

## L'évaluation à fonction diagnostique: définitions et implications

Présenté par:

Catherine Croisetière, UQAM

[croisetiere.catherine@uqam.ca](mailto:croisetiere.catherine@uqam.ca)

En collaboration avec:

Nathalie Loye, Université de Montréal

## Contenu du séminaire

- Contexte de la présentation
- Qui êtes-vous?
- Problématique
- L'évaluation à fonction diagnostique
  - Origine
  - Définitions
  - Caractéristiques
  - Étapes
  - Construction d'un test
  - Exemples d'instruments
- Questions/discussion

## Contexte de la présentation

- Réalisation d'un stage doctoral
- Recension des écrits: manque de clarté concernant la fonction diagnostique de l'évaluation
  - Formative?
  - Destinée à des élèves seulement?
  - Évaluation = test?
- Besoin de développement des évaluations à fonction diagnostique
- Partage des connaissances: document écrit disponible sur demande

Qui êtes-vous?

## Buts de la présentation

- Outiller les étudiants de 2e et 3e cycle, professeurs s'intéressant au sujet
- Discuter de la fonction diagnostique de l'évaluation avec des collègues

## • **PROBLÉMATIQUE**

## Importance de la réussite scolaire

- Niveau de littératie minimal requis afin de fonctionner en société est de plus en plus élevé
- Liens avec décrochage, probabilité d'obtenir un emploi...
- 72 % des élèves obtiennent un diplôme d'études secondaires avant 20 ans (MELS, 2010)

## Pratiques d'évaluation habituelles

- Observations informelles des enseignants, données issues d'instruments non standardisés, usage restreint à une école (Garcia et Desrochers, 1997)
- 59 % des enseignants sont d'avis que les évaluations formatives faites en classe = meilleur moyen de collecter des informations diagnostiques à propos de leurs élèves (Huff et Goodman, 2007)
- Êtes-vous en accord avec ces propos?

## Pistes de réponse

- Évaluations formatives: apportent de judicieuses informations
- Pas toujours bien construites (Huff et Goodman, 2007)
- Peut causer problème si résultat non suffisamment précis, jugement porté à partir de l'instrument non valide
- Peut engendrer interventions inadéquates, pertes de temps (Ricard, 2007)

## Ce que les recherches en disent...

- Quantité et qualité du matériel en lecture à améliorer (Morin et collab., 2006)
- Enseignants manquent de moyens pour aider leur clientèle (Biancarosa, 2005)
- Pratiques d'évaluation ne fournissent pas toujours des informations justes, ne mesurent pas une variété d'attributs cognitifs, ne sont pas autant liées aux enseignements qu'elles le devraient (Huff et Goodman, 2007)

## Modèle RTI

- De plus en plus important
- 3 éléments essentiels: (Shapiro, Solari et Petscher, 2008)
  - **Dépistage universel de tous les élèves pour déterminer lesquels sont à risque/en difficulté**
    - Utilisation de données fiables et valides
  - Programmes d'intervention précoces pour les élèves susceptibles d'éprouver des difficultés en lecture
  - Monitoring fréquent de la progression des élèves

## Importance d'utiliser des tests diagnostiques

- Peuvent contribuer de manière substantielle à l'évaluation formative faite en classe (Roussos et collab., 2007)
- Peuvent contribuer significativement à la démarche d'évaluation d'un élève (Kubiszyn et Borich, 2000) et, ainsi, à sa réussite (Scallon, 2004) par une intervention efficace (Leighton et Gierl, 2007)
- Répondent au besoin d'obtenir des informations qualitatives détaillées à propos du profil d'un élève
- Portrait individuel et de la classe (Maisonneuve, Riopel, Gagnon et Gagnon, 2006)
- Facilitent l'apprentissage (De la Torre, 2009), coup de pouce aux enseignants (Fletcher, 2006)
- Rapprochent l'évaluation de l'enseignement: meilleure compréhension de la performance d'un élève à un test et lient les théories cognitives à l'enseignement (Gierl, Leighton et Hunka, 2007)
- Permettent de tester les théories cognitives portant sur le développement et les problèmes de différentes compétences Cain et Oakhill, 2006)

## Plusieurs modèles de classification diagnostique disponibles dans la littérature

- Facilitent les inférences effectuées à partir des résultats obtenus à un test
- Peuvent optimiser l'apprentissage de l'élève
- Contribuent au processus de décision
- De la Torre, 2009; Rupp, Templin et Henson, 2010

## Développement nécessaire

- Outils d'évaluation standardisés peu nombreux (Cormier, Desrochers et Sénéchal, 2006)
- Peu d'évaluations conçues dans le but de donner une rétroaction de type diagnostique aux différents sujets (Jang, 2009)
- Demande provient de deux sources:
  - concepteurs d'instruments → changement dans les pratiques d'évaluation
  - utilisateurs: → désirent résultats plus détaillés sur la performance de leurs élèves, guider les décisions (Rupp, Templin et Henson, 2010)
- Rapport Knowing what students know (NRC, 2001) nécessité de développer les évaluations à partir de modèles cognitifs
- Étant donné les contraintes de temps, coût, complexité méthodologique, conçues pour une utilisation à grande échelle (Gierl et Leighton, 2007)

## Pour les chercheurs...

Les analyses cognitives des tests en éducation permettent de contribuer aux théories psychométriques:

- Mieux comprendre les scores obtenus aux différents tests
- Mieux comprendre ce qui est mesuré par les tests
- Émergence de nouvelles méthodes et améliorations des tests actuels
- Améliorer les théories sur l'apprentissage
- Gierl, Leighton et Hunka, 2000

## • L'ÉVALUATION DIAGNOSTIQUE



## Origine du mot “diagnostic”

- Provient du grec “diagignoskein”
- Qui signifie:
  - Connaître précisément
  - Décider
  - Se mettre d'accord sur
- D'abord utilisé en médecine et en biologie.
- Définitions varient selon les domaines

## En éducation

- Plutôt récentes
- On privilégiait autrefois les tests de rendement, sélection, placement
- Gagné en popularité depuis 1980-1990
  - Travaux de Embretson (1983)
  - Chapitres de Linn (1989)
- Importance des développements en statistique et en psychologie cognitive

## L'évaluation à fonction diagnostique

- Qu'est-ce qui la différencie des autres types d'évaluation (sommative, de classement, formative)?

De classement: performance de l'élève avant le début de l'instruction, établir le cheminement de l'élève selon ses besoins

Formative: progression des élèves, rétroaction à l'élève et à l'enseignant, guide la planification

Sommative: à la fin d'une unité d'apprentissage, permet à l'élève de passer au niveau suivant, à l'enseignant, de vérifier son efficacité

Diagnostique: informations plus fines, à partir de modèles cognitifs

## Quelques définitions

- Scallon(1988): permet de dépister des cas d'erreurs systématiques afin d'amorcer une intervention de nature corrective
- Legendre (2005): mode d'évaluation qui a pour but d'apprécier les caractéristiques individuelles d'un sujet (style cognitif, style d'apprentissage, intérêt, motivation, maîtrise des préalables, etc.) et de l'environnement pédagogique, lesquelles devraient avoir des influences positives ou négatives sur son cheminement d'apprentissage

## Quelques définitions (suite)

- Desrochers, Simon et Thompson (2011): vise l'identification des élèves à risque d'éprouver des difficultés ou le profilage des forces et des faiblesses des élèves qui éprouvent déjà des difficultés
- Laplante (2011): dans une perspective cognitive, le but de l'évaluation diagnostique est de comprendre quels processus impliqués dans le traitement de la langue écrite sont ou non fonctionnels, et quelles stratégies compensatoires sont utilisées, ceci, afin d'orienter l'intervention pédagogique ou orthopédagogique qui sera privilégiée pour favoriser cet apprentissage

## Quelques définitions (suite et fin)

- Rupp, Templin et Henson (2010) : l'évaluation à fonction diagnostique est un processus systématique visant à obtenir de l'information spécifique à propos des caractéristiques psychologiques d'une personne par l'utilisation de diverses méthodes. Elle vise à justifier, contrôler et optimiser les décisions et leurs impacts. Ce processus inclut 1) la spécification d'une question diagnostique, 2) la sélection d'une méthode diagnostique, 3) l'application et l'évaluation des données obtenues par la méthode choisie, 4) l'interprétation des données et le développement d'un rapport diagnostique, 5) la planification d'une intervention et 6) l'évaluation de l'efficacité de cette intervention.

- Quels éléments ressortent de ces définitions?

- D'après Yang et Embretson (2007), peu importe le domaine ciblé,:
  - Description des caractéristiques distinctives d'une chose ou d'un phénomène
  - Identification de la nature d'une chose ou des causes d'un phénomène
  - Décision ou conclusion, telle que la classification ou le verdict

## Caractéristiques des évaluations et tests diagnostiques

- Les erreurs – source d'intérêt
  - Significations psychologique et statistique de plus en plus reconnues
  - Contribuent à l'identification de stratégies, processus, cheminements particuliers
  - Si un test à choix multiple a bien été rédigé, les réponses erronées peuvent avoir un potentiel de discrimination aussi fort que les bonnes réponses

## Caractéristiques

- Répondants:
  - Ce qui est visé:
    - Élève
    - Groupe
    - Institution
    - Système
  - Plusieurs domaines
  - Les répondants ne sont pas nécessairement visés par la décision

## Caractéristiques

- Moment ciblé:
  - Décision par rapport à un comportement passé
  - Comportement présent
  - Comportement futur
- Objet de l'évaluation
  - Relever et inférer les différences de structures entre les individus (stratégies, processus mentaux)
  - Informations spécifiques
  - Permettent une rétroaction et identification de pistes de remédiation
- Type d'intervention
  - Sélection
  - Support
  - Modification de comportement

## Caractéristiques

- Méthode:
  - Personne qui évalue (répondant ou observateur externe)
  - Instruments variés
  - Données qualitatives et quantitatives: plusieurs scores et profil de performance
  - Longue démarche de collecte de données
    - Doivent faire partie d'un processus à long terme de diagnostic, traitement et évaluation
    - Implique une équipe

## Étapes

- Anamnèse du répondant
- Administrations de divers instruments d'évaluation
- Capacités évaluées en profondeur dans un domaine ciblé
- Rédaction d'un rapport avec pistes d'interventions
- Évaluation du répondant après traitement

## Construction d'un test

- Recours à un bon modèle théorique
  - Grande connaissance des processus, stratégies et habiletés requises lors de la réalisation de la tâche proposée (Gierl, 2007)
- Impossible de diagnostiquer quelque chose que nous ne connaissons pas (Jang, 2005)
- Consultation du modèle lors de chacune des étapes du développement du test

## 4 étapes précises (Jang, 2009)

- Définir le concept évalué par le test
  - Concevoir des tâches spécifiques permettant de donner un diagnostic
  - Développer un système de pointage permettant d'accéder à un profil précis de répondant
  - Optimiser les propriétés de l'instrument afin qu'il soit le plus précis possible
- Ordre non linéaire



## Définir le concept évalué par le test

- Description des mécanismes cognitifs, ou attributs requis pour la réalisation de la tâche
- Spécification des éléments liés aux items ou à la tâche pouvant influencer les mécanismes cognitifs utilisés par le participant lors de la passation du test (Nichols, 1994)
- Pour ce faire,
  - Recension d'écrits
  - Analyses de protocoles verbaux (Leighton et Gierl, 2007)

Un exemple en lecture...  
Attributs requis pour la réalisation de la tâche

Dimension	Caractéristique
Lecteur	Stade initial de l'apprentissage
	Langue maternelle ou langue seconde : lorsque langue seconde, incidence des facteurs motivationnels, expériences préalables, capacité de transfert et compétence globale en langue seconde
	Structures affectives <ul style="list-style-type: none"> <li>➤ motivation</li> <li>➤ intérêts</li> <li>➤ concept de soi</li> <li>➤ peur de l'échec</li> </ul>
	Ensemble de connaissances <ul style="list-style-type: none"> <li>➤ phonologiques</li> <li>➤ syntaxiques</li> <li>➤ sémantiques</li> <li>➤ pragmatiques</li> <li>➤ des structures de texte</li> <li>➤ sur le monde</li> </ul>
	Utilisation des processus de lecture (microprocessus, processus d'intégration, macroprocessus, processus métacognitifs, processus d'élaboration)
	Vitesse de lecture

Éléments liés aux items ou à la tâche pouvant influencer les mécanismes cognitifs utilisés par le participant

- **Texte**
  - intention de l'auteur
  - structure du texte
  - contenu (concepts, structures syntaxiques, longueur des phrases, vocabulaire, connaissances transmises)
  - type de texte (narratif, courant, autre(s))
- **Contexte**
  - Psychologique (intention de lecture : contexte informel ou formel, intérêt)
  - Social (interventions de l'enseignant, interventions des pairs)
  - Physique (bruit, temps, consignes, accès au texte, matériel)
  - Tâche :
    - Type d'évaluation : entrevue, épreuve de closure, observation des mouvements oculaires, questionnaire
    - Format des items : vrai-faux, questions à choix multiple, réponse narrative
      - Pseudo-chance et glissements

## Concevoir des tâches spécifiques

- **Sélection du design**
  - Utilisation d'un tableau de spécifications
  - Nombre suffisant d'items pour mesurer chacun des attributs
- **Administration du test**
  - Consignes
  - Documents remis

## Développement d'un système de pointage

- Peut être facilité par l'utilisation d'une matrice Q

## Exemple de 4 items en mathématiques (Rupp, 2010, p.150)

- 16 profils peuvent être identifiés à partir de 4 attributs, (addition, soustraction, multiplication, division) chacun pouvant être maîtrisé ou non maîtrisé
- Formule:  $2^A = 2^4 = 16$
- Préférable de se centrer sur un petit nombre d'attributs
- Peut être plus précis: par exemple, addition (retenue, deux chiffres, trois chiffres...)

Items	Addition	Soustraction	Multiplication	Division
$2 + 3 - 1 = ?$	1	1	0	0
$4/2 = ?$	0	0	0	1
$5 * 3 - 4 = ?$	0	1	1	0
$8 + 12 = ?$	1	0	0	0

## Optimisation des propriétés de l'instrument

- Paramètres des items
- Validation
- Permet une révision du modèle défini à la première étape

## Contenu

- Collecter également des informations contextuelles, consultations d'autres sources (bulletins...)
- Inclusion de co-variables permet d'améliorer la classification et de donner un diagnostic plus riche

## Forme

- Items susceptibles d'être travaillés en classe
- Concept d'équité (handicap physique, caractéristique linguistique ou culturelle), ne pas favoriser un groupe d'individus aux dépens d'un autre (Simon, 2011)

## Autres éléments importants

- Importance de l'alignement entre le curriculum, l'enseignement et l'évaluation (Leighton et Gierl, 2007)
- Rôle du test = recommander un diagnostic
- Jugement de l'enseignant
- Aucun instrument ne peut tout couvrir (Turcotte et Maisonneuve, 1995)

## Exemples d'instruments d'évaluation diagnostique – au collégial en langue

- TELEC
- Test de dépistage de difficultés en lecture
- Jang (2005)
- Autres instruments

## Test diagnostique de lecture pour le collégial (TELEC)

- Produit par l'équipe de Turcotte (1997) au collège Édouard-Montpetit
- Évalue la motivation à lire et les compétences langagières, cognitives et métacognitives des cégepiens
- 3 documents: fiche d'identification du répondant, cahier de textes, cahier de réponses
- Deux textes (narratif et argumentatif) à la base de huit tâches en lecture
- Consistance interne du test: alpha de Cronbach d'une valeur de 0,77
- 4 niveaux attribués pour chacune des tâches: aucune maîtrise à maîtrise efficace
- Production d'un bulletin diagnostique, décrit forces, faiblesses et suggestions
- Annexe I p.247

## Objets mesurés par le TELEC (Turcotte, 1997)

- Plan affectif (perceptions)  
Évaluation de la perception des élèves à propos de :
  - Difficulté des textes
  - Difficulté des tâches
  - Contrôlabilité des tâches
  - Compétence devant les tâches
  - Compréhension des tâches
  - Valeur des tâches
- Plan langagier - 3 aspects évalués  
Plan discursif (établir le ton d'un texte, le type de texte, le sens d'une image)  
Plan textuel (interpréter la transition entre les paragraphes, établir le sujet et l'idée principale d'un paragraphe, établir l'idée globale d'un texte)  
Plan phrastique (comprendre le sens des mots en contexte, repérer les mots exprimant une idée, utiliser des mots pour exprimer une idée comprise, interpréter un mot de liaison)
- Plan cognitif  
Identification des informations contenues dans un texte  
Inférer des informations non explicites dans un texte  
Résumer le contenu d'un texte  
Saisir la structuration des informations dans un texte
- Plan métacognitif  
Expliciter sa connaissance de certains processus de lecture  
Évaluer sa compréhension de certaines tâches de lecture
- De façon globale  
Vitesse de lecture  
Efficacité de l'élève à comprendre les textes lus

## Test de dépistage de difficultés en lecture

- Riopel, Gagnon, Gagnon et Maisonneuve (2006) au collège Jean-de-Brébeuf
- Extraits de différents textes
- Version longue: 30 items à cinq choix de réponse
- Version abrégée: 20 items à cinq choix de réponse (30 minutes)
- Corrélation entre résultats au test et cote de rendement au collégial 0,52
- Modèle incluant note terminale de français, préparation au test, note terminale en mathématiques, facteur anticipation de l'échec, aspirations scolaires, facteur de croyance explique 82 % de la variance observée dans la cote de rendement au collégial à la fin de la 1<sup>re</sup> session

## Test de dépistage de difficultés en lecture

- Difficultés d'ordre linguistique (énonciation, organisation, progression et cohésion du texte)
- Difficultés liées au raisonnement (structure logique, type de raisonnement, quantificateurs, présupposés)
- Difficultés liées aux connaissances du monde (si ce qui est lu est plausible)
- Exemples d'items (Maisonneuve et al. 2006)



## Thèse de Jang (2005)

- Test en anglais langue seconde (TOEFL)
- Remise d'un bulletin aux enseignants et aux élèves (annexe H)

## Consulter également...

- Mémoire de Ricard (2007)
- Pour le primaire en lecture:
  - Évaluation diagnostique de la lecture (Armand et Sabourin, 1995)
    - fin 1<sup>re</sup> – fin 3<sup>e</sup> – fin 5<sup>e</sup> année
  - Indicateurs dynamiques des savoirs essentiels en lecture (St-Laurent et Giasson, 2009)
  - Trousse d'évaluation GB +
- Pour fin primaire - secondaire
  - Le vol du P.C (Boutard, Claire, Gretchanovsky, 2010)

## Références utiles

- Rupp, A.A., Templin, J. et Henson, R.A. (2010). *Diagnostic Measurement: theory, methods, and applications*. New York, NY: The Guilford Press.
- Leighton, J. P. et Gierl, M. J. (2007). Why cognitive diagnostic assessment? Dans J. P. Leighton et M. J. Gierl (Dir.) *Cognitive diagnostic assessment for education*. New York, NY : Cambridge University Press.

## Autres références

- Armand, F. et Sabourin, S. (1995). *Examen diagnostique de la lecture. Niveau intermédiaire (fin cinquième année)*. Montréal, Québec: Commission scolaire Marie-Victorin.
- Boutard, C., Claire, I. et Gretchanovsky, L. (2010). *Le vol du P.C. : Évaluation fonctionnelle de la lecture chez les sujets de 11 à 18 ans. Isbergues* ; Ortho Édition.
- Chipman, S. F., Nichols, P.D. et Brennan, R. L. (1995). Introduction. Dans P.D. Nichols, S. F. Chipman et R. L. Brennan (Dir.), *Cognitively diagnostic assessment*. New York, NY : Lawrence Erlbaum Associates.
- Desrochers, A., Simon, M. et Thompson, G.L. (2011). Formes et fonctions de l'évaluation de la littératie. Dans M-J Berger et A. Desrochers (Dir.) *L'évaluation de la littératie*. Ottawa, Ontario : Presses de l'Université d'Ottawa.
- Gorin, J. S. (2006). Test design with cognition in mind. *Educational measurement: issues and practice*, 25(4), 21-35.
- Gorin, J. S. (2007). Test construction and diagnostic testing. Dans J. P. Leighton et M. J. Gierl (Dir.) *Cognitive diagnostic assessment for education*. New York, NY : Cambridge University Press.
- Huff, K. Et Goodman, D. P. (2007). The demand for cognitive diagnostic assessment. Dans J. P. Leighton et M. J. Gierl (Dir.) *Cognitive diagnostic assessment for education*. New York, NY : Cambridge University Press.
- Jang, E.E. (2005). *A validity narrative: effects of reading skills diagnosis on teaching and learning in the context of NG TOEFL*. Thèse de doctorat inédite. Champaign, IL: University of Illinois.
- Laveault, D. et Grégoire, J. (2002). *Introduction aux théories des tests en psychologie et en sciences de l'éducation*. Bruxelles, Belgique: de boeck.
- Luecht, R.M. (2007). Using information from multiple-choice distracters to enhance cognitive-diagnostic score reporting. Dans J. P. Leighton et M. J. Gierl (Dir.) *Cognitive diagnostic assessment for education*. New York, NY : Cambridge University Press.
- Nichols, P.D. (1994). A framework for developing cognitively diagnostic assessments, *Review of educational research*, 64(4), 575-603.
- Ricard, C. (2007). *Un test diagnostique qui évalue la compétence en compréhension en lecture chez les élèves en âge de s'inscrire au collégial*. Mémoire de maîtrise inédit, Université du Québec à Montréal, Montréal.
- Riopel, M., Gagnon, A., Gagnon, F. et Maisonneuve, H. (2006). *Production et validation d'un test de lecture visant le dépistage d'élèves en difficulté d'apprentissage*. Montréal, Québec : Collège Jean-de-Brébeuf.
- Scallon, G. (1988). *L'évaluation formative des apprentissages*. Québec, Québec : Les presses de l'Université Laval.
- Simon, M. (2011). Les qualités d'un instrument d'évaluation de la littératie. Dans M-J Berger et A. Desrochers (Dir.) *L'évaluation de la littératie*. Ottawa, Ontario : Presses de l'Université d'Ottawa.
- St-Laurent, L. et Giasson, J. (2009). *Indicateurs dynamiques des savoirs essentiels en lecture*. Québec, Québec: Université Laval. [Document téléchargé le 18 septembre 2009 du site : <http://www1.1lib.us/1lib.us/1lib.us/ca/indisse/>].
- Turcotte, A.G. (1997). *Évolution des compétences et des perceptions comme lecteurs des cégépiens aux études préuniversitaires interprétée à la lumière de l'enseignement collégial*. Thèse de doctorat inédite, Université de Montréal, Montréal.
- Turcotte, A.G. et Maisonneuve, L. (1995). *Un test qui évalue les compétences en lecture : le TELEC*. Association québécoise de pédagogie collégiale (AQPC), Actes du 15<sup>e</sup> colloque de l'AQPC – Pour s'y retrouver, pour se retrouver : le renouveau de l'enseignement collégial, 8 pages.
- Yang, X. et Embretson, S.E. (2007). Construct validity and cognitive diagnostic assessment. Dans J. P. Leighton et M. J. Gierl (Dir.) *Cognitive diagnostic assessment for education*. New York, NY : Cambridge University Press.



## Questions/discussion