

Examining the Predictive Validity of IDAPEL Measures with
ÉCOLE Criterion Assessments for
Determining IDAPEL Benchmark Goals and Cut Points for Risk for
French Language First (FL1) Students

Chantal Dufour-Martel

Roland H. Good

Elizabeth N. Dewey

Rachael Latimer

Dynamic Measurement Group
Technical Report # 13

Suggested Citation

Dufour-Martel, C., Good, R. H., Dewey, B., & Latimer, R. (2012). Examining the Predictive Validity of IDAPEL Measures with ÉCOLE Criterion Assessments for Determining IDAPEL Benchmark Goals and Cut Points for Risk for French Language First (FL1) Students. Technical Report No. #13. Eugene, OR: Dynamic Measurement Group.

Author Note

The authors thank the faculty, staff, students, and parents of participating schools for their effort and support during the course of this study. Correspondence regarding this manuscript should be addressed to Chantal Dufour-Martel, Dynamic Measurement Group, 132 E. Broadway, Suite 636 Eugene, OR 97401; Email: chantaldm@dibels.org.

Examining the Predictive Validity of IDAPEL[®] Measures with ÉCOLE Criterion Assessment for Determining IDAPEL[®] Benchmark Goals and Cut Points for Risk for French Language First (FL1) Students

The purpose of this technical report is to provide a compilation and description of the *Indicateurs dynamiques d'habiletés précoces en lecture* (IDAPEL[®]) (Dufour-Martel, 2003; Dufour-Martel & Good, 2009) benchmark goals and cut points for risk for the IDAPEL composite scores and IDAPEL individual measures. In this report, we provide information about the IDAPEL[®] and ÉCOLE criterion measures, along with descriptive and correlational data. DMG-database IDAPEL[®] participant French-language one (FL1) data from the 2007 – 2008 school year was used for determining the benchmark goals, the cut points for risk, and the IDAPEL Composite Scores. For further information regarding the development of the French-language curriculum-based measures, as well as additional technical reports, please visit: www.dibels.org.

Early screening for French reading difficulty within the French language school context is a challenge because few standardized, validated French language early literacy evaluation tools exist (Cormier, Desrochers, & Sénéchal, 2006). 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, delays in instruction at the appropriate level may occur and may put students at an increased risk for learning 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, seeing as very few exist.

In the English language first (EL1) research literature, a strong evidenced-based body of knowledge exists 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 reading connected text. IDAPEL[®] test materials are of equal difficulty and represent the general curriculum.

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

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, referred to as universal screening, 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 and allow teachers to determine if students are acquiring critical early literacy skills. 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.

An IDAPEL[®] validation study, undertaken with a substantial population of French language first (FL1) student participants in a Canadian province, provides evidence for IDAPEL's reliability and validity for use with this population of students. The battery of tests has also been validated with a smaller group of partial French immersion students. (See Tables 1 and 2).

Method

Participants

The participants in this study were students in kindergarten through second grade from 19 general education classrooms in 5 elementary schools from 3 separate school districts in a French-speaking Canadian province. The data were collected during the 2008-2009 school year. The students were selected from general education classrooms and were receiving French

language reading instruction, including students with disabilities, and students for whom French is a second language, provided they had the response capabilities to participate.

To examine the reliability and validity of the IDAPEL measures, further testing was performed with subsets of the total sample. For validity testing, one subset of 47 students from three participating school sites was administered ÉCOLE selected subtests for grades K through 2. For reliability testing, two-week alternate-form reliability testing was administered at two of the five sites.

The schools involved in the Benchmark Goals Study represent primarily a substantially large French-speaking northern region of the Canadian province of Ontario. 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. Data for this study were captured, primarily, through an extant database (i.e., DIBELS_Net, <https://dibels.net/>). French language student population information and demographic data at the school level were gathered from the Ontario Ministry of Education School Board Profiles website (http://esip.edu.gov.on.ca/english/profiles/quick_info_Comp) for 2007 – 2008 school year. 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	12,000

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

7 secondary

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.

Demographics	Province	Site A	Site B		Site C	
		School	School	School	School	School
		1	2	3	4	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.

Procedures

French teaching staff, school district special services personnel as well as WBTT (Web-Based Teaching Tool) coordinators received an initial 2-day IDAPEL Essential training workshop in the administration and scoring of all six IDAPEL measures in the early fall of the school year. Four-hour webcast refresher training occurred before the winter benchmark collection data point. A review of the materials and administration and scoring rules was provided during this time along with practice with simulated activities. Examiners were provided with opportunities for practice and to ask questions during the training. All data for this study was collected exclusively

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

by trained school personnel from participating school districts and by WBTT coordinators hired by the Learning Disabilities Association of Ontario (LDAO) who spearheaded the research.

Measures

Two sets of measures were used in this study, IDAPEL and ÉCOLE. The first set, IDAPEL[®] (Dufour-Martel, 2003; Dufour-Martel & Good, 2009), is a research-based experimental set of formative French language assessments designed to assess the basic early literacy skills. 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. All IDAPEL measures are individually administered and follow standardized administration procedures and scoring rules. The IDAPEL measure is described first. The second set, ÉCOLE (Échelles de compétences en lecture Desrochers, 2010), is a research-based, normed set of literacy performance indicators. The ÉCOLE measures are non-standardized and individually-administered. For information on the test construction of individual IDAPEL measures, see IDAPEL Technical Report # 12.

Measure Description

Facilité à dénommer des lettres (FDL) is a timed, one minute measure designed to be an indicator of letter naming knowledge. The measures assess a student's fluency in recognizing and articulating letter names. 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

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

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 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. 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) 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

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

presents a 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, /r/ /e/ /g/ /a/ /r/ to receive 5 possible points for the word (one point for each phoneme). 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 /re/ /g/ /ar/ or as /re/ /egar/ indicating a developing sense of phonemic awareness. Alternatively, a student may segment a word at the syllable level and say /re/ /gar/. 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) 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. 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. For example, at the Kindergarten and beginning grade 1 level (form FNM I), sample words are created using VCV, CVC, and CVCV letter-sound patterns (e.g., upa, rur, bémé). The middle of

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

first grade to beginning of second grade form (FNM II) include common letter patterns such as CVCV, CVCVV, CVCVC, CVVCV, CVCVVC, CVCCVC, or CVCVCV (e.g., jonjin, tetou, doivi). 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. The FNM I and FNM II each has over 20 alternate forms for monitoring progress student’s skill on the alphabetic principle. 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.

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

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 reading connected text, 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 focus of the task is for the student 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 passage of text 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. The oral reading fluency component can be used winter of first grade through spring of fifth grade.

For the benchmark assessment, students are asked to read three different grade-level passages for 1 minute each. The scores are the median number of words read correctly and the median number of errors across the three passages. Using the median score from three passages gives the best indicator of student performance over a range of different text and content. The student's accuracy rate is calculated based on the number of words read correctly and the number of errors, using the following formula:

$$Accuracy = 100 \times \frac{\text{median words correct}}{\text{median words correct} + \text{median errors}}$$

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

The passage retell component follows the reading of each passage, provided that the student has read between 20 and 40 words correct per minute on that passage. Passage retell is intended to provide a comprehension check for the FLO assessment, and provides an indication that the student is reading for meaning. 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). There is concern that students who display similar reading behavior will not be identified without a comprehension check. ROR provides an efficient procedure to identify students who are not able to talk about what they have read. Inclusion of ROR explicitly instructs students to be reading fluently for meaning. With a prompted passage retell, it is clear to the student that the intent is to read 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 relevance of the retell to the story 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 3 outlines the core components of early literacy with corresponding IDAPEL[®] measure name.

Table 3

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)

IDAPEL Composite Score (ICS). The IDAPEL Composite Score is a combination of multiple individual IDAPEL-measure raw scores. For example, the Kindergarten middle-of-year composite score is made up of a student’s raw scores on FPS, FDL, FSP, and FNM. In some instances, the score is weighted to stabilize the variability across measures. As an aggregate, the composite score captures the skills a student at a given grade level needs for literacy acquisition. It thus provides a larger and more complete sample of reading behavior than any single measure in isolation. Because the IDAPEL composite score correlates more highly with external outcomes, it was used as a primary internal (i.e., within the IDAPEL assessment system) criterion. For information on how to calculate the grade-level IDAPEL Composite Scores by hand, see IDAPEL Phase 2 training manual.

The scores used to calculate the IDAPEL Composite Score vary by grade and time of year therefore, it is important to note that the composite score generally cannot be used to directly measure growth over time or to compare results across grades or times of year.

However, because the logic and procedures used to establish benchmark goals are consistent across grades and times of year, the percent of students at or above benchmark can be compared.

Criterion measure. ÉCOLE 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, verifying measure data entry, and for merging IDAPEL[®] benchmark data with criterion data in order to create a master data file. The criterion sampling goal was 50 students per grade for all grades, Kindergarten to the 2nd grade. As a result of attrition, the ÉCOLE sample is smaller than the initial IDAPEL sample.

The ÉCOLE Composite Score. The ÉCOLE Composite score is also a combination of multiple individual ÉCOLE subtest raw scores. Similarly, subtests were grouped by predictor skill. In Kindergarten, CPPA encompasses the following raw scores subtests: Phonemic Elision (PHE), Grapheme Identification (GRS), and First Phoneme Identification (IIP). For first grade, CPPA includes several subtests: Grapheme Identification (GRS), Oral Word Reading (W-Read), Oral Reading of Non-Words (NW-Read), Word-Picture Match (CWP), Sentence-Picture Match (CSP), and Text Read 'Marie' (TXT Read). ÉCOLE correlations falling within the .5 to .6 ranges were used as criterion predictor measures. For a description of ÉCOLE subtests, see IDAPEL Technical Report # 12.

Description of the Data

Sample Size. A total of 510 students participated in the IDAPEL FL1 (French Language First) Benchmark Goal study. There were 238 participants in kindergarten, 206 in first grade, and 143 in second grade. Scores were recorded during the 2008-2009 and 2009-2010 school years,

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

consecutively, and data is available for students with complete across year data for both years. There were 39 students with recorded complete data in both kindergarten and first grade, and 26 students with recorded data in both first grade and second grade. There were a total of 156 students with recorded data for the ÉCOLE measures; 80 students in kindergarten, 51 students in first grade, and 25 students in second grade.

Invalid Scores. A score is deemed 'invalid' if it meets certain criteria on an individual-measure basis. For all measures, an invalid score is primarily defined as a score that is higher than the maximum number of points possible for that particular measure. An invalid score is secondarily defined as any score that is not mathematically possible within the measure. For FNM, FLO, and ACC (accuracy), there are pairs of scores that can be invalid. An explanation of the rules and procedures for handling invalid scores follows.

FNM/NMR. Each nonsense word is worth between three and four points (total number of 'letter sounds'). If a student reads the word as a whole word, then the student receives the maximum number of points for that word. Regardless of the total number of nonsense words attempted, there is a minimum-score and a maximum-score possible for each combination of correct letter sounds (FNM) score and whole words read (NMR). The minimum is found by calculating $(NMR \times 3)$. The maximum is found by calculating $(NMR \times 4)$. Normally, invalid pairs of FNM and NMR scores would result in the NMR score set to missing, leaving the FNM score intact. But in this dataset, there were only a few of these pairs of scores and it was impossible to tell which score was the possible error implying a mathematical error in the booklet. Therefore, both scores in each invalid pair were set to missing.

FLO/ACC. In this dataset, the number of errors a student made on FLO was not available. Thus, for ACC, we evaluated the invalid-score candidate based on the possibility that it was mis-

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

entered. The rules follow: if FLO=0, and ACC=missing, then ACC was set to 0; if FLO>0, and ACC=missing, then ACC was left missing; if FLO=missing and ACC was anything but missing, then ACC was set to missing.

ROR. If a student's FLO=0, then the student's ROR score was set to missing.

FPS. There were 6 invalid scores in beginning of year, and 6 invalid scores in middle of year in kindergarten.

FNM/NMR. There was one pair of invalid middle of year scores in kindergarten, three pairs of invalid end of year scores in kindergarten, and there was one pair of invalid scores in beginning of year second grade. These scores were set to missing.

Outliers. There were 46 outliers in the FL1 dataset. There were 20 outliers in kindergarten (FPS BOY = 1, FDL BOY = 1, FPS MOY = 3, FNM MOY = 1, NMR MOY = 8, FNM EOY = 2, NMR EOY = 4); 21 in first grade (FDL BOY = 1, FNM BOY = 4, NMR BOY = 3, FNM MOY = 2, NMR MOY = 4, FLO MOY = 2, ROR MOY = 1, FLO EOY = 3, ROR EOY = 1); and five in second grade (FLO BOY = 2, FLO MOY = 1, FLO EOY = 1, ROR EOY = 1; BOY = beginning of year; MOY = middle of year; EOY = end of year). Because none of these scores were invalid, they were allowed to remain in the dataset.

Preparation for Data Analysis

Raw scores were used in the analysis. The scores were not normally distributed. Instead, scores from NMR had large spikes at zero: 75% of kindergarten students recorded scores of zero on NMR at middle of year, and 50% at end of year; 25% of first-grade students recorded scores of zero at beginning and middle year, and 10% at end of year. The scores from FLO accuracy were negatively skewed in every grade in which they were administered. For all other measures in kindergarten through second, scores were positively skewed. The second grade scores were

less positively skewed than in kindergarten and first-grade. The scores from the ÉCOLE measures were normally distributed; all non-parametric tests about the mean (Shapiro-Wilks), which are appropriate for small sample sizes, returned non-significant p -values.

Scores for IDAPEL Composite Score were assumed to be normal. Normality tests (Kornolgorov-Smirnov and Anderson-Darling) returned non-significant p -values for kindergarten in middle of year, first grade in middle of year and end of year, and second grade in end of year. Tests from the other times of year (kindergarten beginning and end of year, first grade beginning of year, and second year beginning and middle of year) returned some evidence that the distributions were not normally distributed (p -values ranged from .01 to .04) due to extra-variability in the data.

Cohorts. Students were categorized into two separate groups (cohorts) to evaluate the IDAPEL measures. The first cohort contains students that recorded at least one score in both kindergarten and first grade; this is cohort k-1. The second cohort contains students that recorded at least one score in both first and second grade; this is cohort 1-2. Each cohort contains scores from all IDAPEL measures and ÉCOLE measures. The purpose was to evaluate growth from the first year of recorded data to the second year for students that were present in both years of the FL1 benchmark goal study.

Research Questions

This technical report is focused on answering three specific questions:

1. What are the within- and across-year levels of adequate growth on student reading performance as measured by IDAPEL?

2. What is the predictive and concurrent validity of IDAPEL[®] assessments with regard to Échelles de compétences en lecture (ÉCOLE), a criterion measure of reading proficiency that includes comprehension?
3. What levels of performance on IDAPEL[®] measures (i.e., benchmark goals and cut-points for risk) would allow for an accurate prediction of a student's likelihood of reaching subsequent benchmark goals?

Results

This section presents the results of various analyses on the IDAPEL measures in relation to each other, the IDAPEL Composite Score, and the external criterion, ÉCOLE. Descriptive statistics and correlations by cohorts for individual IDAPEL measures and the IDAPEL Composite Score are presented. For the reader, a strong correlation falls at .70 and above, a moderate-strong correlation range falls between .50 and .69, a moderate correlation falls between .30 and .49. Next, within- and across-year levels of growth on student reading performance as measured by IDAPEL are presented, followed by criterion-related validity of IDAPEL. Following IDAPEL validity, information about our rationale, levels of support, and design specification in regards to developing the IDAPEL benchmark goals and cut-points for risk for each measure and for the IDAPEL Composite Score (ICS) is presented. This is followed by a table summarizing IDAPEL Composite Scores descriptive statistics, then a table presenting percentage of students reaching the benchmark goal on the IDAPEL Composite Score and for each measure. A discussion section following the results answers the research questions posed above.

Descriptive Statistics. Table 4 reports descriptive statistics for students from the two cohorts in the FL1 study. There is growth for all measures across all times of year and grade levels. The following Tables: 5, 6, 7, 8, 9, and 10 report correlations between IDAPEL measures across grade level.

From Table 5, kindergarten ICS maintains moderate-strong to strong correlational relationships with first-grade ICS from beginning of year in kindergarten to end of year in first grade ($r = .56$ to $.84$). Individual kindergarten IDAPEL measures FPS, FDL, FSP, and FNM maintain moderate to strong correlational relationships with beginning-of-year first-grade outcomes ($r = .48$ to $.85$) and middle-of-year first-grade outcomes ($r = .33$ to $.79$), and small to strong correlational relationships with end-of-year first-grade outcomes ($r = .22$ to $.70$).

From Table 6, first grade ICS maintains moderate-strong to strong correlational relationships with second grade ICS from beginning of year in first grade to end of year in second-grade ($r = .68$ to $.89$). Individual first-grade IDAPEL measures FDL, FNM, and FLO maintain moderate-strong to strong correlational relationships with beginning-of-year second-grade outcomes ($r = .57$ to $.89$), middle-of-year second-grade outcomes ($r = .66$ to $.89$), and end-of-year second-grade outcomes ($r = .63$ to $.85$). For FSP, the correlational relationships with ICS, FNM, NMR, and FLO in first grade is trivial to small ($r = .16$ to $.41$) and not statistically significant.

From Tables 7, 8, 9, and 10, the within-grade correlational relationships for IDAPEL measures FDL, FPS, FNM, and FLO are strong within each time of year and moderate-strong to strong across times of year. Within-grade correlational relationships for NMR and ROR in second grade are moderate to strong. For FSP, correlational relationship within kindergarten is moderate-strong to strong across all times of year (Table k1_kcorr) and moderate to strong in

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

first grade beginning and middle of year (Table k1_1corr). The correlational relationship within first grade at middle and end of year is small to moderate, and many correlations are not significant at the alpha = .05 level, which implies that FSP does not predict later reading outcomes past first grade beginning of year.

Table 4

Descriptive Statistics by Cohort for IDAPEL and ÉCOLE for the School Years 2008-2009 and 2009-2010.

Measure by Time of Year	Descriptive Statistics by Cohort					
	Cohort K-1			Cohort 1-2		
	N	Mean	SD	N	Mean	SD
<i>Kindergarten</i>						
Beginning of Year						
FPS	38	14.76	9.79	--	--	--
FDL	38	16.61	12.82	--	--	--
ICS	38	31.37	19.41	--	--	--
Middle of Year						
FPS	38	22.63	12.83	--	--	--
FDL	38	21.97	14.06	--	--	--
FSP	38	13.92	11.62	--	--	--
FNM	38	17.24	9.59	--	--	--
NMR	37	0.24	0.76	--	--	--
ICS	38	75.76	41.67	--	--	--
End of Year						
FDL	34	29.26	15.07	--	--	--

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

FSP	34	22.91	15.10	--	--	--
FNM	34	24.68	14.28	--	--	--
NMR	33	0.18	0.77	--	--	--
ICS	34	73.26	35.64	--	--	--

ÉCOLE Administered at End of Year

cppa0e	37	56.59	21.36	--	--	--
total0e	37	80.46	35.59	--	--	--

First Grade

Beginning of Year

FDL	39	41.77	17.99	26	40.23	17.37
FSP	39	37.44	17.85	26	25.19	14.35
FNM	39	35.10	25.56	26	35.46	21.94
NMR	28	4.64	4.51	21	1.95	3.60
ICS	39	102.54	46.20	26	96.58	40.67

Middle of Year

FSP	38	35.21	14.08	26	35.08	10.94
FNM	38	56.21	33.49	26	54.50	27.95
NMR	37	10.05	8.57	25	4.52	6.55
FLO	38	39.08	26.14	26	43.81	22.80
ACC	38	83.21	24.29	26	91.58	8.44
ROR	37	3.97	4.43	26	17.15	8.31
ICS	37	176.84	93.44	25	189.20	64.43

End of Year

FSP	37	44.46	16.73	26	39.58	11.89
-----	----	-------	-------	----	-------	-------

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

FNM	37	82.68	36.30	26	93.31	38.20
NMR	36	16.72	10.14	21	7.67	9.76
FLO	37	61.05	32.55	26	67.12	27.31
ACC	37	87.92	22.52	26	96.50	4.19
Measure by Time of Year	N	M	SD	N	M	SD
ROR	36	4.83	6.98	26	20.85	7.27
ICS	36	240.22	107.76	21	256.95	65.98
ÉCOLE Administered at End of Year						
cppa0e	--	--	--	25	190.16	39.39
total0e	--	--	--	25	239.76	51.69

Second Grade

Beginning of Year

FNM	--	--	--	26	98.19	34.72
NMR	--	--	--	26	20.69	9.78
FLO	--	--	--	26	70.96	27.51
ACC	--	--	--	25	95.08	8.30
ROR	--	--	--	26	22.23	8.77
ICS	--	--	--	25	226.04	70.33

Middle of Year

FLO	--	--	--	26	77.31	25.27
ACC	--	--	--	26	96.65	8.15
ROR	--	--	--	26	26.62	9.04
ICS	--	--	--	26	199.15	48.44

End of Year

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

FLO	--	--	--	26	87.88	28.77
ACC	--	--	--	26	95.88	5.10
ROR	--	--	--	26	26.31	10.08
ICS	--	--	--	26	225.19	66.74

Note. Cohorts include students with recorded scores for at least one measure in both grade levels. ICS = IDAPEL Composite Score; 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; FLO = Facilité en lecture orale; ROR = rapport oral du récit. ÉCOLE was administered at the end of school Year 2008-2009. cppa0e = Conscience phonologique/ Principe alphabétique kindergarten end of year; total0e = total ÉCOLE score kindergarten end of year.

Table 5

Correlations between Kindergarten and First-Grade IDAPEL Measures for Cohort K-1 for the School Years 2008-2009 and 2009-2010.

Kindergarten Measure by Time of Year for School Year 2008 - 2009	First-Grade IDAPEL Measure by Time of Year for School Year 2009-2010							
	Beginning of Year							
	ICS	FDL	FSP	FNM	NMR	FLO	ACC	ROR
Beginning of Year								
ICS	.80 (38)	.79 (38)	.72 (38)	.68 (38)	.62 (27)	--	--	--
FPS	.70 (38)	.70 (38)	.70 (38)	.60 (38)	.45* (27)	--	--	--
FDL	.68 (38)	.66 (38)	.55 (38)	.57 (38)	.53** (27)	--	--	--
Middle of Year								
ICS	.80 (38)	.73 (38)	.81 (38)	.71 (38)	.62 (27)	--	--	--
FPS	.68 (38)	.63 (38)	.74 (38)	.61 (38)	.37† (27)	--	--	--

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

FDL	.73 (38)	.67 (38)	.73 (38)	.63 (38)	.59** (27)	--	--	--
FSP	.61 (38)	.60 (38)	.64 (38)	.54 (38)	.62 (27)	--	--	--
FNM	.75 (38)	.63 (38)	.67 (38)	.70 (38)	.59** (27)	--	--	--
NMR	.40* (37)	.36* (37)	.43** (37)	.42** (37)	.61 (26)	--	--	--

End of Year

ICS	.84 (34)	.85 (34)	.85 (34)	.68 (34)	.56** (25)	--	--	--
FDL	.83 (34)	.85 (34)	.73 (34)	.69 (34)	.67 (25)	--	--	--
FSP	.60 (34)	.59 (34)	.77 (34)	.48** (34)	.23† (25)	--	--	--
FNM	.79 (34)	.80 (34)	.79 (34)	.66 (34)	.44* (25)	--	--	--
NMR	.14† (33)	.13† (33)	.26† (33)	.12† (33)	.31† (24)	--	--	--

Middle of Year

ICS	FDL	FSP	FNM	NMR	FLO	ACC	ROR
-----	-----	-----	-----	-----	-----	-----	-----

Beginning of Year

ICS	.59 (36)	--	.43** (37)	.47** (37)	.24† (36)	.69 (37)	.53 (37)	.42** (36)
FPS	.53 (36)	--	.50** (37)	.39* (37)	.14† (36)	.60 (37)	.51** (37)	.41* (36)
FDL	.50** (36)	--	.28† (37)	.42* (37)	.26† (36)	.60 (37)	.41* (37)	.33* (36)

Middle of Year

ICS	.73	--	.50**	.61	.47**	.75	.65	.39*
-----	-----	----	-------	-----	-------	-----	-----	------

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

	(37)		(38)	(38)	(37)	(38)	(38)	(37)
FPS	.69 (37)	--	.49** (38)	.58 (38)	.47** (37)	.71 (38)	.65 (38)	.41* (37)
FDL	.66 (37)	--	.43** (38)	.61 (38)	.47** (37)	.66 (38)	.56 (38)	.29† (37)
FSP	.48** (37)	--	.46** (38)	.38* (38)	.25† (37)	.55 (38)	.42** (38)	.21† (37)
FNM	.69 (37)	--	.33* (38)	.54 (38)	.43** (37)	.68 (38)	.64 (38)	.44** (37)
NMR	.30† (36)	--	.18† (37)	.27† (37)	.22† (36)	.38* (37)	.16† (37)	.00† (36)

End of Year

ICS	.79 (33)	--	.61 (33)	.72 (33)	.47** (33)	.81 (33)	.69 (33)	.36* (32)
FDL	.77 (33)	--	.55 (33)	.76 (33)	.48** (33)	.79 (33)	.62 (33)	.34† (32)
FSP	.54** (33)	--	.62 (33)	.45** (33)	.25† (33)	.56 (33)	.55 (33)	.42* (32)
FNM	.74 (33)	--	.51** (33)	.69 (33)	.42* (33)	.79 (33)	.61 (33)	.29† (32)
NMR	.12† (32)	--	.30† (32)	.10† (32)	.06† (32)	.12† (32)	.12† (32)	.00† (31)

End of Year

ICS	FDL	FSP	FNM	NMR	FLO	ACC	ROR
-----	-----	-----	-----	-----	-----	-----	-----

Beginning of Year

ICS	.56 (35)	--	.39* (36)	.47** (36)	.49** (35)	.63 (36)	.46** (36)	.35* (35)
FPS	.54 (35)	--	.48** (36)	.49** (36)	.47** (35)	.58 (36)	.39* (36)	.34* (35)
FDL	.43** (35)	--	.22† (36)	.34* (36)	.39* (35)	.51** (36)	.40* (36)	.28† (35)

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Middle of Year

ICS	.69 (36)	--	.49** (37)	.63 (37)	.63 (36)	.72 (37)	.58 (37)	.33† (36)
FPS	.67 (36)	--	.47** (37)	.64 (37)	.65 (36)	.69 (37)	.53 (37)	.35* (36)
FDL	.60 (36)	--	.38* (37)	.54 (37)	.55 (36)	.63 (37)	.55 (37)	.25† (36)
FSP	.50** (36)	--	.52** (37)	.50** (37)	.42* (36)	.55 (37)	.36* (37)	.14† (36)
FNM	.61 (36)	--	.29† (37)	.50** (37)	.57 (36)	.63 (37)	.55 (37)	.41* (36)
NMR	.33† (35)	--	.30† (36)	.38* (36)	.29† (35)	.38* (36)	.13† (36)	-.08† (35)

End of Year

ICS	.75 (31)	--	.57 (32)	.73 (32)	.74 (31)	.77 (32)	.58 (32)	.30† (31)
FDL	.74 (31)	--	.48** (32)	.71 (32)	.72 (31)	.78 (32)	.57 (32)	.24† (31)
FSP	.53** (31)	--	.47** (32)	.52** (32)	.53** (31)	.52** (32)	.41* (32)	.50** (31)
FNM	.70 (31)	--	.55** (32)	.68 (32)	.70 (31)	.73 (32)	.49** (32)	.21† (31)
NMR	.17† (30)	--	.24† (31)	.15† (31)	.14† (30)	.17† (31)	.10† (31)	-.03† (30)

Note. Pair-wise sample sizes reported in parenthesis. ICS = IDAPEL Composite Score; FPS = Facilité à reconnaître la premier son; FDL = Facilité à dénommer des lettres; FSP = Facilité à segmenter les phonèmes; FNM = Facilité à lire des non-mots; FLO = Facilité en lecture orale; ACC = FLO Accuracy; ROR = Rapport oral du récit. Unless marked, correlations are significant, $p < .001$; '**' $p < .01$, '*' $p < .05$, '†' Not Significant.

Table 6

Correlations between Kindergarten and First-Grade Measures for Cohort 1-2 for the School

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Years 2008-2009 and 2009-2010.

First Grade Measure by Time of Year for School Year 2008 - 2009	Second-Grade IDAPEL Measure by Time of Year for School Year 2009-2010					
	Beginning of Year					
	ICS	FNM	NMR	FLO	ACC	ROR
Beginning of Year						
ICS	.73 (25)	.68 (26)	.51** (26)	.76 (26)	.52** (25)	.55** (26)
FDL	.80 (25)	.73 (26)	.58** (26)	.81 (26)	.62 (25)	.58** (26)
FSP	.24† (25)	.19† (26)	.20† (26)	.24† (26)	.16† (25)	.37† (26)
FNM	.63 (25)	.60** (26)	.40* (26)	.67 (26)	.41* (25)	.42* (26)
NMR	.44* (21)	.43† (21)	.40† (21)	.51* (21)	.28† (21)	.20† (21)
Middle of Year						
ICS	.87 (24)	.80 (25)	.57** (25)	.88 (25)	.71 (24)	.60** (25)
FSP	.27† (25)	.30† (26)	.33† (26)	.30† (26)	.06† (25)	.09† (26)
FNM	.79 (25)	.75 (26)	.51** (26)	.84 (26)	.57** (25)	.54** (26)
NMR	.31† (24)	.32† (25)	-.04† (25)	.49* (25)	.25† (24)	.08† (25)
FLO	.76 (25)	.68 (26)	.46* (26)	.82 (26)	.55** (25)	.59** (26)
ACC	.83 (25)	.65 (26)	.54** (26)	.77 (26)	.82 (25)	.71 (26)
ROR	.58** (25)	.48* (26)	.11† (26)	.64 (26)	.43* (25)	.56** (26)
End of Year						
ICS	.89 (21)	.81 (21)	.84 (21)	.79 (21)	.78 (21)	.76 (21)
FSP	.32† (25)	.30† (26)	.30† (26)	.30† (26)	.24† (25)	.20† (26)
FNM	.67 (25)	.68 (26)	.54** (26)	.57** (26)	.52** (25)	.51** (26)
NMR	.14† (21)	.05† (21)	.08† (21)	.03† (21)	.26† (21)	.25† (21)
FLO	.84 (25)	.77 (26)	.44* (26)	.89 (26)	.57** (25)	.61 (26)

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

ACC	.88 (25)	.72 (26)	.60** (26)	.77 (26)	.90 (25)	.73 (26)
ROR	.57** (25)	.45* (26)	.15† (26)	.59** (26)	.49* (25)	.61 (26)
Middle of Year						
	ICS	FNM	NMR	FLO	ACC	ROR
Beginning of Year						
ICS	.68 (26)	--	--	.74 (26)	.37† (26)	.65 (26)
FDL	.75 (26)	--	--	.78 (26)	.46* (26)	.68 (26)
FSP	.28† (26)	--	--	.30† (26)	.12† (26)	.42* (26)
FNM	.57** (26)	--	--	.64 (26)	.27† (26)	.55** (26)
NMR	.42† (21)	--	--	.51* (21)	.20† (21)	.28† (21)
Middle of Year						
ICS	.84 (25)	--	--	.89 (25)	.54** (25)	.60** (25)
FSP	.16† (26)	--	--	.18† (26)	.09† (26)	.15† (26)
FNM	.76 (26)	--	--	.82 (26)	.46* (26)	.54** (26)
NMR	.38† (25)	--	--	.47* (25)	.17† (25)	.31† (25)
FLO	.75 (26)	--	--	.82 (26)	.41* (26)	.59** (26)
ACC	.85 (26)	--	--	.83 (26)	.63 (26)	.63 (26)
ROR	.63 (26)	--	--	.67 (26)	.33† (26)	.63 (26)
End of Year						
ICS	.85 (21)	--	--	.87 (21)	.63** (21)	.67 (21)
FSP	.25† (26)	--	--	.26† (26)	.13† (26)	.21† (26)
FNM	.67 (26)	--	--	.66 (26)	.45* (26)	.63 (26)
NMR	.16† (21)	--	--	.12† (21)	.15† (21)	.19† (21)
FLO	.82 (26)	--	--	.90 (26)	.43* (26)	.73 (26)

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

ACC	.91 (26)	--	--	.86 (26)	.78 (26)	.65 (26)
ROR	.62 (26)	--	--	.63 (26)	.34† (26)	.70 (26)
End of Year						
	ICS	FNM	NMR	FLO	ACC	ROR
Beginning of Year						
ICS	.77 (26)	--	--	.76 (26)	.58** (26)	.69 (26)
FDL	.82 (26)	--	--	.78 (26)	.69 (26)	.70 (26)
FSP	.41* (26)	--	--	.29† (26)	.31† (26)	.52** (26)
FNM	.63 (26)	--	--	.67 (26)	.42* (26)	.56** (26)
NMR	.48* (21)	--	--	.62** (21)	.32† (21)	.25† (21)
Middle of Year						
ICS	.78 (25)	--	--	.85 (25)	.67 (25)	.49* (25)
FSP	.25† (26)	--	--	.23† (26)	.14† (26)	.31† (26)
FNM	.69 (26)	--	--	.80 (26)	.56** (26)	.43* (26)
NMR	.31† (25)	--	--	.41* (25)	.23† (25)	.13† (25)
FLO	.73 (26)	--	--	.79 (26)	.56** (26)	.54** (26)
ACC	.78 (26)	--	--	.77 (26)	.75 (26)	.53** (26)
ROR	.60** (26)	--	--	.61 (26)	.43* (26)	.56** (26)
End of Year						
ICS	.88 (21)	--	--	.86 (21)	.80 (21)	.66** (21)
FSP	.33† (26)	--	--	.29† (26)	.26† (26)	.31† (26)
FNM	.75 (26)	--	--	.63 (26)	.66 (26)	.71 (26)
NMR	.11† (21)	--	--	.08† (21)	.16† (21)	.06† (21)
FLO	.84 (26)	--	--	.88 (26)	.61 (26)	.73 (26)

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

ACC	.80 (26)	--	--	.76 (26)	.87 (26)	.51** (26)
ROR	.60** (26)	--	--	.53** (26)	.47* (26)	.61 (26)

Note. Pair-wise sample sizes reported in parenthesis. ICS = IDAPEL Composite Score; FPS = Facilité à reconnaître la premier son; FDL = Facilité à dénommer des lettres; FSP = Facilité à segmenter les phonèmes; FNM = Facilité à lire des non-mots; FLO = Facilité en lecture orale; ACC = FLO Accuracy; ROR = Rapport oral du récit. Unless marked, correlations are significant, $p < .001$; '**' $p < .01$, '*' $p < .05$, '†' Not Significant.

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Table 7

Within Grade Correlations for IDAPEL Kindergarten Measures from Cohort K-1

Measure by Time of Year	2	3	4	5	6	7	8	9	10	11	12	13	14
Beginning of Year													
1. ICS	.81 (38)	.89 (38)	.83 (37)	.66 (37)	.80 (37)	.62 (37)	.81 (37)	.29† (37)	.87 (33)	.86 (33)	.65 (33)	.82 (33)	.18† (33)
2. FPS	--	.46 (38)	.75 (37)	.72 (37)	.53 (37)	.67 (37)	.69 (37)	.36* (37)	.78 (33)	.68 (33)	.78 (33)	.72 (33)	.23† (33)
3. FDL		--	.70 (37)	.45 (37)	.81** (37)	.43 (37)	.71 (37)	.17† (37)	.73 (33)	.79 (33)	.38* (33)	.69 (33)	.09† (33)
Middle of Year													
4. ICS			--	.90 (38)	.91 (38)	.81 (38)	.82 (38)	.42** (37)	.89 (33)	.82 (33)	.78 (33)	.83* (33)	.35 (32)
5. FPS				--	.70 (38)	.71 (38)	.70 (38)	.35* (37)	.83 (33)	.68 (33)	.81 (33)	.80 (33)	.25† (32)
6. FDL					--	.64 (38)	.76 (38)	.30† (37)	.80 (33)	.82 (33)	.61 (33)	.72 (33)	.29† (32)
7. FSP						--	.44 (38)	.68 (37)	.73 (33)	.64 (33)	.75 (33)	.69 (33)	.65 (32)
8. FNM							--	.09† (37)	.81 (33)	.76 (33)	.63 (33)	.76 (33)	.06† (32)

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

9. NMR	--	.34 (32)	.30* (32)	.38* (32)	.35* (32)	.44† (32)
End of Year						
10. ICS		--	.92 (34)	.81 (34)	.95 (34)	.23† (33)
11. FDL			--	.63 (34)	.82† (34)	.16† (33)
12. FSP				--	.73* (34)	.40 (33)
13. FNM					--	.23† (33)
14. NMR						--

Note. Pair-wise sample sizes reported in parenthesis. ICS = IDAPEL Composite Score; FPS = Facilité à reconnaître la premier son; FDL = Facilité à dénommer des lettres; FSP = Facilité à segmenter les phonèmes; FNM = Facilité à lire des non-mots; FLO = Facilité en lecture orale; ACC = FLO Accuracy; ROR = Rapport oral du récit. Unless marked, correlations are significant, $p < .001$; '**' $p < .01$, '*' $p < .05$, '†' Not Significant.

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Table 8

Within Grade Correlations for IDAPEL First Grade Measures from Cohort K-1

Measure by Time of Year	2	3	4	5	6	7	8	9	10
Beginning of Year									
1. ICS	.91 (39)	.79 (39)	.94 (39)	.83 (28)	.86 (37)	.32† (38)	.79 (38)	.57 (37)	.92 (38)
2. FDL	--	.70 (39)	.75 (39)	.61 (28)	.78 (37)	.39* (38)	.71 (38)	.44** (37)	.85 (38)
3. FSP		--	.68 (39)	.70 (28)	.67 (37)	.50** (38)	.60 (38)	.40* (37)	.64 (38)
4. FNM			--	.91 (28)	.80 (37)	.12† (38)	.76 (38)	.56 (37)	.88 (38)
5. NMR				--	.68 (26)	.26† (27)	.66 (27)	.55** (26)	.71 (27)
Middle of Year									
6. ICS					--	.40* (37)	.94 (37)	.83 (37)	.92 (37)
7. FSP						--	.32* (38)	.15† (37)	.31† (38)
8. FNM							--	.83 (37)	.86 (38)
9. NMR								--	.66 (37)
10. FLO									--
	11	12	13	14	15	16	17	18	19
Beginning of Year									

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

1. ICS	.68 (38)	.36* (37)	.81 (36)	.35* (37)	.74 (37)	.74 (36)	.88 (37)	.64 (37)	.24† (36)
2. FDL	.65 (38)	.30† (37)	.78 (36)	.44** (37)	.74 (37)	.70 (36)	.83 (37)	.61 (37)	.18† (36)
3. FSP	.66 (38)	.28† (37)	.61 (36)	.43** (37)	.56 (37)	.51** (36)	.62 (37)	.60 (37)	.18† (36)
4. FNM	.53 (38)	.32† (37)	.70 (36)	.18† (37)	.63 (37)	.65 (36)	.81 (37)	.48** (37)	.23† (36)
5. NMR	.53** (27)	.09† (26)	.59** (26)	.27† (26)	.56** (26)	.50** (26)	.66 (26)	.47* (26)	-.11† (25)
Middle of Year									
6. ICS	.83 (37)	.33* (36)	.95 (35)	.42* (36)	.85 (36)	.90 (35)	.95 (36)	.74 (36)	.20† (35)
7. FSP	.55 (38)	.23† (37)	.41* (36)	.68 (37)	.32† (37)	.28† (36)	.38* (37)	.55 (37)	.20† (36)
8. FNM	.66 (38)	.24† (37)	.87 (36)	.34* (37)	.78 (37)	.84 (36)	.88 (37)	.60 (37)	.12† (36)
9. NMR	.62 (37)	.24† (36)	.75 (35)	.25† (36)	.62 (36)	.74 (35)	.71 (36)	.55 (36)	.10† (35)
10. FLO	.68 (38)	.27† (37)	.86 (36)	.38* (37)	.80 (37)	.82 (36)	.94 (37)	.58 (37)	.18† (36)
11. ACC	--	.38* (37)	.83 (36)	.49** (37)	.73 (37)	.72 (36)	.75 (37)	.87 (37)	.26† (36)
12. ROR		--	.29† (35)	.00† (36)	.17† (36)	.18† (35)	.32† (36)	.31† (36)	.94 (36)
End of Year									
13. ICS			--	.40* (36)	.96 (36)	.94 (36)	.96 (36)	.83 (36)	.14† (35)
14. FSP				--	.37* (37)	.29† (36)	.38* (37)	.51** (37)	-.07† (36)
15. FNM					--	.94 (36)	.89 (37)	.72 (37)	.05† (36)

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

16. NMR	--	.89 (36)	.68 (36)	.07† (35)
17. FLO		--	.72 (37)	.19† (36)
18. ACC			--	.19† (36)
19. ROR				--

Note. Pair-wise sample sizes reported in parenthesis. ICS = IDAPEL Composite Score; FPS = Facilité à reconnaître la premier son; FDL = Facilité à dénommer des lettres; FSP = Facilité à segmenter les phonèmes; FNM = Facilité à lire des non-mots; FLO = Facilité en lecture orale; ACC = FLO Accuracy; ROR = Rapport oral du récit. Unless marked, correlations are significant, $p < .001$; '**' $p < .01$, '*' $p < .05$, '†' Not Significant.

Table 9

Within Grade Correlations for IDAPEL First Grade Measures from Cohort 1-2

Measure by Time of Year	2	3	4	5	6	7	8	9	10
Beginning of Year									
1. ICS	.89(26)	.56**(26)	.91(26)	.65**(21)	.69 (25)	.51**(26)	.62 (26)	.24†(25)	.69 (26)
2. FDL	--	.40*(26)	.68 (26)	.49*(21)	.73(25)	.39*(26)	.66 (26)	.27†(25)	.70(26)
3. FSP		--	.38†(26)	.17†(21)	.07†(25)	.53**(26)	-.01†(26)	-.15†(25)	.14†(26)
4. FNM			--	.68 (21)	.64 (25)	.40*(26)	.58**(26)	.29†(25)	.63 (26)
5. NMR				--	.65**(20)	.51*(21)	.71 (21)	.15†(20)	.67 (21)

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Middle of Year

6. ICS					--	.21†(25)	.95(25)	.55**(25)	.93(25)
7. FSP						--	.22†(26)	-.21†(25)	.23†(26)
8. FNM							--	.53**(25)	.84(26)
9. NMR								--	.39†(25)
10. FLO									--

	11	12	13	14	15	16	17	18	19
--	----	----	----	----	----	----	----	----	----

Beginning of Year

1. ICS	.70(26)	.51**(26)	.82(21)	.40*(26)	.77(26)	-.05†(21)	.75(26)	.65 (26)	.51**(26)
2. FDL	.79(26)	.54**(26)	.86(21)	.39†(26)	.76(26)	.16†(21)	.74(26)	.71(26)	.60**(26)
3. FSP	.22†(26)	.14†(26)	.40†(21)	.34†(26)	.50**(26)	-.07†(21)	.21†(26)	.27†(26)	.20†(26)
4. FNM	.59**(26)	.44*(26)	.64**(21)	.26†(26)	.64 (26)	-.14†(21)	.71(26)	.52**(26)	.41*(26)
5. NMR	.45*(21)	.44*(21)	.47†(17)	.57**(21)	.30†(21)	-.27†(17)	.51*(21)	.40†(21)	.03†(21)

Middle of Year

6. ICS	.86(25)	.74(25)	.78(20)	.38†(25)	.56**(25)	.08†(20)	.86(25)	.82(25)	.57**(25)
7. FSP	.11†(26)	.03†(26)	.23†(21)	.59**(26)	.31†(26)	-.57**(21)	.20†(26)	.15†(26)	-.12†(26)
8. FNM	.74(26)	.70(26)	.74 (21)	.36†(26)	.49*(26)	.02†(21)	.82(26)	.74(26)	.50**(26)

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

9. NMR	.34†(25)	.49*(25)	.19†(20)	-.07†(25)	.04†(25)	.39†(20)	.47*(25)	.37†(25)	.40*(25)
10. FLO	.76(26)	.75(26)	.69 (21)	.31†(26)	.55**(26)	.05†(21)	.84(26)	.69(26)	.57**(26)
11. ACC	--	.64 (26)	.86(21)	.42*(26)	.66 (26)	.19†(21)	.71(26)	.82(26)	.60**(26)
12. ROR		--	.65**(21)	.30†(26)	.48*(26)	.24†(21)	.79(26)	.56**(26)	.82(26)
End of Year									
13. ICS			--	.35†(21)	.90(21)	.32†(21)	.85(21)	.84(21)	.63**(21)
14. FSP				--	.42*(26)	-.42†(21)	.24†(26)	.29†(26)	.11†(26)
15. FNM					--	.15†(21)	.68 (26)	.64 (26)	.55**(26)
16. NMR						--	.14†(21)	.19†(21)	.47*(21)
17. FLO							--	.71(26)	.72(26)
18. ACC								--	.59**(26)
19. ROR									--

Note. Pair-wise sample sizes reported in parenthesis. ICS = IDAPEL Composite Score; FPS = Facilité à reconnaître la premier son; FDL = Facilité à dénommer des lettres; FSP = Facilité à segmenter les phonèmes; FNM = Facilité à lire des non-mots; FLO = Facilité en lecture orale; ACC = FLO Accuracy; ROR = Rapport oral du récit. Unless marked, correlations are significant, $p < .001$; '**' $p < .01$, '*' $p < .05$, '†' Not Significant.

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Table 10

Within Grade Correlations for IDAPEL Second Grade Measures from Cohort 1-2

Measure by Time of Year	2	3	4	5	6	7		
Beginning of Year								
1. ICS	.91 (25)	.92 (25)	.91 (25)	.85 (25)	.71 (25)	.93 (25)		
2. FNM	--	.78 (26)	.84 (26)	.62* (25)	.48 (26)	.81 (26)		
3. NMR		--	.60 (26)	.65* (25)	.48 (26)	.63 (26)		
4. FLO			--	.65 (25)	.69 (26)	.86 (26)		
5. ACC				--	.66 (25)	.85 (25)		
6. ROR					--	.69 (26)		
	8	9	10	11	12	13	14	
Beginning of Year								
1. ICS	.92 (25)	.70 (25)	.72 (25)	.89 (25)	.89 (25)	.85 (25)	.60** (25)	
2. FNM	.81 (26)	.57** (26)	.61 (26)	.79 (26)	.82 (26)	.72 (26)	.52** (26)	
3. NMR	.57 (26)	.53** (26)	.42* (26)	.60** (26)	.61 (26)	.60** (26)	.36† (26)	
4. FLO	.95 (26)	.47* (26)	.76 (26)	.90 (26)	.95 (26)	.70 (26)	.70 (26)	
5. ACC	.74 (25)	.83 (25)	.53** (25)	.72 (25)	.64 (25)	.90 (25)	.36† (25)	

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

6. ROR	.71 (26)	.43* (26)	.73 (26)	.71 (26)	.69 (26)	.59** (26)	.65 (26)
Middle of Year							
7. ICS	.93 (26)	.82 (26)	.77 (26)	.91 (26)	.87 (26)	.91 (26)	.65 (26)
8. FLO	--	.58** (26)	.75 (26)	.94 (26)	.96 (26)	.80 (26)	.70 (26)
9. ACC		--	.43 (26)	.58 (26)	.49 (26)	.86† (26)	.27** (26)
10. ROR			--	.82 (26)	.75 (26)	.61 (26)	.87 (26)
End of Year							
11. ICS				--	.95 (26)	.86 (26)	.83 (26)
12. FLO					--	.75 (26)	.72 (26)
13. ACC						--	.53 (26)
14. ROR							--

Note. Pair-wise sample sizes reported in parenthesis. ICS = IDAPEL Composite Score; FPS = Facilité à reconnaître la premier son; FDL = Facilité à dénommer des lettres; FSP = Facilité à segmenter les phonèmes; FNM = Facilité à lire des non-mots; FLO = Facilité en lecture orale; ACC = FLO Accuracy; ROR = Rapport oral du récit. Unless marked, correlations are significant, $p < .001$; '**' $p < .01$, '*' $p < .05$, '†' Not Significant.

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Criterion-Related Validity. To evaluate the predictive and concurrent validity of IDAPEL measures in kindergarten and first grade, ÉCOLE was administered at the end of the school year 2008-2009. The results are presented in Table 11.

FPS. The predictive validity for beginning and middle-of-year FPS: The correlation of FPS with ÉCOLE is moderate-strong at beginning of year assessment, and strong at middle of year assessment. Correlations range from .69 (beginning of year) to .73 (middle of year, CAPP). Concurrent validity is not available, because FPS is not administered at end-of-year assessment.

FDL. The predictive and concurrent validity of kindergarten FDL: The correlation of FDL with ÉCOLE CAPP K level at the three benchmarks, beginning, mid and end of year, are all strong and range from .72 (beginning of year), to .78 (middle of year), and .85 (end of year). First grade correlation is strong at .80.

FSP. The predictive and concurrent validity of kindergarten FSP: Correlations for kindergarten PSF with ÉCOLE CAPP for middle and end of year are moderate-strong, and range from .62 (beginning of year) to .69 (end of year). Correlations for first-grade FSP with ÉCOLE CAPP are small, and range from .26 (beginning of year) to .21 (middle of year), but moderate .43 at end of year.

FNM-NSC. The predictive and concurrent validity of kindergarten FNM-NSC: Correlations with ÉCOLE CAPP are strong, and range between .85 (beginning of year) and .81 (end of year). In first grade, FNM-NSC correlations with ÉCOLE CAPP are moderate-strong to strong, and range from .53 (beginning of year) to .80 (middle of year), to .65 (end of year). Correlational relationships are strong across kindergarten and first grades and time of year.

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

FNM-NMR. The predictive and concurrent validity of kindergarten FNM-NMR: Correlations with ÉCOLE CAPPa are small and range from .24 (middle of year) to .19 (end of year). First grade correlations with ÉCOLE CAPPa are moderate at the beginning and middle of the year, .47 (beginning of year) to .35 (middle of year), and small at end of year .11.

FLO Words Correct, Accuracy and ROR: The concurrent and predictive validity results of these three measures fall within similar standards for validity and within similar ranges across the different time points of the year. Correlations with ÉCOLE CAPPa middle of year are, respectively all strong: .76 (FLO), .83 (ACC), and .62 (ROR). Similarly, for end of year, correlation are all strong: .81 (FLO), .88 (ACC), and .59 (ROR).

IDAPEL Composite Score (ICS). At the Kindergarten level, ICS is calculated from different measure scores at beginning (FPS and FDL), middle of year (FPS, FDL, FSP, and FNM), and end of year (FDL, FSP, FNM). The concurrent and predictive validity of ICS with ÉCOLE CAPPa across year at the Kindergarten level are both strong: predictive validity at .82 (beginning of year), and .85 (middle of year) and concurrent validity at .88 (end of year). At the first grade level, ICS is again calculated from different measure scores at beginning (FDL, FPS and FNM), and respectively at middle of –m and end of year (FNM, NMR, and FLO). For first grade, predictive correlations with ÉCOLE CAPPa across year remain strong: .70 (beginning of year), .86 (middle of year), and concurrent validity at .86 (end of year).

Table 11

Predictive and Concurrent Validity of IDAPEL Measures with Selected ÉCOLE Scores

IDAPEL Measure by Grade and Time of Year	ÉCOLE Measure by Grade (end of year)			
	Kindergarten		First Grade	
	CPPA	Total	CPPA	Total

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Predictive Validity - Beginning of Year

ICS	.82 (36)	.86 (36)	.70 (25)	.78 (25)
FPS	.69 (36)	.74 (36)	--	--
FDL	.72 (36)	.75 (36)	.80 (25)	.81 (25)
FSP	--	--	.26† (25)	.32† (25)
FNM	--	--	.53** (25)	.66 (25)
NMR	--	--	.47* (21)	.52* (21)

Predictive Validity - Middle of Year

ICS	.85 (36)	.86 (36)	.86 (24)	.86 (24)
FPS	.73 (36)	.73 (36)	--	--
FDL	.78 (36)	.76 (36)	--	--
FSP	.62 (36)	.70 (36)	.21† (25)	.26† (25)
FNM	.85 (36)	.78 (36)	.80 (25)	.80 (25)
NMR	.24† (35)	.36* (35)	.35† (24)	.32† (24)
FLO	--	--	.76 (25)	.76 (25)
ACC	--	--	.83 (25)	.82 (25)
ROR	--	--	.62 (25)	.57** (25)

Concurrent Validity - End of Year

ICS	.88 (32)	.88 (32)	.86 (20)	.87 (20)
FDL	.85 (32)	.86 (32)	--	--
FSP	.69 (32)	.70 (32)	.43* (25)	.39† (25)
FNM	.81 (32)	.82 (32)	.65 (25)	.69 (25)
NMR	.19† (31)	.18† (31)	.11† (20)	.06† (20)
FLO	--	--	.81 (25)	.81 (25)

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

ACC	--	--	.89 (25)	.88 (25)
ROR	--	--	.64 (25)	.59** (25)

Note. Pair-wise sample sizes reported in parenthesis. ICS = IDAPEL Composite Score; FPS = Facilité à reconnaître la premier son; FDL = Facilité à dénommer des lettres; FSP = Facilité à segmenter les phonèmes; FNM = Facilité à lire des non-mots; FLO = Facilité en lecture orale; ACC = FLO Accuracy; ROR = Rapport oral du récit. CPPA = Conscience Phonologique/ Principe Alphabétique; Total = ÉCOLE total score. Unless marked, correlations are significant, $p < .001$; '**' $p < .01$, '*' $p < .05$, '†' Not Significant.

IDAPEL Benchmark Goals and Cut-Points for Risk

Development and Rationale. The IDAPEL benchmark goals and cut points for risk were developed from data collected in this study. The benchmark goals are empirically derived, criterion-referenced target scores that represent adequate reading progress, and are based on the predictive validity of a score on a measure at a particular point in time, compared to later IDAPEL measures and external outcome assessments.

In the 2011 National Assessment of Educational Progress (NAEP), 33% of students who score below the level of reading skills are judged to be ‘Basic’, and 67% of students who score below the level are judged to be ‘Proficient’. According to NAEP, “Basic denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at a given grade.” Thus, students who score at the 40th percentile or above on a high-quality, nationally norm-referenced test are likely to be rated Basic or above on the NAEP and can be considered to have adequate reading skills. While to date, there are no criterion-reference norms for ÉCOLE, in this study, we used the 40th percentile or above on the ÉCOLE as *one approximation* of adequate reading skills. Our intent is to develop generalizable benchmark goals and cut points that will be relevant and appropriate for a wide variety of reading outcomes, across a wide variety of regions,

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

and for diverse groups of students. The IDAPEL Composite Score was also used as a criterion to support the research behind the goals.

Benchmark Goals / Core Support. The benchmark goal indicates a level of skill where the student is likely to achieve the next IDAPEL benchmark goal or reading outcome. If a student achieves a benchmark goal, then the odds are in favor of that student achieving later reading outcomes if he/she receives generally effective, research-based instruction from a core classroom curriculum.

Cut-Point for Risk / Intensive Support. The cut points for risk indicate a level of skill below which the student is unlikely to achieve subsequent reading goals without receiving additional, targeted instructional support. Students with scores below the cut point for risk are identified as likely to need *intensive support*. Intensive support refers to interventions that incorporate something more or something different from the core curriculum or supplemental support.

Strategic Support. Between a benchmark goal and a cut point for risk is a range of scores where the student's future performance is harder to predict. Students with scores below the benchmark goal but above the cut point for risk are identified as likely to need *strategic support*. Strategic support refers to carefully targeted additional support in the skill areas where they are having difficulty.

IDPAEL Composite Score : Step-by-Step Procedures. Our fundamental logic for developing the IDAPEL Composite Score benchmark goals and cut points for risk was to begin with the external outcome goal and work backward in that step-by-step process. We used ÉCOLE at the end of the year with a level of performance that would represent adequate reading skills. Next we specified the benchmark goal and cut point for risk on the end of year IDAPEL

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Composite Score with respect to the end-of-year external criterion. Then, using the IDAPEL Composite Score end-of-year goal as an internal criterion, we established the benchmark goals and cut points for risk on the middle-of-year IDAPEL Composite Score. Finally, we established the benchmark goals and cut points for risk on the beginning-of-year IDAPEL Composite Score using the middle-of-year IDAPEL Composite Score as an internal criterion.

Once the benchmark goals and cut points for risk were established for the IDAPEL Composite Score, they were used to establish the specific goals and cut points for risk for each individual IDAPEL measure. The same step-by-step procedures were used for the individual measures. Table 15 outlines descriptive statistics for IDAPEL Composite Scores for the two cohorts in the FL1 study.

Design Specifications. The primary specification for the IDAPEL benchmark goals was to establish a level of skill where students scoring at or above benchmark have favorable odds (80%-90% probability estimated from a logistic regression) of achieving subsequent reading outcomes. In between the benchmark goal and the cut point for risk is a level of skill where the odds are about even (40%-60% probability estimated from a logistic regression) of achieving subsequent reading outcomes. The primary specification for an IDAPEL cut point for risk is a level of skill where students scoring below that level have low odds (10%-20% probability estimated from a logistic regression) of achieving subsequent reading outcomes.

A secondary consideration in establishing benchmark goals and cut points for risk was based on an examination of marginal probability of categorizing students into levels of support inconsistent from predictor to criterion.

To illustrate the concept of marginal probability and how the goals and cut-points were derived, Table 12 reports an example of the marginal probability for the IDAPEL Composite

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Score for first grade in middle of year. The end-of-year IDAPEL Composite Score is used to determine the benchmark goal and cut-point for risk for the middle-of-year IDAPEL Composite Score. The total number of students whose scores were above the 40th percentile for the end-of-year IDAPEL Composite Score is 89, approximately 61% of the total. Of those students, 79 students had scores that were also above the 40th percentile on the middle-of-year IDAPEL Composite Score. There were some students who crossed support groups from middle to end of year, but the goal was to derive the benchmark goals and cut-points for risk to keep the marginal probability for each time-of-year approximately equivalent.

Table 12

Example Marginal Probability Table for First-Grade Middle of Year IDAPEL Composite Score

End of Year Level-of-Support Groups	Middle of Year Level-of-Support Groups			Row Total	Marginal Row Percent
	Intensive	Strategic	Core		
Core	1	9	79	89	61%
Strategic	5	15	8	28	19%
Intensive	25	3	2	30	20%
Column Total	31	27	89	147	
Marginal Column Percent	21%	18%	61%		

Note. Core Support indicates the number of students whose scores were above the 40th percentile, i.e., benchmark status; Strategic Support indicates the number of students whose scores were between the 20th and 40th percentile; Intensive Support indicates the number of students whose scores were below the 20th percentile, i.e., the cut-point for risk. ICS benchmark goals: 125 (middle of year) and 203 (end of year). ICS cut-points for risk: 88 (middle of year) and 146 (end of year). For IDAPEL Composite Score Benchmark Goals, see Details pages.

In addition to the primary and secondary considerations in establishing benchmark goals and cut points for risk, we also considered a number of issues including:

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

- The pattern of student performance in the scatter plot. We tried to establish goals where students scoring at or above benchmark on the predictor were mostly also at or above benchmark on the criterion.
- The receiver operator characteristic (ROC) curve analysis. A large area under curve (AUC) is desirable in ROC analysis and indicates a good trade-off of sensitivity and specificity.
- Other metrics for decision utility including sensitivity, specificity, negative predictive power, positive predictive power, percent accurate classification, and Kappa.
- The overall pattern of benchmark goals and cut points for risk across measures and grades. In addition, we considered the theoretical relations between core components of early literacy in our model.

No single concern was used in isolation from other concerns. Frequently we had to balance disparate concerns to obtain a satisfactory compromise. For example, increasing the benchmark goal might result in a better match of marginal probability, but might compromise the predicted probability in the logistic regression analysis. Alternatively, a lower benchmark goal might work better for the beginning of year to middle of year analysis, but perform more poorly in the middle of year to end of year analysis. In other cases, the logistic regression analysis did not fit the data well, and consequently the role of the logistic regression analysis was discounted in establishing the benchmark goals and cut points for risk. The benchmark goals and cut points for risk represent our best balance of all the considerations identified here.

Linking IDAPEL Score Levels to Likely Need for Support. A key point in this discussion of probability is that the student's outcome is unknown and not fixed at the time of the initial screening. Instead, the outcome is the result of both the student's initial skills and the targeted,

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

differentiated instruction and intervention that are provided as a direct result of the screening information. The link between the probability of achieving subsequent early literacy goals, IDAPEL score levels, and likely need for support is summarized in Table 13. For all students, those who are at or above benchmark, below benchmark, and well below benchmark, our charge is to provide adequate support so they all achieve subsequent early literacy goals.

Table 13

Odds of Achieving Subsequent Early Literacy Goals, Score Levels, and Likely Need for Support.

Odds of achieving subsequent early literacy goals	Score Level	Likely need for support to achieve subsequent early literacy goals
80% to 90%	At or Above Benchmark <i>scores at or above the benchmark goal</i>	Likely to Need Core Support
40% to 60%	Below Benchmark <i>scores below the benchmark goal and at or above the cut point for risk</i>	Likely to Need Strategic Support
10% to 20%	Well Below Benchmark <i>scores below the cut point for risk</i>	Likely to Need Intensive Support

Levels of Support. Table 14 separates students into groups based on the newly proposed benchmark goals and cut-points for risk for the IDAPEL measures.

Each grade recorded data for first-grade measures, thus their levels of support can be directly compared. From Table 14, the proportion of students at benchmark on most first-grade measures is similar across cohorts, that is, more than 70% of the students reached benchmark on ICS, FNM, FLO, and ACC. Scores on FSP, ROR, and NMR were disparate across cohorts. In general, students in Cohort K-1 performed better on FSP, ROR, and NMR than their Cohort 1-2

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

counterparts. For FSP, the proportion of first-grade students between cohorts in the Intensive Support category is not significantly different (10% for Cohort K-1, 8% in Cohort 1-2), but the remaining students were split disparately into the Strategic (21% in Cohort K-1 and 58% in Cohort 1-2) and Core Support (69% in Cohort K-1 and 35% in Cohort 1-2) groups. We see similar results for NMR. For ROR, the disparity is reversed; more students are at benchmark in Cohort 1-2 than for Cohort K-1.

Table 14

Percentage of Students in Each Level of Support Category by Cohort

Measure by Time of Year	Level of Support Groups by Cohort					
	Cohort K-1			Cohort 1-2		
	Intensive	Strategic	Core	Intensive	Strategic	Core
<i>Kindergarten</i>						
Beginning of Year						
ICS	26.32	23.68	50.00	--	--	--
FPS	18.42	26.32	55.26	--	--	--
FDL	28.95	15.79	55.26	--	--	--
Middle of Year						
ICS	28.95	18.42	52.63	--	--	--
FPS	21.05	21.05	57.89	--	--	--
FDL	26.32	15.79	57.89	--	--	--
FSP	28.95	31.58	39.47	--	--	--
FNM	18.42	21.05	60.53	--	--	--
End of Year						
ICS	23.53	8.82	67.65	--	--	--

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

FDL	17.65	17.65	64.71	--	--	--
FSP	20.59	32.35	47.06	--	--	--
FNM	26.47	2.94	70.59	--	--	--
<hr/>						
<i>First Grade</i>						
Beginning of Year						
ICS	20.51	5.13	74.36	26.92	11.54	61.54
FSP	10.26	20.51	69.23	7.69	57.69	34.62
FNM	20.51	15.38	64.10	11.54	26.92	61.54
NMR	21.43	21.43	57.14	61.90	9.52	28.57
Middle of Year						
ICS	21.62	2.70	75.68	4.00	8.00	88.00
FNM	18.42	13.16	68.42	7.69	11.54	80.77
NMR	18.92	8.11	72.97	32.00	20.00	48.00
FLO	21.05	2.63	76.32	3.85	7.69	88.46
ACC	21.05	5.26	73.68	11.54	3.85	84.62
ROR	24.32	64.86	10.81	0.00	11.54	88.46
End of Year						
ICS	19.44	5.56	75.00	4.76	9.52	85.71
FNM	13.51	10.81	75.68	11.54	11.54	76.92
NMR	5.56	11.11	83.33	23.81	42.86	33.33
FLO	16.22	8.11	75.68	3.85	11.54	84.62
ACC	18.92	18.92	62.16	3.85	15.38	80.77
ROR	75.00	13.89	11.11	0.00	11.54	88.46
<hr/>						

Second Grade

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Beginning of Year

ICS	--	--	--	8.00	12.00	80.00
FNM	--	--	--	8.00	4.00	88.00
NMR	--	--	--	3.85	15.38	80.77
FLO	--	--	--	7.69	7.69	84.62
ACC	--	--	--	7.69	7.69	84.62
ROR	--	--	--	7.69	11.54	80.77

Middle of Year

ICS	--	--	--	3.85	11.54	84.62
FLO	--	--	--	3.85	7.69	88.46
ACC	--	--	--	7.69	11.54	80.77
ROR	--	--	--	3.85	15.38	80.77

End of Year

ICS	--	--	--	11.54	7.69	80.77
FLO	--	--	--	11.54	7.69	80.77
ACC	--	--	--	7.69	7.69	84.62
ROR	--	--	--	11.54	34.62	53.85

Note. Approximate sample sizes: Cohort K-1 = 38; Cohort 1-2 = 26. ICS = IDAPEL Composite Score; FPS = Facilité à reconnaître la premier son; FDL = Facilité à dénommer des lettres; FSP = Facilité à segmenter les phonèmes; FNM = Facilité à lire des non-mots; FLO = Facilité en lecture orale; ACC = FLO Accuracy; ROR = Rapport oral du récit.

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

Table 15

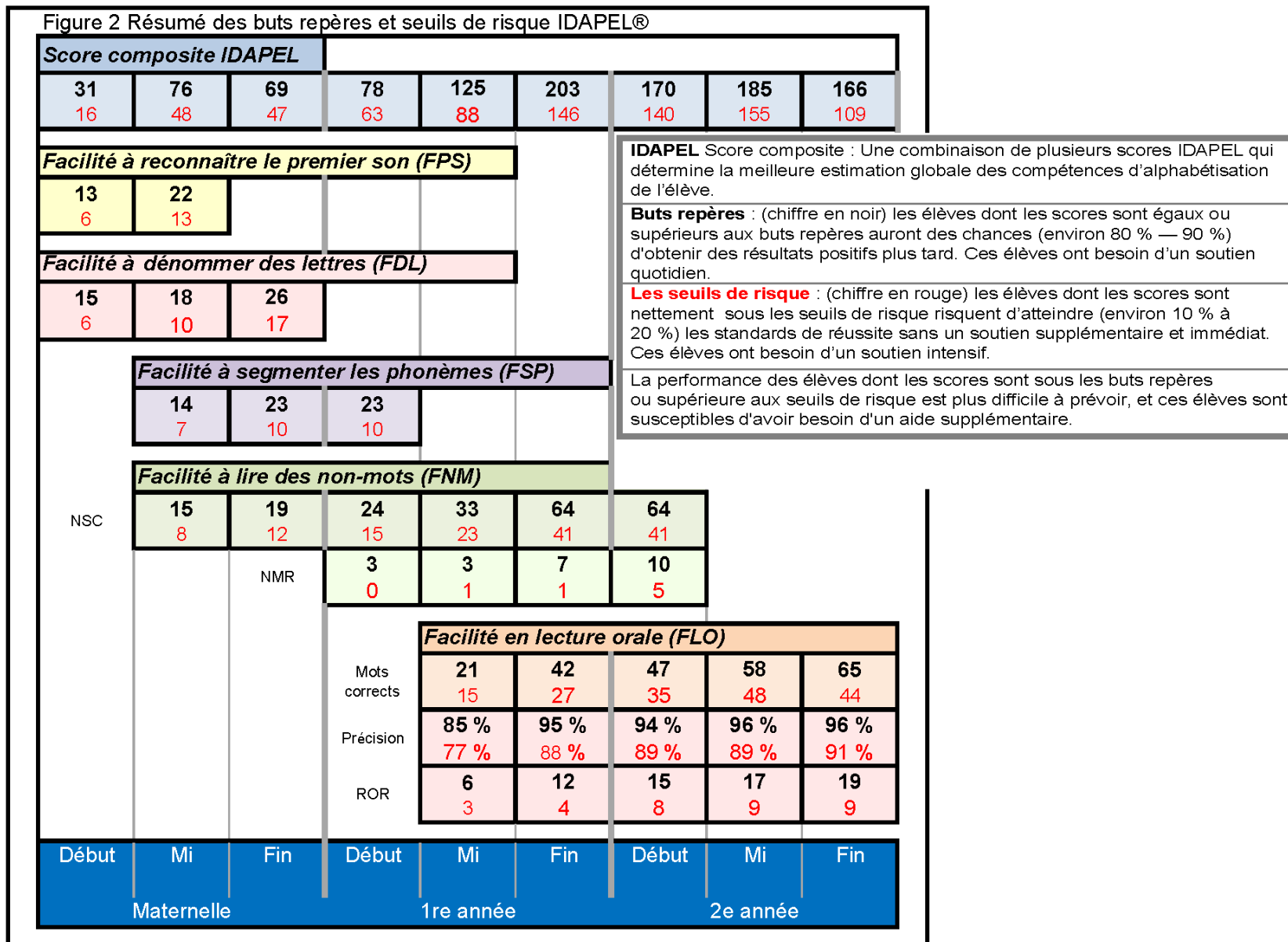
Descriptive Statistics by Cohort for IDAPEL Composite scores for the School Years 2008-2009 and 2009-2010.

ICS by Time of Year	Descriptive Statistics by Cohort					
	Cohort K-1			Cohort 1-2		
	N	Mean	SD	N	Mean	SD
<i>Kindergarten</i>						
Beg.	38	31.37	19.41	--	--	--
Mid.	38	75.76	41.67	--	--	--
End	34	73.26	35.64	--	--	--
<i>First Grade</i>						
Beg.	39	102.54	46.20	26	96.58	40.67
Mid.	37	176.84	93.44	25	189.20	64.43
End	36	240.22	107.76	21	256.95	65.98
<i>Second Grade</i>						
Beg.	--	--	--	25	226.04	70.33
Mid.	--	--	--	26	199.15	48.44
End	--	--	--	26	225.19	66.74

Note. Cohorts include students that recorded scores for at least one measure in both grade levels. ICS = IDAPEL Composite Score; FPS = Facilité à reconnaître la premier son; FDL = Facilité à dénommer des lettres; FSP = Facilité à segmenter les phonèmes; FNM = Facilité à lire des non-mots; FLO = Facilité en lecture orale; ROR = Rapport oral du récit. ÉCOLE criterion subtests were administered at the end of school Year 2008-2009.

Figure 16

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)



Discussion

In order to predict, as accurately as possible with IDAPEL, a student's likely success of future reading outcomes, this study developed a rationale and a step-by-step procedure for elaborating levels of performance, (i.e., benchmark goals and cut-points for reading risk) for each of the IDAPEL measures and for each grade kindergarten, first and second. Our secondary task was to evaluate the reliability and validity of IDAPEL with an internal (IDAPEL composite score) and external (ÉCOLE) criterion measure. This discussion section details the results from both analyses.

1. What are the within- and across-year levels of adequate growth on student reading performance as measured by IDAPEL?

There is evidence for substantial and adequate within- and across-year skill growth in all grades and in all areas of early literacy skills acquisition as measured by IDAPEL which is what we expected. The one exception is with FSP whereby no evidence of growth is seen beyond middle of year first grade. Data confirms that students have achieved maximum scores by beginning of year, which further validates our recommendation that no further administration of the measure occur beyond beginning of year first grade unless there is need for concern with a given student. Given the data outcomes, we decided not to include benchmark goals on this measure for middle and end of year in first grade. As a whole, the IDAPEL measures capture adequately well, student performance on early literacy skills as evidenced by strong levels of skills growth-across year and across grades.

2. What is the predictive and concurrent validity of IDAPEL® assessments with regard to Échelles de compétences en lecture (ÉCOLE), a criterion measure of reading proficiency that includes comprehension?

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

To answer this question, we examined the correlational relationships between the IDAPEL measures and ÉCOLE criterion measure scores that contribute to the grade-level CPPA (conscience phonologique/principe alphabétique) score. The concurrent and predictive validity for IDAPEL Composite Score (ICS) is strong across all grades and time of year. All kindergarten measures (with the exception of the NMR score within the FNM measure) are moderate-strong to strong. In first grade, FNM, FLO, ACC, and ROR are moderate-strong to strong. The small correlational relationship with FSP in first grade provides further evidence that administration of that measure is not necessary at the middle of year, unless a student is at-risk on that skill. Our data indicates that higher scores on this measure in the middle of year do not satisfactorily predict later reading outcomes. One possible explanation is that students have moved beyond this skill to the next foundational literacy skill, the alphabetic principle. The validity in kindergarten is strong for FDL and moderate-strong to strong for FSP, but is no longer significant by first grade.

3. What levels of performance on IDAPEL[®] measures (i.e., benchmark goals and cut-points for risk) would allow for an accurate prediction of a student's likelihood of reaching subsequent benchmark goals?

The benchmark goals/cut points for risk for the individual IDAPEL measures in addition to the benchmark goals and cut points for risk for the IDAPEL Composite Score provide important information about the likelihood of students achieving later reading goals. These benchmark goals and cut points for risk are found in Figure 16. The development of these goals can be found in the results section of this technical report, beginning on page 43.

Summary. As a whole within and across years, the IDAPEL measures capture adequately well, student performance on early literacy skills as evidenced by strong levels of skills growth-across

RUNNING HEAD: IDAPEL Benchmark Goals and Cut Points for Risk (FL1)

year and across grades. The one exception is with FSP whereby no evidence of growth is seen beyond middle of year first grade. The validity of IDAPEL kindergarten measures remain moderate-strong to strong for FDL, FPS, FSP, and FNM-NSC. FNM-NMR does not appear to be as strong with these cohorts of kindergarten students. Similarly, the validity of first grade IDAPEL measures FNM, FLO, ACC, and ROR are moderate-strong to strong, and not so for FSP. The concurrent and predictive validity of FLO words correct and accuracy is well supported by ÉCOLE, and affirms that IDAPEL FLO is a robust measure of reading skill. Similarly, the IDAPEL Composite Score (ICS) correlates well with ÉCOLE, is strong across all grades and time of year, and is an excellent indicator of overall reading skill.

Limitations

Key issues contribute to the study limitations. First and foremost, given small sample size, this study needs to be replicated with a larger set of students such as to confirm the benchmark goals. Benchmark goals should be re-evaluated with an independent group of students in order to increase the diversity of the sample. While we are confident that the measures adequately capture within- and across-year skill growth, we are also keen to continue IDAPEL measure research such as to further explore the need to adjust individual measure and to continue validating both measures and benchmark goals. Another key question for future research would be to examine whether or not the benchmark goals are functioning as intended for all grade levels.

A second limitation concerns the cohorts themselves. We know that cohort K-1 received instruction during both kindergarten and first grade year as we were able to capture raw scores on all measures across the two consecutive school years. The level of instruction is reflected in the percentage of students reaching the benchmark goal at beginning, middle and end of year.

What we don't know is if whether or not cohort 1-2 attended kindergarten or even received instruction at the kindergarten level. Despite this unknown, we see that the percentage of students reaching the benchmark goal is high across the years of the study from middle of year first grade through the end of year in second grade.

Implications for Research and Practice

Future research should include repeating this study with a sample that is more representative of the wide range of student performance across all measures to allow for greater generalizability of results, and to confirm whether or not the benchmark goals are functioning as intended for all grade levels represented in this study.

References

- Cormier, P., Desrochers, A., & Sénéchal, M. (2006). Validation et consistance interne d'une batterie de tests pour l'évaluation multidimensionnelle de la lecture en français. *Revue des sciences de l'éducation* 32, 1, 205-225.
- 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).
- Desrochers, A. (2011). *ÉCOLE: épreuves de compétence en lecture*. Manuscript in preparation, School of Psychology, University of Ottawa, Ottawa, Canada.
- 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. (2003). *Assessing French Reading Skills of Elementary French Immersion Students: Utility of DIBELS in French*. Unpublished doctoral dissertation, University of Oregon, Eugene.
- Dufour-Martel, C., & Good, R. H. (2009). Investigating the psychometric properties of three French language early reading measures, *Effective Education*, 1, 1, 87-100.
- 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.

- 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.
- Kaminski, R. A., & Good, III, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review* 25, 215-227.
- 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
- 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.
- National Research Council. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- 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.
- Simonart, G. (1998). *Échelle d'Apprentissages Scolaires primaires (A) (ÉCHAS)*. Association Francophone des Agents Psycho-Médico-sociaux, Belgique.
- Tindal, G., & Marston, D. (1990). *Classroom-based assessment: Testing for teachers*. Columbus, OH: Merrill.

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