Slide 1

The use of formative assessment for educational decision making in cross-cultural contexts

Chantal Dufour-Martel, Dynamic Measurement Group

Elizabeth Schaughnency, University of Otago

Elana Weinberger, Yeshiva University

Scott J. Goldberg, Yeshiva University

Kristen MacConnell, Nido de Aquillas, Santiago

Roland H. Good, Dynamic Measurement Group

Dynamic
Measurement Group
Supporting School Success One Step at a Time

1

Slide 2

Chantal Dufour-Martel, Ph.D.

Investigating the psychometric properties of *IDAPEL* French-language early literacy measures with students learning to read in French

2

Slide 3

IDAPEL Indicateurs Dynamiques d'Habiletés Précoces en Lecture

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)

3

Slide 4

Purpose & Design

Study explored the reliability and validity of the measures with French language first students

We used a correlational research design

4

Criterion: École (A. Desrochers, U of Ottawa)

IDAPEL Measures	ÉCOLE Measures
Facilité à reconnaître le premier son (FPS)	Isolement du phonème initial
Facilité à segmenter les phonèmes (FSP)	Segmentation phonémique
	Élision phonémique
Facilité à dénommer des lettres (FDL)	Lecture de sons simples et complexes
Facilité à lire des non-mots (FNM)	Lecture orale de pseudo mots
Facilité en lecture orale (FLO)	Lecture orale de mots
	Test de lecture orale « Marie »
Rappel oral du Récit (ROR)	Épreuve de compréhension mots-dessins
	Épreuve de compréhension phrases-dessins

Method

Participants included Kindergarten, first and second grade French-speaking students assessed at three periods during the school year

A subset was assessed at two-week intervals using alternate forms on 2 measures

Another sub-set of student participants assessed on criterion ÉCOLE at end of year

Participants

Grade level	Number	Mean age (End of school year)
Kindergarten	138	5.9 years (71 months)
First Grade	110	6.9 years (82.2 months)
Second Grade	50	7.9 years (95 months)

298 students
47 % girls
53 % boys

- included all students who were receiving French-language reading instruction, including students with disabilities.

- From 19 general education classrooms in 5 elementary schools from 3 separate school districts (urban, suburban and rural) in a Canadian province where English is majority language and French a minority language.

Data collection

35 French teaching staff received 2 day essential training in measure administration with mid-year refresher training

Collected and entered data at three time points during school year (fall, winter, spring)

Alternative form reliability collected by trained support personnel

Criterion measure data collected by University of Ottawa doctoral students

Analyses

Addressed four issues:

1. Measure sensitivity for skill growth
2. Alternate form reliability
3. Construct validity with criterion measures
4. Predictive validity beginning to end of year

1. Verification of skill growth across year

Analyses:

Growth patterns observed from mean scores verified through ANOVA with Bonferonni correction

Kindergarten and Grade 1 analysis carried out with sufficient statistical power ($\geq$ .92) except for Grade 2 ROR; power insufficient ($\leq$ .16) .

Omnibus tests on 3 benchmark data points were followed with multiple comparisons; if significant, mean scores differed significantly from Fall to Winter, and from Winter to Spring (p < .001).

Results:

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 three time points 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.

2. Alternate form reliability

All coefficients, except for 1, range between .53 and .81.

Most of coefficients approach Salvia et al.'s (2007) reliability criterion for important 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 not for screening decisions.

The word recoding (NMR) coefficient at the end of Grade 1 is low (i.e. .50); (possibility is that students do not have the skill EOY)

2. Alternate form reliability for FLO

Reliability information can be drawn from the data for FLO and ROR since three forms were taken at different time points in grades 1 and 2.

Scores on the different forms of (FLO) taken closely in time are highly correlated with one another (range from .80 to .98).

Results indicate most measures meet Salvia et al.'s reliability standard for important educational decisions.

The one subset of measures that deviates from this pattern is ROR measured at the end of Grade 1. The lower correlations (.66, .75, .66) may reflect problems with the measure or with the teacher assessment procedure.

4. Construct validity Analyses

Most IDAPEL measures that tapped phonological awareness, decoding or reading skills were positively and significantly correlated with conceptually related ÉCOLE measures

See two exceptions at the end of year (EOY) grade 1

- Ceiling effects on FSP - reasonable to the extent that most students have mastered the skill (no BM goal for middle of year nor EOY)

- For NMR, distribution of scores show significant floor effects (our recommended BM goal is 3 recoded words EOY). Floor effects may have to do with instruction ?

13

4. Predictive validity: beginning to end of year

GK and G1 correlations are reported

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.

Similar pattern in G1; significant correlations range from .19 to .74; largest coefficients are found between early FDL and FNM and the late measures of oral reading (e.g. FLO, text Marie, Words Read).

An unexpected finding for NMR taken in the fall is not significantly correlated with the same measure taken in the spring

Overall, the results strongly suggest that FDL, FSP, and FNM are useful predictors of future performance in reading as indexed by multiple outcome indicators

14

Conclusions

1. Measure sensitivity for skill growth	1. 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).
2. Alternate form reliability	2. The reliability of most measures is satisfactory for most screening decisions and some for making important educational decisions at the individual student level.
3. Construct validity with criterion measures	3. IDAPEL measures were positively and significantly correlated with ÉCOLE measures. EOY G1outcomes on FSP were less strongly correlated for obvious reasons; EOY G1 outcomes on NMR is reflected in BM goal.
4. Predictive validity beginning to end of year	4. GK: early FPS and FDL were found to be strongly correlated with EOY measures of phonological awareness, letter knowledge and reading. G1: FDL, FNM and FLO were found to be the most potent and consistent predictors of year-end measures of single-word and sentence reading.

15

Conclusions

There is 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 student standing.

16

4

Chantal Dufour-Martel
chantaldm@dibels.org

17