

WEBDÉTECTE ET LE SOUS-CLASSEMENT INTENTIONNEL AU TEST DE CLASSEMENT EN ANGLAIS

Gilles RAÏCHE, professeur - Université du Québec à Montréal

RÉSUMÉ

Le personnel enseignant des départements de langues des cégeps et collèges du réseau collégial québécois est à la recherche, depuis longtemps, de moyens simples et efficaces pour, sinon contrer, du moins amenuiser le comportement de sous-classement aux tests de classement en anglais langue seconde de la part des étudiants. Plusieurs enseignants ont cru que la solution au problème du sous-classement résidait dans l'élaboration d'un test de classement plus efficace dans l'appréciation du niveau d'habileté de l'étudiant. Malheureusement, comme le démontrent les résultats consécutifs à l'élaboration par Laurier, Froio, Pearo et Fournier (1998) d'un test de classement en anglais langue seconde amélioré, le TCALS II, les étudiants du réseau collégial continuent à tenter de se sous-classer à ce test. D'autres solutions doivent donc être envisagées. Nous sommes à leur quête depuis plusieurs années (Raïche, 2000, 2002a; Raïche et Blais, 2003).

Nous présentons ici les étapes, ainsi que les résultats, d'un programme de recherche qui propose une solution au problème (Raïche, 2002b) et qui a mené au développement d'une application logicielle mise à la disposition de tous les collèges publics et privés du réseau collégial québécois en 2005. Les diverses phases du programme de recherche ont été soutenues financièrement par le programme d'aide à la recherche sur l'enseignement et l'apprentissage (PAREA) du ministère de l'Éducation du Québec, la Fondation pour l'innovation du Nouveau-Brunswick (FINB), le Conseil de recherches en sciences humaines du Canada (CRSH), l'Université de Moncton et l'Université du Québec à Montréal. Il a été réalisé en collaboration avec le Collège de l'Outaouais et le Conseil national de recherches du Canada (CNRC).

OBJECTIF SPÉCIFIQUE DU PROGRAMME DE RECHERCHE

Tenant compte du grand nombre d'administrations de tests de classement dans chacun des collèges, soit plus de 1 000 par an à l'intérieur d'un collège de taille moyenne, il est nécessaire d'utiliser une formule simple pour dépister les tentatives de sous-classement. Selon nous, l'analyse de la cohérence des réponses à chacune des questions permet de détecter des patrons de réponses plutôt étranges et ainsi peu probables. Certains auteurs se sont intéressés à cette problématique et ont proposé des indices numériques de scores plutôt suspects à un test (Cizek, 1999; Drasgow, Levine et Zickar, 1996; Meijer et Sijtsma, 2001; Reise et Flannery, 1996). C'est l'une des piste des plus prometteuses à envisager pour pallier ce problème. Maintes mesures d'ajustement inadéquat ont été proposées. Ces mesures ne s'adressent toutefois pas au problème spécifique du sous-classement. Habituellement, elles visent plutôt à déterminer de façon globale le mauvais ajustement des réponses au test sans se soucier de la nature du problème: copie, réponse au hasard, mauvaise utilisation de la feuille de réponses, etc. Nous avons voulu évaluer l'application de ces indices numériques à la détection du comportement de sous-classement dans le réseau collégial.

Pour nous permettre de répondre à cet objectif, nous devons, à une phase 1, identifier les stratégies utilisées par les étudiants pour tenter de se sous-classer au test de classement en anglais, langue seconde, utilisé à l'intérieur du réseau collégial québécois. À cette fin, nous avons organisé une rencontre avec des étudiants du Collège de l'Outaouais. Par la suite, à la phase 2, nous avons sélectionné et développé des indices de détection de patrons de réponses suspects qui sont mis à l'essai à partir d'une simulation par ordinateur. Ces indices sont ensuite appliqués (phase 3) à l'identification des étudiants des deux premiers tours du SRAM inscrits au Collège de l'Outaouais en 2002-2003 dont le patron de réponses au test de classement en anglais, langue seconde, est plutôt suspect. Enfin, à la phase 4, une application logicielle, *WebDéetecte*, est développée pour permettre aux collèges participants de réaliser la détection.

A. PHASE 1: LES STRATÉGIES DE SOUS-CLASSEMENT DES ÉLÈVES

Seize étudiants, qui avaient participé à l'administration du TCALS en 2001-2002, inscrits à un cours d'*Anglais langue seconde* de niveau 102 au Collège de l'Outaouais (604-EWG-03) du même groupe-classe ont participé à une entrevue de groupe. Il s'agissait d'un cours adapté à la famille des programmes rattachés aux Sciences sociales et aux Techniques humaines, regroupement, bien sûr, propre au Collège de l'Outaouais.

Par 14 fois, les étudiants ont indiqué qu'ils répondraient au hasard au test. Choisir la mauvaise réponse, avec diverses variantes, a été proposé par dix étudiants: c'est ce que nous nommons, plus loin, la stratégie de réponses inversées. Plusieurs ont souligné qu'ils utiliseraient cette dernière stratégie seulement s'ils connaissaient la bonne réponse. Malheureusement, puisque nous avons recueilli ces informations par écrit, nous n'avons pas pu découvrir si ces étudiants adopteraient une autre stratégie lorsqu'ils ne connaissent pas la bonne réponse. Six étudiants opteraient pour omettre de répondre à certaines questions, tandis que trois choisiraient de donner la bonne réponse aux questions faciles et la mauvaise réponse aux questions difficiles. Dans ce dernier cas, on peut se demander comment les étudiants peuvent juger du niveau de difficulté des questions. Certains ont indiqué qu'ils donneraient le même choix de réponse à plusieurs questions du test.

Certaines stratégies proposées ne nécessitent pas l'application d'une procédure bien complexe pour dépister le comportement de sous-classement. C'est le cas notamment des stratégies où l'étudiant omet sciemment de répondre à des questions. Seules quelques recommandations aux collèves, faciles à appliquer, sont alors nécessaires. Toutefois, les stratégies qui ont été le plus fréquemment identifiées, telles que de répondre au hasard ou de choisir la mauvaise réponse, nécessitent la mise en œuvre de procédures de dépistage plus sophistiquées. Les suites à cette étude ont d'ailleurs consisté à élaborer des procédures de dépistage de ces deux dernières stratégies.

B. PHASE 2: UNE SIMULATION D'ÉTUDIANTES ET D'ÉTUDIANTS QUI TENTENT DE SE SOUS-CLASSER

Puisque à partir des données réelles d'une cohorte étudiante nous ne pouvons pas connaître à l'avance quels sont les étudiants qui ont cherché à se sous-classer, il est impossible de mettre à l'épreuve les indices de détection directement avec ces étudiants. Seulement une simulation des patrons de réponses normaux et aberrants nous permet de vérifier le pouvoir de détection des différents indices.

Par ordinateur, nous avons simulé des patrons de réponses dans différentes conditions. Nous avons produit:

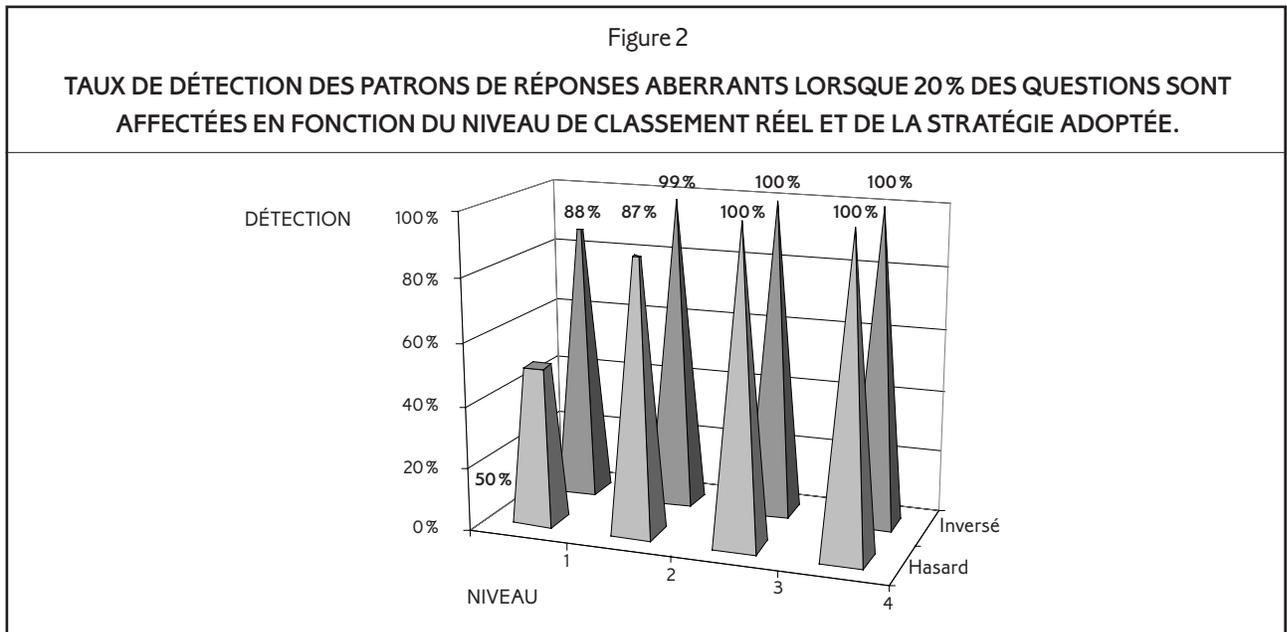
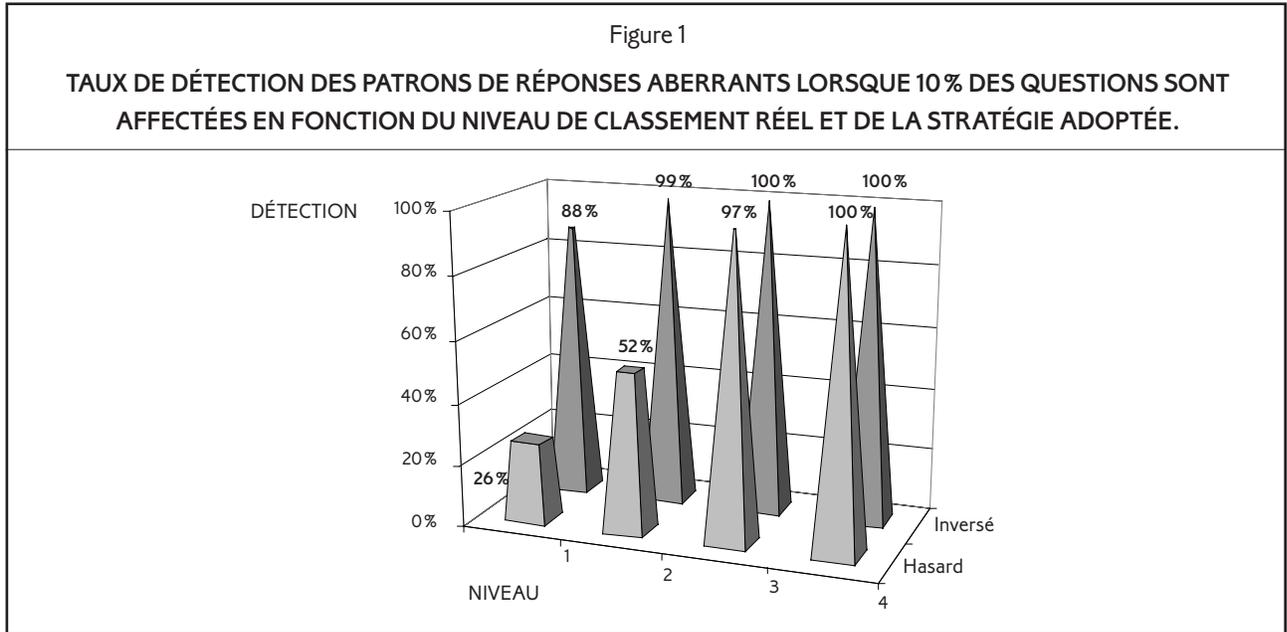
1. des patrons de réponses totalement normaux;
2. des patrons de réponses où 10 % des premières questions reçoivent des réponses produites au hasard;
3. des patrons de réponses où 20 % des premières questions reçoivent des réponses produites au hasard;
- 4) des patrons de réponses où 10 % des premières questions reçoivent des réponses inversées;
5. des patrons de réponses où 20 % des premières questions reçoivent des réponses inversées.

La Figure 1 nous renseigne sur l'efficacité de l'indice retenu, I_z , pour identifier les cas de sous-classement lorsque seulement 10 % des réponses sont soit au hasard, soit inversées. Quand aussi peu que 10 % des réponses sont données au hasard, le taux de détection des cas réels de sous-classement est égal ou supérieur à 97 % pour les étudiants des niveaux 3 et 4, et ce, quelle que soit la stratégie adoptée.

L'indice I_z se montre donc très efficace en ce qui a trait aux étudiants dont le niveau d'habileté en anglais, langue seconde, est élevé. Ce sont justement ces étudiants que nous espérions pouvoir identifier par cette procédure. Le résultat est supérieur à ce que nous avions espéré.

Lorsque 20 % des réponses sont données au hasard ou inversées, ces pourcentages augmentent considérablement. C'est ce qu'illustre la Figure 2. Même au niveau 1, au moins la moitié des cas de sous-classement sont détectés, tandis qu'aux niveaux 2, 3 et 4 le taux de détection est toujours supérieur ou égal à 87 %, quelle que soit la stratégie adoptée.

L'indice I_z constitue donc un indicateur très efficace pour nous permettre de détecter l'utilisation d'une stratégie de réponses au hasard ou inversée. De plus, son efficacité est à son maximum chez les étudiants qui, selon nous, risquent le plus souvent de désirer se sous-classer, soit ceux dont le niveau d'habileté en anglais est élevé.



C. PHASE 3: UNE APPLICATION AUX 1^E ET 2^E TOURS DU SRAM DE LA COHORTE 2002-2003 AU COLLÈGE DE L'OUTAOUAIS

Forts de ces résultats, nous avons appliqué l'indice I_z à la détection des cas potentiels de sous-classement chez les étudiants de la cohorte 2002-2003 admis et inscrits au Collège de l'Outaouais au premier et au second tours du Service d'admission du Montréal métropolitain (SRAM).

Nous avons appliqué l'indice I_z aux résultats des étudiants regroupés des premier et second tours du SRAM. Ces étudiants étaient au nombre de 1361. De ces 1361, l'indice I_z a permis d'identifier 141 étudiants dont le patron de réponses est douteux, soit 10,36 % de la cohorte. Cela nous donne une idée du pourcentage d'étudiants que cette démarche nous permet de détecter.

D. PHASE 4: LE DÉVELOPPEMENT D'UNE APPLICATION EN LIGNE DE DÉTECTION – *WEBDÉTECTE*

Comme nous l'avions planifié dès le début du programme de recherche, à la suite de ces résultats, une application logicielle a été développée. Nous offrons maintenant aux cégeps et collèges qui acceptent de participer à l'amélioration de l'utilisation de l'indice I_z au dépistage des étudiants qui tentent de se sous-classer une solution logicielle disponible directement sur Internet: *WebDétekte*. La Figure 3 affiche le portail des applications en *E-évaluation* réalisées par notre groupe de recherche où on peut accéder à *WebDétekte*. La Figure 4, présente, pour sa part, la page d'accueil de l'application.



Figure 3
PORTAIL DES APPLICATIONS RÉALISÉES
EN E-ÉVALUATION

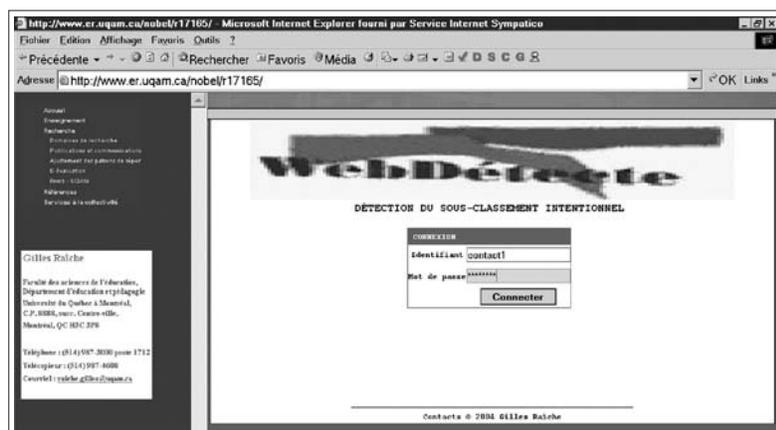


Figure 4
PAGE D'ACCUEIL DE WEBDÉTECTE

Actuellement, l'application ne s'adresse qu'aux cégeps qui utilisent la version longue du TCALS II (85 questions). Ceux-ci peuvent télécharger directement les résultats de leurs étudiants sur la plateforme et recevoir immédiatement une liste de ceux et celles qui auraient pu tenter de se sous-classer. Les collègues peuvent choisir d'utiliser un taux de fausses détections de 1 % ou de 5 %. Il faut cependant souligner que la détection ne garantit pas le dépistage des seuls étudiants qui auraient tenté de se sous-classer, puisque l'indice utilisé est à tout usage et indique uniquement que le patron de réponses est inusité. Ceux et celles qui auraient cherché à avoir un résultat plus élevé, par exemple, même s'ils sont peu nombreux, sont aussi détectés.

L'utilisation du logiciel est sans frais et devrait le demeurer tant que les collèges, qu'ils soient privés ou publics, acceptent que les chercheurs responsables du programme de recherche conservent les résultats non nominatifs de ces étudiants dans la base de données. De cette façon, il sera possible de continuer le développement de l'application et d'améliorer les stratégies de dépistage.

LES SUITES AU PROJET

Plus nos travaux sur la détection des patrons de réponses inappropriés avancent, plus les pistes de recherche s'accumulent, et des étudiants réalisent leur projet de recherche aux cycles supérieurs sur ce thème à l'UQAM. Voilà la liste de quelques travaux en cours ou en préparation :

1. l'étude fine des comportements de sous-classement volontaire dans le but de permettre des simulations plus réalistes de ceux-ci et d'élaborer des indices spécifiques à ces comportements ;
2. l'extension des fonctionnalités de WebDétecte aux versions courtes du TCALS II ;
3. l'étude des stratégies à utiliser par les professeurs de langues à la suite de la détection des étudiants ;
4. le développement d'une version informatisée et adaptative du TCALS II ;
5. le développement d'une modélisation qui permettrait d'obtenir une estimation rapprochée du niveau d'habileté de l'étudiant, malgré le fait que son patron de réponse soit inapproprié ;
6. le développement d'une adaptation en anglais du TCALS II.

BIBLIOGRAPHIE

CIZEK, G. J., *Cheating on Test: How to do it, Detect it, and Prevent it*, Mahwah, Laurence Erlbaum Associates, 1999.

DRASGOW, F., M.V. LEVINE et M.J. ZICKAR, « Optimal Identification of Mismeasured Individuals », *Applied Measurement in Education*, 9(1), 1996, p. 47-64.

FOURNIER, P., *Pour un test incontestable: rapport de recherche sur les tests de classement en anglais (langue seconde) au collégial*, Québec, Ministère de l'Éducation, Direction générale de l'enseignement collégial, 1992.

HAMBLETON, R.K., H. SWAMINATHAN et H. J. ROGERS, *Fundamentals of Item Response Theory*, Newbury Park, Sage, 1991.

LAURIER, M., L. FROIO, C. PAERO et M. FOURNIER, *Test de classement d'anglais langue seconde au collégial*, Montréal, Collège de Maisonneuve, 1998.

MEIJER, R.R. et K. SIJTSMA, « Methodology Review: Evaluating Person Fit », *Applied Psychological Measurement*, 25(2), 2001, p. 107-135.

RAÎCHE, G., *Le sous-classement des élèves aux tests de classement en anglais langue seconde au Collège de l'Outaouais*, Gatineau, Collège de l'Outaouais, 2000.

RAÎCHE, G., « Pour dépister les étudiants qui cherchent à sous-performer au test de classement en anglais langue seconde dans le réseau collégial québécois », *Bulletin de l'Association pour la recherche au collégial*, 15(3), 2002a, p. 8-9.

RAÎCHE, G., *Rater un examen, c'est payant ! Le dépistage du sous-classement au test de classement en anglais, langue seconde, au collégial*, Gatineau, Collège de l'Outaouais, 2002b.

RAÎCHE, G. et J.-G. BLAIS, *The Distribution of Person-fit Indices Conditional on the Estimated Proficiency Level and the Detection of Underachievement at a Placement Test*, Communication présentée au 68^e congrès annuel de la Psychometric Society tenu à Caliri, Sardaigne, 2003.

REISE, S. P. et W. P. FLANNEREY, « Assessing Person-fit on Measures of Typical Performance » *Applied Measurement in Education*, 9(1), 1996, p. 9-26.