrative purpose range (i.e. .69) but falls short for screening decisions. The word recoding (NMR) measure at the end of Grade 1 is inadequate in its present form and requires some improvement.

Additional reliability information can be drawn from the data for FLO and ROR since three forms of these measures were taken at different time points in Grade 1 and Grade 2. The zero-order correlations among the different forms of these tests are reported in Table 12.

Table 12

IDAPEL FLO Passage-Level Correlations Within Time of Year for Grades 1 and 2

Passage Combination	FLO by Grade		ROR by Grade	
	First	Second	First	Second
Beginning of Year				
1, 2	--	.97 (47)	--	.82 (11)
1, 3	--	.96 (47)	--	.87 (11)
2, 3	--	.97 (47)	--	.90 (11)
Middle of Year				
1, 2	.98 (79)	.94 (35)	.80 (49)	.88 (35)
1, 3	.95 (79)	.94 (35)	.77 (48)	.81 (35)
2, 3	.95 (89)	.97 (35)	.87 (48)	.80 (35)
End of Year				
1, 2	.95 (86)	.98 (36)	.66 (58)	.86 (36)
1, 3	.94 (85)	.98 (36)	.75 (57)	.86 (36)
2, 3	.96 (85)	.98 (36)	.66 (57)	.89 (36)

Note. Pair-wise sample sizes are indicated in parenthesis. FLO = Facilité en lecture orale; ROR = Rapport oral du récit. BOY = Beginning of Year, MOY = Middle of Year, EOY = End of Year. All correlations are significant, *p* < .001.

The data presented in Table 12 indicate that the scores on the different forms of oral reading fluency (FLO) and retell fluency (ROR) taken closely in time are typically highly correlated with one another. Most correlation coefficients range from .80 to .98. The one subset of measures that deviates from this pattern is ROR measured at the end of Grade 1. The lower correlations may reflect problems with the measures themselves or with the teacher assessment procedure. These measures need further attention and they may require improvement. Despite these qualifications, this set of results indicate that most of these measures meet Salvia et al.'s reliability standard for making important educational decisions concerning individual students.

Relationships Between the IDAPEL and the ÉCOLE Measures

The present study provided us an opportunity to investigate the construct validity of the IDAPEL measures by investigating their relationships with other tests designed to measure the same constructs. We examined concurrent, criterion-related validity of the IDAPEL measures by reviewing at student performance on the IDAPEL assessments along with their performance on the criterion ÉCOLE measures both taken at the end of the school year. These relationships were assessed by computing the Pearson Product-Moment correlation among conceptually related scores for K and Grade 1. Grade 2 data were discarded from these analyses because the student sample size was insufficient for estimating reliability coefficients. Relevant correlations are reported in table 13.

Table 13

Concurrent Criterion-Related Validity of IDAPEL and ÉCOLE Measures in Kindergarten and Grade 1.

IDAPEL Measure	ÉCOLE Measure	Kindergarten		Grade 1	
		n	*r*	*n*	*r*
FSP	IIP	49	.54**	37	-.09
FSP	PHS	49	.54**	37	.18
FSP	PHE	49	.50**	37	.25
FDL	GRS	49	.68**		
FNM	GRS	49	.65**	37	.52**
FNM	WRead	49	.62**	37	.58**
FNM	NWRead	49	.68**	37	.44**
FNM	TXTRead	48	.70**	37	.63**
NMR	WRead	49	.33*	37	.17
NMR	NWRead	49	.28*	37	.10
NMR	TXTRead	48	.32*	37	.21
FLO	GRS			37	.71**
FLO	WRead			37	.87**
FLO	NWRead			37	.66**
FLO	TXTRead			37	.74**
FLO	WPM			37	.60**
FLO	SPM			37	.68**
ROR	WRead			26	.61**
ROR	NWRead			26	.52**
ROR	TXTRead			26	.66**
ROR	WPM			26	.60**
ROR	SPM			26	.63**

Note. FSP: Facilité à segmenter les phonèmes; FDL : Facilité à dénommer les lettres; FNM : Facilité à lire les nonmots; FLO : Facilité en lecture orale; ROR : Rapport oral du récit; IIP : Identification of initial phoneme; PHS : Phonemic Segmentation; PHE : Phonemic Elision; GRS : Grapheme Sounding; WRead : Word reading; NWRead : Nonword reading; TXTRead : Text reading; WPM : Word-Picture matching; SPM : Sentence-picture matching. Symbol * means $p < .05$; ** means $p < .01$.

By design, IDAPEL assessments measure specific reading-related skills and the fluency of their deployment. By contrast, a large proportion of the ÉCOLE measures used in this study are self-paced and focused on response accuracy only. Therefore, several of the correlation coefficients reported in Table 13 exclude the common variance that would be attributable to method similarity (i.e., in relating two fluency measures). Although this exclusion was expected to lower the observed correlation coefficients, most were found to be equal or above .50.

Most IDAPEL measures that tapped phonological awareness, decoding or reading skills were positively and significantly correlated with conceptually related ÉCOLE measures. The two exceptions to this pattern are found among the measures used at the end of Grade 1. For instance, FSP was not significantly correlated with any of the phonological awareness measures of the ÉCOLE battery in Grade 1; whereas a clear linear relationship was observed at the end of Kindergarten. Since the same ÉCOLE measures were used in both grades, the evidence suggests that the FSP test at the end of Grade 1 needs further examination and possibly some improvement. The second exception to the general pattern pertains to NMR. However, as noted earlier, the reliability of both FSP and NMR have been found to be a cause for concern. The magnitude of the mean scores and standard deviations reported in Tables 7 and 9 seem to rule out the possibility of a ceiling effect in Grade 1 performance on these measures. These results suggest that the FSP and NMR measures used at the end of Grade 1 require closer analysis.

Predicting Year-End Outcome

The IDAPEL measures were primarily designed to track student progress over the school year. They could also serve to predict future success or failure in learning to read. This study permits us to explore the predictive validity of the IDAPEL measures by examining the

relationship between measures taken at the beginning of the school year and those taken at year

end. In this section we report for Kindergarten and Grade 1 the correlations between measures

taken at the beginning and the end of the school year. The correlations that pertain to

Kindergarten and Grade 1 are reported in tables 14 and 15, respectively.

Table 14
Correlations Among the Measures Taken at The Beginning and
The End of Kindergarten

Measures taken at year end	Measures taken in the Fall			
	FPS		FDL	
	r	*n*	*r*	*n*
FDL	.61**	84	.76**	84
FSP	.55**	84	.50**	91
FNM	.60**	84	.67**	84
NMR	.40**	84	.45**	84
IIP	.55**	79	.41**	79
PHS	.53**	79	.36**	79
PHE	.56**	79	.57**	79
GRS	.57**	79	.61**	79
WRead	.59**	79	.50**	79
NWread	.58**	79	.46**	79
TXTread	.58**	65	.65**	65

*Note.*FSP: Facilité à segmenter les phonèmes; FDL : Facilité à dénommer les
lettres; FNM : Facilité à lire les nonmots; FLO : Facilité en lecture orale;
ROR : Rapport oral du récit; IIP : Identification of initial phoneme; PHS :
Phonemic Segmentation; PHE : Phonemic Elision; GRS : Grapheme
Sounding; WRead : Word reading; NWRead : Nonword reading; TXTRead :
Text reading. Symbol * means $p < .05$; ** means $p < .01$.

The pattern of these results is clear and consistent: measures of FPS and FDL taken at the

beginning of Kindergarten are positively and significantly correlated with all measures taken at

year end (i.e., IDAPEL as well as ÉCOLE measures). All coefficients except for one are equal to

or greater than .40. These results strongly suggest that FPS and FDL are useful predictors of future performance in Kindergarten and they could be used to screen children at risk for difficulties in phonological analysis, letter learning, letter-string decoding, and single word reading.

The general pattern observed in the Grade 1 data is similar to that observed in the Kindergarten data: Measures taken at the beginning of the school year are positively and significantly correlated with most measures taken at year end. The significant correlations range from .19 to 74. The largest coefficients are found between early FDL and FNM and the late measures of oral reading (e.g. FLO, WRead). An unexpected finding is that the measure of NMR taken in the fall is not significantly correlated with the same measure taken in the spring. The dubious reliability of this measure has been noted earlier in this report. We now provide further evidence that it requires further analysis and adjustment. Overall, the results strongly suggest that FDL, FSP, and FNM are useful predictors of future performance in reading as indexed by multiple outcome indicators. According to Hopkins (2002) standards for predictive validity, correlations smaller than .09 indicate a very small relationship, .10 to .29 a small relationship, .30-.49 a moderate relationship, .50-.60 a moderate-strong relationship, and above .70 a strong relationship. The coefficients reported in Table 15 cover the entire range of these nominal categories. For the purpose of predicting future reading performance, FDL and FNM clearly are the best IDAPEL indicators.

Table 15

Correlations Among the Measures Taken at The Beginning and The End of Grade 1

Measures taken at year end	Measures taken in the Fall							
	FDL		FSP		FNM		NMR	
	r	*n*	*r*	*n*	*r*	*n*	*r*	*n*
FSP	0.17	86	.59**	86	.19*	86	0.14	86
FNM	.63**	86	.28**	86	.64**	86	.47**	86
NMR	.28**	86	.30**	86	.29**	86	0.07	86
FLO	.64**	86	.33**	86	.74**	86	.60**	86
ROR	.42**	58	0.06	58	.25*	58	.25*	58
IIP	0.04	51	0.05	51	.29*	51	.23*	51
PHS	0.16	51	.41**	51	.24*	51	.25*	51
PHE	.37**	51	.32*	51	.51**	51	.47**	51
GRS	.62**	51	.43**	51	.46**	51	.38**	51
WRead	.63**	51	.28*	51	.70**	51	.60**	51
NWread	.57**	51	0.16	51	.47**	51	.40**	51
TXTread	.68**	51	.24*	51	.47**	51	.36**	51
WPM	.48**	51	0.21	51	.41**	51	.37**	51
SPM	.51**	51	.35**	51	.42**	51	.31*	51

Note. FSP: Facilité à segmenter les phonèmes; FDL: Facilité à dénommer les lettres; FNM: Facilité à lire les nonmots; NMR: Nombre de mots recodés; FLO: Facilité en lecture orale; ROR: Rapport oral du récit; IIP: Identification of initial phoneme; PHS: Phonemic Segmentation; PHE: Phonemic Elision; GRS: Grapheme Sounding; WRead: Word reading; NWRead: Nonword reading; TXTRead: Text reading; WPM: Word-Picture matching; SPM: Sentence-picture matching. Symbol * means $p < .05$; ** means $p < .01$.

Discussion

The primary purpose of the present study was to investigate the psychometric properties and usefulness of IDAPEL. Four general issues were considered: a) How sensitive are these indicators to changes in reading-related abilities over a school year? b) How reliable are these measures? c) How do these measures relate to an independent set of measures that tap the same constructs? And d) Can the measures taken at the beginning of a school year predict reading-related performance at year end? The results reported here provide fairly clear answers to these questions.

The main strategy for addressing the first question consisted of verifying if statistically significant change was detected from one measurement point to the next. The results indicate that most IDAPEL measures are sensitive to the change that occurred between the three time points at which data were collected (i.e. Fall, Winter, and Spring). The only measure that showed non-significant sensitivity is ROR in Grade 2. Grade 2 test results must be interpreted with caution as its sample size was exceedingly low, resulting in deflated statistical power. The same analyses were carried out on the ÉCOLE data to verify that these measures were also sensitive to changes in knowledge and skills from one year end to the next. Except for IIP, all measures were sensitive to change from the end of Kindergarten to the end of Grade 1. From Grade 1 to Grade 2, only measures of reading were found to be sensitive to skill growth; measures of phonological awareness and alphabetic knowledge were not. Again, the low sample size in Grade 2 may have prevented us from ascertaining real differences. Beyond these analyses Grade 2 data were not considered further.

The reliability of IDAPEL measures was investigated via the administration of alternate forms of the same test over a two-week time interval. Test-retest correlations and the application of

the Spearman-Brown Prophesy Formula confirmed that the reliability of most IDAPEL measures is satisfactory for mostly screening decisions and some for making important educational decisions at the individual student level. Weaker results were found for NMR and FSP in Grade 1. Although the current state of FSP is within the range of acceptability, NMR does require further attention and improvement. A similar analysis indicated that the correlations among alternate forms of FLO and ROR are typically very high ($r \geq .80$), but lower at the end of Grade 1 (r between .66 and .75).

Construct validity was examined by inspecting the linear relationship between the IDAPEL measures and conceptually related ÉCOLE measures, both taken at the end of Kindergarten and Grade 1. Most IDAPEL measures were positively and significantly correlated with conceptually related ÉCOLE measures. FSP and NMR in Grade 1 were less strongly correlated with the outcome measures but this pattern is not entirely surprising since we found evidence that the reliability of these particular measures are on the low side of the spectrum. Such a condition is expected to lower the upper bound of their correlation with other measures and this is what was observed.

One desirable function of IDAPEL would be to help screen children at risk for future reading difficulty. Although the ideal time for such a screening operation is still an open question, there is fairly strong agreement that it should be done early, either in Kindergarten or early in Grade 1 (Fréchette & Desrochers, 2011; Rathvon, 2004). A required feature of screening tools is demonstrated predictive validity. To explore this feature we examined the correlation between all IDAPEL measures taken at the beginning of the school year and all measures taken at year end. In Kindergarten, early FPS and FDL were found to be strongly correlated with year-end measures of phonological awareness, letter knowledge and reading. In Grade 1 FDL and FNM were found to

be the most potent and consistent predictors of year-end measures of single-word and sentence reading.

To sum up, the present study permitted us to investigate several aspects of the IDAPEL measures: their sensitivity to change in student skills, their reliability, their construct validity, and their predictive validity. Even though some of these measures displayed some weaknesses, we report clear initial empirical evidence that the IDAPEL battery can serve several useful purposes (e.g. progress monitoring, screening children at risk of reading difficulty) and justify important screening decisions on individual students' standing.

As with any empirical study, continued study replication with similar sets of students remains important. The importance and need to replicate and extend findings with larger samples of student participants as well as with additional criterion measures cannot be understated.

References

Adams, M.J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA:

MIT Press.

Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a

difference in early word recognition and developmental spelling? *Reading Research*

Quarterly, 24(1), 49-66.

Borzone de Manrique, A., & Signorini, A. (1994). Phonological awareness, spelling, and reading

abilities in Spanish-speaking children. *British Journal of Educational Psychology 64*, 429-

439.

Blachman, B. A. (2000). Phonological Awareness. In M. L. Kamil, P. B. Mosenthal, P. D.

Pearson, & R. Barr (Eds.), *Handbook of Reading Research,* (Vol. III, pp. 483-502). New

Jersey: Lawrence Erlbaum.

Carnine, D., Silbert, J., & Kame'enui, E. (1997). *Direct Instruction Reading* (3rd Ed). New York:

Merrill.

Comeau, L., Cormier, P., Grandmaison, E., & Lacroix, D. (1999). A longitudinal study of

phonological processing skills in children learning to read in a second language. *Journal of*

Educational Psychology 91(1), 29-43.

Content, A. (1993). Le rôle de la médiation phonologique dans l'acquisition de la lecture. In J. P.

Jaffre, L. Sprenger-Charolles, & M. Fayol (Eds.), *Lecture-écriture: Acquisition* (pp. 80-96).

Paris: Nathan.

Cossu, G., Shankweiler, D., Liberman, I. Y., Katz, L., & Tola, T. (1988). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics 9*, 1-16.

Daoust, F., Laroche, L., & Ouellet, L. (1996). *SATO. Système d'analyse de texte par ordinateur: version 4.0: manuel de référence.* Montréal, Service d'analyse de textes par ordinateur (SATO).

Demont, E. (1994). Conscience phonologique, conscience syntaxique. In G. Gregoire & B. Piérart (Eds.), *Evaluer les troubles de la lecture: Les nouveaux modèles théoriques et leurs implications diagnostiques* (pp. 195-208). Gembloux, Belgique: De Boeck Université.

Desrochers, A. (2011). *ÉCOLE: épreuves de compétence en lecture*. Manuscript in preparation, School of Psychology, University of Ottawa, Ottawa, Canada.

Desrochers, A., Cormier, P., & Thompson, G. (2005). La sensibilité phonologique et l'apprentissage de la lecture. La Parole, 34-35-36, 113-138.

Desrochers, A., & Glickman, V. (2008). *National Strategy for Early Literacy. Reading assessment policy research: Understanding elementary reading assessment options.* London, ON: Canadian Language and Literacy Research Network. Retrieved July 10, 2009, from http://docs.cllrnet.ca/NSEL/Current/NSEL_ElementaryReading.pdf

Desrochers, A., Simon, M., & Thompson, G. L. (2011). Formes et fonctions de l'évaluation de la littératie. In M. J. Berger & A. Desrochers (Eds.), *L'évaluation de la littératie* (pp. 29-61). Ottawa, Canada : Presses de l'Université d'Ottawa.

Dewitz, P., & Dewitz, P. K. (2003). They can read the words, but they can't understand: Refining comprehension assessment. *Reading Teacher, 56,* 422–435.

Dufour-Martel, C., & Good, R. H. (2009). Investigating the psychometric properties of three French language early reading measures, *Effective Education, 1,* 1, 87-100.

Fréchette, S., & Desrochers, A. (2011). Le dépistage des élèves à risque de difficulté en lecture. In M. J. Berger & A. Desrochers (Eds.), *L'évaluation de la littératie* (pp. 81-116). Ottawa, Canada : Presses de l'Université d'Ottawa.

Fuchs, L. S., & Deno, S. L. (1991). Curriculum-based measurement: Current applications and future directions. *Exceptional Children, 57,* 466-501.

Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice, 13,* 204-219.

Fuchs, L.S., Fuchs, D., Hosp, M.K., and Jenkins, J.R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. Scientific Studies of Reading, 5, 239-256.

Good, R. H., III, Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*(3), 257-288.

Grandmaison, E. (1996). *Longitudinal relationships among phonological awareness, verbal working memory, lexical access, and reading achievement in English-speaking children placed in French immersion.* Conference papers presented at Quebec City, Quebec, Canada.

Hamilton, C., & Shinn, M.R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review*, *32*, 228–240.

Hart, B. & Risley, T. R. (1995). Meaningful Differences in the Everyday Experience of Young

American Children. Brookes Publishing

Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: a valuable assessment

tool for reading teachers. *The Reading Teacher, 59, 7*, 636-644.

Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual

primer on coefficient alpha. Measurement and Evaluation in Counseling and Development.

Retrieved on-line 2/16/2011 http://www.highbeam.com/doc/1G1-80370171.html

Hopkins, W. G. (2002). A scale of magnitudes for the effect statistics. In *A review of statistics*.

Retrieved July 30, 2010 from http://www.sportsci.org/resource/stats/effectmag.html.

Huang, H. S., & Hanley, R. J. (1995). Phonological awareness and visual skills in learning to read

Chinese and English. *Cognition 54*, 73-98.

Kaminski, R. A., & Good, III, R. H. (1996). Toward a technology for assessing basic early literacy

skills. *School Psychology Review 25*, 215-227.

Lecocq, P. (1991). *Apprentissage de la lecture et dyslexie* [Learning reading and dyslexia].

Liege: Mardaga.

Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A grade-level lexical database

from French elementary school readers. *Behavior Research Methods, Instruments, &

Computers, 36*(1), 156-166

Lundberg, I. (1994). Reading difficulties can be predicted and prevented: A Scandinavian

perspective on phonological awareness and reading. In C. Hulme & M. Snoling (Eds.),

Reading development and dyslexia (pp. 180-199). London: Whurr.

McGill-Franzen, A. (1987). Failure to learn to read: Formulating a policy problem. *Reading

Research Quarterly*, 22(4), 475-490.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd Ed.). New York: McGraw-Hill.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Report of the subgroups.* Bethesda, MD: National Institute of Child Health and Human Development.

Perfetti, C.A., & Zhang, S. (1996). What it means to learn to read. In M. F. Graves, B. M. Taylor, & P. van den Broek (Eds.), *The first R: Every child's right to read* (pp. 37-61). New York: Teachers College Press.

Porpodas, C. (1993). The relation between phonemic awareness and reading and spelling of Greek words in the first school year. In M. Carretero, M. Pope, R. J. Simons, & J. I. Pozo (Eds.), *Learning and instruction* (vol. 3, pp. 203-217). Oxford: Pergamon Press.

Rathvon, N. (2004). *Early reading assessment: A practitioner's handbook*. New York: Guidford.

Saint-Laurent, L. & Giasson, J. (2010). *Épreuves informatisées des indicateurs des savoirs essentiels en lecture pour élèves francophones au début du primaire (INDISSE).* Manuscript under review, Faculté des sciences de l'éducation, Université Laval, Québec, Canada.

Salvia, J., Ysseldyke, J., & Bolt, S. (2007). *Assessment in special and inclusive education* (10th Ed). Boston: Houghton Mifflin Company.

Schargel, F. P., & Smink, J. (2001). *Strategies to help solve our school dropout problem.* Larchmont, NY: Eye on Education.

Shinn, M.R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children*. New

York: Guilford.

Shinn, M.R. (Ed.). (1998). *Advanced applications of Curriculum-Based Measurement*. New York:

Guilford.

Schneider, W., Kuspert, P., Roth, E., Vise, M., & Marx, H. (1997). Short- and long-term of

training phonological awareness in kindergarten: Evidence from two German studies.

Journal of Experimental Child Psychology 66, 311-340.

Simonart, G. (2006). *Échelle d'apprentissages scolaires primaires (ÉCHAS)*. Namur, Belgique :

Association francophone des agents psycho-médico-sociaux.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.) (1998). *Preventing reading difficulties*

in young children. Washington, DC: National Academy Press.

Statistique Canada[a] (2006). Étude: Alphabétisme et minorités de langue officielle. Ottawa :

Gouvernement du Canada. Retrieved October, 2009, from

http://www12.statcan.ca/census-recensement/index-fra.cfm

Statistics Canada[b] (2006). Community Profiles from the 2006 Census: Ottawa. Retrieved April,

2010

from http://www12.statcan.ca/census-recensement/2006/dp-pd/prof/92-

591/index.cfm?Lang=E

Tindal, G., & Marston, D. (1990). *Classroom-based assessment: Testing for teachers*. Columbus,

OH: Merrill.

Vacca, J. S. (2008). Crime can be prevented if schools teach juvenile offenders to read. *Children*

and Youth Services Review, 30, (9), 1055-1062.

Yopp, Hallie Kay. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, Vol. 23(2),* 159–177.

Yopp, H. K. (1992). Developing phonemic awareness in young children. *Reading Teacher*, *45* (9), 696-703.