

*Diagnostic cognitif et métacognitif
au seuil de l'université*

**Les Éditions de l'Université
de Liège**

31, Boulevard Frère-Orban
4000 LIÈGE (Belgique)

Courriel : editions.univ.liege@skynet.be

Tél. (inter + 32) - (0)4 - 254 37 98

Fax (inter + 32) - (0)4 - 254 37 98

© 2003



Tous droits de reproduction,
d'adaptation et de traduction
réservés pour tous pays.

Imprimé en Belgique.

Couverture : Roland Manigart (LabSET – ULg)

Mise en page : Nelly Saenen (STE-ULg)

Colette Pruniaux-Degotte (Cefal)

D/2003/8886/7

ISBN 2-930322-61-6

Coordinateur : D. LECLERCQ

*Diagnostic cognitif et métacognitif
au seuil de l'université*

**Le projet MOHICAN
mené par les 9 universités
de la Communauté Française Wallonie Bruxelles**

Le terme MOHICAN explicité

Au sens strict, MOHICAN est l'acronyme de trois mots : Monitoring (observations, recueil d'indices), Historique (suivi d'une cohorte pendant un an) et CANDidature (4000 étudiants entrant dans les universités de la Communauté française de Belgique). Nous utilisons l'expression « le premier » pour exprimer le vœu que cette opération soit suivie par d'autres, de la même inspiration, modifiées à la lumière des enseignements tirés de la présente recherche.

Ce terme se veut aussi métaphorique. Il évoque l'observation bienveillante, par des scrutateurs (les Mohicans) réputés pour leur capacité d'interpréter des indices subtils, d'un phénomène qui se répète année après année. Il s'agit de la transition (de la migration) de milliers d'étudiants de l'enseignement secondaire vers l'université, un défi difficile à relever pour la majorité puisque l'on sait que 40% seulement atteindront le but en un an. La première année constitue une sorte de désert à traverser, et l'on sait que ceux qui disposent de plus de « réserves » dès le départ ont plus de chances que les autres d'y arriver.

Conscient de ce péril, l'observateur mohican n'a pas pour objectif de voler à la place des oiseaux migrateurs, ce qui serait impossible et inutile. Il veille cependant à ménager, dans ce pays aride qu'est la première année de candidature, des possibilités d'escalas où ceux qui en ont besoin trouvent les ressources (l'eau et la nourriture) pour poursuivre leur voyage. Encore faut-il qu'ils soient conscients de la nécessité de faire escale (information précoce et diagnostique) et aient le courage de passer du temps à « faire le plein ». A chacun ses responsabilités...

Pourquoi une image de sextant ?

Ce sextant sur la page de couverture veut symboliser le type d'évaluation, de mesure qui a été faite dans ce projet MOHICAN. En marine en effet, le sextant sert au navigateur à déterminer sa position lors d'un voyage dont il connaît avec précision la destination. Et cette position va, très normalement, évoluer. Il s'agit de « faire le point », pour le dépasser, pour avancer et non pas pour « fixer » une valeur des capacités de chaque personne.

C'est ce qui a été voulu dans la recherche MOHICAN : faire avec chaque étudiant « le point » sur plusieurs matières.

Le présent livre, lui, est destiné à fournir des données plus globalisées, et non des résultats de personnes ou de sections. A ce niveau global, c'est le pays qui dispose de données fournies par le même sextant, pour se positionner.

La mise en page est de **Nelly Saenen**, secrétaire au STE-ULg et au LabSET-ULg. Nous la remercions pour la quantité et la qualité de son travail.

Les données MOHICAN initiales ont été recueillies par formulaires de lecture optique de marques grâce au **SMART-ULg**, le Système d'Aide à la Réalisation de Tests, dirigé par Jean Luc Gilles (jl.gilles@ulg.ac.be).

Ce projet a été possible grâce à une convention de recherche confiée au CIUF (à son Groupe de Travail « Réussites en Candidatures ») par le Ministère de l'Enseignement Supérieur, de l'Enseignement de Promotion sociale et de la Recherche scientifique.

Préface

Le faible taux de réussite des étudiants de première candidature ne laisse personne indifférent. Il était légitime que ceux qui ont pour mission de former ces étudiants se soucient de cette situation.

Des études ont déjà été consacrées à cette question et le lien entre certaines variables et la réussite des étudiants a été établi de façon quantitative. Intégrant ces variables prédictives, un modèle global a été élaboré. L'apport de MOHICAN à ce modèle consiste à lui injecter des variables complémentaires comme l'origine socio-culturelle des étudiants, leurs préacquis, leurs méthodes d'apprentissage et leurs efforts individuels.

Le travail qui est développé dans cet ouvrage revêt encore un autre intérêt : il a testé les populations d'étudiants de toutes les institutions universitaires de la Communauté française avec les mêmes épreuves.

L'étude a encore été poussée plus loin. Il convenait en effet de dépasser le stade de la prédiction en fournissant, aux étudiants comme aux enseignants universitaires, des informations susceptibles d'améliorer les chances de réussite en première candidature. Pour moi, cet aspect est essentiel. La prédiction ne doit pas être mise en œuvre pour soutenir ou orienter une quelconque sélection, elle doit viser la guidance et encourager la réussite des étudiants.

Le problème de la réussite en candidature nous préoccupe tous. L'étude MOHICAN apporte un éclairage utile sur certaines facettes à ce problème. A ce titre, elle constitue une étape intéressante dans la recherche des solutions.

La route risque d'être encore longue avant d'atteindre des résultats satisfaisants. Raison de plus pour continuer à chercher. En cette matière, comme en d'autres, le chemin se trace en marchant.

*Françoise DUPUIS
Ministre de l'Enseignement supérieur,
de l'Enseignement de promotion sociale
et de la Recherche scientifique*

Le premier des MOHICANS

Une recherche-action de Monitoring Historique des Candidatures

D. Leclercq (Ed.)

SOMMAIRE

Partie I : Vues transversales

Introduction	13
Le projet MOHICAN – Check-up : son origine, ses intentions, ses méthodes	
<i>D. Leclercq, C. Conti, J.-M. De Ketele, M. Delhaxhe, P. Dupont, J.-P. Lambert, J.-P. Lambotte, B. Noël, M. Romainville et J.-L. Wolfs</i>	
Chapitre premier	33
Connaissance partielle, analyse spectrale et métacognition	
<i>D. Leclercq</i>	
Chapitre 2	51
Les feedbacks des check-up MOHICAN aux étudiants et aux enseignants	
<i>D. Leclercq, F. Georges, J.-L. Gilles, M. Pirson et F. Simon</i>	
Chapitre 3	67
Résultats check-up par check-up et corrélations entre eux	
<i>D. Leclercq, F. Georges, C. Dupont et P. Detroz</i>	
Chapitre 4	105
Incorporation des feedbacks MOHICAN dans les interviews individuels	
<i>B. Noël et S. Goemare</i>	
Chapitre 5	109
La consultation du feedback MOHICAN : Fréquences, intérêts et implications	
<i>M. Delforge, B. Diricq et P. Dupont</i>	
Chapitre 6	117
L'information des feedbacks pour les professeurs. L'opération VOCA PROFS à l'ULg ou Estimation par les professeurs des résultats des étudiants au check-up de Vocabulaire	
<i>D. Leclercq, J.-L. Gilles et C. Dupont</i>	

Chapitre 7	127
Liens entre caractéristiques de départ (dont les résultats aux check-up) et les réussites en 1 ^{re} candidature	
<i>P. Detroz et D. Leclercq</i>	

Chapitre 8	155
Changer de méthode d'étude suite aux feedbacks : l'opération RESSAC (Résultats d'Épreuves Standardisées au Service de l'Apprentissage en Candidatures)	
<i>D. Leclercq, P. Detroz, C. Dupont et J.-L. Gilles</i>	

Partie II : L'analyse check-up par check-up

Chapitre 9	173
Analyses psychométriques des questions des 10 check-up MOHICAN. Vue d'ensemble	
<i>D. Leclercq et J.-L. Gilles</i>	

Chapitre 10	181
Analyses éduométriques et indices métacognitifs appliqués aux questions des 10 check-up MOHICAN	
<i>D. Leclercq et M. Poumay</i>	

Chapitre 11	191
Vocabulaire	
<i>M. Montballin et D. Leclercq</i>	

Chapitre 12	205
Syntaxe	
<i>J.-M. Defays et D. Leclercq</i>	

Chapitre 13	213
Compréhension de textes	
<i>Ph. Hougardy et D. Leclercq</i>	

Chapitre 14	217
Géographie et lecture de graphiques, cartes et tableaux	
<i>F. Orban et D. Leclercq</i>	

Chapitre 15	227
Mathématiques	
<i>M. Lebrun, J. Lega et D. Leclercq</i>	
Chapitre 16	239
Physique	
<i>P. Chapelle et D. Leclercq</i>	
Chapitre 17	245
Chimie	
<i>A. Cornelis et D. Leclercq</i>	
Chapitre 18	251
Biologie	
<i>J.-C. Verhaege et D. Leclercq</i>	
Chapitre 19	257
Connaissances en arts et culture	
<i>D. Leclercq et A. Lamblin</i>	
Chapitre 20	269
Connaissances générales en histoire, Actualité, Economie	
<i>D. Leclercq et F. Georges</i>	

Partie III : Synthèse et Conclusions

Chapitre 21	281
Synthèse des résultats	
<i>D. Leclercq</i>	
Chapitre 22	309
Conclusions générales et perspectives sur les rapports entre MOHICAN et la réussite en 1 ^{re} année	
<i>D. Leclercq</i>	

Introduction

Le projet MOHICAN – Chek-up : son origine, ses intentions, ses méthodes

*D. Leclercq (ULg), C. Conti (FPMs), J.-M. Deketele (UCL), M. Delhaxhe (ULg),
P. Dupont (UMH), J.-P. Lambert (FUSL), J.-P. Lambotte (FUSAGx),
B. Noel (FUCAM), M. Romainville (FUNDP), J.-L. Wolfs (ULB).*

Introduction

Sommes-nous vraiment entrés dans la « société de la connaissance » ? Qu'elle existe ne semble plus faire de doute tant, dans notre société tertiaisée, la communication et l'information jouent d'importants rôles économiques et sociaux.

L'université est au cœur de ce mouvement, comme lieu de production, de transmission et de critique de la connaissance.

La communauté universitaire que constituent les enseignants et les étudiants s'est-elle dotée des concepts, des procédures et des moyens adéquats et suffisants pour mener à bien les processus d'enseignement et d'apprentissage, c'est-à-dire pour atteindre des objectifs de réussite, dans leurs aspects qualitatif et quantitatif ? Pas suffisamment encore. La présente étude veut contribuer à cette instrumentation essentiellement au niveau conceptuel. On constatera, à côté de la part consacrée aux données objectives, une part aussi importante accordée à la subjectivité des étudiants dans leur regard sur leurs propres connaissances. Cette importance se justifie à nos yeux par le fait que c'est aux étudiants qu'il incombe d'apprendre et de prendre les décisions dans ce sens, sur base de leur auto-cognition. Et ce d'autant plus que leur cursus universitaire et leur formation continuée ultérieure tout au long de leur vie professionnelle exigeront d'eux de plus en plus de servuction (le vocable francophone pour « self-service ») : leur prise en charge de leur propre formation. Enfin, de nombreuses recherches¹ montrent que, parmi les facteurs cognitifs et affectifs qui influencent l'apprentissage, « c'est la métacognition² qui est la plus fortement pondérée ».

La présente étude comporte des facettes multiples susceptibles d'intéresser diversement les lecteurs. Certains s'intéresseront à la façon dont les check-up ont été conçus, d'autres, à la façon dont les résultats sont présentés, d'autres aux résultats eux-mêmes comme points de repères, d'autres encore à la façon dont les résultats ont été communiqués, consultés, exploités... et enfin à l'impact sur la réussite que des procédures ont pu avoir.

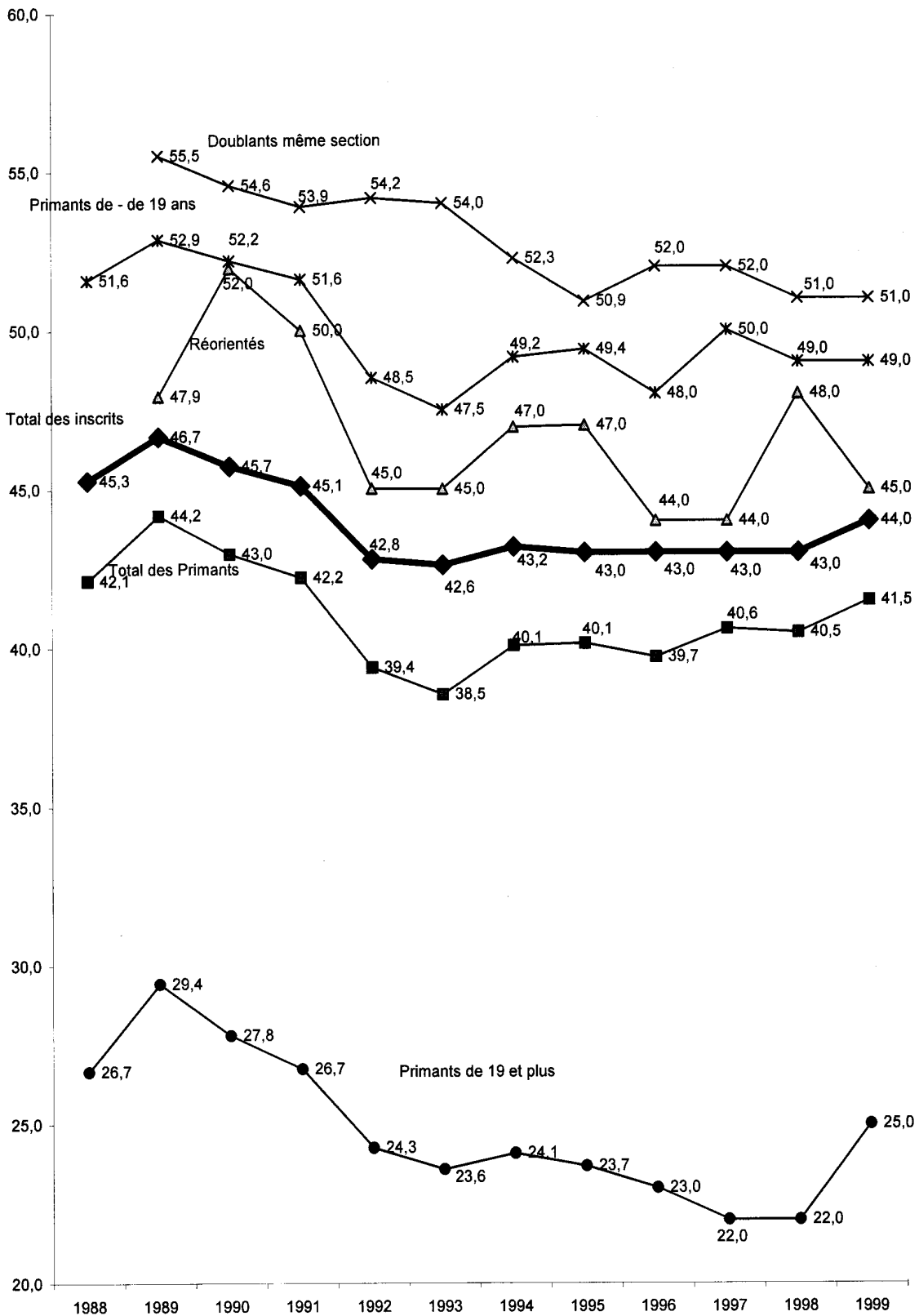
Tous ces points ne sont pas traités de façon égale tant les données et les sites de l'opération MOHICAN étaient multiples. Certains de ces points sont illustrés par un seul cas. C'est que, malgré l'abondance des chiffres, ce sont surtout des questions qualitatives qui sont abordées ; avec des données pour nourrir et illustrer le débat plutôt que le conclure. Nous avons voulu fournir des repères tout en restant conscients qu'ils sont marqués non seulement par l'époque, mais aussi par le public, par les méthodes et par les contenus.

Ceux qui souhaiteraient s'en tenir à cette approche qualitative pourraient lire, dans un premier temps, la présente introduction (le projet MOHICAN, ses motivations et ses méthodes), le chapitre 1 (L'analyse spectrale), les check-up de leur choix (dans la Partie II), et enfin les chapitres 21 (Synthèse) et 22 (Conclusions et perspectives).

¹ Tardif (1992, p. 46) cite l'analyse de Wang, Haertel et Walberg (1990), qui, pour aboutir à cette conclusion, ont étudié 150 publications traitant de cette question.

² En bref « Connaissance de sa connaissance », mais voir une définition plus précise dans le chapitre 1 (Glossaire).

Taux de réussite en première candidature



1. Les motivations du projet MOHICAN

1.1 La mission du Groupe de travail « Réussite » du CIUF

En 1991, un Groupe de travail « Réussite en candidatures » a été instauré par le CIUF (Conseil Interuniversitaire de la Communauté Française), sous la présidence de D. Leclercq (ULg), et composé d'un représentant par université. Actuellement les membres sont J.-M. Deketele (UCL), M. Delhaxhe (ULg), P. Dupont (UMH), C. Conti (FPMs), J.-P. Lambert (FUSL), J.-P. Lambotte (FUSAGx), B. Noel (FUCAM), M. Romainville (FUNDP), J.-L. Wolfs (ULB). La mission de ce groupe est d'étudier le problème du trop faible taux de réussite en première candidature (aux alentours de 40% pour les primants et de 50% pour les répétants) et de faire des propositions pour améliorer la situation. Ce groupe a produit deux rapports³ avant la présente étude.

1.2 La confirmation de la longue « stagnation » des taux de réussite

Le graphique ci-contre montre que les taux de réussite en première année des universités sont très stables depuis plus de dix ans et tournent autour des valeurs moyennes approximatives suivantes :

- 40% pour les primants,
- 50% pour les doublants de même section ;
- 45% pour tous les étudiants confondus ;

- 25% pour les 19 ans et plus ;
- 50% pour les 18 ans et moins

Ces données doivent être intégrées dans un **modèle global simplifié**. Nous proposons ce modèle en section 3 ci-après.

Ensuite, en section 4, nous présenterons une synthèse du modèle prédictif à partir des variables étudiées par la recherche « La population étudiante » coordonnée par Droesbeke et al. (2001).

Enfin, en section 5, nous présenterons les apports de MOHICAN à ce modèle.

³ Leclercq et al., (1997). *Franchir le cap des candis*, Bruxelles : CIUF.
Boxus et al., (1993). *Statistiques de réussites et Innovations pédagogiques*, Bruxelles, CIUF.

1.3 Que penser de la stagnation des Taux de Réussite depuis 20 ans ?

Pour répondre à cette question, il nous paraît nécessaire de faire des hypothèses (dans la colonne de droite) sur l'évolution de 8 variables (dans la colonne de gauche) relevant de la Société (S), des Antécédents des étudiants (A), des universités (U) et des performances des étudiants (E) à l'entrée et à la sortie de la première année universitaire.

Selon nos hypothèses, au cours de ces 20 dernières années,

... en ce qui concerne...	... il s'est produit...
(S1) ... les conditions socio-économico-culturelles globales,	<ul style="list-style-type: none"> – ... une diminution du taux d'étudiants financièrement à l'aise pendant leurs études⁴ – ... une diminution de la confiance des étudiants dans les études comme assurance d'obtenir un emploi (les 30 glorieuses sont derrière nous) – ... une diminution de l'intérêt pour certaines études (les sciences par exemple, dont la situation actuelle est décrite dans l'Eurobaromètre 2001) – ... une diminution de l'homogénéité de la formation dans le secondaire suite à une diversification des politiques de l'enseignement selon les réseaux et les écoles
(S2) ... la diversité et le nombre des étudiants entrants,	... une diminution de l'homogénéité des entrants, pour la raison ci-dessus et par la loi d'omnivalence des diplômes
(S3) ... les conditions matérielles de travail,	... une diminution du financement par étudiant (l'allocation de fonctionnement des universités est restée stable depuis 1972 alors que dans le même temps, le nombre d'étudiants a été affecté d'un facteur 1,5).
(A1,A2,A3) ... les Antécédents des étudiants ⁵ ,	... une augmentation du taux de filles inscrites à l'université
(U1) ... l'existence d'un examen d'entrée en Sciences Appliquées,	... un statu quo
(U2) ... les exigences de sortie universitaires en 1 ^{re} année	... un statu quo
(E1) ... l'existence d'un taux de réussite supérieur aux primants chez les doublants et les réorientés ⁶ ,	... un statu quo
(E2) ... les compétences des étudiants à l'entrée,	... une controverse sans données objectives permettant de trancher. En effet, certains pensent «ils ne savent plus rien de ce qui est essentiel» ; d'autres «ils savent autre chose, mais que l'université ne reconnaît pas», d'autres encore «ce qu'il savent maintenant les prépare mieux à réussir».

Etant donné les hypothèses ci-dessus (d'une évolution défavorable dans les facteurs S ou « sociétales »), **notre réponse** à la question posée dans le titre « *Que penser de la stagnation des Taux de Réussite depuis 20 ans ?* » était, au départ de la présente étude, « **les institutions et les encadrants universitaires ont accompli un tour de force en réussissant à maintenir un taux de réussite équivalent à celui des années précédentes, alors que tant de facteurs négatifs auraient dû, en toute logique, l'entraîner vers le bas** ».

Cependant, cette conclusion reste tout aussi hypothétique car nous n'avions pas la preuve que les exigences universitaires de sortie (U2) sont inchangées, et nous ne disposions pas d'une mesure de « l'état d'entrée » (E2). Nos successeurs, eux, auront au moins, sur quelques-uns de ces points, des repères grâce à MOHICAN.

⁴ Dans le rapport d'activités de l'AGP-SAE-CF (Adm. Gén. Personnel, Service des Allocations d'Etudes de la Communauté française) de 1999, alors que le taux de «cas» considérés comme «spéciaux» était de 3% en 1974 (p. 10), il est passé à 16% en 1999 (p.12). Dans l'étude de Geron et Vandekerke (Les conditions de vie des étudiants. Chiffres-clés et chiffres phares, CIUF, juin 1999), 12,8% des étudiants universitaires exercent une activité rémunérée régulière. D'autres (37%) ont aussi travaillé mais comme jobistes, tout spécialement durant l'été, les conditions légales permettant à l'employeur qui engage un jeune pendant 30 jours maximum de bénéficier d'un rabatement de salaire de 40%. Entre 1990 et 1999, les allocations d'études ont subi un plafonnement débouchant sur un déficit de financement de 17,5%. Au cours des années 2000 et 2001, la Ministre actuelle a effectué un rattrapage substantiel de deux fois 7%.

⁵ Dreesbeke et al, 2001

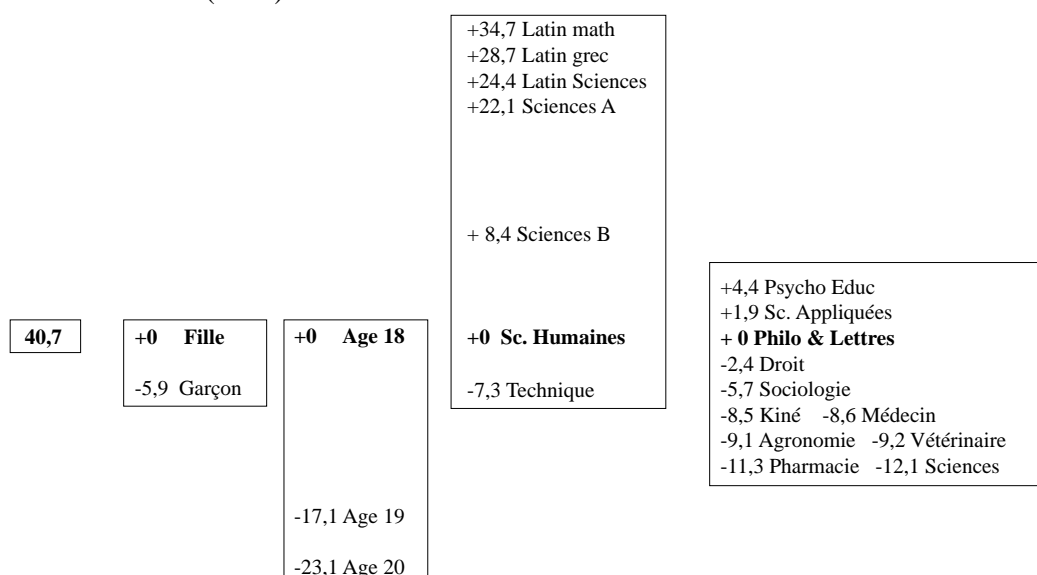
⁶ Leclercq et al., «Franchir le cap des candis», Bruxelles : CIUF, 1997.

1.4 Les variables prédictives de la réussite à l'Université

Dans la lutte contre l'échec scolaire par « la pédagogie de la maîtrise », Benjamin Bloom (1969) a distingué les variables non changeables des variables changeables par les enseignants. C'est à des variables de ce deuxième type que le projet MOHICAN est consacré. Néanmoins, nous commencerons par rappeler ce que l'on sait de variables moins changeables, du moins par les enseignants des universités.

On connaît assez bien l'évolution des taux de réussite au cours des dernières années (voir graphique 1). On connaît aussi les liens entre la réussite et certaines caractéristiques des étudiants. Dans leur ouvrage « La population étudiante », Dreesbeke et al. (2001, p. 201) établissent même, à partir de cohortes d'étudiants regroupant 41700 entrants depuis 1993-1994 (p. 194), une équation de prédiction par laquelle on peut estimer la probabilité de réussite (ou les pourcentages de chances) d'un (e) étudiant (e) quand on en connaît certaines caractéristiques.

L'équation part de la moyenne globale, qui est de 40,7% pour les primants de cette année-là, en lui ajoutant des chances (en %) ou en lui en retirant selon les variables.

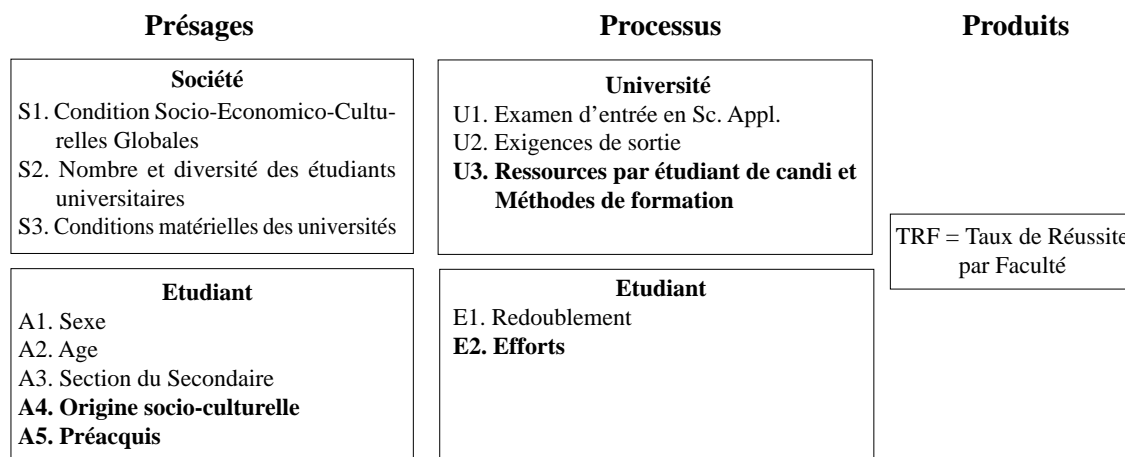


On peut donc calculer qu'une fille de 18 ans issue de Sciences humaines dans le secondaire, inscrite en Philo et lettres a 40,7% de chances de réussir (plus exactement, lors de l'année 1999-2000, 40,7% des personnes dans cette situation ont réussi). Partant de cette même moyenne globale (40,7%), on peut calculer les chances d'un garçon (-5,9%) de 19 ans (-17,1%) issu de Latin grec dans le secondaire (+28,7%) et inscrit en droit (-2,4%), c'est-à-dire 43,4%.

Les dés sont-ils jetés à ce point ? Les chances de succès sont-elles immuables étant donné les conditions de départ ? Nous savons qu'il n'en est rien et que d'autres variables, qui n'ont pas été prises en compte dans le modèle ci-dessus, pourraient jouer un rôle prédictif (explicatif) elles aussi. Ainsi (1) les stratégies, méthodes et ressources déployées par les formateurs, (2) des données socio-biographiques de l'étudiant telles que l'origine socio-culturelle ou les différences d'acquis à l'entrée, (3) la manière dont l'étudiant, une fois entré dans la formation universitaire, apprend et exerce son nouveau métier (Coulon, 1997). En particulier, Romainville (2000, 46-73) explique que ce métier exige que l'étudiant (1) réussisse son processus d'affiliation au monde universitaire ; (2) s'affilie aux sous-cultures de ce monde et notamment aux genres de textes qui caractérisent ces sous-cultures ; (3) développe une conception de l'apprentissage et des croyances épistémiques qui soient compatibles avec celles de ses enseignants-chercheurs ; (4) mette en œuvre une approche d'apprentissage en profondeur des cours ; (5) adopte une méthodologie du travail universitaire particulière, en cohérence avec les points précédents. D'où nos ajouts au modèle.

1.5 Les apports de MOHICAN au modèle global

Le modèle suivant organise les variables en 3 blocs, connus sous l'expression « les 3 P » : les Présages (connaissables avant l'entrée), les Processus (ce qui s'est VRAIMENT passé durant les études) et les Produits (les résultats). Les variables apportées par MOHICAN figurent en gras : U3, A4, A5, E2).



MOHICAN injecte dans le modèle les variables suivantes :

- **(A4) : origine socio-culturelle** : lors du check-up MOHICAN d'octobre 1999, le diplôme des parents a été recueilli et il sera possible de calculer leur liaison avec la réussite en première année de candidature (chap. 7B).
- **(A5) : les préacquis** : cette vue est utile non seulement pour juger de l'impact de ces préacquis sur la réussite (voir chapitre 7) mais pour disposer d'un point de repère pour les années ultérieures (à condition qu'il subsiste alors suffisamment d'acquis communs jugés intéressants).
- **(U3) : les méthodes de formation** : exploitant ces données. Une d'entre elles, appelée RES-SAC : Résultats d'Epreuves Standardisées au Service de l'Apprentissage en Candis a fait l'objet d'une analyse d'impact. Il s'agit de la communication des données à des fins formatives pour provoquer une auto-régulation par les étudiants de leurs méthodes d'apprentissages (voir chapitre 8).
- **(E2) : les efforts individuels des étudiants** car (le chapitre 8 le montre), si certaines initiatives de l'institution sont utiles pour obtenir une amélioration, elles ne sont pas suffisantes, dans la mesure où elles ne semblent avoir d'effet que chez une portion seulement des étudiants : ceux qui décident d'en tirer parti. Par ailleurs, nous sommes bien conscients du caractère modeste de telles approches au regard de la complexité des processus liés aux « méthodes de travail et stratégies d'apprentissage » des étudiants dont on trouvera un développement dans Wolfs (1998).

1.6 Commentaires sur chacune de ces variables prédictives

(A4) – L'origine socio-culturelle des étudiants

Cette variable n'est pas recueillie systématiquement dans toutes les universités de la Communauté française de Belgique et n'est pas intégrée dans la banque de données du Conseil des Recteurs Francophones (CreF).

Nous faisons l'hypothèse que les étudiants ont d'autant plus de probabilités de réussite que leurs parents possèdent des diplômes élevés (universitaires notamment). Une des variables causales pourrait être qu'en moyenne, les revenus de ces parents sont supérieurs, dispensant ainsi les étudiants de travailler parallèlement à leurs études, pour gagner leur vie (cf. étude de de Kerchove et Lambert, 1995, citée in Leclercq, 1998, p. 44). Une autre raison pourrait être une série de conditions plus favorables (disponibilité de ressources). Cette variable sera étudiée au chapitre 7.

(A5) – Les différences de préacquis

A égalité de sexe, d'âge, de filière du secondaire, de section d'inscription à l'université, bref toutes ces caractéristiques restant égales, les étudiants peuvent différer entre eux par les acquis antérieurs, et cette différence peut influencer (en plus ou en moins) la réussite en fin de première candidature. Nous faisons l'hypothèse que les types d'acquis qui sont les plus liés à la réussite diffèrent de section en section (par exemple, les compétences à l'entrée en Syntaxe risquent d'être plus liées à la réussite en première candidature en Philosophie et Lettres et en Droit qu'en première candidature en médecine). Ces liens seront étudiés au chapitre 7, sans perdre de vue que les check-up MOHICAN comportaient peu de questions portant sur les compétences transversales.

(U3) – Les méthodes des formateurs

Certains enseignants présentent à leurs étudiants dès le début de l'année, dans un « engagement pédagogique », leurs objectifs, méthodes, y compris d'évaluation ; d'autres non. Certains entraînent leurs étudiants au type d'examen qu'ils auront à affronter ; d'autres non. Certains organisent des interrogations « partielles » dispensatoires de matières ; d'autres non. Certains conçoivent des livres extrêmement clairs, illustrés, facilitant la compréhension ; d'autres non. Certains donnent des cours oraux très motivants, où les étudiants sont très participatifs⁷ ; d'autres non. Et ce qui vient d'être dit pour un (e) enseignant (e) peut l'être pour toute une section, toute une faculté, voire même toute une université. Cette variable n'est pas étudiée dans MOHICAN.

(E2) – Les efforts faits par chaque étudiant

Certains étudiants assistent aux cours oraux, d'autres non. Certains préparent à domicile ces cours oraux ; d'autres non. Certains les exploitent à domicile, en remettant de l'ordre dans leurs notes, sans délai. d'autres non. Certains lisent des documents supplémentaires, consultent leurs Unités de Documentation. d'autres non. Certains consacrent de nombreuses heures à l'étude (aspect quantitatif) et étudient en profondeur, ne mémorisant jamais du non-compris ; d'autres non. Certains posent des questions, participent aux « répétitions », se font conseiller par les experts des méthodes de travail mis à leur disposition par les universités ; d'autres non.

Nous faisons l'hypothèse que l'implication de l'étudiant dans ses études a un impact sur sa réussite. Cette variable est étudiée, hélas sur un seul cas (RESSAC) au chapitre 8.

Le projet MOHICAN tentera d'apporter des éclairages sur les liens entre certaines de ces variables et la réussite, mais de façon incomplète étant donné le caractère exploratoire de cette recherche pionnière puisqu'une telle entreprise (tester les étudiants entrant dans les universités avec les mêmes épreuves) n'avait jamais eu lieu en Belgique.

⁷ Leclercq, 1998, p. 166-167 conceptualise cette notion par le TOPACMER : le Taux d'Occupation et de Participation Active Complète Manifeste Exprimée et Répétée.

2. Le projet

Le projet MOHICAN (MONitoring HIstorique des CANDidatures) s'est donné pour **objectifs** de

- rassembler des données fiables (dépassant les on-dits et les rumeurs) ;
- fournir concrètement et de façon précoce aux étudiants d'abord, aux encadrants ensuite, des informations à visée formative (utilisables en vue d'améliorer les chances de réussite).

Parce que des impératifs de faisabilité ont prévalu sur d'autres considérations, l'ensemble des épreuves CHECK-UP MOHICAN

- **n'est pas un relevé exhaustif** des compétences jugées pertinentes à l'entrée de diverses actions. Par exemple, sont absentes des matières spécifiques telles que la statistique, l'orthographe, la littérature, et surtout des compétences transversales telles que l'esprit de synthèse, la prise de notes, l'esprit critique, la capacité d'apprendre, etc.
- **n'est pas un bilan des acquis de l'enseignement secondaire**. Celui-ci a ses objectifs propres qui viennent d'ailleurs d'être spécifiés sous forme de « compétences terminales ». Ces spécifications n'étaient pas disponibles au moment de la conception des check-up MOHICAN ;
- **n'est pas une série des épreuves parfaitement ciblées** pour des sections particulières puisque ces épreuves sont les mêmes pour tous : l'épreuve de mathématique est la même pour les ingénieurs que pour les autres sections ; les épreuves de français sont les mêmes pour les romanistes que pour les autres étudiants ; l'épreuve de vocabulaire n'est pas « orientée » vers chaque discipline.

Les épreuves CHECK-UP MOHICAN

- **sont dix épreuves séparées**, conçues par dix auteurs (ou groupes d'auteurs) différents, issus des 9 universités de la Communauté, sur la base de leurs observations (dans leurs institutions) des difficultés rencontrées par leurs étudiants de *1^{re} candidature*. Souvent, ces observations sont étayées par des résultats à des épreuves ou par des enquêtes auprès des enseignants ;
- **sont des « coups de sonde »** dans un univers de connaissance bien plus large. C'est tout spécialement le cas du vocabulaire (seulement 45 questions !), les connaissances artistiques, les repères historiques, socio-économiques, l'actualité.
- des épreuves donnant, plus qu'une image des connaissances des étudiants, *une image de leur regard sur leurs connaissances* : surestimations, sous-estimations, confiance, prudence et réalisme.
- **sont** des épreuves dont les résultats s'expriment en concepts et *indices nouveaux* pour refléter soit des états de connaissance des étudiants, soit des propriétés des questions et des check-up entiers.

Des épreuves check-up MOHICAN *ultérieures* **pourraient / devraient** :

- être une reconduction de la présente opération, mais **ciblée** pour des sections particulières
- être le résultat d'une **concertation avec l'enseignement secondaire**, sur base des compétences terminales maintenant disponibles. Ce dialogue avait été amorcé par la participation de l'inspection de l'enseignement secondaire à deux moments-clés du projet MOHICAN. Le contexte a changé. D'une part, le secondaire a maintenant de ses compétences terminales ; d'autre part, l'université dispose des résultats « MOHICAN ». Le dialogue entre ces deux instances devrait être systématisé pour optimiser la transition entre les deux créneaux d'enseignement, notamment via...
 - une *définition* par l'enseignement universitaire, outre des compétences spécifiques visées à chaque faculté et chaque section, des *compétences transversales qu'il vise à*

développer au cours du premier cycle (les candis), et ce dans la perspective du premier cycle universitaire en trois ans défini par l'accord de Bologne ;

- une *enquête systématique sur les besoins ressentis* en candidatures, dans diverses sections.
- tenir compte des **renouveaux pédagogiques en cours** dans l'enseignement universitaire
- approfondir l'étude de l'impact des **différences socio-culturelles**.
- porter sur les langues **étrangères** et sur des **habiletés « transversales »** telles que la capacité d'apprendre.

La présente opération comporte donc bien des lacunes qui devraient être comblées par d'autres opérations. Il ne s'agit donc, dans notre esprit, que du premier des MOHICANs à venir. L'opération MOHICAN apporte des réponses à une série de questions telles que sa faisabilité (son acceptabilité, ses délais de mise en œuvre, son respect de la confidentialité promise), de la valeur informative des feedbacks pour les enseignants et les étudiants, de la consultation par les étudiants de leurs feedbacks personnalisés, de l'exploitation institutionnelle à des fins d'amélioration et l'impact de ces interventions. Ces questions font l'objet de chapitres spécifiques et les réponses sont synthétisées au chapitre 21.

3. L'engagement des membres du comité de pilotage

Suite aux analyses qu'il a présentées dans son rapport *Franchir le cap des candis* (CIUF, Bruxelles, 1997), le Groupe de Travail « Réussite en candis » considère, dans le débat actuel sur les limitations à l'entrée et au doublement, que

- Une partie des dés ont déjà été jetés par l'origine socio-économique familiale et par le cursus du secondaire.
- La responsabilité des étudiants en est atténuée dans la mesure où une part importante de leur parcours leur a échappé.
- En outre, même leurs choix libres étaient des choix faits alors qu'ils étaient mineurs.

Les membres du Groupe de Travail pilotant la recherche MOHICAN partagent les conclusions suivantes de de KERCHOVE et LAMBERT (1996)⁸ :

- (1) Toute mesure radicale (telle que le numerus clausus ou l'examen d'entrée)... ne ferait que consacrer les handicaps socio-culturels de départ.
- (6) Les étudiants abordent l'enseignement supérieur avec des chances très inégales selon leur capital socio-culturel de départ... Une certaine période d'adaptation et d'orientation à l'entrée de l'enseignement supérieur apparaît donc nécessaire pour une partie importante des étudiants sous peine d'accentuer encore l'« écrémage social » et de réduire l'efficacité du système global en termes d'allocation optimale selon les aptitudes de chacun.
- (9) Un renforcement des conditions d'accueil (évaluation, remédiation, préparation à une réorientation) en 1^{re} année peut être considéré, de ce point de vue, comme un « investissement » hautement rentable pour la collectivité.

L'opération CHECK-UP a pour but de fournir des REPERES aux étudiants et aux institutions dans une perspective formative. Le groupe s'opposera à toute dérive qui aboutirait à donner à ce dispositif et ces instruments une visée sélective.

Toute la démarche ne prend son sens que si elle est suivie de remédiation mise en œuvre par les institutions et fréquentée par les étudiants.

⁸ Les numéros (1, 6 et 9) dans le cadre sont ceux de la publication originale.

4. Phases du projet MOHICAN

4.1 Procédure

Phase 1 : la réalisation d'épreuves « CHECK-UP » dans 10 domaines

Sur base d'objectifs précisés pour chacun des check-up, ces épreuves ont été créées par des experts des 9 universités et revues par le groupe de travail. Deux séances-clés ont été organisées. L'inspection de l'enseignement secondaire et les délégués des étudiants du supérieur ont été conviés à la première (décembre 1998 au CIUF) au cours de laquelle ont été exposées les actions de ce type, moins amples et éparées, menées par diverses universités. Les représentants de l'Inspection de l'enseignement secondaire ont aussi participé à la seconde séance de travail (mai 1999 à Liège) consacrée à l'examen des épreuves check-up avant leur utilisation.

4.2 Format des questions

Les épreuves étaient composées⁹ de QCM à 5 solutions et, le plus souvent deux solutions supplémentaires « Aucune » et « Toutes » dites aussi « Solutions Générales » (ou SG) car elles concernaient *toutes* les questions de l'épreuve. Pour 5 des dix épreuves, certaines des réponses correctes étaient en effet « Aucune » ou « Toutes ». Ces épreuves sont marquées du signe *.

4.3 Modalités de réponses

Les étudiants devaient accompagner leur réponse par l'un des 6 degrés de certitude suivants :

0%, 20%, 40%, 60%, 80%, 100%.

Il s'agit, bien évidemment, par cette méthode

1. d'obtenir des résultats plus nuancés, par exemple de distinguer les connaissances
 - a) *utilisables* (réponses correctes avec un Degré de Certitude supérieur à 50 %);
 - b) *dangereuses* (réponses incorrectes avec un Degré de Certitude supérieur à 50 %);
 - c) *inutilisables* (réponses correctes ou non avec un Degré de Certitude inférieur à 50 %).
2. de donner aux étudiants une idée de leur réalisme dans l'appréciation de leur degré de qualité de leur réponse (se sous-estiment-ils ? Se surestiment-ils ?), c'est-à-dire un aspect de leur métacognition.

Epreuve		Auteur (s)	Nb Q		Destinés aux étudiants de
Vocabulaire	*	M. Monballin	45	73	Toutes les facultés
Syntaxe et logique	*	J.M. Defays	12		Toutes les facultés
Compréhension de texte	*	Ph. Hougardy	6		Toutes les facultés
Lecture de cartes, graphiques, tableaux en géographie		F. Orban	10		Toutes les facultés
Mathématique		M. Lebrun & J. Legas	22	50	Sciences, médecine, ingénieurs
Physique		P. Chapelle	10		Sciences, médecine, ingénieurs
Chimie		A. Cornélis	8		Sciences, médecine, ingénieurs
Biologie		J.C. Verhaeghe	10		Sciences, médecine, ingénieurs
Connaissances artistiques	*	D. Leclercq & A. Lamblin	25	50	Sciences humaines
Repères historiques, économiques, actualité	*	D. Leclercq & F. Georges	25		Sciences humaines

⁹ A quelques exceptions près (16 sur 45 en Vocabulaire, 2 sur 10 en Géographie, 2 sur 22 en mathématiques, 3 sur 10 en Biologie, soit 23 exceptions sur 173 questions.

5. L'administration des check-up

Les check-up ont tous été administrés la même semaine entre le 4 et le 8 octobre 1999, sur une durée de trois heures, via des formuLOMs (formulaires destinés à la Lecture Optique de Marques), dans 8 des 9 universités de la Communauté française.

Ces universités sont :

UCL = Université catholique de Louvain

ULg = Université de Liège

FUNDP = Facultés Universitaires Notre Dame de la Paix à Namur

UMH = Université de Mons-Hainaut

FUCAM = Facultés Universitaires Catholiques de Mons

FUSL = Facultés Universitaires Saint Louis à Bruxelles

FPMs = Facultés polytechniques de Mons

FUSAGx = Facultés Universitaires Agronomiques de Gembloux.

L'ULB, Université Libre de Bruxelles, qui administrait une telle épreuve (comme d'ailleurs la plupart des institutions universitaires), depuis plusieurs années, a préféré ne pas proposer à ses étudiants l'épreuve MOHICAN en supplément de celle qu'elle administrait par ailleurs. Les représentants de l'ULB ont cependant participé de bout en bout au pilotage de l'opération MOHICAN.

Le plan **d'échantillonnage** des étudiants testés figure sur la page qui suit. On voit que tous les étudiants n'ont pas été testés, mais des étudiants d'une même section (ex. : kiné) ont été testés dans plus d'une université, afin de permettre l'élaboration des résultats globaux par section sans possibilité d'identification des universités concernées conformément à la tradition du CIUF.

Un code chiffré a été attribué (1 = Philo... 9 = Agronomie) aux universités. Les sections à l'intérieur des facultés ont reçu un code-lettre (en Faculté des Sciences, A pour math, B pour physique, etc.). Lorsque la faculté ne comporte pas de section en candidature, le code est 0.

Tous les étudiants ne se sont pas identifiés quant à leur section dans une faculté ; nous leur avons attribué le code X.

La reproduction des 10 questionnaires et des formuloms (pour les 4000 étudiants), la lecture optique de ceux-ci une fois remplis, la correction informatique des réponses, la production de feedbacks (un par épreuve) pour chacun des 400 étudiants, pour chacune des 51 questions, ont été réalisés par le SMART-ULg (Système Méthodologique d'Aide à la Réalisation de Tests à l'Université de Liège), sous la coordination de son directeur, J.-L. Gilles.

Université et code	Faculté et code	Section et code	Voc.	Synt.	Comp.	Lect. Doc géo	Hist. Soc. Eco	Arts	Math	Bio	Chim	Phys.	
TOTAUX			3905	3790	3469	3738	1418	1399	2539	2526	2522	2520	
UCL 1	Philo 1	A (Histoire)	42	41	37	39	42	42					
		B (Classique)	6	6	6	6	6	6					
		C (Romane)	95	91	78	84	94	94					
		D (Germanique)	38	34	30	30	37	37					
		E (Histoire de l'Art et Archéologie)	57	52	47	50	55	54					
	Droit 2	A (Droit)	184	187	186	186	192	185					
		B (Sciences Politiques)	11	11	11	11	15	15					
		X	56	43	42	42	40	40					
	Sciences 5	A (Mathématiques)	35	33	33	32				37	37	36	37
		B (Physique)	35	36	35	36				33	33	33	33
		C (Chimie)	27	28	28	28				24	24	24	24
		D (Biologie)	54	58	58	58				60	59	60	59
		F (Géographie)	9	11	11	11				12	12	12	12
		G (Vétérinaire)	82	77	77	77				78	77	77	77
		X	88	98	97	98				99	98	98	98
	Médecine 7	A (Médecine)	211	188	184	185				212	211	211	210
		B (Biomédicale)	27	21	20	19				28	28	28	28
C (Sciences dentaires)		43	41	40	39				43	43	43	43	
D (Pharmacie)		80	82	80	80				84	84	83	83	
E (Education Physique)		52	49	49	49				53	52	51	52	
F (Kinésithérapie)		121	120	120	120				121	121	121	119	
G (Santé Publique)		3	2	2	2				2	2	2	2	
X		57	66	62	62				55	54	53	54	
Agronomie 9	0	131	124	124	124				117	117	117	117	
ULG 3	Droit 2	A (Droit)	191	174	30	169	162	162					
		B (Sciences Politiques)	58	43	6	41	8	8					
		X	58	66	6	63	38	38					
	Psycho 3	0	241	229	229	229	228	227					
	Sc. Appl. 6	A (Ingénieur civil)	144	148	148	148				98	97	98	98
		B (Architecture)	8	9	9	9				3	3	3	3
		X	92	80	80	79				112	112	112	112
	Médecine 7	A (Médecine)	160	154	154	154				157	158	158	158
		B (Biomédicale)	38	39	38	39				37	37	37	37
		C (Sciences dentaires)	12	8	8	8				13	13	13	13
D (Pharmacie)		78	73	72	73				77	77	77	77	
E (Education Physique)		2	2	2	2				2	2	2	2	
F (Kinésithérapie)		14	11	11	11				15	15	15	15	
X		9	13	13	13				7	7	7	7	
Vété 8	0	277	269	259	264				264	264	264	264	
FUNDP 4	Sciences 5	A (Mathématiques)							40	40	40	40	
		B (Physique)							12	12	12	12	
		C (Chimie)							31	29	31	31	
		D (Biologie)							49	48	46	44	
		E (Sciences géologiques)							11	11	11	11	
		F (Géographie)							16	15	14	15	
		H (Pharmacie)							69	67	67	67	
		X							19	18	18	18	
UMH 5	Psycho 3	0	145	144	144	144	144	137					
FUCAM 6	Eco 4	B (Sciences de gestion)	149	142	137	142	146	145					
		D (Sciences Politiques)	32	30	29	30	29	28					
FUSL 7	Eco 4	A (Sciences économiques)	23	24	24	24	21	21					
		B (Sciences de la gestion)	17	17	16	17	12	12					
		C (Ingénieur de la gestion)	47	48	47	48	44	44					
		D (Sciences Politiques)	26	32	31	31	27	27					
		E (Sociologie)	10	12	12	12	12	12					
		F (Information et Communication)	28	31	31	31	32	32					
		X	31	26	25	26	31	30					
FPMs 8	Sc. Appl. 6	A (Ingénieur civil)	161	154	148	152			157	157	157	157	
		B (Architecture)	18	18	17	18				17	17	17	17
		X	20	21	20	21				17	17	17	17
FUSAGx 9	Agronomie 9	0	265	258	254	256			257	257	256	256	

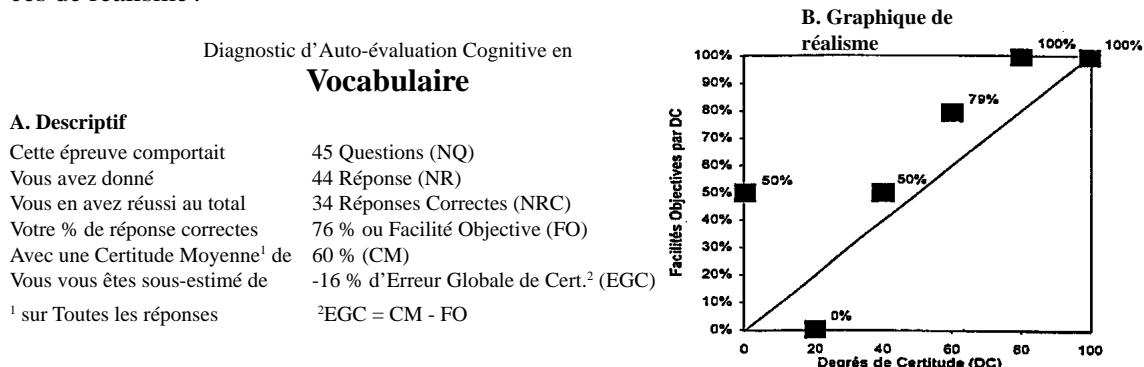
6. Les feedbacks aux étudiants

Le tout premier objectif de l'opération MOHICAN était de fournir aux étudiants un descriptif le plus précis possible de leurs résultats aux check-up.

En octobre, MOHICAN a fourni à chacun des 4000 étudiants testés son enveloppe scellée (confidentielle) contenant une feuille de résultats par épreuve (8 feuilles pour les étudiants de sciences, médecine et ingénieurs, 6 feuilles pour les autres). Conformément aux objectifs, ces informations étaient

- personnalisées* (décrivant le plus diagnostiquement possible LEUR situation);
- détaillées* (leurs réponses confrontables aux réponses correctes et aux critères de correction rendus avec leurs questionnaires, leur graphique de réalisme et leur spectre des performances pour chaque test);
- normées* (leur permettant de se situer par rapport à la moyenne des étudiants de leur section);
- précoces* (à un moment, fin octobre 1999, permettant la réaction, c'est-à-dire des modifications, des remédiations);
- confidentielles* (aucune valeur individuelle nominative n'a été communiquée aux encadrants).

Voici une partie d'un feedback « étudiant » pour l'épreuve de vocabulaire, le graphique et les indices de réalisme :



Ce feedback, donné pour chacun des check-up, est auto-descriptif et ne fait pas référence au groupe. Par contre, dans le tableau spectral ci-dessous, la colonne FO (Facilité Objective) permet à l'étudiant de connaître le pourcentage de réussite pour chaque question par les étudiants de sa section (information normative) et de situer chacune de ses réponses sur le **spectre de qualité** des performances (information critérielle).

R = votre Réponse

D. Tableau spectral

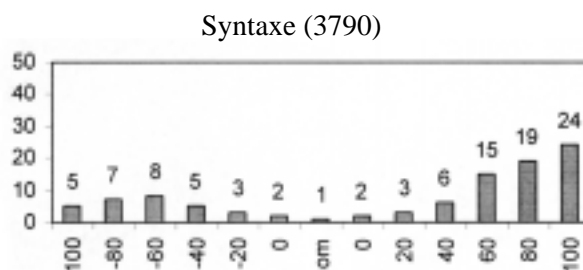
RC	R	Vos Réponses Incorrectes (RI)					OM	Vos Réponses Correctes (RC)					FO	CONTENUS DES QUESTIONS
		-100	-80	-60	-40	-20		0	20	40	60	80		
2	2	01										100	73	corrobores
1	1	02					0					100	28	prérogatives
3	3	03										100	35	stigmatisé
4	4	04										100	54	sous-jacentes
6	5	05			-60								4	talentes
4	4	06							60				76	divergent
3	3	07							60				37	dénégations
1	3	08			-60								57	discrédit
4	4	09							40				35	a inféré
6	5	010										100	50	inhérentes
5	3	011				-40							55	clivages
6	2	012			-60								14	à l'instar de
5	5	013											42	étudé
2	2	014											68	judicieuse
4	4	015											83	réfute

Les modalités d'interprétation de ces données sont détaillées au chapitre 2.

7. Les feedbacks aux enseignants

Un deuxième objectif de MOHICAN était d’informer les enseignants, donc les sections, des résultats des étudiants aux divers check-up, mais de façon anonyme pour éviter tout effet oedipien de la prédiction ou tout phénomène de halo¹⁰. Il s’agit donc de la répartition des résultats du groupe, tous étudiants confondus.

En novembre 1999, MOHICAN a fourni aux 51 sections différentes les résultats statistiques de la section, (en moyenne pour 6 tests, et, au total, 348 relevés statistiques différents) dont le *spectre des réponses* pour chaque question de chaque test. Voici ci-contre un exemple de spectre des réponses pour l’ensemble des 12 questions du check-up « Syntaxe » sur 3790 étudiants. On constate que 24 % des réponses ont été données correctement avec la certitude maximale (100%). C’est ce que nous appelons la Maîtrise totale. On constate aussi que 19% l’ont été avec 80% de certitude, mais que 5% l’avaient été incorrectement avec 100% de certitude (Méprise totale).



Les modalités détaillées de lecture et d’interprétation de ces résultats sont fournies au chapitre 2 ; les analyses check-up par check-up aux chapitres 3 et 11 à 20.

8. Le feedback à la Communauté éducative

Ce feedback est disponible sur le site du CIUF (<http://www.ciuf.be/coordination/mohican.htm>). On y trouvera les statistiques globales, sans distinction de section ni d’université. Elles sont décrites test par test au chapitre 3 ci-après. Ce sont ces statistiques qui pourraient servir d’ancrage à des comparaisons lors d’années ultérieures, notamment les valeurs relatives aux indices métacognitifs (Prudence, Confiance, Discriminance, Réalisme par Calibration, Erreur Moyenne de Certitude).

La consultation des feedbacks par les étudiants a fait l’objet d’une enquête (voir chapitre 5).

La valeur informative pour les enseignants a fait l’objet d’une expérience-sondage (voir chapitre 6).

La confidentialité promise aux étudiants a été respectée. Dans aucune université, les professeurs n’ont été mis au courant des résultats **individuels**. Une personne par université (le membre du Groupe de Travail qui pilote l’opération) est détentrice des informations permettant de suivre la trace des étudiants et de faire des liens entre ces check-up, la réussite et bien d’autres renseignements combinés. On a pu ainsi calculer les relations entre les Check-up MOHICAN et la réussite, et de façon différentielle selon les sections. C’est l’objet du chapitre 7.

Dans certaines sections de certaines universités, des opérations formatives ont été entreprises (voir chap. 4). Le chapitre 8 est consacré à l’opération RESSAC qui a combiné les résultats MOHICAN et ceux d’interrogations dispensatoires (« partiels ») de janvier pour provoquer une modification des stratégies d’études.

Le chapitre 21 présente une synthèse des résultats et le chapitre 22 des conclusions et perspectives.

¹⁰ Leclercq, D. (Ed.), (1998), 55-62.

9. Composition du groupe de travail « Réussite en candidatures »

Président :

Dieudonné LECLERCQ
Professeur à l'ULg
Technologie de l'Education
Boulevard du Rectorat, 5
4000 LIEGE
D.Leclercq@ulg.ac.be

Membres :

Bernadette NOEL
Chargée de cours aux F.U.C.A.M.
Chaussée de Binche, 151
7000 MONS
bernadette.noel@fucam.ac.be
Suppléant : Sabine GOEMAERE

Jean-Marie DE KETELE
Professeur à l'UCL
Pédagogie expérimentale
Place Cardinal Mercier, 10
Jean-marie.DeKetele@psp.ucl.ac.be
1348 LOUVAIN-LA-NEUVE

José-Luis WOLFS
Chargé de cours à l'ULB
Avenue F. Roosevelt, 50
1050 BRUXELLES
jwolfs@ulb.ac.be

Jean-Paul LAMBERT
Vice-Recteur des F.U.S.L.
Bd du Jardin Botanique, 43
1000 BRUXELLES
lambert@fusl.ac.be

Marc ROMAINVILLE
Chargé de cours aux F.U.N.D.P.
Département Education et Technologie
Rue de Bruxelles, 61
5000 NAMUR
marc.romain@fundp.ac.be
Suppléant : Sandrine BIEMAR

Calogero CONTI
Professeur à la F.P.Ms
Service de mécanique rationnelle
Boulevard Dolez, 31
7000 MONS
Calogero.Conti@fpms.ac.be
Suppléant : Bertrand DIRICQ

Michel DELHAXHE
Assistant à l'Ulg
Responsable du Service «Guidance Etude»
Allée du 6 août, B11
4000 LIEGE
mdelhaxhe@ulg.ac.be

Pol DUPONT
Professeur à l'U.M.H.
Méthodologie et formation
Place du Parc, 20
7000 MONS
Plo.Dupont@umh.ac.be
Suppléant : Muriel DELFORGE

Jean-Pierre LAMBOTTE
Professeur à la F.U.S.A.Gx
Unité de mathématique
Passage des déportés, 2
5030 GEMBLOUX
lambotte.jp@fsagx.ac.be
Suppléant : Vincent MINET

Auteurs et co-auteurs des 10 check-up :

Pierre CHAPELLE
Faculté Polytechnique de Mons
Rue de Houdain, 9
7000 MONS

André CORNELIS
ULg au Sart-Tilman
Allée de la Chimie, 17
Bât. B6
4000 LIEGE

Andre.Cornelis@ulg.ac.be

Jean-Marc DEFAYS

ULg
Place du XX Août, 7
Bât. A1
4000 LIEGE

jmdefays@ulg.ac.be

François GEORGES

LabSET – Ulg
B9 – Sart-Tilman
4000 LIEGE

Fr.Georges@ulg.ac.be

Philippe HOUGARDY

ULB
Avenue F. Roosevelt
1050 BRUXELLES

phougardy@admin.ulb.ac.be

André LAMBLIN

Inspecteur
Rue de Saint-Amand, 165,
7600 Peruwelz

Marcel LEBRUN

LEGA Jacques
UCL
Grand Rue, 54
1348 LOUVAIN-LA-NEUVE

lebrun@ipm.ucl.ac.be

lega@fyam.ucl.ac.be

Michèle MONBALLIN

FUNDP
Département Langues et Littératures romanes
Rue de Bruxelles, 61
5000 NAMUR

michele.monballin@fundp.ac.be

Bernadette NOEL

Chaussée de Binche, 151
7000 MONS

noel@message.fucam.ac.be

Françoise ORBAN

Rue de Bruxelles, 61
5000 NAMUR

francoise.orban@fundp.ac.be

Jean-Claude VERHAEGHE

ULB
Rue de la Gare, 81
5670 TREIGNES

jcverhae@ulb.ac.be

Chercheurs de l'Université de Liège impliqués dans MOHICAN :

GILLES Jean-Luc, DETROZ Pascal et PIRSON Marc (du **SMART** : Service Méthodologique d'Aide à la Réalisation de Tests – ULg) : www.smart.ulg.ac.be

POUMAY Marianne, GEORGES François et DUPONT Chantal (du **LabSET** : Laboratoire de Soutien à l'Enseignement Télématique – ULg) : www.ulg.ac.be/labset

SIMON Fabienne (du **CRIFA** : Centre de Recherche sur l'Instrumentation, la Formation et l'Apprentissage – ULg) : www.crifa.fapse.ulg.ac.be

Dactylographie et mise en page : Nelly SAENEN (STE-ULg)

10. Bibliographie

- Boxus et al. (1993). Groupe de travail « Réussite en candidatures ». *Rapport de travail*, Bruxelles : Conseil interuniversitaire francophone.
- Bloom, B. (1979). *Caractéristiques individuelles et apprentissage scolaire*. Bruxelles : Labor.
- Boxus, E. et al. (1993). *Statistiques de réussites et innovations pédagogiques en candidatures*. Bruxelles : CIUF.
- Coulon A., (1997), *Le métier d'étudiant*, Paris : PUF.
- de Kerchove A.-M et Lambert, J-P. (1996). Le libre accès à l'enseignement supérieur en Communauté française. Quelques données de base pour un pilotage du système. *Reflets et Perspectives*, Tome 35, 4^e trim 1996, 4, 453-468.
- Droesbeke, J.-J. (Ed.) (2001). *La population étudiante. Descriptions, évolution, perspectives*. Bruxelles : Editions de l'Université de Bruxelles. Collection Statistique et Mathématiques Appliquées.
- Geron, J. et Vandekerke, M. (1999). *Les conditions de vie des étudiants. Chiffres et chiffres phares*. Bruxelles : CIUF.
- Leclercq et al., Groupe de travail « Réussite en candidatures », *Franchir le cap des candis*, Bruxelles : Conseil interuniversitaire francophone, 1997.
- Leclercq, D. (Ed) (1998). *Pour une pédagogie universitaire de qualité*. Sprimont : Mardaga.
- Romainville, M. (2000). *L'échec dans l'université de masse*, Paris : L'Harmattan
- Tardif J. (1992), *Pour un enseignement stratégique. L'apport de la psychologie cognitive*, Montréal : Logiques-Ecoles
- Wang, M.-C., Haertel, G. D. et Walberg, H.J. (1990), What influences learning ? A content analysis of review literature. *Journal of Educational Research*, 84 (1), 30-44.
- Wolfs, J.-L (1998). *Méthodes de travail et stratégies d'apprentissage. Du secondaire à l'université*. Recherche-Théorie-Application, Bruxelles : De Boeck.

Chapitre premier

Connaissance partielle¹¹, analyse spectrale et métacognition

D. Leclercq, Président du Groupe de Travail CIUF « Réussite en candidatures »

¹¹ On trouvera, dans le glossaire, la définition de ces termes.

1. Introduction ou glossaire alphabétique

Les concepts qui suivent s'inscrivent dans l'idée générale de **connaissance partielle**, développée par De Finetti (1965), Van Naerssen (1965), Shuford et al. (1966), Leclercq (1983, 1985, 1990), Leclercq et Bruno (1993) et bien d'autres et mesurée grâce aux **degrés de certitude** (DC) que les étudiants sont invités à fournir en plus de chaque réponse à une question. Nous prolongerons cette idée en distinguant la **Maîtrise** (les Réponses Correctes avec DC supérieur à 50 %), la **Mi-connaissance** (les Réponses Correctes avec DC < 50 %), la **Méconnaissance** (Réponses Incorrectes ou RI avec DC < 50 %) et la **Méprise** (RI avec DC > 50 %). Nous considérons que Mi-connaissance et Méconnaissance constituent des **connaissances inutilisables**.

Ces quatre situations peuvent être situées sur un axe unique, celui de la qualité spectrale des réponses (Jans et Leclercq, 1998). Elles permettent aussi de calculer des indices de **Confiance** et de **Prudence**, relatifs et absolus (voir chapitre 10, ci-après).

Parmi toutes les façons de recueillir des degrés de certitude (Leclercq et Gilles, 2000), nous avons veillé à adopter l'une des **consignes admissibles** (selon Shuford, 1996). La **consigne minimale** consiste à demander si l'on est « plus sûr » que 50 % ou « moins sûr » que 50 %.

Nous avons adopté ici une consigne que nous appelons « Multiples de 20 % » : les étudiants doivent choisir entre 6 degrés : **0 %, 20 %, 40 %, 60 %, 80 % et 100 %**

Outre le score classique (0/1), on peut attribuer à chaque réponse +DC un **score spectral** allant de -100 à +100, ce qui, au total d'une épreuve donne une **Moyenne Spectrale** (MS). L'**attitude des étudiants** vis-à-vis des degrés de certitude dépend des scores qui en découlent ; les check-up MOHICAN n'étant pas sanctionnés, ce problème se pose moins que lors d'examens.

Pour chaque étudiant, pour chaque question et pour chaque test, on peut calculer divers indices objectifs, subjectifs ou combinant les deux. Au Taux d'Exactitude (TE) ou **Facilité Objective** (FO), s'ajoute la **Certitude Moyenne** (CM) ou **Facilité Subjective** (FS), celle-ci pouvant se décomposer en FS des réponses Correctes (**FSC**) et FS des réponses Incorrectes (**FSI**). Si FSC est inférieure à 50 %, on parle d'un **manque de confiance** et si FSI est supérieure à 50 %, d'**imprudence**. Dans une QCM, la différence entre la FSC et la FSI du distracteur-vedette est appelée **Fracture subjective** ou **Discriminance** (terme nouveau).

Le produit de la Facilité Objective d'une question et de la Facilité Subjective des réponses correctes (FSC) donne une « **Surface de compétence** » et le produit de la Difficulté Objective par FSI donne une **Surface d'incompétence**, la différence des deux déterminant une **Surface de Maîtrise**.

Le **réalisme** de l'étudiant se calcule de diverses façons :

- Par son Erreur Moyenne de Certitude (**EMC**).
- Par son Erreur Moyenne Absolue de Certitude (**EMAC**).
- Par son indice de Réalisme par Calibration.
- Par les valeurs de Confiance, Prudence, Discriminance.

Certains de ces indices ont surtout leur pertinence au niveau de populations d'étudiants, d'autres en vue de comparer des scores à des pré et des post-tests, d'autres encore pour caractériser des épreuves ou des questions et d'autres enfin des étudiants individuellement.

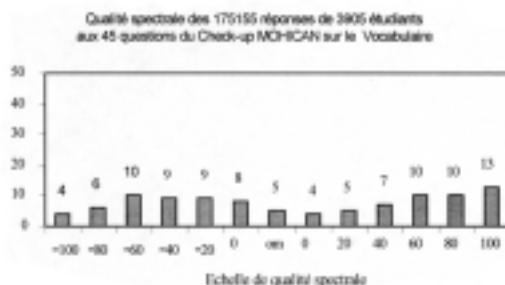
Dans MOHICAN, nous n'avons pas calculé toutes les valeurs pour tous les indices mentionnés dans le présent glossaire : nous avons fait des choix. Nous espérons que ceux-ci seront féconds. Nous avons voulu donner ici une vue d'ensemble des concepts utilisés.

Analyse spectrale des réponses

Les diverses «qualités» des réponses peuvent être rangées dans des catégories qui forment une suite sans variation brusque (d'où le nom «spectre»), et pour lesquelles il n'existe pas de convention générale quant aux noms à leur donner. Nous en proposerons ci-après pour former une taxonomie des qualités spectrales des réponses.

Le **score spectral d'une réponse** (SSR) donnée avec un degré de certitude de 80% est -80 si la réponse est incorrecte et +80 si elle est correcte. La **Moyenne Spectrale** (MS) des réponses d'un étudiant à un tests de NQ questions est la moyenne de ses NQ scores spectraux.

Voici la **distribution spectrale** des qualités des réponses de 3905 étudiants aux 45 questions du check-up de Vocabulaire. Les nombres représentent des pourcentages (arrondis à l'unité), calculés donc sur 175.725 réponses et 175.725 degrés de certitude (DC). On appelle hémispectre gauche la distribution (de -100 à -0) à gauche de l'omission et l'hémispectre droit celle (de 0 à 100 %) qui est à droite de l'omission.



Une taxonomie de 4 zones (arbitraires) de qualité spectrale¹² a été définie comme suit (voir chapitre 10, Leclercq et Poumay).

Nous décidons d'appeler « **Assurée** » une réponse accompagnée d'un degré de certitude supérieur à 50% (valeur que nous appellerons « **seuil** ») et « **Douteuse** » une réponse accompagnée d'un degré de certitude inférieur à ce seuil de 50%. La valeur du seuil est arbitraire. La partie centrale du spectre regroupe toutes les réponses Doubteuses (et les Omissions), bref la connaissance « inutile » (voir à ce terme) parce que l'individu n'osera pas s'y fier; aux extrémités: la connaissance utilisable (voir à ce terme). De gauche à droite du spectre 4 types de performances (plus l'omission) sont identifiables :

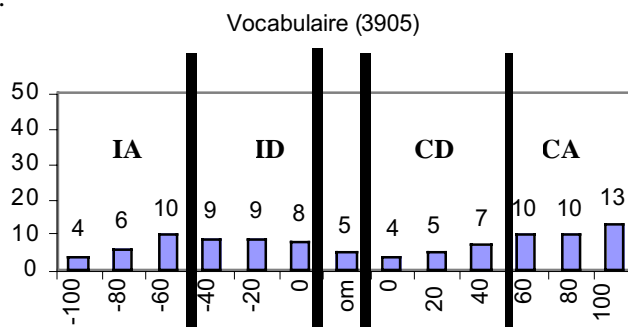
IA = Incorrectes Assurées ou **Méprise** ou **méconnaissance nuisible**, voire dangereuse ou Imprudence (certitude > au seuil de 50%), ici 20% des réponses.

ID = Incorrectes Doubteuses ou **Méconnaissance** ou **ignorance avouée** ou Prudence (cert< au seuil de 50%), ici 26% des réponses.

OM = Omissions

CD = Correctes Doubteuses ou **Mi-connaissance** ou connaissance **partielle** ou Manque de confiance (cert< au seuil de 50%), ici 16% des réponses.

CA = Correctes Assurées ou **Maîtrise** ou **connaissance fiable** ou Confiance (cert> au seuil de 50%), ici 33% des réponses.



ID et CD peuvent aussi être appelées « **connaissance inutilisable** » (Hunt, 1993).

¹² plus l'omission

Attitude des étudiants face aux degrés de certitude

L'attitude des étudiants vis-à-vis de la méthode de recueil des Degrés de Certitude (DC) dépend de ce que l'on en fait. Il n'est pas rare que de nouveaux systèmes sont détestés tout simplement parce qu'ils sont plus défavorables (en points) aux étudiants que les anciens. Dans le testing MOHICAN, ce problème ne se posait pas car les réponses n'étaient associées à aucun barème de tarifs, bien qu'il en existe¹³.

Calibration : voir Réalisme par Calibration.

Certitude : voir Introduction du présent chapitre.

Cognitif

Qui a rapport aux processus d'acquisition et de gestion de la connaissance, du savoir et aux processus intellectuels en général, les distinguant des processus affectifs, émotionnels ou sensori-moteurs.

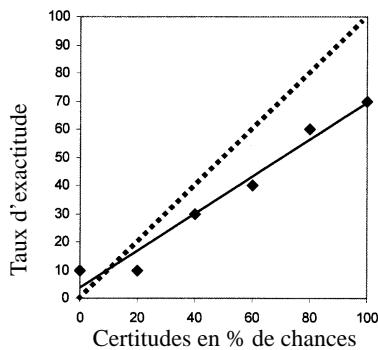
Cohérence dans l'utilisation des degrés de certitude

Quand un étudiant réussit d'autant mieux que ses certitudes sont élevées, on parle de cohérence (interne) de l'étudiant. Il peut cependant être en décalage avec la réalité (réalisme externe ou calibration). C'est pourquoi on calcule deux indices différents. La cohérence est une corrélation (représentée par un nuage de points, ici 6 points), d'autant plus élevée (proche de la valeur +1) que ces points « se mettent sur » (s'ajustent à) une ligne droite. Mais l'individu n'est réaliste (par calibration) que si cette ligne droite est la diagonale !!!

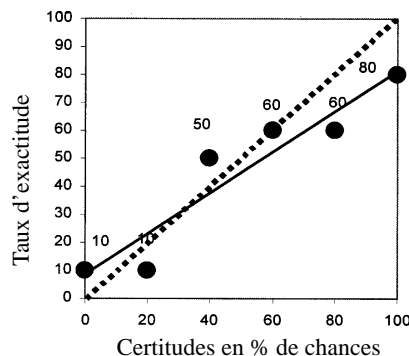
La cohérence est mesurée par la corrélation entre les Taux d'exactitude des divers degrés de certitude (en verticale) et ces degrés eux-mêmes (horizontalement). La Cohérence peut donc aller de -1 à +1. L'ennui de cet indice est qu'il ne dit rien de la proximité des Taux d'exactitude à la diagonale (bref, il ne dit rien de la qualité de la calibration).

Cela apparaît clairement dans l'exemple A du graphique gauche ci-dessous : les pointillés représentent la diagonale et la droite d'ajustement du nuage de points est en trait plein. Cette droite passe beaucoup mieux au milieu du nuage que ne le fait la diagonale ; la corrélation et donc la cohérence vaut **0,98**, mais le réalisme (par calibration) est beaucoup plus faible. En effet, alors que la certitude moyenne est de 50 %, l'exactitude moyenne n'est que de 36,6 %. L'Erreur Moyenne de Centration (EMC) vaut donc 13,13 % (surestimation moyenne de 13,13 %) et l'Erreur Moyenne Absolue de Centration (EMAC) vaut 16,6 %. Donc le réalisme par calibration vaut $100 - 16,6 = 83,4$.

Exemple A où chacun des 6 degrés de certitude a été donné 10 fois dans un test de 60 questions



Exemple B où chacun des 6 degrés de certitude a été donné 10 fois dans un test de 60 questions



¹³ Leclercq, 1998, 177-178

Par contre, dans l'exemple B du graphique de droite, la cohérence n'est que de 0,95 mais le réalisme par calibration est meilleur. En effet, alors que la certitude moyenne vaut 50 %, l'exactitude moyenne est de 45 %. L'Erreur Moyenne de Certitude (EMC) vaut donc 5 % et l'EMAC, 11,66 %. Le réalisme par calibration vaut donc **88,33**.

On peut donc dire que ce sujet à ce test a été très cohérent avec lui-même, mais beaucoup moins en accord avec la réalité. Ce désaccord est représenté par les distances à la diagonale. Dans MOHICAN, nous calculons les deux : le réalisme par calibration et l'Erreur Moyenne de centration. On trouvera les distributions de leurs valeurs pour les étudiants, test par test, dans le chapitre 3 et, dans le chapitre 7, leur corrélation avec la réussite de l'année.

Confiance (Indice de)

Certitude Moyenne des réponses correctes. On parle aussi de **Facilité Subjective des réponses Correctes (FSC)**. Quand cette **FSC** est en dessous d'un certain seuil (ex. : 50 %), nous disons **ici** qu'il y a « Confiance ». Il est justifié de fixer un seuil différent selon les matières et même selon les questions dans chaque matière.

Connaissance partielle

La connaissance partielle est liée à la quantification de l'**information**. Dans leur célèbre ouvrage « Theory of communication », Shannon et Weaver (1949) donnent, comme définition du mot « *Information* » : « *ce qui réduit l'incertitude* », ou plus exactement ce qui réduit le nombre de possibilités entre lesquelles on hésite¹⁴. Ils proposent une unité d'information : le **bit**, c'est-à-dire l'événement informatif qui réduit le nombre de possibilités ou l'incertitude à la moitié de ce qu'elle(s) étai(en)t, 2 bits la (les) réduisant au quart de ce qu'elle(s) étai(en)t, 3 bits au huitième, etc.

Le concept de connaissance partielle fonde théoriquement des méthodes de testing développées par De Finetti (1965), Van Naerssen (1966), Van Lenthe (1993), Shufford (1966), Bruno (1993), Leclercq (1983 et 1993). Ces méthodes ont été décrites dans Leclercq et Gilles (2000). Pour MOHICAN, nous avons privilégié l'une d'entre elles « la consigne M20 ou les Multiples de 20 ». Selon cette optique, *apprendre n'est pas passer d'une connaissance nulle à une connaissance totale, mais d'une connaissance partielle à une connaissance moins partielle*.

Consignes admissibles de recueil des degrés de certitude

Il existe bien des variantes dans les consignes possibles pour recueillir la connaissance partielle d'étudiants répondant aux questions d'une épreuve. Leclercq et Gilles (2000) en ont décrit et critiqué plusieurs. Shufford et al. (1966) ont démontré, dans leur article « *Admissible Probability Measurement Procedures* » que les seules méthodes acceptables consistent à demander aux testés d'accompagner chacune de leurs réponses d'un degré de certitude tenant compte du **pourcentage de chances** qu'a, à ses yeux, la réponse fournie d'être correcte (il s'agit de post-diction, alors qu'une prédiction consisterait à fournir ses chances de fournir une réponse correcte à une question non encore connue sur un thème connu).

Consigne minimaliste de recueil des degrés de certitude

Cette consigne demande aux testés de distinguer entre « sûr(e) » et « non sûr(e) », le premier cas signifiant « au-delà de 50% » et le second signifiant « en deçà de 50% » de chances d'avoir fourni la réponse correcte. Nous n'avons pas utilisé cette consigne, mais la consigne M20.

¹⁴ Cette définition découle de celle de l'entropie qui exprime la « dispersion » des probabilités dans les possibilités. Ainsi, si on considère que dans une QCM la solution correcte est une des 5 solutions proposées mais que l'étudiant n'a AUCUNE préférence pour l'une d'entre elles ni AUCUNE raison de penser que l'une d'entre elles est moins probable que les autres, alors l'entropie (ou dispersion ou incertitude ou ignorance) est maximale.

Consigne M20 ou « les Multiples de 20% » de recueil des degrés de certitude

Cette consigne a été proposée pour MOHICAN par Leclercq et Gilles :

« Choisissez votre Degré de Certitude (DC)
ou pourcentage de chances parmi l'un des 6 suivants :
0% 20% 40% 60% 80% 100% ».

Cette consigne (ou échelle) correspond à ce que De Finetti (1965, 102-109) appelle le « *five stars system* » où l'étudiant peut accompagner sa réponse d'une à cinq étoiles, chacune « pesant » 20% de certitude (l'absence d'étoile constituant une sixième possibilité). Le nombre de degrés de certitude est limité à six dans la mesure où Leclercq (1983, 1993) a montré la limitation de la capacité d'adultes non entraînés à distinguer plus de degrés (le « *magical number seven* » de G. Miller (1956) s'observe aussi dans ce domaine).

Corrélation

Degré de liaison, de co-variation entre deux variables (par exemple les résultats des mêmes étudiants à deux tests différents : les plus forts à un test sont-ils aussi les plus forts à l'autre et avec les mêmes « distances » entre étudiants ?

Degrés de certitude

Voir ci-avant l'introduction du présent chapitre.

Diagnostic (le premier mot du titre du présent ouvrage)

Détermination de la nature d'une maladie (en médecine), des forces et de faiblesses (en éducation). Le diagnostic étiologique est orienté vers les causes et le diagnostic pronostic vers l'évolution probable de la situation selon l'influence des interventions (thérapie, remédiation, etc.). Dans son « Dictionnaire de l'Evaluation et de la Recherche en Education », G. de Landsheere (1979, 172) définit un test diagnostique comme « Un test servant à localiser les difficultés d'apprentissage en vue d'y remédier ». Les résultats d'un tel test sont forcément détaillés ; un synonyme signalé par G. de Landsheere pour « *Diagnostic Test* » est « *Analytical Test* ».

Difficulté Objective (DO)

La difficulté Objective **d'un test pour une personne** est le Taux d'erreurs (omissions comprises) à ce test, bref **100% – FO** (voir **Facilité Objective**). La Difficulté Objective **d'une question pour une population** de sujets est le taux d'erreurs (omissions comprises) à cette question par cette population.

Discriminance (Indice de)

Différence entre l'Indice de Confiance et l'Indice de Prudence, donc entre deux certitudes moyennes. Voir à ces termes (Confiance – Prudence) et à Fracture Subjective.

Erreur Moyenne de Centration (EMC = FS-FO)

Cet indice est calculé par la différence entre la Facilité Subjective (FS) moyenne ou Certitude Moyenne (CM) et la Facilité Objective (FO) moyenne ou Taux d'Exactitude pour l'ensemble du test. Si cette différence est négative, l'étudiant s'est **SOUS estimé** : sa performance moyenne est meilleure que sa prédiction moyenne (en termes de moyenne de ses certitudes associées à ses différentes réponses). Si EMC est positif, cela signifie que l'étudiant s'est, en moyenne, **SURestimé**. L'inconvénient de l'EMC est que cet indice permet que des surestimations soient compensées par des sous-estimations. Quand celles-ci s'équivalent, on a l'impression (fausse) que l'auto-estimation est parfaite. On peut comparer cette situation à celle d'une personne qui arriverait à ses rendez-vous dans la moitié des cas avec 30 minutes d'avance et dans l'autre moitié des cas avec 30 minutes de retard et qui conclurait « je suis toujours parfaitement à l'heure en moyenne ».

On a donc ressenti le besoin de développer des indices (dont EMAC) tenant compte de l'ampleur de toutes les erreurs, en valeur absolue, pour qu'elles ne puissent se compenser.

Erreur Moyenne Absolue de Certitude (EMAC)

Pour obtenir EMAC, on commence par calculer le Taux d’Exactitude (TE) ou Taux de Réussite ou encore Facilité Objective (FO) pour chacun des degrés de certitude utilisés par regroupement de toutes les réponses ayant été fournies avec le même degré de certitude. Ex. : Si 10 réponses ont été données avec la certitude 60 %, et que 7 d’entre elles sont correctes, on écrira que TE de 60 % = 70 %. L’Erreur Absolue de Certitude de ce degré de certitude est la différence entre le degré lui-même (par ex 60%) et TE (par exemple 70%) soit une EAC de 10%, valeur, qui hélas **ne donne pas le sens** de cette erreur **absolue** : par surestimation ou sous-estimation ? L’indice ne le dit pas. EMAC est la moyenne (pondérée par le nombre de réponses sur lesquelles chaque TE a été calculé) des EAC. C’est EMAC qui sert à calculer l’indice de **Réalisme par Calibration** (voir à ce terme).

Facilité objective (FO) ou Taux d’Exactitude

La Facilité Objective **d’une question** pour une population de sujets est le pourcentage de Réponses Correctes ou %RC à cette question par cette population. La Facilité objective **d’un test** pour une personne est son Taux d’exactitude ou %RC à ce test. La Facilité Objective Moyenne d’un groupe de personnes à un test est la Moyenne des Taux d’exactitude de cette population aux NQ questions de ce test.

Facilité Subjective (FS) ou Certitude Moyenne

L’exactitude (0/1) d’une réponse à une question constitue sa **Facilité Objective** (FO) et le degré de certitude qui y est associé constitue la **Facilité Subjective** (FS) d’une Question «q» pour **un** Etudiant «e» (valeurs A).

Ce qui vient d’être dit à propos d’une question peut s’appliquer à l’ensemble des questions (donc du **test** entier) auquel **un** étudiant «e» a répondu (valeurs C). On parle alors de la Facilité Objective de l’épreuve ou du test (le Taux de Réponses Correctes au test) pour cet étudiant et de la Facilité Subjective (la moyenne de ses degrés de certitude aux diverses questions de ce test) pour ce même étudiant.

	E_1		E_e				E_{ne}	
Q_1								
Q_q			A					B
Q_{nq}								
			C					D

Ces mêmes indices peuvent être calculés **pour une question** «q» sur **tous les étudiants** d’une section, d’une faculté, d’une université, d’une communauté (valeurs B). Ils peuvent aussi être calculés **pour un test**, par la moyenne des réponses de **tous** les étudiants à **toutes** les questions (valeurs D).

Facilité Subjective des réponses Correctes (FSC) ou Indice de Confiance

C’est la moyenne des certitudes accompagnant les réponses correctes. Voir **Facilité Objective**.

Facilité Subjective des réponses Incorrectes (FSI) ou Indice de Prudence

C’est la moyenne des certitudes accompagnant les réponses incorrectes. Voir **Difficulté Objective**.

Fracture Subjective (Fract S) ou Indice de Discriminance

L'indice FractSQ est l'**écart** (la fracture) entre FSC et FSI, entre l'Indice de Confiance et l'Indice de Prudence. Voici les valeurs moyennes pour les 10 check-up MOHICAN 1999 :

Total	Vocab	Syntaxe	Comp.T	Comp.G	Math	Phys	Chim	Biol	Arts	Hist-éco
FSC	62,9	74,2	67,3	61,3	80	71,4	68,4	66,1	64,3	70
FSI	43,9	60	55,7	49,4	54,8	44,5	41	45,4	38,7	35,3
Fract	18,9	14,2	11,6	12	25,2	26,9	27,4	20,7	25,6	34,7

On constate que la **Fracture Subjective** moyenne **la plus élevée** est celle du check-up **Connaissances artistiques (33%)**. Une explication pourrait être que dans ce type de matière, quand on connaît, on en est bien conscient car on a souvent le soutien puissant de la mémoire épisodique d'événements vécus tels que avoir lu un livre, vu un film, visité un musée, etc., et donc on avance un degré de certitude élevé. Par contre, quand on ne connaît pas, contrairement à d'autres matières, on dispose de peu de moyens de « raisonnement » pour trouver la solution correcte, et, soit on omet (c'est un des deux check-up, avec la lecture de documents, qui a le taux d'omissions le plus élevé : 7%), soit on répond, mais avec un Degré de Certitude faible. C'est le check-up dont la Facilité Subjective moyenne des réponses Incorrectes, FSI, est de loin la moins élevée : 30%.

Les fractures subjectives (Indices de Discriminance) qui suivent en importance sont celles des check-up de **mathématique (27%)**, de physique (28%) et de chimie (28%). Pour les mathématiques, une explication pourrait être qu'il s'agit du contenu pour lequel les étudiants ont le plus de **possibilités de « vérification logique »** interne, ce qui débouche sur des Facilités Subjectives des Réponses Correctes (FSC) ou Confiances très élevées. C'est en effet pour les mathématiques que la FSC moyenne est la plus élevée (78%). Cette caractéristique a son revers : c'est aussi la cause, mais heureusement avec un effet moins ample, de Facilités Subjectives des Réponses Incorrectes (FSI) élevées elles aussi ; le check-up de Mathématique a, après le check-up de Syntaxe, la deuxième valeur moyenne de FSI (51%). On peut aussi calculer la Fracture subjective entre la réponse correcte et le distracteur-vedette (qui a été le plus choisi).

Voici comment se répartissent les fractures subjectives (ou Indices de Discriminance) dans les 173 questions des 10 check-up :

	NQ	Aber- rante (néga- ve)	Insigni- fiante 0 à 4,9%	Très faible 5 à 9,9%	Faible 10 à 14,9%	Modeste 15 à 19,9%	Nette 20 à 29,9%	Elevée 30 à 39,9%	Très élevée >=40%
Vocabulaire	45	4	4	5	13	7	8	4	0
Syntaxe	12	1	0	5	3	3	0	0	0
Compr. Texte	6	2	0	2	0	1	1	0	0
Compr. Géo	10	1	1	2	2	2	2	0	0
Math	22	1	0	2	5	3	8	2	1
Physique	10	0	1	0	2	3	3	1	0
Chimie	8	0	1	0	2	2	2	1	0
Biologie	10	2	0	3	2	1	2	0	0
Conn. artistiques	25	1	2	2	5	4	4	0	7
Hist-Actual. Econ.	25	1	2	3	3	1	9	4	2
	173	13	11	24	37	27	39	12	10
		7,5%	6,3%	13,8%	21,3%	15,6%	22,5%	12,7%	5,7%

On est frappé par la présence de fractures subjectives aberrantes (négatives), insignifiantes (proches de 0) ainsi que par la rareté des fractures très élevées.

Inutilisables (connaissances)

Voir utilisables

Maîtrise

Voir **Qualité spectrale des réponses**

Maîtrise totale

C'est le pourcentage de réponses correctes avec la certitude la plus élevée (100%). Voir **Méprise totale**.

Méconnaissance

Voir **Qualité spectrale des réponses**

Mi-connaissance

Voir **Qualité spectrale des réponses**

Méprise

Voir **Qualité spectrale des réponses**

Méprise totale

C'est le pourcentage de réponses incorrectes avec le degré de certitude le plus élevé (100%). Voici ces pourcentages pour les dix check-up :

Total	Vocab	Syntaxe	Comp.T	Comp.G	Math	Phys	Chim	Biol	Arts	Hist-éco
Maîtrise Totale	13	24	13	14	36	28	22	17	24	18
Méprise Totale	4	5	8	8	7	4	4	5	5	5

Faut-il être satisfait ou non du fait que les valeurs moyennes de Maîtrise totale valent de 13 % à 36 % et celles de la méprise totale de 4 à 8% ? La réponse, à notre avis, varie de check-up en check-up et surtout de question à question (voir chapitres 11 à 20).

Métacognition

« Domaine qui regroupe

1. les connaissances introspectives conscientes qu'un individu particulier a de ses propres états et processus cognitifs
2. les capacités que cet individu a de délibérément contrôler et planifier ses propres processus cognitifs en vue de la réalisation d'un but ou d'un objectif déterminé »

(Gombert, 1990, 27 ; cité par Noël, Romainville et Wolfs (1995,49)).

Moyenne Spectrale (MS) des réponses d'un sujet à un test à NQ questions

Cet indice est obtenu par la somme des **scores spectraux** de ce sujet. La MS peut être calculée pour un ensemble de sujets ayant répondu à un même test. Dans l'exemple du check-up Vocabulaire MOHICAN (voir Qualité spectrale des réponses), **MS du test** =

$$= ((4 * -100) + (6 * -80) + (10 * -60) + (9 * -40) + (9 * -20) + (5 * 20) + (7 * 40) + (10 * 60) + (10 * 80) + (13 * 100)) / 100$$
$$= (-400 -480 -360 -180 + 100 + 280 + 600 + 800 + 1300) / 100 = (-1420 + 3080) / 100 = 1660 / 100 = 16,60.$$

On peut calculer cette valeur pour chaque étudiant à chaque épreuve. On peut aussi calculer cet indice MS pour chaque question, pour chaque étudiant, ou pour un groupe d'étudiants à une question particulière.

Quand elle est calculée à partir des mêmes données, la MS est égale à la Surface de maîtrise (voir plus loin).

Nuance ou Discriminance

C'est la différence (en %) entre la FSC (Facilité Subjective des Réponses Correctes) ou Certitude Moyenne attribuée aux réponses correctes et la FSI ou Certitude Moyenne attribuée aux réponses incorrectes.

Omission

Absence de réponse à une question. Dans le système recourant aux degrés de certitude, les **Omissions** sont, tant pour les étudiants que pour les évaluateurs, une « zone aveugle » dont quasi rien ne peut être dit, dont aucune information en retour pour l'apprentissage ne peut être tirée, sauf que les étudiants sont incapables (mais à quel degré ?) de fournir la réponse correcte. Si un étudiant décide d'omettre à toutes les questions, nous ne saurons rien -et lui non plus- de sa connaissance partielle. Et lorsque, plus tard, il fournira la réponse correcte, nous saurons, évidemment, qu'il y a eu « progrès », mais nous ne pourrons le quantifier subtilement. C'est pourquoi nous invitons les étudiants à répondre plutôt qu'à omettre, quitte à accompagner leur réponse du degré de certitude 0.

Prudence (Indice de...) ou FSI

Certitude Moyenne (ou Facilité Subjective moyenne) des réponses Incorrectes. Quand cet indice (**FSI**) est **inférieur à 50 %**, nous disons **ici** qu'il y a Prudence. Il serait justifié de fixer un seuil différent selon les matières et même selon les questions dans chaque matière. Voir Confiance (Indice de).

Réalisme par Calibration d'une personne à un test

Le Réalisme est la conformité entre les estimations subjectives et la réalité. Ce concept peut être opérationnalisé par plusieurs indices (ex. : Prudence, Confiance, Discriminance ou EMC). Le Réalisme par Calibration est calculé par la formule **100-EMAC**, c'est-à-dire la perfection (100%) moins l'Erreur Moyenne Absolue de Certitude (voir EMAC). Pas plus qu'EMAC cet indice ne dit si la non perfection (100%) est due à plus de surestimation que de sous-estimation ou l'inverse. On peut aussi exprimer le réalisme par l'ensemble des trois indices de Prudence, Confiance, Discriminance.

Pour permettre de comprendre en profondeur le Réalisme par Calibration, nous ouvrons quatre rubriques de (1) à (4).

(1) Réalisme par Calibration. Limites de confiance (ou Moustaches) des taux d'Exactitude

L'erreur moyenne absolue de certitude (EMAC) et de là le Réalisme par Calibration (100-EMAC), procède par regroupement de réponses (ici par exemple afin de calculer des Taux de Réussite). Plus le nombre de réponses sur lesquelles un taux de réussite est calculé est élevé plus le Taux d'Exactitude est fiable, plus son erreur de mesure est réduite, plus cette valeur, calculée sur un échantillon de réponses, représente la valeur que l'on obtiendrait si le nombre des réponses avait été très grand. Ainsi, si l'étudiant a réussi 8 réponses sur 10 (ces 10 réponses n'étant que l'échantillon recueilli grâce à ce test), le pourcentage « 80% » a pour erreur de mesure $\sqrt{pq/N}$ (le signe racine couvrant le tout). Donc ici $\sqrt{(0,8*0,2)/10}$ ou $\sqrt{0,16/10}$ ou $\sqrt{0,016}$ soit 0,126. La valeur de 80% (FO ou Taux de Réponses Correctes avec certitude 60%) doit donc être « assortie » en plus comme en moins de 12,6% pour définir la zone (donc de 67,4% à 92,6%) dans laquelle la vraie valeur se trouve avec une certaine probabilité.¹⁵

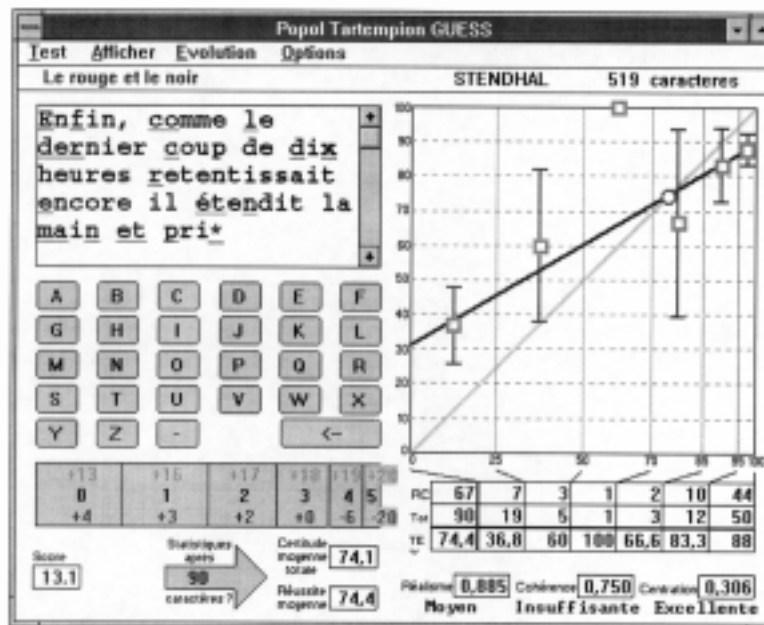
¹⁵ En fait, nous avons porté de part et d'autre de la valeur observée (80%) UNE FOIS l'erreur de mesure (12,6%). Si l'on fait (comme souvent en psychométrie) l'hypothèse d'une distribution gaussienne des erreurs de mesure, cela nous garantit 64% de chances que la vraie valeur soit dans cet intervalle. Nous décidons de travailler avec ce niveau d'incertitude, dans une perspective d'évaluation FORMATIVE. Quand il s'agit de recherches orientées vers des conclusions ou des mesures orientées vers les sanctions (examens par exemple), les psychologues et pédagogues décident d'habitude, pour exclure plus de chances de se tromper, d'assortir la valeur observée de DEUX (plus exactement 1,96) FOIS l'erreur de mesure pour s'assurer que l'intervalle (allant ici de 54,8% à 105,2%) aie 95% de comprendre la vraie valeur. Si l'on veut mettre 99% de chances de son côté, alors il faut multiplier l'erreur de mesure par 2,58, ce qui donne 32,5% (et l'intervalle deviendrait ici 47,5% - 112,5%).

Avec une Facilité Objective de 80% là où la Facilité Subjective était annoncée à 60%, nous pouvons dire à l'étudiant que sur ces 10 réponses, il s'est sous-estimé (car l'intervalle de confiance va de 67,4% à 92,6%, ce qui ne recouvre PAS la valeur attendue : 60%).

Si la Facilité Objective avait été 70% (7 réussites sur 10), l'erreur de mesure aurait été $\sqrt{(0,7*0,3)/10}$, soit $\sqrt{0,21/10}$, soit $\sqrt{0,021}$, soit 0,144. L'intervalle de confiance irait alors de 55,6% à 74,4%, ce qui couvre la FS (Facilité Subjective) annoncée (60%) et ne nous permet pas de conclure à une sous-estimation.

(2) Réalisme par Calibration. Entraînement des étudiants

L'exemple suivant présente les 90 premières réponses (dont 67 correctes) au jeu proposé par le logiciel GUESS (Leclercq et Gilles, 1994) d'entraînement à l'utilisation de degrés de certitude (ces degrés étant ici définis par les 6 intervalles suivants 0%-25%, 25%-50%, 50%-70%, 70%-85%, 85%-95%, 95%-100%) différents de la consigne utilisée dans MOHICAN. Ce logiciel propose à l'étudiant de deviner les lettres successives d'un texte tiré au hasard dans son réservoir de textes (ici un extrait de «Le rouge et le Noir» de Stendahl).



(3) Réalisme par Calibration. Son graphique

C'est un nuage de 6 points (les TE à la verticale de chacun des degrés de certitude). Comme TE et les degrés de certitude sont exprimés dans une même unité (pourcentages, de réussite pour l'un, de chances de réussite pour l'autre), le réalisme parfait se produit quand tous les TE sont sur la diagonale.

On constate les «moustaches» ou «limites de confiance» de part et d'autre des 6 pourcentages ou Taux de Réponses Correctes ou TE (Taux d'Exactitude): 36,8% ; 60% ; 100% ; 66,6% ; 83,3% ; 88%). On voit que les deux premiers TE (ou %RC), les plus faibles, à gauche, même assortis vers le haut et vers le bas de leur **limite de confiance** NE recourent PAS la diagonale. Dès lors, nous sommes autorisés pour ces deux degrés de certitude à dire à cet étudiant qu'il s'y sous-estime. A l'inverse, pour le dernier degré de certitude à droite (le plus élevé), ses limites de confiance elles non plus NE recourent PAS la diagonale et nous sommes là aussi autorisés à dire à cet étudiant qu'avec ce degré de certitude il se surestime. Pour les trois autres degrés de certitude, nous ne pouvons pas conclure à une erreur d'estimation, ou plus exactement nous laissons à l'étudiant le «bénéfice du doute» parce que les moustaches ou intervalles de confiance recouvrent la diagonale.

(4) Réalisme par Calibration. Son calcul

Les nombres de l'exemple qui suit ont été volontairement conçus pour faciliter les calculs. Imaginons une épreuve de 60 questions (peu importe qu'elles soient à choix multiple ou à réponse construite) et que 60 réponses ont été données, dont 10 avec certitude 0%, 10 avec 20%, 10 avec 40%, 10 avec 60%, 10 avec 80% et 10 avec 100% de certitude (ligne 2).

Les Nombres de Réponses Correctes (ligne 3) et les Taux d'Exactitude (ligne 4) étaient les suivants pour les différents degrés de certitude :

Certitude ou Facilité Subjective (FS)	0%	20%	40%	60%	80%	100%	Total	Moy
Nbre de Réponses (NR)	10	10	10	10	10	10	60	
Nbre de Réponses Correctes (NRC)	3	1	4	4	6	9	27	45%
Taux d'Exactitude (TE) ou Facil Obj (FO)	10%	10%	30%	40%	60%	70%		
Erreur de Certitude (FS-TE)	-10%	10%	10%	20%	20%	30%		
TEC = Tot d'Erreur de Cert * Nbre de R	-100	100	100	200	200	300	800	13,3
TEAC (TEC en valeur absolue)	100	100	100	200	200	300	1000	16,6

EMAC

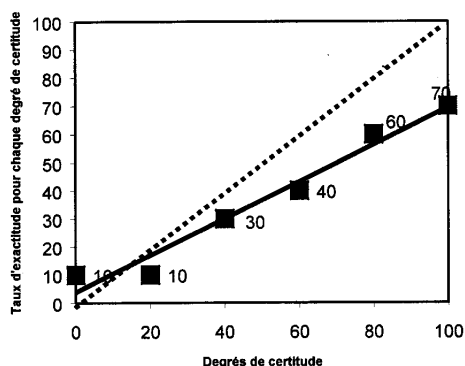
On peut calculer les Erreurs de Certitude (ligne 5), puis leur Total par DC (ligne 6) et enfin leur total en valeur absolue (ligne 7).

La Moyenne **EMC** (TEC / NR, ici 800/60) vaut 13,3% ce qui constitue une sur-estimation de la difficulté moyenne de **13,3%**.

La Moyenne **EMAC** (TEAC / NR, ici 1000 / 60) vaut 16,6, ce qui signifie une Erreur moyenne (sans que l'on puisse dire dans quel sens) de **16,6%** chaque fois que l'on estime.

Le Réalisme par Calibration valant 100-EMAC, ici le réalisme par Calibration est **83,4**.

Exemple de démonstration où chacun des 6 degrés de certitude a été donné 10 fois dans un test de 60 questions



Voici les moyennes et les écarts-types des **réalismes** par Calibration aux 10 check-up MOHICAN (voir détails chapitre 4 car ces valeurs sont calculées sur des nombres différents pour chaque check-up) :

	Vocabulaire	Syntaxe	Compréhension	Compréh. Géographie	Math	Physique	Chimie	Biologie	Arts	Histoire Actu-Eco
Moyenne	79	76	67	77	78	75	74	72	80	78
Ecart-Type	10	14	16	12	13	14	14	13	10	9

Solutions Générales dans une QCM

Ces solutions sont valables pour toutes les questions d'une épreuve constituée de QCM. Elles peuvent être implicites quand elles ne sont pas reproduites dans chaque question. En voici quatre : Aucune, Toutes, Manque, Absurdité. On peut en imaginer d'autres.

Spectre de qualité des réponses

Voir **Analyse spectrale des réponses**

Surfaces de Compétence (SurfC)

Cette expression (surface) vise à rappeler que cet indice est obtenu par la multiplication de deux proportions ou pourcentages l'un par l'autre. SurfC est la Facilité Objective (ou FO Taux de réussite) multiplié par la Facilité Subjective des Réponses correctes (FSC) ou la Certitude Moyenne (CM) attribuée à ces réponses correctes. Pour le check-up de Vocabulaire, SurfC = 49 * 60 = **29,4**

$$\text{Surface de Compétence (SurfC)} = (\text{FO} \times \text{FSC}) / 100$$

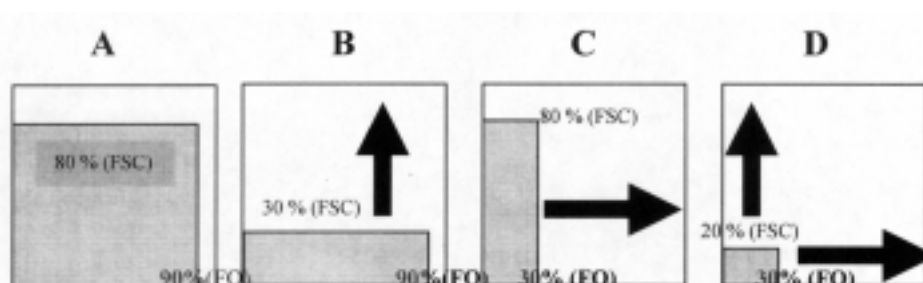
Facilité Objective (FO) x Facilité Subjective des réponses Correctes (FSC),
le tout divisé par 100.

Quand toutes les réponses correctes ont été données avec une certitude maximale (100%), alors SurfC = FO. Dans la plupart des cas, le SurfC est inférieur à FO. *L'intérêt de cet indice SurfC* est qu'il attire l'attention sur le fait qu'être Compétent, ce n'est pas seulement donner la réponse correcte, c'est la donner avec une certitude élevée (du moins quand il existe une réponse correcte qui fait l'unanimité des spécialistes).

Voici 4 situations (ou résultats) «repères» pour 5 questions fictives : A, B, C et D.

Les FO (Facilités Objectives ou Taux d'exactitude) sont portés en abscisse (axe des x).

Les FSC (Facilités Subjectives des réponses Correctes) ou Certitudes Moyennes des réponses Correctes sont portées en ordonnées (axe des y). La Surface de Compétences (SurfC) est représentée par la surface grisée.



La question **A** (Haute certitude sur base large) a été réussie par un pourcentage élevé (90%) d'étudiants qui ont donné leur réponse avec une certitude moyenne élevée (80%), ce qui donne une SurfC de **81%** (de surface grisée).

La question **B** (Basse certitude sur base large) a été réussie par un pourcentage aussi élevé (90%) d'étudiants, mais qui étaient beaucoup moins sûrs de leur réponse en moyenne (FSC = 30%). La SurfC vaut donc **27%** (de surface grisée). C'est par l'élévation de la certitude moyenne (cf la flèche vers le haut) de ceux qui ont bien répondu que l'amélioration globale de la performance à la question B se fera essentiellement.

La question **C** (Haute certitude sur base étroite) a été réussie par un faible pourcentage (30%) d'étudiants, mais, à cette question, ces étudiants étaient très sûrs de leur réponse (correcte) : 80%, ce qui donne une SurfC de **24%** (de surface grisée). C'est par l'élargissement du taux de réponses correctes ou de la Facilité Objective (cf. la flèche vers la droite) que l'amélioration globale de la performance à la question B se fera essentiellement.

La question **D** (Basse certitude sur base étroite) a été réussie par très peu d'étudiants (30%) très peu sûrs de leur réponse (FSC = 20%), a une SurfC = **6%**. C'est par l'amélioration des deux aspects à la fois que se fera l'amélioration globale (cf. les deux flèches).

Voici les types de Surfaces de Compétence pour les 173 questions de 10 check-up :

	NQ	Base étroite (moins de 50% de réussite)		Base large (plus de 50%)	
		Certitude basse (moins de 50% de certitude moyenne)	Certitude haute (plus de 50% de certitude moyenne)	Certitude basse (<50% de Cert. Moy)	Certitude haute (>50% de <Cert. Moy.)
Vocabulaire	45	10	14	0	21
Syntaxe	12	0	2	0	10
Compr. Textes	6	1	2	0	3
Géo et cartes	10	1	3	1	5
Math	22	0	3	0	19
Physique	10	1	2	0	7
Chimie	8	0	5	0	3
Biologie	10	1	5	0	4
Histoire-Actu-Eco	25	5	6	1	13
Conn. Artistiques	25	5	10	1	9
Total	173	24	52	3	94
soit	100 %	13,8%	30%	1,7%	54,3%

Donc, dans cette opération, sur les 173 questions, la Surface de Compétence est pour une bonne moitié des questions supérieure à 25% puisque ET la base (la Facilité Objective ou le Taux de réponses correctes) et la hauteur (la certitude moyenne des réponses correctes ou FSC) sont toutes deux supérieures à 50%... et que la Surface de compétence est la multiplication de l'un par l'autre.

La deuxième situation, en fréquence, sont les questions dont la base est étroite (moins de 50% de réussite), mais la hauteur est élevée (certitude moyenne supérieure à 50%). Il s'agit de 30% des questions. Il est assez normal que les étudiants qui ont réussi, même s'ils sont moins de la moitié, aient une certitude élevée dans leur réponse (correcte).

En troisième lieu viennent les questions (13,8% seulement) dont la base (le taux de réponses correctes) et la hauteur (la certitude moyenne de ces réponses correctes) est basse elle aussi. Cette situation est plus problématique : pourquoi, quand des étudiants ont fourni la réponse correcte, ont-ils une certitude moyenne inférieure à 50% ?

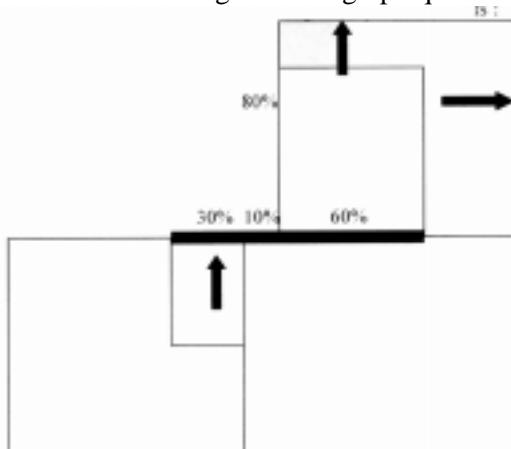
Enfin sont exceptionnelles les questions à large base (réussies par plus de la moitié des étudiants) mais dont la certitude moyenne de ceux qui ont bien répondu n'atteint pas 50%. Elles sont plus problématiques encore.

Dans les chapitres qui suivent, le lecteur aura l'occasion de rencontrer ces questions des deux derniers types et d'essayer d'apporter, s'il le veut, ses propres explications.

Surface d'Incompétence (SurfI)

C'est la Difficulté Objective (ou Taux de Non réussite) multiplié par la Facilité Subjective des Réponses Incorrectes (FSI) : $SurfI = DO * FSI$.

C'est la surface en bas à gauche du graphique ci-dessous :



L'amélioration de la situation peut venir de la diminution de DO ou de FSI (la flèche vers le haut). On observera que la somme de FO, OM et DO (représentée par la barre horizontale épaisse) fait toujours 100%. Pour le check-up de Vocabulaire, la SurfI moyenne est $(45 * 40) / 100 = 18\%$ et pour le check-up de Physique la SurfI moyenne = $13,1\%$.

Surface de Maîtrise (SurfM)

Cet indice est calculé par $SurfM = SurfC - SurfI$.

Pour le check-up de Vocabulaire MOHICAN, SurfM moyenne vaut $29,4\% - 18\% = 11,4\%$ et pour le check-up de Physique MOHICAN, la SurfM moyenne vaut $44,9\% - 13,1\% = 31,8\%$.

On pourrait s'étonner de ne pas voir apparaître les **Omissions** dans ce calcul de SurfM. En fait, cette non-incorporation dans les formules est logique, car elles interviennent « en creux ». En effet, plus leur nombre est élevé, plus elles empêchent d'obtenir une SurfC parfaite. Mais à l'inverse, elles constituent une partie des difficultés qui ont une Facilité Subjective des réponses Incorrectes (FSI) nulle, et donc ne contribuent pas non plus à gonfler la Surface d'Incompétence. On pourrait (ce que nous n'avons pas fait) incorporer aux Erreurs les Omissions comme si elles étaient des erreurs avec certitude 0 %.

Ces valeurs sont calculables par étudiant. Elles ne sont utiles que dans une perspective d'amélioration où des valeurs successives de ces taux peuvent être comparées, ce qui n'est pas le cas dans l'opération MOHICAN, mais aurait pu l'être si un enseignant avait reposé les mêmes questions après remédiation. Voici les valeurs des SurfC, SurfI (valeurs négatives) et SurfM moyennes (**en %**) pour les 10 check-up :

Total	Vocab	Syntaxe	Comp.T	Comp.G	Math	Phys	Chim	Biol	Arts	Hist-éco
N =	3905	3790	3469	3738	2539	2520	2522	2526	1399	1418
OM	5	1	6	8	3	3	5	3	7	5
FO =	49	69	44	55	66	65	55	49	49	52
FSC =	60	73	68	62	78	69	66	63	63	59
SurfC =	29,4	50,4	29,9	34,1	51,5	44,9	36,3	30,9	30,9	30,7
DO =	45	29	51	36	32	32	41	48	44	44
FSI =	40	56	48	40	51	41	38	43	30	46
SurfI =	-18	-16,2	-24,5	-14,4	-16,3	-13,1	-15,6	-20,6	-13,2	-20,2
SurfM	11,4	34,2	5,4	19,7	35,2	31,8	20,7	10,3	17,7	10,5
FractS	20	17	20	22	27	28	28	20	33	13
MS	11,4	33,4	1,14	24	34,9	32,3	21,6	11,1	19,0	14,2

La Moyenne Spectrale (voir à ce terme), quand elle est calculée sur les mêmes données de base, est **égale à SurfM** (voir **Surface de Maîtrise**). Ce n'est pas le cas ici car les FO moyennes sont les moyennes des FO des questions (en plus à partir de valeurs arrondies à l'unité) et les MS moyennes sont les moyennes des MS des étudiants. On voit, néanmoins, la proximité des valeurs obtenues.

Dans les tableaux, nous affectons SurfI d'un signe négatif pour rappeler que plus il grandit, plus l'incompétence grandit. Cet indice se justifie pour ceux qui pensent que la prudence doit être encouragée parce que l'ignorance reconnue est moins grave qu'une erreur tenue pour vraie avec une certitude élevée, de telles connaissances constituant des dangers. Comme disait Mark Twain, «Ce n'est pas ce que nous ignorons qui nous cause des problèmes, c'est ce que nous savons... mais qui est faux».

On pourrait trouver des SurfM négatives, ce qui n'est le cas pour aucun des check-up pris dans leur ensemble (donc en moyenne), mais cela se produit évidemment pour des questions particulières et des étudiants particuliers à certains check-up. Cet indice se justifie pour ceux qui pensent qu'une mesure du niveau de qualité d'une performance à un check-up doit tenir compte de la SurfC et de la SurfI, et que cette combinaison (la SurfM) est la mesure de qualité la plus valide sur le plan conceptuel (*construct validity*). On verra (au chapitre 7) que cet indice a aussi une validité prédictive.

Utilisables et Inutilisables (réponses ou connaissances)

Cette distinction est due à Hunt (1993).

Dans le graphique suivant,

IA = Réponses Incorrectes Assurées (certitude >50 %).

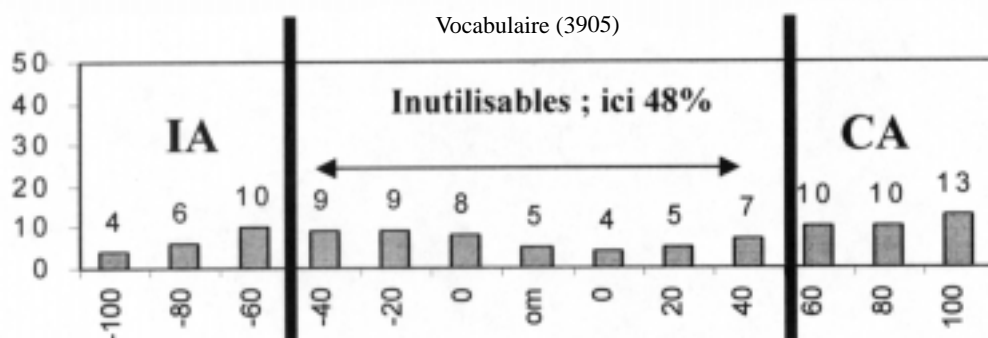
ID = Réponses Incorrectes Douteuses (certitude <50 %).

CD = Réponses Correctes Douteuses (certitude <50 %).

CA = Réponses Correctes Assurées (certitude >50 %).

Les **D** (ID + CD) pour lesquelles la certitude est inférieure à 50% constituent les connaissances **INUTILISABLES** car trop peu assurées.

Les **A** (IA + CA) pour lesquelles la certitude est supérieure à 50% constituent les connaissances utilisables, mais les IA sont potentiellement dangereuses.



Vigilance Cognitive

Capacité de détecter les incohérences, les absurdités, les manques de données, etc. Les Solutions Générales dans les QCM sont conçues pour entraîner et tester cette capacité.

2. Bibliographie

- Bruno, J. (1993). Using Testing to Provide Feedback to Support Instruction : A Reexamination of the Role of Assessment Organizations, in D. Leclercq & J. Bruno (Eds). *Item Banking : Interactive Testing and Self-Assessment*. NATO ASI Series, Berlin : Springer Verlag, pp. 190-209.
- De Finetti, B. (1965). Methods for discriminating levels of partial knowledge concerning a test item, *British Journal of Mathematical and Statistical Psychology*. 18, 87-123.
- Jans, V., Leclercq, D. (1998). Mesurer l'effet de l'apprentissage à l'aide de l'analyse spectrale des performances. In In Depover et Noël (Eds.), *L'évaluation des compétences et des processus cognitifs, Modèles, pratiques et contextes*.
- Gombert, J.-E. (1990), *Le développement métalinguistique*, Paris : PUF.
- Hunt, D. (1993). Human self assessment : theory and application to learning and testing, in D. Leclercq & J. Bruno (Eds). *Item banking : self-assessment and interactive testing*. NATO ARW, F112, Berlin : Springer Verlag, 177-189.
- Leclercq, D. (1983). Confidence marking, its use in testing, 106. in Postlethwaite & Choppin, *Evaluation in Education*, vol. 6, 161-287, Oxford : Pergamon Press.
- Leclercq, D. (1985), Computer-managed Testing (943-944) & Confidence Marking (967-969), in T. Husen & N. Postlethwaite, *The International Encyclopedia of Education*, Oxford : Pergamon.
- Leclercq, D. (1990), Computer-managed Testing (453-454) & Confidence Marking (343-345), in H. Walberg & G. Haertel, *The International Encyclopedia of Educational Evaluation*, Oxford : Pergamon.
- Leclercq, D. et Bruno, J. (Eds) (1993), *Item banking : self-assessment and interactive testing*, NATO ARW, F112, Berlin : Springer Verlag.
- Leclercq D. & Gilles, J.-L. (2000). Techniques de mesure dans l'auto-évaluation. Dix techniques d'auto-estimation de la qualité des réponses, in G. Figari & M. Achouche, *L'activité évaluative réinterrogée. Regards scolaires et socioprofessionnels*, Bruxelles : De Boeck, 134-146.
- Leclercq, D. (Ed. 1998). *Pour une pédagogie universitaire de qualité*. Sprimont : Mardaga.
- Leclercq, D. et al. (1993). Validity, reliability and acuity of self-assessment in educational testing, in D. Leclercq & J. Bruno (Eds). *Item banking : self-assessment and interactive testing*, NATO ARW, F112, Berlin : Springer Verlag, 114-131.
- Leclercq, D. et Gilles, J.-L., GUESS, Un logiciel pour s'entraîner à l'auto-estimation de sa compétence cognitive. Actes du 3^e Colloque International ESIEE, Marne-La-Vallée, décembre 1994, 137-158.
- Miller, G.-A. (1956). The Magical Number Seven, plus or minus two. *Psychological Review*, Vol. 63, 81-97.
- Noël B., Romainville, M. et Wolfs, J-L, (1995). La métacognition : facettes et pertinence du concept en éducation, *Revue Française de pédagogie*, n°112, juillet-août-septembre 1995, 47-56.
- Shannon, C.E. & Weaver, W. (1949), *The Mathematical Theory of Communication*, University of Illinois Press.
- Shuford, E., Albert, A. & Massengill, N. (1966). Admissible probability measurement procedures. *Psychometrika*, 31,125-145.
- Van Naerssen, R.F., A scale for the measurement of subjective probability, *Acta Psychologica*, 1965, 20, 2, 159-166.

Chapitre 2

Les feedbacks des check-up Mohican aux étudiants et aux enseignants

D. Leclercq, ULg, Président du Groupe de Travail CIUF « Réussite en candidatures »

F. Georges, Chercheur au LabSET, Université de Liège

J.-L. Gilles, Directeur du SMART, Université de Liège

M. Pirson, Informaticien au SMART, Université de Liège

F. Simon, Chercheur au CRIFA, Université de Liège

1. Aux étudiants

Comment lire et interpréter le Rapport CHECK-UP 1999 de chaque épreuve

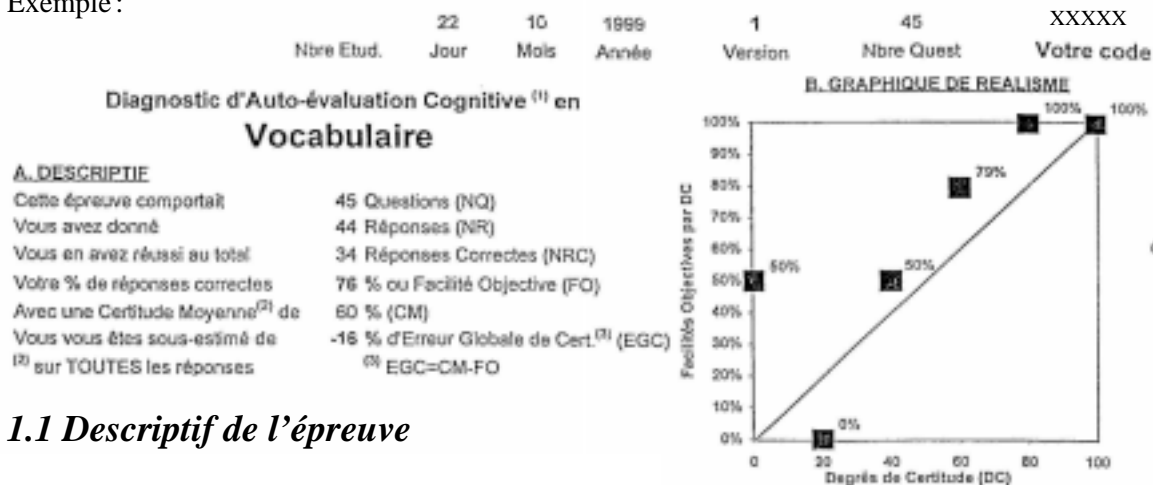
Remarques :

- L'importance (le **poids**) de chaque question peut varier d'une section à l'autre. Ce sont vos encadrants qui jugent de cette importance.
- Aucune des épreuves ne permet de tirer des conclusions sur TOUTE la matière. Il s'agit chaque fois de « coups de sonde ». Il vous appartient d'estimer s'ils sont révélateurs de l'ensemble de vos connaissances dans ces matières. Les scores **totaux** ne sont que des **repères** pour vous situer par rapport au groupe. Ce sont surtout les résultats PAR QUESTION qui ont du sens.

Identification

Sont imprimés successivement : votre université, votre faculté (et section éventuellement), le nombre d'étudiants de votre faculté ayant présenté l'épreuve, la date d'impression des résultats (jour, mois, année), la version de votre questionnaire (tous ont reçu les mêmes questions, mais pas dans le même ordre, d'où les numéros différents de formulaires), le nombre de questions de l'épreuve et votre Code CHECK-UP (Ne révélez pas votre code si vous voulez maintenir la confidentialité).

Exemple :



1.1 Descriptif de l'épreuve

- Le nombre de questions de l'épreuve (NQ).
- Le nombre de réponses que vous avez données (NR).
- Votre nombre total de réponses correctes (NRC).
- La facilité objective (FO) soit votre % de réponses correctes à l'épreuve.
- Votre Certitude Moyenne (CM) en % (calculée sur TOUTES les réponses).
- Votre Erreur Globale de Certitude (EGC) en %, soit CM-FO. L'idéal est 0 (erreur nulle). Une Erreur Globale de Certitude (EGC) négative représente une SOUS-estimation, une EGC positive représente une SUREstimation. Attention ! Une EGC proche de 0 peut être le résultat de la compensation de surestimations par des sous-estimations. Ainsi, dans l'exemple ci-dessus, l'étudiant (e) s'est sous-estimé (e) pour tous les degrés de certitude, sauf pour 20%, où il (elle) s'est surestimé (e).

1.2 Votre graphique de réalisme par Calibration

Dans le graphique de réalisme par Calibration (ci-contre), un carré noir (6 au maximum car il y a 6 degrés de certitude possibles) présente votre pourcentage de réponses correctes pour un degré de certitude donné. Cela s'appelle aussi la Facilité Objective (FO) ou Taux d'Exactitude (TE) par degré de certitude. Un étudiant (e) parfaitement « calibré » réaliste par Calibration aurait tous ses petits carrés sur la diagonale. En cas de sous-estimation, ils sont au-dessus de la diagonale, et en cas de surestimation, en-dessous de la diagonale.



Réalisme parfait



Sous-estimation



Surestimation

1.3 Détails du calcul de votre Réalisme par Calibration (pour que vous puissiez vérifier)

C. DÉTAILS DU CALCUL DE VOTRE RÉALISME

	NC	NU	DC	POC	EC	NU*EC
Vous avez réussi 1 fois sur 2 la certitude 0% soit à 50, 0% de réussite, donc avec	1	2	0%	50, 0%	-50, 0%	d'EC (total = 100)
Vous avez réussi 0 fois sur 1 la certitude 20% soit à 0, 0% de réussite, donc avec	0	1	20%	0, 0%	20, 0%	d'EC (total = 20)
Vous avez réussi 2 fois sur 4 la certitude 40% soit à 50, 0% de réussite, donc avec	2	4	40%	50, 0%	-10, 0%	d'EC (total = 40)
Vous avez réussi 23 fois sur 29 la certitude 60% soit à 79, 3% de réussite, donc avec	23	29	60%	79, 3%	-19, 3%	d'EC (total = 560)
Vous avez réussi 3 fois sur 3 la certitude 80% soit à 100, 0% de réussite, donc avec	3	3	80%	100, 0%	-20, 0%	d'EC (total = 60)
Vous avez réussi 5 fois sur 5 la certitude 100% soit à 100, 0% de réussite, donc avec	5	5	100%	100, 0%	0, 0%	d'EC (total = 0)
Votre Total (pondéré) d'Erreurs de Certitude (TEC) est (en valeur Absolue, sans compter le sens des erreurs) = 780						
Votre Erreur Moyenne (Absolue) de Certitude est 18 = TEC/NR=(780/44)=EMAC						
Votre réalisme (100-EMAC) est donc de 82 (100-17,7)						

Vérifier

Pour **chaque degré de certitude** (de 0% à 100%) apparaissent, en colonnes successives :

- votre nombre de réponses correctes (NC) et votre nombre d'utilisations (NU) de ce degré de certitude (DC),
- votre pourcentage de réussite (ou Facilité Objective FOC),
- votre Erreur de Certitude avec ce degré (EC), positive ou négative,
- le sous-total (en valeur absolue) des Erreurs de Certitude (TEC) *pour ce degré*, calculé par la formule $NU \times EC$.

La somme de ces Sous-Totaux d'Erreurs de certitude donne le Total (absolu) d'Erreurs de Certitude (TEAC) et l'Erreur Moyenne Absolue de Certitude (EMAC = TEC / NR).

On peut ainsi calculer le **Réalisme** par Calibration par la formule **100 – EMAC**.

Un indice de **réalisme** par Calibration de 80 signifie une « Erreur Moyenne Absolue de Certitude » (positive ou négative) de 20% (100-80) dans l'estimation de ses exactitudes par ses certitudes. Etant donné le petit nombre de questions, une ou deux erreur (s) peuvent à elles seules peser énormément sur l'indice de Réalisme par Calibration. Considérez donc l'ensemble de vos graphiques et de vos indices de réalisme par Calibration.

D. TABLEAU SPECTRAL

RC	R	OI	Vos Réponses Incorrectes (RI)						Vos Réponses Correctes (RC)						FO	CONTENUS DES QUESTIONS
			-100	-80	-60	-40	-20	-0	0	20	40	60	80	100		
2	2	01												100	73	corroborer
1	1	02							0						28	prérogatives
3	3	03												100	35	stigmatisé
4	4	04													54	sous-jocentes
6	5	05			-60										4	latentes
4	4	06									60				76	divergent
3	3	07									60				37	dénégations
1	3	08			-60										57	disacrédit
4	4	09								40					35	a l'instar de
5	5	010												100	50	inhérentes
5	3	011				-40									55	divages
6	2	012			-60										14	à l'instar de
5	5	013									60				42	élué
2	2	014									60				68	judicieuse
4	4	015									60				83	réfute
3	3	016									60				58	précarité
6	5	017			-60										21	présomptions
1	1	018									60				60	partiale
1	1	019										60			80	habileté
6	6	020									60				22	exhaustif
3	3	021									60				66	griets
2	0	022						0							42	émanciper
1	1	023									60				61	subordonner
1	1	024									60				66	asservir
2	2	025									60				51	afranchir
1	1	026									60				72	sonnerie
3	2	027						0							52	invective
1	1	028									60				79	ordre
3	3	029										60			76	addition
2	2	030										60			81	permis
3	3	031									60				62	inhibé
1	1	032									60				77	proscrit
2	2	033									60				81	égal
1	1	034									60				81	retarder
3	3	035									60				68	contester
1	1	036									60				61	sjourner
2	2	037									60				53	anticiper
2	4	038					-20								45	probité
1	2	039			-60										18	réactionnaire
3	3	040									60				53	exhorter
2	2	041									60				70	arbitraire
4	4	042												100	66	empirique
3	2	043				-40									29	dépédition
1	1	044									60				37	allégation
4	2	045			-60						60				30	subversion

Vos rép. Incorrectes = 12 soit 24%		Vos rép. Correctes = 24 soit 24%	
Celles du groupe = 46%		Celles du groupe = 54%	
Cert. Moy des R Incorrectes = 46%		Cert. Moy des R Correctes = 65%	
Celles du groupe = 51%		Celles du groupe = 71%	
Erreur		Exactitude	
ignorée	reconnue	peu assurée	assurée
Méconnaissance		Ignorance	Connaissance
totale	partielle	partielle	totale
Savoirs			
dangereux	inutiles	utiles	

Interprétations possibles des positions spectrales des réponses (avec certitudes)

1.4 Tableau spectral de vos performances

Haut du tableau :

- **Colonne 1** : les réponses correctes attendues (RC)
- **Colonne 2** : vos réponses (R)
- **Colonne 3** : la liste des questions, de Q1 à Q... (Les numéros des questions sur votre feuille de résultats correspondent aux numéros qu'elles portent sur votre formulaire de questions. Les étudiants qui ont eu un autre formulaire ont reçu une feuille de résultats avec les questions présentées dans un autre ordre).
- **Colonne 4** : vos réponses incorrectes positionnées par leur certitude en (**pourcents négatifs**) : la position de ces résultats doit être interprétée par la légende du bas du tableau (Erreur, Méconnaissance, savoirs dangereux, etc.).
- **Colonne 5** : vos omissions (parfois il n'y en a pas)
- **Colonne 6** : vos réponses correctes positionnées par leur certitude en (**pourcents positifs**) : la position de ces résultats doit être interprétée par la légende du bas du tableau (Exactitude, Connaissance, savoirs utilisables, etc.).
- **Colonne 7** : la **Facilité Objective** (FO en %) de la question **pour le groupe** (son taux de réussite) dont le nombre (d'étudiants) est rappelé.
- **Colonne 8** : le contenu des questions.

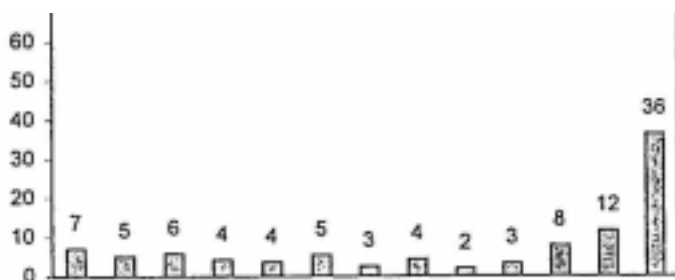
Dans la *distribution spectrale*, plus vos notes (degrés de certitude en gras) sont **positionnées** à droite, mieux c'est. Les notes très à gauche invitent à considérer pourquoi vous vous êtes trompé(e) avec une certitude élevée.

Dans le **bas du tableau spectral**, cinq lignes :

- **Ligne 1** : vos nombres et pourcentages de réponses correctes et incorrectes comparés à ceux du groupe (votre section).
- **Ligne 2** : votre indice de Prudence, c'est-à-dire vos certitudes moyennes (idéalement les plus faibles possibles) avec Réponses incorrectes et votre indice de Confiance, c'est-à-dire vos certitudes moyennes (idéalement les plus élevées possibles) avec Réponses Correctes, à nouveau comparées à celles du groupe (votre section).
- **Ligne 3** : à gauche, les deux types d'erreurs : ignorée (certitude supérieure à 50%) et recon nue ou avouée (certitude inférieure à 50%) et, à droite, les deux types d'exactitude : peu assurée (certitude inférieure à 50%) et assurée (certitude supérieure à 50%).
- **Ligne 4** : la même chose exprimée en d'autres termes : **Méconnaissance** (totale ou partielle) ou erreurs, **Ignorance** (Omissions), **Connaissance** (partielle ou totale) ou Réponses correctes.
- **Ligne 5** : La même chose à nouveau, en termes de savoirs **dangereux** (parce que certains bien qu'erronnés, on les utilisera), de savoirs **inutilisables** (qu'ils soient corrects ou incorrects, comme la certitude est inférieure à 50%, on n'osera pas les utiliser), de savoirs **utilisables** (car supérieurs en certitude à 50%).

2. Aux enseignants de la section...

- L'importance (le **poids**) de chaque question peut varier d'une section à l'autre. Ce sont les encadrants de chaque section qui jugent de cette importance et la communiquent à leurs étudiants.
- Aucune des épreuves ne permet de tirer des conclusions sur TOUTE la matière. Il s'agit chaque fois de « coups de sonde ». Il vous appartient d'estimer dans quelle mesure ils sont révélateurs de l'ensemble des connaissances des étudiants dans ces matières. Les scores **totaux** ne sont que des **repères**. Ce sont surtout les résultats PAR QUESTION qui ont du sens.
- **SVP : Rassurez** les étudiants qui auraient des résultats faibles et s'inquiéteraient excessivement.



Comment lire et interpréter les Résultats CHECK-UP de chaque épreuve par section ou faculté

Identification de l'épreuve et du groupe d'étudiants

Sont imprimés successivement : votre université, votre faculté et section (éventuellement), la date d'impression des résultats (jour, mois, année), le nombre de questions de l'épreuve, le nombre d'étudiants de votre faculté (ou section) ayant présenté l'épreuve.

Dans l'exemple qui est fourni, nous avons maquillé tout indice permettant d'identifier l'université et la section.

Considérons, pour simplifier les raisonnements ultérieurs qu'il y avait 100 étudiants, si bien que les nombres d'étudiants et les pourcentages seront, *dans cet exemple*, les mêmes.

2.1 Graphique spectral de la distribution des réponses à l'ensemble de l'épreuve

Comment le lire :

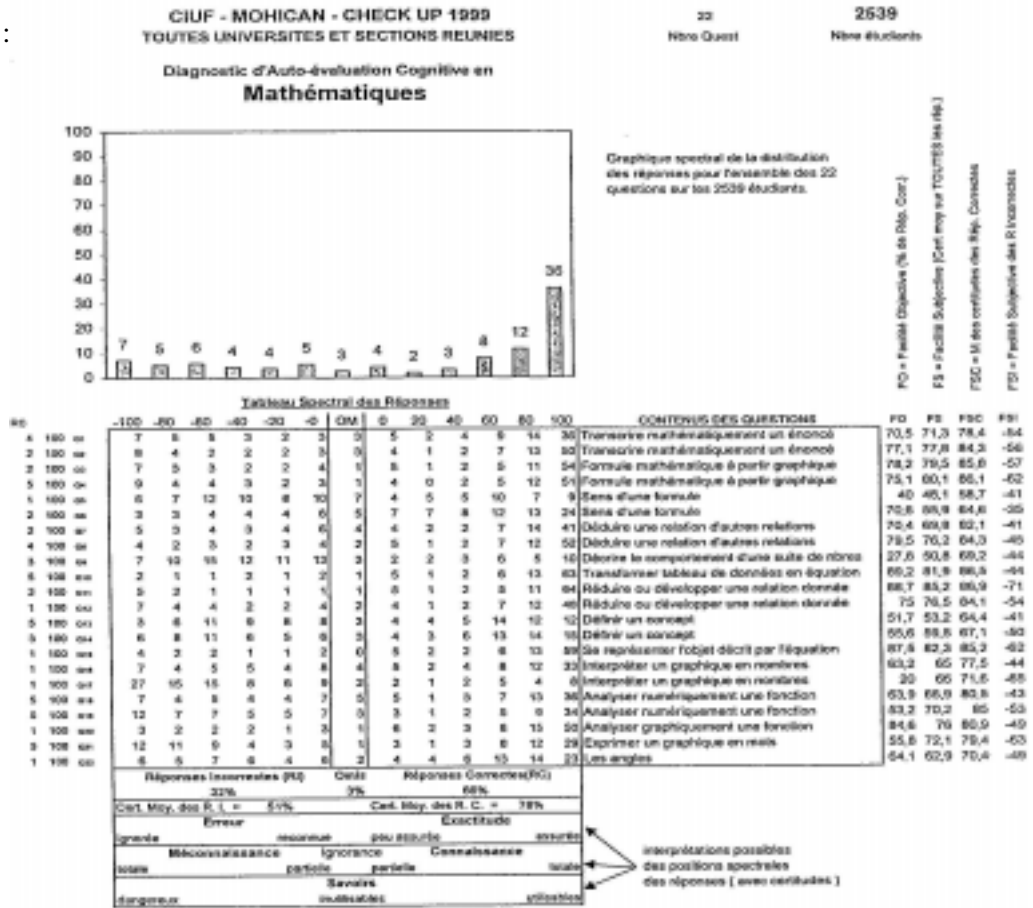
- L'axe horizontal représente le spectre des qualités des performances. La pire est à l'extrême gauche et est notée -100 pour rappeler qu'il s'agit des réponses Incorrectes avec certitude 100 %. La meilleure est à l'extrême droite (réponses Correctes avec 100 %). Les six catégories de gauche constituent l'hémispectre gauche, les six de droite l'hémispectre droit. On s'attend à ce que leurs distributions soient en « J ».
- Chaque bâtonnet est surmonté du pourcentage (ARRONDI à l'unité).
- Les 3 bâtonnets de gauche représentent les taux (%) de Réponses Incorrectes (RI) avec les 3 Degrés de Certitude (DC) les plus élevés, de gauche à droite : 100%, 80% et 60%, c'est-à-dire supérieurs à 50%. Ce sont les pires des façons de se tromper, et cela constitue la **connaissance nuisible**, voire dangereuse ou encore la **Méprise** (ici **18%**, c'est-à-dire 7%+5%+6%).
- Les 3 bâtonnets à leur droite représentent les taux de Réponses Incorrectes avec certitudes faibles (40%, 20% et 0%) ainsi que l'omission qui constitue l'**ignorance reconnue ou avouée** ou encore la **Méconnaissance**. Ici, **16%** (4%+4%+5%+3%).
- Les trois bâtonnets qui suivent représentent les Réponses Correctes (RC) avec les 3 Degrés de Certitude (DC) les plus faibles (0%, 20%, 40%), c'est-à-dire la **connaissance partielle** ou encore **Mi-connaissance**. Ici **9%** (4%+2%+3%).
- Enfin les trois bâtonnets de droite représentent les Réponses Correctes (RC) avec les 3 Degrés de Certitude (DC) les plus élevés (60%, 80%, 100%), c'est-à-dire la **connaissance assurée** ou encore la **Maîtrise**. Ici 56% (8%+12%+36%).
- La meilleure des façons de fournir une réponse correcte est représentée par le bâtonnet de l'extrême droite du graphique (ici 21%) puisqu'il s'agit de réponses accompagnées du degré de certitude 100%, ou encore **Maîtrise totale**.
- Le bâtonnet du milieu représente le taux d'**OMISSIONS** (non réponses), ici **3%**.
- L'idéal serait une absence de bâtonnets dans les 7 premières positions, ou alors, une concentration des réponses incorrectes sur les certitudes les plus faibles (20% et, encore mieux : 0%), avec une **forme en J de la distribution dans les réponses incorrectes** (dans l'hémispectre gauche).
- De même, l'idéal serait, pour les réponses correctes une **courbe « en J »** des sommets des bâtonnets de cet hémispectre, ce qui signifierait que les **réponses correctes** se concentrent sur les valeurs élevées des degrés de certitude (80% et, mieux encore, 100%). On a bien une telle courbe en effet dans l'exemple ci-contre pour l'hémispectre droit.

Des améliorations de performances consisteraient non seulement en un transfert de réponses de la gauche (incorrectes) ou du centre (omissions) vers la droite du graphique (correctes), mais aussi en un déplacement des incorrectes vers plus de doute (de réserve) et des correctes vers plus de certitude (d'assurance). On peut donc mesurer des modifications subtiles de capacités, ce qui est bien nécessaire car l'apprentissage ne progresse pas toujours « de rien à tout », ce que les distributions, très étalées, des réponses à ces épreuves, montrent. Si la connaissance était affaire de Tout ou Rien, il n'y aurait que deux bâtonnets : ceux des 2 extrêmes.

Nous vous invitons à discuter de ces divers types de performances avec les étudiants, en vous étant assuré(e) qu'ils ont bien compris les concepts et sont à même d'interpréter leurs graphiques, pour lesquels ils ont reçu les mêmes explications que ci-dessus.

2.2 Le tableau spectral des Réponses

Exemple :



De gauche à droite, la **signification des colonnes** est la suivante, question par question :

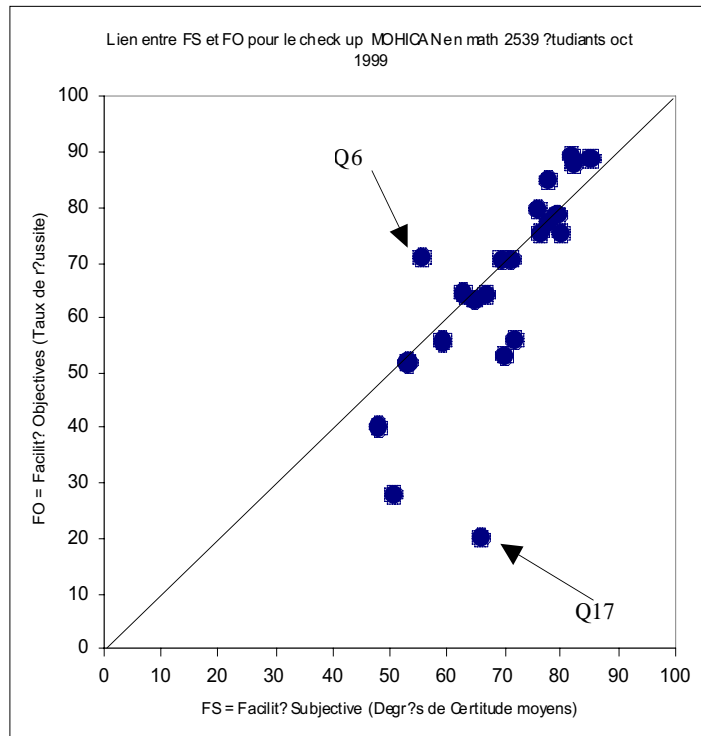
- **Colonne 1** : la RC ou Réponse Correcte attendue
- **Colonne 2** : le total (en principe 100%) des taux (%) de réponses répartis dans les 13 positions.
- **Colonne 3** : le numéro de la Question (précédé de la lettre Q) dans la forme 1 de l'épreuve (ce n'est pas le même numéro dans la forme 2)
- **Colonnes « -100 » à « -0 »** : Le taux de Réponses Incorrectes (RI) avec les degrés de certitude 100 (codage : « -100 »), 80%, 60%, 40%, 20% et 0% pour cette question
- **Colonne « OM »** : taux d'Omissions
- **Colonnes « 0 » à « 100 »** : taux de Réponses Correctes avec les Degrés de Certitude correspondants
- **Colonne « Contenus des Questions »** : sans commentaire
- **Colonne « FO »** : Facilité Objective ou Taux (%) de Réponses Correctes (RC) pour cette question pour la section considérée (ici l'exemple est celui de toutes les universités et sections réunies)
- **Colonne « FS »** : Facilité Subjective ou Certitude Moyenne (les omissions non prises en compte)
- **Colonne « FSC »** : Moyenne des Certitudes des seules Réponses Correctes (Indice de Confiance).
- **Colonne « FSI »** : Moyenne des Certitudes des seules Réponses Incorrectes (sans les omissions) (Indice de Prudence).

Que pouvons-nous en tirer ?

1-L'amplitude de variation (AV) des Certitudes Moyennes ou Facilités Subjectives (FS) est bien plus étroite (de 48,1 à 85,2, soit une AV de 37,1 seulement) que celle des Taux d'Exactitude ou Facilités Objectives (FO) (quasi la moitié, comme le montre le graphique) car il y a des « surprises » pour certaines questions mal réussies (ex : Q17). **La corrélation entre FS et FO est élevée (0,76)**, mais des surestimations sont **compensées** par des sous-estimations.

Sur les 22 questions, FS est 15 fois très proche de FO (moins de 5% d'écart en moyenne), ce qui indique en général un réalisme par question (ce qui ne signifie pas un réalisme de chaque étudiant).

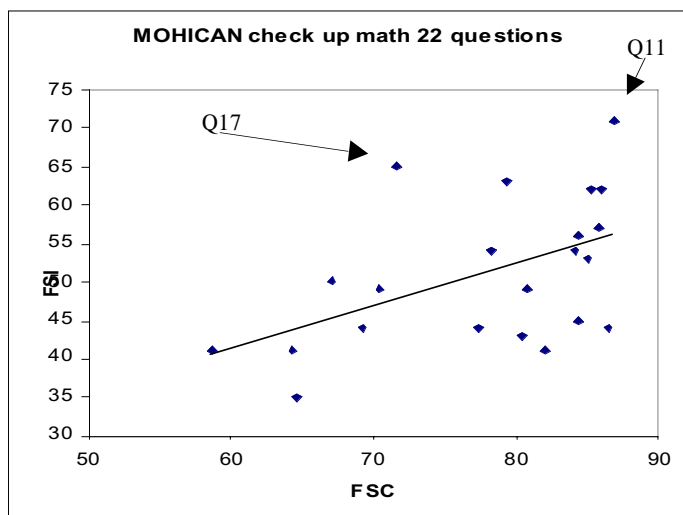
Pour 5 questions Q5, Q9, Q17, Q19 et Q21, FO est inférieur à FS, ce qui indique, pour ces 5 questions une surestimation. Enfin, pour la question Q6, FO est supérieur de 14,7% à FS, ce qui indique pour cette question une sous-estimation. Les surestimations (plus nombreuses que les sous-estimations) se concentrent sur certaines questions (pour Q17 elle est de 46% !), dont il faut donc discuter avec les étudiants pour rectifier les erreurs.



2-Les FSC (Facilités Subjectives des Réponses Correctes) ou indices de Confiance sont proches les uns des autres, quelles que soient les questions : les FSC vont de 58,7% à 86,9% (AV = 28,2%).

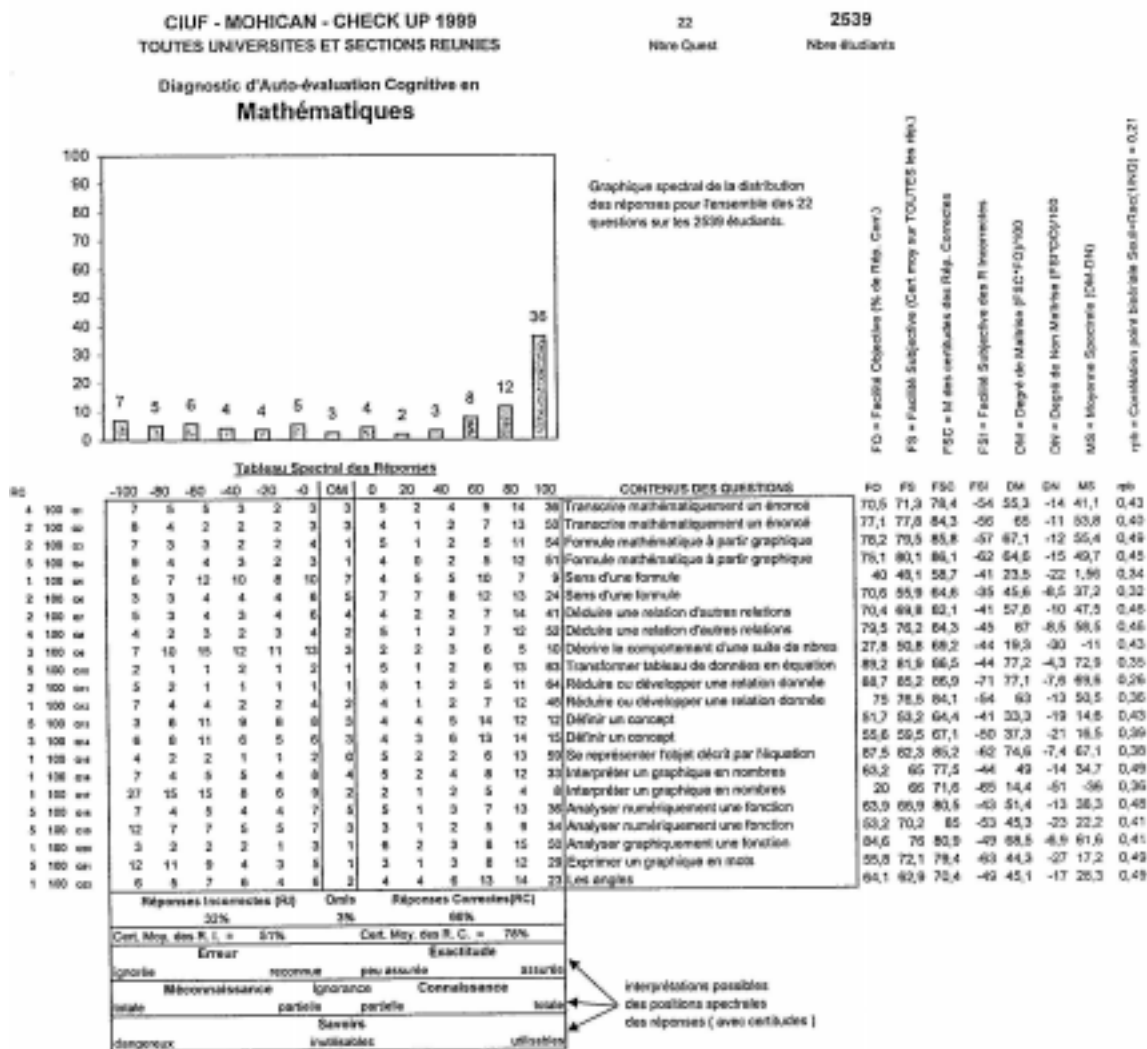
Les FSI (Facilités Subjectives des Réponses Incorrectes) ou indices de Prudence sont aussi proches les uns des autres : les FSI vont de 35% à 71% (AV (ou Ampleur de Variation) = 36%).

Leur corrélation est de 0,50, ce qui signifie que certaines erreurs se font avec une certitude anormalement élevée (ex : Q11 et Q17), ce dont il faut discuter avec les étudiants.



Toutes les valeurs de Prudence ou de FSI supérieures à 50% (à peu près la moitié des questions sont dans ce cas) constituent un problème : il n'est pas normal que la certitude moyenne de réponses erronées soit supérieure à 50% ! Notons que cette valeur-repère ou seuil pourrait être différente

d'une matière à l'autre, et même d'une question à l'autre. Ainsi, en éducation du patient à sa maladie, on se fixe comme valeur repère le seuil de 90% de certitude pour accompagner les réponses correctes, et dans l'information de prévention du SIDA, 100%.
Faisons deux pas de plus dans l'analyse (les 5 dernières colonnes).

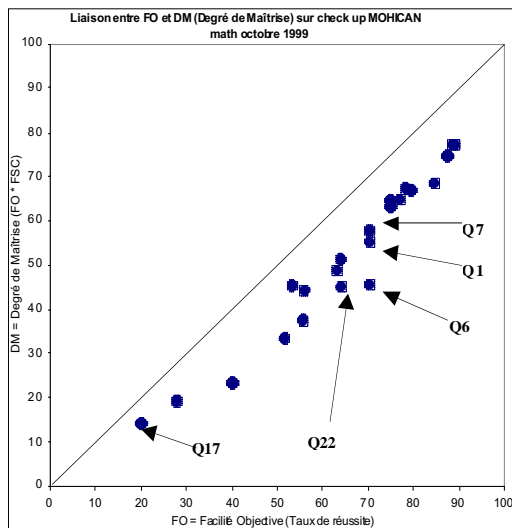


- **Colonne «DM» : Degré de Maîtrise ou Surface de Compétence (SurfC)** soit FO x FSC. Il s'agit donc du taux d'exactitude, nuancé par la certitude moyenne (ou l'indice de confiance) avec laquelle les réponses correctes ont été fournies. La valeur de SurfCDM est égal ou inférieur à FO.
- **Colonne «DN» : Degré de Non Maîtrise, ou Surface d'Incompétence (SurfI)**, soit DO x FSI (Difficulté Objective valant 100- (FO+OM)). Il s'agit donc du taux d'erreur nuancé par l'indice de Prudence. On peut en effet se tromper avec une certitude faible ou forte... et ce n'est pas la même chose !
- **Colonne «MS» ou Moyenne Spectrale**, calculée par DM-DN (ou SurfC – SurfI) ou **Surface de Maîtrise**. Elle ne tient pas compte des omissions. MS ou SurfM est inférieure ou égale à SurfC. Il s'agit d'un indice «résumant» l'information contenue dans DM et DN, ou SurfC et SurfI, au risque de cacher certaines différences. En effet, une même valeur de MS ou SurfM peut résulter de diverses combinaisons de DM et DN, de SurfC et SurfI !!

Que pouvons-nous en tirer ? (Exemple sur le Check-up de Mathématiques)

1. L'indice SurfC¹⁶ (ou DM) est proche de l'indice FO ($r=0,76$), mais tient compte du fait que pour certaines questions, la certitude des réponses correctes est plus faible que pour d'autres. Ainsi Q1, Q6 et Q7 ont des FO équivalentes : 70,5%, 70,6% et 70,4%. Cependant, leurs FSC sont différentes : respectivement 78,4%, 64,5% et 82,1%.

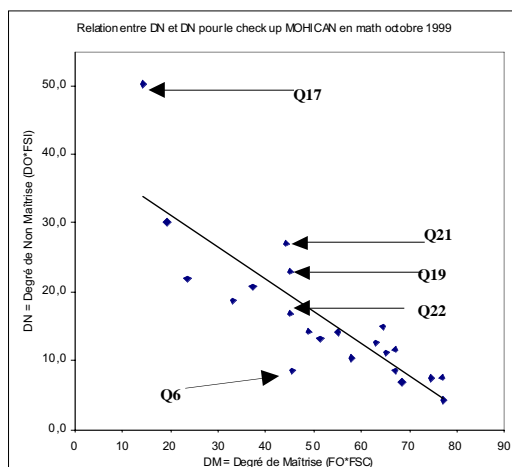
Le parallélisme entre FO et SurfC (ou DM) indique que même quand le taux de réponses correctes (FO) est très faible (ce qui est le cas pour la question Q17 (réussie à 20% seulement), la certitude moyenne ou Facilité Subjective moyenne de ceux qui ont fourni une réponse Correcte (FSC) est presque aussi élevée que pour les autres questions. Ainsi la FSC de Q17 est 71,6% et de 70,4% pour Q22 dont la FO vaut 64,1%



2. L'indice SurfI va de -7,4 à -51.

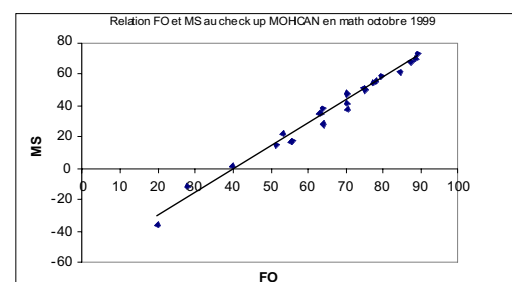
Plusieurs questions peuvent avoir le même Degré de Maîtrise ou SurfC, tout en ayant des Degrés d'Incompétences (ou de Non Maîtrise ou SurfI) différents. C'est le cas pour les questions Q6, Q19, Q21 et Q22 dont le DM est de 45 environ. Par contre, leurs DI sont respectivement 8,5 ; 23 ; 27 et 17. Il apparaît donc qu'à égalité de Degré de Maîtrise (SurfC), Q6 est la question des quatre qui est l'objet de la plus faible INCOMPÉTENCE et Q21 de la plus grande incompétence.

La Q17 atteint le record d'incompétence pour l'épreuve (son DN = 51). Sa position tout à fait atypique sur le graphique indique qu'il importerait que le professeur et les étudiants comprennent en profondeur la nature d'une erreur d'une telle ampleur et ses causes.



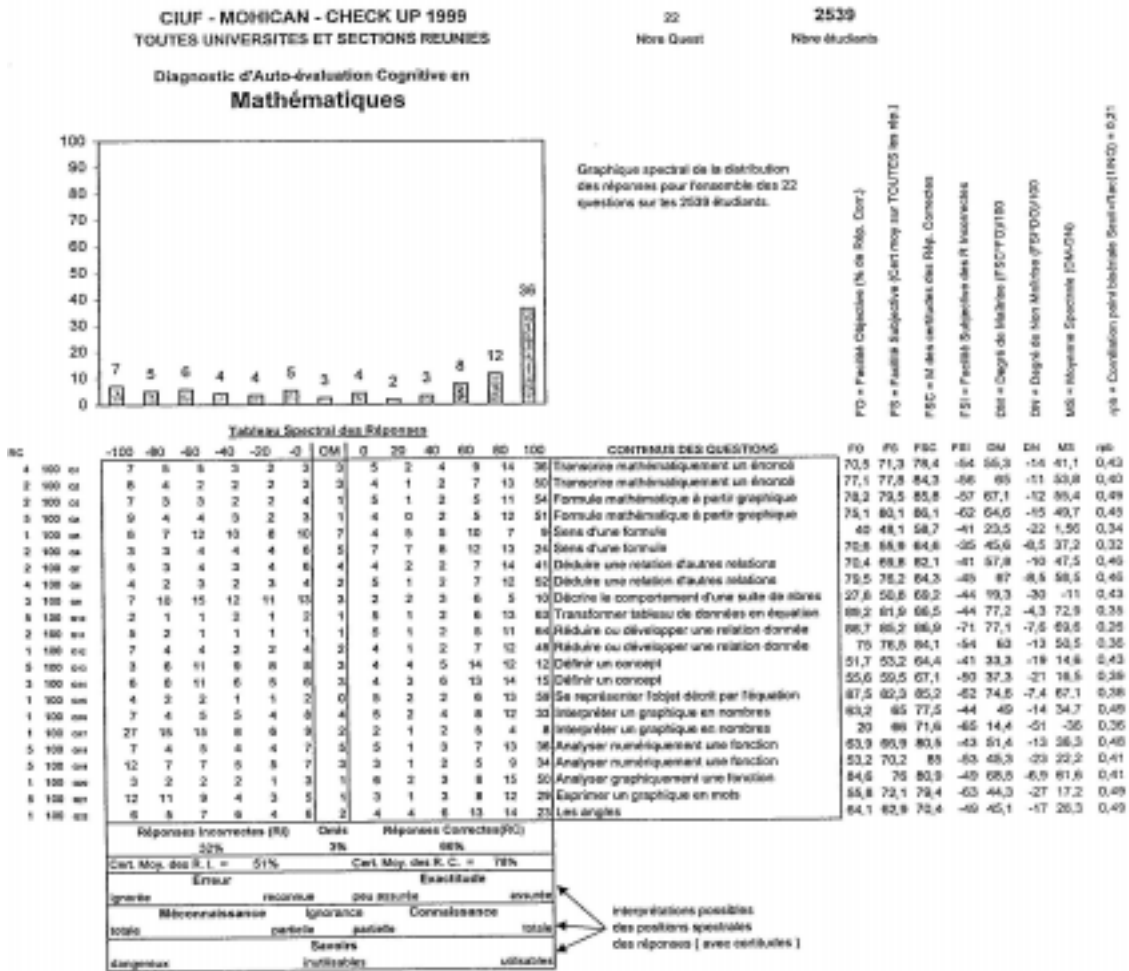
3. L'indice MS (SurfM) d'une question est plus étendu, plus subtil (ici il va de -36 à 72,9 soit une Amplitude de Variation de 98,9) que FO qui va, lui, de 20 à 89,2, soit une amplitude de variation (AV) de 69,2. La corrélation entre MS (ou SurfM) et FO est très élevée (ici =0,99) dans ce cas-ci, mais ne l'est pas toujours.

Ces indices par question sont aussi calculables par étudiant, pour qui ils sont précieux pour nuancer la profondeur de la maîtrise... et de l'incompétence.



¹⁶ Rappelons que la Surface de Compétence (SurfC) est le produit de la Facilité Objective (FO) et de la Facilité Subjective moyenne de ces réponses correctes : SurfC = FO * FSC Ou FO*Confiance.

3. Les indices de cohérence interne des questions (les 2 dernières colonnes)



La colonne « rpb » ou corrélation point bisériale fournit la valeur de la corrélation entre le choix ou non (0/1) d'une solution (ici n'a été repris dans cette colonne que le choix de LA solution correcte) à cette question et le score au total de l'épreuve (sur les NQ questions). Comme le score (0/1) à chaque question « contribue » au total, il est donc « automatiquement » corrélé avec ce total. La valeur de cette « **corrélation automatique** » due au « recouvrement » est calculable par la formule « $1/\sqrt{NQ}$ » où NQ est le nombre de questions de l'épreuve. Pour une épreuve comptant 22 questions, $\sqrt{22}$ vaut 4,69. La corrélation automatique vaut donc $1/4,69 = 0,21$. Nous nous attendons donc à ce que les rpbis des réponses correctes aient une valeur égale ou supérieure à 0,21. C'est le cas pour TOUTES les 22 questions du check-up de mathématique.

Les 5 dernières lignes (du bas) du tableau indiquent :

Ligne 1 : De gauche à droite, les taux de réponses incorrectes, omises, correctes.

Ligne 2 : La certitude moyenne (CM) ou Facilité Subjective des réponses Incorrectes (FSI) et des correctes (FSC). Lignes 3, 4, 5 : trois façons de lire les mêmes données.

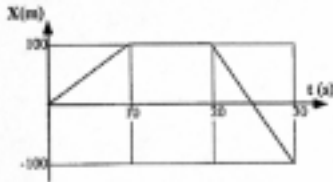
Ces indices rpbis sont fournis pour donner aux enseignants des repères sur la « qualité » de chaque question. Ils sont le résumé des valeurs détaillées question par question et solution par solution (voir chapitres 11 à 20, check-up par check-up).

4. Illustration

Illustration de l'application de ces indices à deux questions (Q17 et Q18) d'une même épreuve (passée par les mêmes étudiants) : l'épreuve de mathématique (Lebrun et Lega).

Q17 est réussie par peu d'étudiants (20%), mais ceux-ci ont une certitude assez élevée (71,6%) alors que Q18 est trois fois mieux réussie (63,9%) et ces 63,9 % étudiants ont une certitude moyenne de 9% à peine plus élevée (80,5%) que les 20% d'étudiants qui ont réussi Q17.

Question 17 : La position X (exprimée en mètres) d'un mobile est donnée en fonction du temps t (exprimé en secondes) sur le graphique suivant,



que vaut la vitesse du mobile au temps. $t = 25$ s ?

Surface de Compétence		14,30%	%	r.bis	CM
1. -20 m/s		20	0,36	71,6	
2. 0 m/s		50,3	-0,35	66,5*	
3. 5 m/s		1,5	-0,09	45,3	
4. 10 m/s		2,9	-0,01	54,6*	
5. 20 m/s		17,9	0,18	74*	
6. aucune		5,7	-0,07	46,5	
7. toutes		0,1	-0,02	50	

Un cinquième des étudiants (20%) ont réussi, mais ce sont de bons étudiants ($r.bis = 0,36$), avec une certitude moyenne assez bonne (71,6%), donc avec **confiance**. La moitié des étudiants (50,3%) ont choisi le distracteur 2, avec une certitude élevée pour un distracteur (66,5%), donc avec **imprudence**, ce qui rend la **fracture subjective** ténue : 5,1 % pour ces des étudiants faibles ($r.bis = -0,35$). Le distracteur 5, par lequel on fait une erreur de signe, est choisie par près d'un cinquième des étudiants (17,9%), des étudiants plutôt bons ($r.bis = 0,18$) et une certitude moyenne (74%) tellement élevée pour un distracteur, donc une telle **imprudence**, que la **fracture subjective** aberrante entre ce distracteur et la réponse correcte est négative (-2,4%) !!!

Question 18 : Le tableau ci-dessous fournit quelques points des deux paraboles. Identifiez la ou les valeur(s) de X correspondant à d'éventuels points d'intersection des deux courbes.

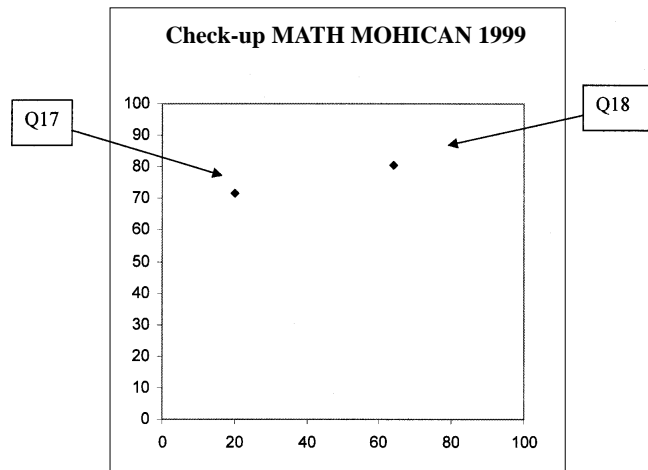
$$Y_1 = -X^2 + 3X - 12 \quad \text{et} \quad Y_2 = X^2 + X - 36$$

X	-4	-3	-2	-1	0	1	2	3	4	5
Y1	-40	-30	-22	-16	-12	-10	-10	-12	-16	-22
Y2	-24	-30	-34	-36	-36	-34	-30	-24	-16	-6

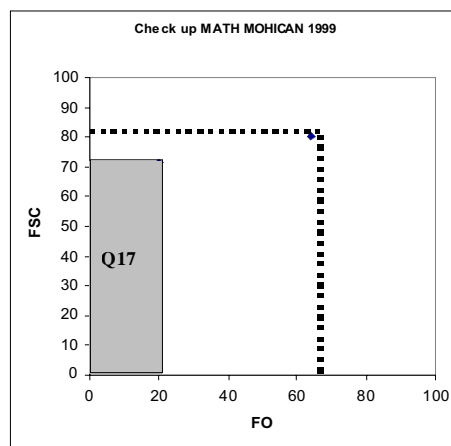
Surface de Compétence		51,40%	%	r.bis	CM
1. $X = -3$		11,6	-0,17	63,5*	
2. $X = -3$ et $X = 2$		2,6	-0,17	44,8	
3. $X = 2$ et $X = 4$		2,9	-0,15	20,5	
4. $X = 2$		1,1	-0,12	29	
5. $X = -3$ et $X = 4$		63,9	0,48	80,5	
6. aucune		11,8	-0,19	49,6	
7. toutes		0,6	-0,05	30,7	

Près des deux tiers des étudiants (63,9%) ont choisi la solution correcte, de très bons étudiants ($r.bis = 0,48$) avec une certitude moyenne élevée (80,5%), donc avec une **confiance élevée**. Deux distracteurs (le 1 et le 6.) se disputent la vedette (choisis par 11,6% et 11,8% des étudiants), par des étudiants du même niveau de faiblesse ($r.bis = -0,17$ et $-0,19$), mais avec des certitudes moyennes assez différentes : 63,5% (**Imprudence**) et 49,6%. Cette dernière certitude moyenne est probablement plus faible que l'autre parce qu'elle porte sur le choix du distracteur "6. Aucune". La **fracture subjective** avec ce distracteur-vedette est de 30,9 % (élevée).

Traduites graphiquement, les valeurs FO (Facilité Objective) et FSC (Indice de Confiance) permettent de situer Q17 et Q18 comme suit :



Les Surfaces de Maîtrise, elles, se représentent comme suit :



Ceci permet mieux de visualiser les Degrés de Maîtrise de ces deux questions.

5. Conclusion du chapitre

Les concepts, formules, indices et graphiques qui ont été présentés dans ce chapitre peuvent apparaître comme beaucoup d'innovations. Nous les pensons nécessaires. Dans cette « société de la connaissance », il est temps que l'on traite celle-ci avec les degrés de précisions, de profondeur et d'intensité requis pour donner des feedbacks précis non seulement aux apprenants, mais aussi aux enseignants et aux pilotes du système éducatif.

Chapitre 3

Résultats **Check-up par check-up** **et corrélations entre eux**

D. Leclercq, ULg, Président du Groupe de Travail CIUF « Réussite en candidatures »

J.-L. Gilles, Directeur du SMART, Université de Liège

F. Georges, Chercheur au LabSET, Université de Liège

C. Dupont, Chercheur au LabSET, Université de Liège

P. Detroz, Chercheur au SMART, Université de Liège

1. Présentation des résultats

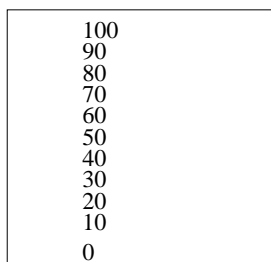
Les résultats concernant chacun des 10 check-ups sont concentrés sur deux pages se faisant face et structurées en quatre colonnes.

Dans la **colonne 1** (de gauche) sont présentés successivement

- Les critères de choix ou de création des questions.
- Les trois **valeurs** statistiques **classiques** pour un test :
 - le Pourcentage de Réponses Correctes (ici le pourcentage MOYEN pour tous les étudiants et toutes les questions) ou **%RC** ou FO (Facilité Objective)
 - le Pourcentage d'Erreurs
 - le Pourcentage d'Omissions
- L'analyse spectrale avec
 - le **graphique spectral** (la répartition ou le profil) de la qualité des réponses allant de -100 (Réponses Incorrectes avec certitude 100) à +100 (Réponses Correctes avec certitude 100).
 - les indices de **prudence**, de **confiance**, de **discriminance**
 - l'indice de Maîtrise Totale (Les Réponses Correctes avec 100% de certitude),
 - l'indice de Méprise Totale (Les Réponses Incorrectes avec certitude 100%)
 - l'indice de Facilité Subjective des Réponses Correctes (**FSC**) et des Réponses Incorrectes (**FSI**)
 - l'indice de fracture subjective
 - les connaissances utiles, inutilisables et dangereuses.

Dans la colonne 2 : L'échelle des taux de réussite

Dans ce type de représentation, chaque question occupe sur l'échelle la position correspondant à son indice de facilité Objective (FO) ou Taux de réussite.



Les **colonnes 3 et 4** présentent les distributions de 4 indices sur les étudiants testés :

MS = Moyennes Spectrales

FO = Facilités Objectives

EMC = Erreur de Centration

Réal = Réalisme par Calibration

Ces indices sont calculés pour les **Filles** et les **Garçons**, et l'Ampleur de l'Ecart (**AE**) entre ces deux groupes est indiqué quand il est statistiquement significatif.

L'AE est calculé par la différence des deux moyennes (Celle des filles moins celle des garçons ou l'inverse selon le sens de la différence) divisée par l'écart-type de la distribution totale.

Ces indices sont aussi calculés pour les étudiants âgés de **18 ans et -** et ceux âgés de **19 ans et +**. Enfin, dans une grille, une série de **corrélations** entre divers indices relatifs à ce check-up.

2. Les nombres de base des calculs

La **distribution spectrale et les indices FO et FS** ont été calculés sur un maximum d'étudiants, même quand certaines données manquaient, par exemple leur âge, leur sexe ou leur résultat à une autre épreuve. Ainsi, pour le check-up Vocabulaire, ce nombre est 3905.

Les **corrélations intertests** ont été calculées sur tous les étudiants pour lesquels nous disposions de résultats à tous les check-up prévus, même si nous ignorions le sexe ou l'âge. Pour l (épreuve de vocabulaire, ce nombre est 3801.

Les **valeurs par sexe** ont été calculées sur les étudiants pour lesquels ces données étaient disponibles. Pour le vocabulaire, ce nombre est 3724 (1720 garçons et 2004 filles).

Les **valeurs par âges** ont été calculées sur les étudiants pour lesquels ces données étaient disponibles. Pour le vocabulaire, ce nombre est 3713 (2395 de 18 ans et moins ; 1318 de 19 ans et plus).

Nous ne commenterons pas tous les résultats (plus de 80 valeurs numériques par test), fournissant aux lecteurs des **données brutes** permettant le cas échéant d'élaborer des hypothèses.

3. Le format QCM des questions

Les questions des 10 check-up se présentent sous le format « QCM » (Questions à Choix Multiple) pour deux raisons fondamentales :

- a) Nous voulions *donner* en moins d'un mois *un feedback très détaillé* à 4000 étudiants sur 7 check-up en moyenne, soit 28 000 feuilles de feedback personnalisées (confidentielles) « distribuées » dans 61 sections de 8 universités. Dans ce contexte, il s'est avéré que seuls les QCM permettraient de traiter la quantité de données voulue, avec la rapidité voulue, avec l'objectivité voulue et dans les limites (non voulues) des budgets.
- b) Nous voulions *habituer les étudiants* à une forme de QCM sollicitant la vigilance cognitive et se rapprochant des Questions Ouvertes.

4. L'effet des solutions « AUCUNE » et « TOUTES »

4.1 Une volonté

Le groupe de travail CIUF pilotant le projet MOHICAN a voulu ajouter aux solutions traditionnelles des QCM les deux solutions *générales* AUCUNE et TOUTES dans les consignes des 10 épreuves, 5 seulement contenant effectivement des questions dont la réponse correcte était AUCUNE ou TOUTES, appelées en bref questions à RC « AUCUNE » et questions à RC « TOUTES ».

4.2 Une surprise ?

Ces solutions ont pu désarçonner les étudiants non habitués. En effet, la plupart des questions à RC « AUCUNE » se retrouvent au bas de l'échelle des réussites dans les check-up de syntaxe, de vocabulaire, de connaissances artistiques.

4.3 A y regarder de plus près

On pourrait juger inélégant de « piéger » les étudiants avec de telles solutions générales inhabituelles. Nous avons au contraire estimé que les étudiants doivent s'y mettre *le plus tôt possible* (à l'entrée à l'Université, c'est déjà bien tard). D'ailleurs, ces questions **dont la solution correcte est AUCUNE** sont-elles si « vicieuses » que cela ? En fait, elles demandent, il est vrai, de *rejeter toutes* les solutions incorrectes. Ne pas le faire, *accepter l'une d'elles est-il pardonnable* ou « acceptable » ? Nous laissons le lecteur juge sur base de la distribution des réponses¹⁷ (en %) sur trois questions de Vocabulaire et une question d'Histoire.

Vocabulaire			Histoire
A l'instar de	Exhaustif	Latent	Date de la bombe atomique sur Hiroshima
35 % à l'insu de	34 % AUCUNE	52 % ancien	37 % 1946
21 % le contraire de	20 % critique	20 % profond	25 % 1943
19 % AUCUNE	14 % pessimiste	12 % AUCUNE	21 % AUCUNE
9 % maladroitement	12 % accentuant	6 % patent	6 % 1955
7 % exagérément	8 % élogieux	3 % douloureux	4 % 1950
6 % en désobéissant	6 % ironique	2 % Toutes	4 % 1940
0,6 % Toutes	0,4 % Toutes	4 % Omissions	0,4 % Toutes
3 % Omissions	6 % Omissions		3 % Omissions
Sur 3905 étudiants	Sur 3905 étudiants	Sur 3905 étudiants	Sur 1399 étudiants

Tout spécialement pour les trois questions de Vocabulaire, il nous paraît *impensable* de choisir n'importe laquelle des solutions proposées comme synonyme. Il *faut avoir le courage de les rejeter toutes*, seule preuve de la compréhension de ces termes. Ceci montre aussi que les QCM peuvent être construites selon des **consignes plus subtiles** que la consigne classique « *l'une des solutions proposées est correcte* », consigne exécrationnelle car encourageant excessivement le « *vogelpik* ». Ce terme néerlandais signifie « *blind guessing* » ou « choix aveugle », qu'il ne faut pas confondre avec « *educated guessing* » ou « *informed guessing* », c'est-à-dire « choix en état de certitude non maximale », ou en état de « connaissance partielle ».

¹⁷ Résultats présentés dans chaque colonne par ordre décroissant de « popularité » des solutions.

Vocabulaire

(Auteur M. Monbalin, FUNDP) : 8 universités – 3905 étudiants

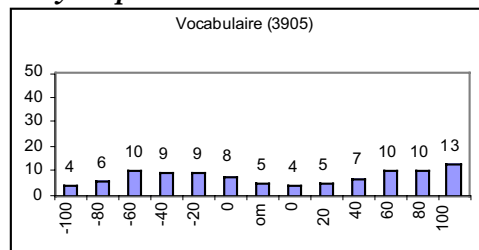
Choix des questions

Les 45 termes de l'épreuve ont été choisis sur base d'une enquête de l'auteur auprès d'enseignants universitaires indiquant quels mots étaient mal compris. On ne s'étonnera donc pas que les résultats confirment le faible taux de compréhension, l'intérêt du check-up étant d'informer précisément chaque étudiant (car un mot compris par l'un ne l'est pas par l'autre) et chaque section (car le phénomène peut être plus accentué dans certaines sections, et surtout, les mots importants ne sont pas forcément les mêmes dans toutes les sections).

Indices globaux (voir spectre ci-dessous)

Le taux de réussite moyen (FO) est 49%, avec 45% d'erreurs et 5% d'omissions, sur 3905 étudiants testés.

Analyse spectrale



Qualité spectrale : IA-ID-OM-CD-CA = 20-26-5-16-33

Confiance Collective 62,9 %

Prudence Collective 43,9 %

Discriminance Collective 18,9 %

Maîtrise totale (RC 100 %) 13 %

Méprise totale (RI 100 %) 4 %

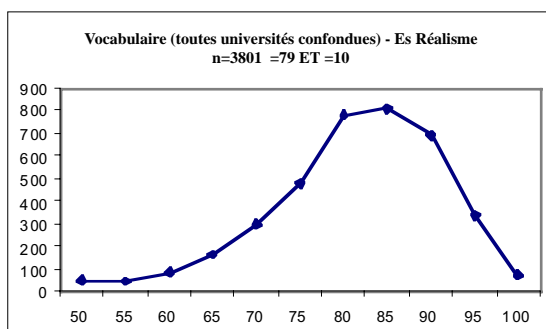
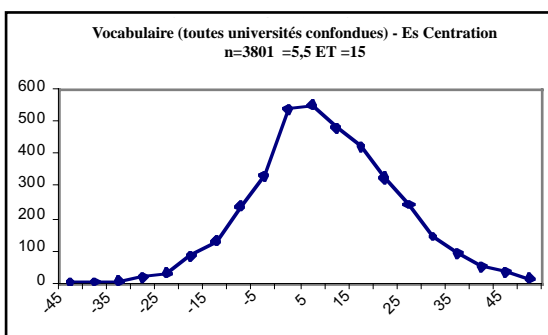
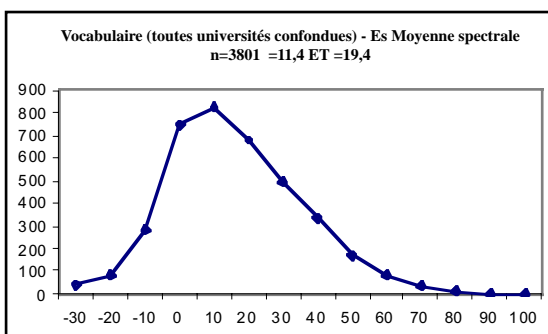
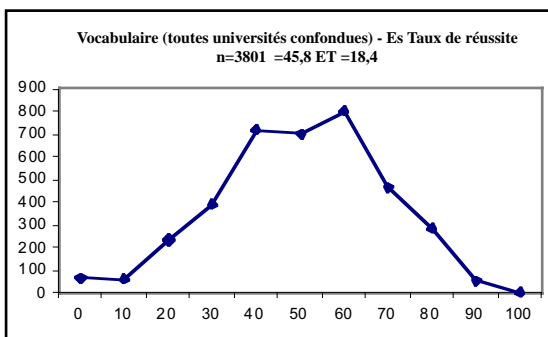
A titre de comparaison, le taux de certitude moyen ou Facilité Subjective des réponses Correctes (FSC) vaut 80 % en Mathématique (mais ce ne sont pas tout à fait les mêmes étudiants et surtout pas les mêmes questions). Par contre, le taux moyen de certitude ou Facilité Subjective des réponses Incorrectes (FSI) est de 55 % en Math, ce qui est plus regrettablement élevé ! **La fracture subjective (FSC-FSI) en mathématique vaut 27%**, ce qui est mieux qu'en vocabulaire.

33% des connaissances sont **utiles** (réponses correctes avec plus de 50% de certitude) et **20%** des connaissances sont **dangereuses** (réponses incorrectes avec plus de 50% de certitude) ; **47%** sont **inutilisables**.

Voici les termes classés par ordre (approximatif) décroissant de réussite (dont le % est à gauche du mot) :

100 %	
90	84,3 réfuter
80	
	76,7 ordre 75,5 permis
	74,1 divergent 73,9 addition
	71,0 judicieux 71,1 proscrit
70	70,6 habilité
	69,8 légal 69,4 retarder
	67,7 asservir 66,6 corrobore
	63,8 sommation
	61,1 inhibé
60	60,6 contester
	58,2 discrédit
	57,1 arbitraire
	54,1 sous-jacent 54,0 ajourner
	51,2 empirique
50	50,1 subordonner
	49,2 invectives 48,1 griefs
	47,9 affranchir 46,6 précarité 46,7 anticiper
	45,3 clivages 44,4 partielle 43,9 inhérentes
	41,3 émanciper 41,3 présomptions (AUCUNE)
40	39,9 dénégations
	35,7 exhorter 35,6 allégation
	34,6 inféré 34,5 exhaustif (AUCUNE)
	32,6 éludé 31,8 stigmatisé 31,6 subversion
	31,1 déperdition
30	
	28,4 prérogatives 27,2 probité
20	
	19,3 réactionnaire
	18,6 à l'instar de (AUCUNE)
	11,9 latentes (AUCUNE)
10	
0%	

Vocabulaire



Taux de Réponses Correctes

	n	μ		n	μ
Garçons	1720	48,6*	18-	2395	46,1
Filles	2004	43,6*	19+	1318	45,6

L'AE en faveur de ces garçons vaut +0,27

Moyenne Spectrale

	n	μ		n	μ
Garçons	1720	14,3*	18-	2395	11,5
Filles	2004	9,1*	19+	1318	11,5

L'AE en faveur de ces garçons vaut +0,26

Erreur Moyenne de Centration (Idéal = 0)

	n	μ		n	μ
Garçons	1720	6,6*	18-	2395	4,5
Filles	2004	4,4*	19+	1318	7,1

L'AE de l'Erreur en faveur de ces garçons vaut +0,15, ce qui signifie que les filles se surestiment moins que les garçons.

Réalisme par Calibration (Idéal = 100)

	n	μ		n	μ
Garçons	1720	78,4	18-	2395	78,5
Filles	2004	79,1	19+	1318	79,2

Corrélation avec	Vocab.	Syntaxe	Compr.	Géo.	Math.	Phys.	Chimie	Biologie	Art	Histoire
En Moyenne spectr.	-----	0,49*	0,32*	0,39*	0,37*	0,28*	0,27*	0,28*	0,42*	0,45*
En Taux de Réussite	-----	0,52*	0,37*	0,34*	0,39*	0,34*	0,30*	0,30*	0,47*	0,48*
En Erreur de Centr.	-----	0,62*	0,49*	0,49*	0,50*	0,45*	0,37*	0,46*	0,43*	0,48*
En Réalisme Calibr.	-----	0,35*	0,19*	0,24*	0,26*	0,18*	0,14*	0,20*	0,24*	0,27*
N	3801	3526	3236	3478	2110	2114	2127	2129	1228	1241
nombre de questions	45	12	6	10	22	10	8	10	25	25

* L'astérisque indique les différences ou corrélations significatives à $p < 0,05$. Les autres sont Non Significatives et non commentées.

Syntaxe

(Auteur J.-M. Defays, ULg) : 8 universités – 3790 étudiants

Choix des questions

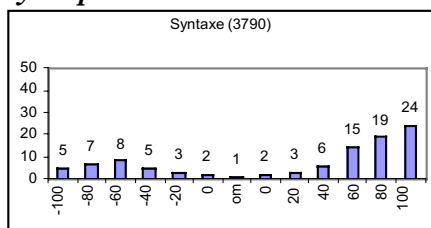
Les 12 questions de l'épreuve ont été choisies parmi un plus grand nombre de questions du test mis au point par l'auteur à l'ULg et validé depuis des années.

En fait, il s'agit de 6 contenus, chacun étant représenté par deux questions. On remarquera que les taux de réussite sont parfois assez différents pour les deux questions d'un même contenu. Cependant, les deux moins bons résultats sont obtenus pour une question dont la réponse correcte est « aucune » (31,8%) et pour une question dont la réponse correcte est « toutes » (48,9%). La forme des questions a donc pu jouer, ce qui reste néanmoins étonnant (les étudiants ne devraient pas être désorientés par des formes aussi simples de questionnement inhabituel).

Indices globaux

Le taux de réussite moyen est **69%**, avec 29% d'erreurs et 1% d'omissions, sur 3790 étudiants testés.

Analyse spectrale



Qualité spectrale : IA-ID-OM-CD-CA = **20-10-1-11-58**

Confiance Collective **74,2 %**

Prudence Collective **60 %**

Discriminance Collective **14,2 %**

Maîtrise totale (RC 100 %) **24 %**

Méprise totale (RI 100 %) **5 %**

En comparaison, en Mathématique FSC (Confiance Collective) vaut 80 % et FSI (Prudence collective) vaut de 55 % !

L'hémispectre des réponses correctes se présente comme une courbe en J, ce dont on doit se réjouir, mais l'hémispectre de gauche (réponses incorrectes) est, lui, très loin d'être une courbe en J, hélas.

58% des connaissances sont **utiles** (réponses correctes avec plus de 50% de certitude) et **20%** des connaissances sont **dangereuses** (réponses incorrectes avec plus de 50% de certitude) ; **22%** seulement sont **inutilisables**.

Voici les termes classés par ordre (approximatif) décroissant de réussite (dont le % est à gauche du mot) :

100 %

90,4 articulations logiques

90

87,6 inférences

86,7 articulations logiques

85,4 conjonctions

80 79,6 inférences

70 69,9 pronom relatif

69,2 pronoms

63,3 ponctuation

60

58,5 pronoms

58,6 ponctuation

50 48,9 conjonctions (TOUTES)

40

31,8 pronom relatif (AUCUNE)

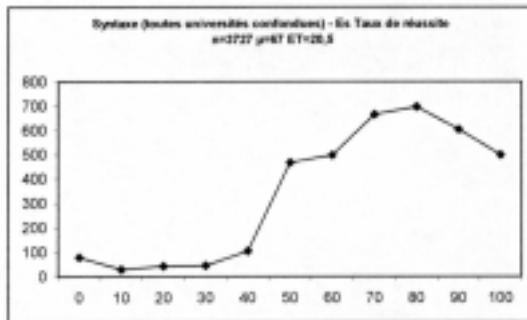
30

20

10

0%

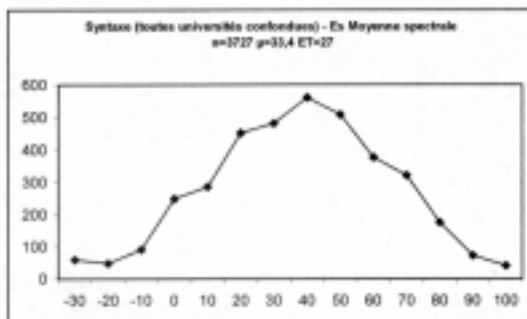
SYNTAXE



Taux de Réponses Correctes (max = 100%)

	n	μ		n	μ
Garçons	1706	66,0*	18-	2363	69,2*
Filles	1961	68,3*	19+	1297	63,6*

Aucune Ampleur d'Effet (AE) n'atteint 0,07 (en positif ou négatif).



Moyenne Spectrale (min = -100, max = 100)

	n	μ		n	μ
Garçons	1706	32,1*	18-	2363	35,4*
Filles	1961	34,8*	19+	1297	30,2*

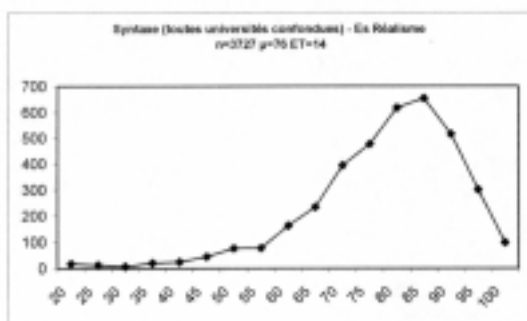
Aucune Ampleur d'Effet (AE) n'atteint 0,07 (en positif ou négatif).



Erreur Moyenne de Centration (Idéal = 0)

	n	μ		n	μ
Garçons	1706	3,7*	18-	2363	0,1*
Filles	1961	-0,1*	19+	1297	4,7*

Aucune Ampleur d'Effet (AE) n'atteint 0,07 (en positif ou négatif).



Réalisme par Calibration (Idéal = 100)

	n	μ		n	μ
Garçons	1706	75,9	18-	2363	75,7
Filles	1961	76,6	19+	1297	76,5

Aucune Ampleur d'Effet (AE) n'atteint 0,07 (en positif ou négatif).

Corrélation avec	Vocab.	Syntaxe	Compré.	Géo.	Math.	Phys.	Chimie	Biologie	Art	Histoire
En Moyenne spectr.	0,49*	-----	0,30*	0,34*	0,29*	0,20*	0,19*	0,19*	0,34*	0,26*
En Taux de Réussite	0,52*	-----	0,41*	0,38*	0,34*	0,26*	0,22*	0,21*	0,38*	0,25*
En Erreur de Centr.	0,62*	-----	0,49*	0,45*	0,51*	0,44*	0,35*	0,41*	0,40*	0,43*
En Réalisme	0,35*	-----	0,13*	0,23*	0,26*	0,20*	0,13*	0,11*	0,23*	0,23*
n	3526	3727	3401	3664	2079	2082	2094	2097	1234	1250
nombre de questions	45	12	6	10	22	10	8	10	25	25

* Corrélations significatives marquées à $p < 0,05$

Compréhension de texte

(Auteur Ph. Hougardy, ULB) : 8 universités : 3469 étudiants

Choix des questions

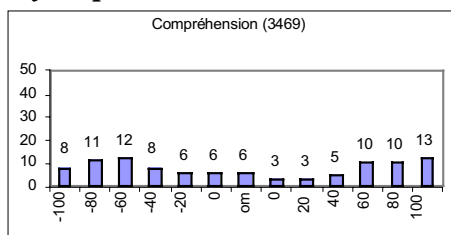
L'épreuve a été réduite à 6 questions seulement parce que la lecture du texte (une page en petits caractères) prenait déjà beaucoup de temps.

Indices globaux

Le taux de réussite moyen est **44%**, avec 51% d'erreurs et 6% d'omissions, sur 3469 étudiants testés.

C'est l'épreuve la moins bien réussie en moyenne sur les dix check-up MOHICAN.

Analyse spectrale



Qualité spectrale : IA-ID-OM-CD-CA = **31-20-6-11-33**

Confiance Collective 67,3 %
Prudence Collective 55,7 %
Discriminance Collective 11,6 %
Maîtrise totale (RC 100 %) 13 %
Méprise totale (RI 100 %) 8 %

En comparaison, en Mathématique FSC (Confiance Collective) vaut 80 %, FSI (Prudence Collective) vaut 55 % et la fracture spectrale (Discriminance Collective) 25 % !

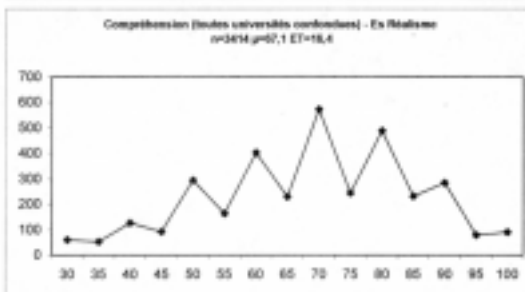
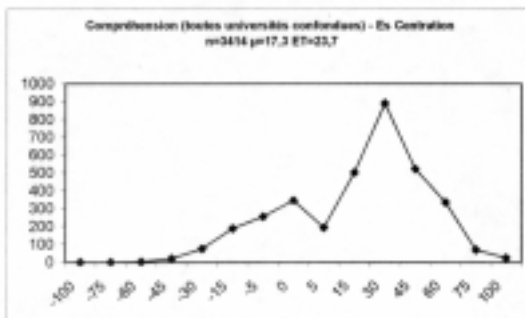
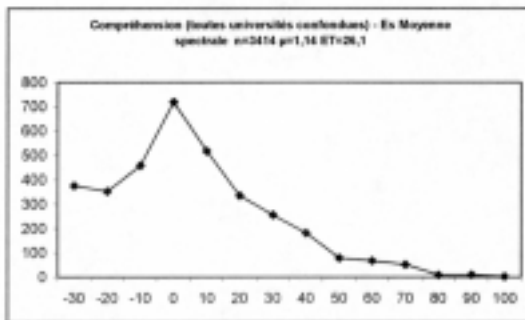
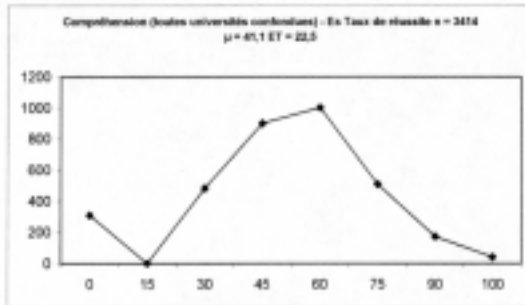
L'hémispectre de droite (les réponses correctes) est légèrement en J, comme espéré, mais celui de gauche est « en cloche », hélas.

33% des connaissances correctes sont **utiles** et **31%** **dangereuses**, ce qui est énorme. **36%** sont **inutilisables**.

Voici les termes classés par ordre (approximatif) décroissant de réussite (dont le % est à gauche du mot) :

100 %	
90	
80	
	75,6 définition de mot
	74,4 interprétation
70	
60	
50	50,5 donner un titre
40	
30	
	26,8 interprétation
20	
	17,8 caractéristique de concept (TOUTES)
	16,3 caractéristique de concept (TOUTES)
10	
0%	

COMPREHENSION DE TEXTE



Taux de Réponses Correctes

	n	μ
Garçons	1595	41,9
Filles	1768	40,6

	n	μ
18-	2173	42,1*
19+	1180	39,9*

L'AE en faveur de ces 18- vaut +0,14

Moyenne Spectrale

	n	μ
Garçons	1595	1,8
Filles	1768	0,2

	n	μ
18-	2173	1,1
19+	1180	0,9

Erreur Moyenne de Centration (Idéal = 0)

	N	μ
Garçons	1595	17,9
Filles	1768	16,6

	N	μ
18-	2173	16,5
19+	1180	18,4

Surestimation gigantesque (de 17 % en moyenne, mais avec un mode de 35 % !!) aussi bien chez les filles que les garçons.

Réalisme par Calibration (Idéal = 100)

	N	μ
Garçons	1595	66,5*
Filles	1768	67,8*

	N	μ
18-	2173	67,2
19+	1180	67,1

L'AE en faveur de ces filles vaut +0,08

Corrélation avec	Vocab.	Syntaxe	Compré.	Géo.	Math.	Phys.	Chimie	Biologie	Art	Histoire
En Moyenne spectr.	0,32*	0,30*	-----	0,25*	0,1912*	0,14*	0,17*	0,21*	0,22*	0,17*
En Taux de Réussite	0,37*	0,41*	-----	0,36*	0,2590*	0,22*	0,22*	0,21*	0,25*	0,16*
En Erreur de Centr.	0,49*	0,49*	-----	0,39*	0,4039*	0,33*	0,28*	0,34*	0,28*	0,33*
En Réalisme	0,19*	0,13*	-----	0,09*	0,1245*	0,04*	0,08*	0,14*	0,07*	0,08*
n	3236	3401	3414	3378	2040	2043	2057	2058	1066	1081
nombre de questions	45	12	6	10	22	10	8	10	25	25

* Corrélations significatives marquées à $p < 0,05$

Géographie et Lecture de cartes, graphiques, documents

(Auteur F. Orban, FUNDP) : 8 universités – 3738 étudiants

Choix des questions

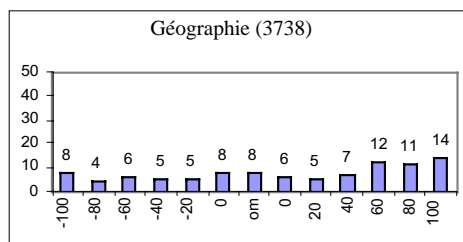
L'épreuve de « géographie » a été choisie (cela aurait pu être celle d'histoire ou d'économie) pour tester la capacité à lire des tableaux de nombres, des graphiques, des cartes (de géographie). Seules 4 questions (délocalisation, mégalopole, Andes, coordonnées géographiques) testent purement les connaissances géographiques.

On constate qu'une même tâche (lecture de carte) peut être très différemment réussie selon les questions (de 93% à 37%). Il en irait probablement de même si on avait diversifié plus les questions de « lecture de graphiques » ou si l'on avait posé des questions de « construction de graphiques ».

Indices globaux

Le taux de réussite moyen est 55%, avec 36% d'erreurs et 8% d'omissions, sur 3738 étudiants testés.

Analyse spectrale



Qualité spectrale : IA-ID-OM-CD-CA = 18-18-8-18-37

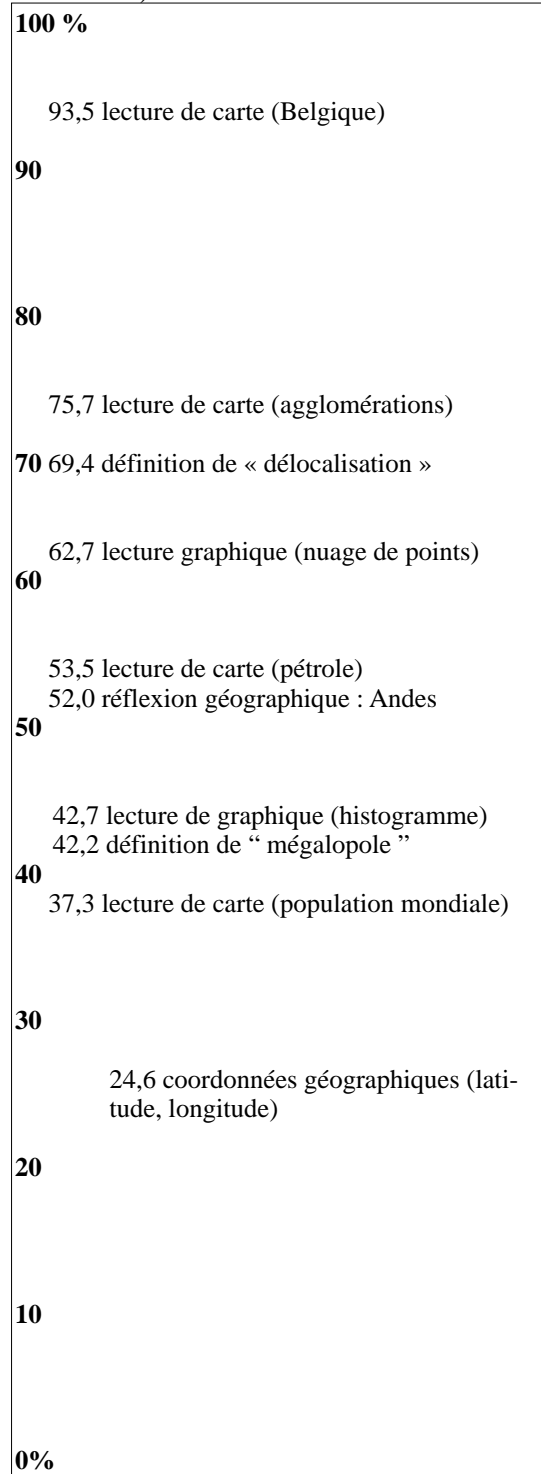
Confiance Collective 61,5 %
Prudence Collective 49,4 %
Discriminance Collective 12 %
Maîtrise totale (RC 100 %) 14 %
Méprise totale (RI 100 %) 8 %

En comparaison, en Mathématique FSC (Confiance Collective) vaut 80 %, FSI (Prudence Collective) vaut 55 % et la fracture spectrale (Discriminance Collective) 25 % !

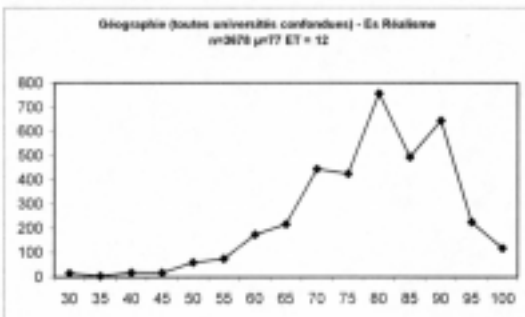
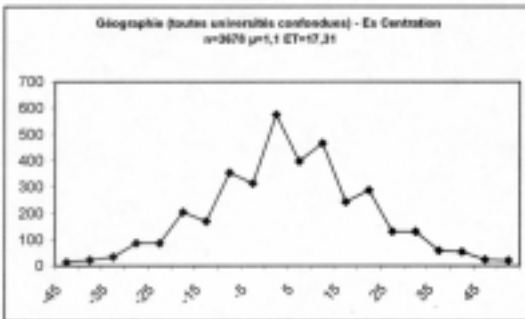
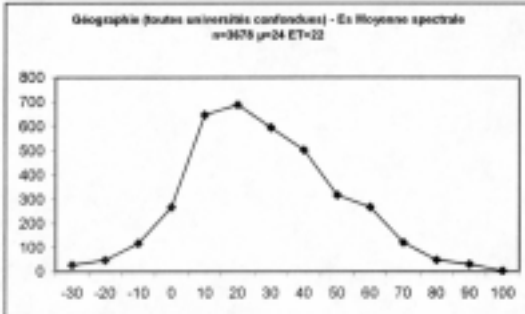
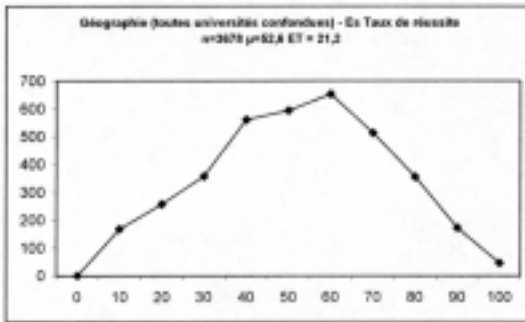
L'hémispectre de droite (les réponses correctes) est légèrement en J, comme espéré, mais celui de gauche est « plat », hélas.

37% des connaissances correctes sont **utilisables** et 18% **dangereuses**, ce qui est beaucoup. 45% sont, hélas, inutilisables.

Voici les... classés par ordre (approximatif) décroissant de réussite (dont le % est à gauche du mot) :



Géographie – Lecture de documents



Taux de Réponses Correctes (max = 100%)

	n	μ
Garçons	1687	57,9*
Filles	1935	48,1*

	N	μ
18-	2343	54,2*
19+	1268	49,7*

L'AE en faveur de de ces garçons vaut +0,45 ce qui est très net

L'AE en faveur de ces -18 ans vaut +0,21 différence modérée

Moyenne Spectrale (min = -100, max = +100)

	n	μ
Garçons	1687	29,1*
Filles	1935	19,7*

	N	μ
18-	2343	24,9*
19+	1268	22,4**

L'AE en faveur de ces garçons vaut +0,43 Très nette supériorité

L'AE en faveur de ces 18- vaut +0,11 très légère supériorité

Erreur Moyenne de Centration (Idéal = 0)

	n	μ
Garçons	1687	1,6
Filles	1935	0,6

	N	μ
18-	2343	0,2*
19+	1268	2,6*

L'AE en faveur de ces 19+ vaut +0,13

Erreur de centration quasi nulle (proche de 0) pour la moyenne de tous !!

Réalisme par Calibration (Idéal = 100)

	n	μ
Garçons	1687	76,6
Filles	1935	77,1

	N	μ
18-	2343	77,5*
19+	1268	76,6*

L'AE de ces 19+ vaut +0,07

Réalisme très légèrement supérieur pour ces 19 ans et + (plus faibles de 4,5 % en moyenne que ces 18 ans et -).

Corrélation avec	Vocab.	Syntaxe	Compr.	Géo.	Math.	Phys.	Chimie	Biologie	Art	Histoire
En Moyenne spectr.	0,39*	0,34*	0,25*	-----	0,38*	0,30*	0,24*	0,15*	0,30*	0,37*
En Taux de Réussite	0,45*	0,38*	0,36*	-----	0,44*	0,36*	0,27*	0,20*	0,38*	0,41*
En Erreur de Centr.	0,49*	0,45*	0,39*	-----	0,46*	0,38*	0,35*	0,36*	0,34*	0,39*
En Réalisme	0,24*	0,23*	0,09*	-----	0,17*	0,17*	0,09*	0,08*	0,19*	0,21*
n	3478	3664	3378	3678	2049	2052	2067	2067	1232	1248
nombre de questions	45	12	6	10	22	10	8	10	25	25

* Corrélations significatives marquées à $p < 0,05$

Mathématiques

(Auteurs M. Lebrun et J. Lega, UCL) : 8 universités – 2539 étudiants

Choix des questions

Les 22 questions de l'épreuve constituent des « traductions » d'un mode d'expression mathématique à un autre :

F = Formule (donc symbolique)

N = Numérique (une équation ou une série de nombres ou un tableau de nombres)

M = Mots (une définition ou un problème)

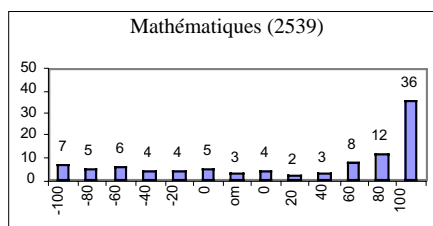
G = Graphique

Par exemple, l'expression $N \rightarrow G$ signifie : « A partir d'une série de Nombres, créer (ou choisir) le Graphique correspondant » ; l'expression $N \rightarrow N$ consiste, par exemple, à résoudre une équation (on reste dans le numérique).

Indices globaux

Le taux de réussite moyen est 66 %, avec 32 % d'erreurs et 3 % d'omissions, sur 2539 étudiants testés.

Analyse spectrale



Qualité spectrale : IA-ID-OM-CD-CA = 18-13-3-9-56

Confiance Collective 80 %
Prudence Collective 54,8 %
Discriminance Collective 25,2 %
Maîtrise totale (RC 100 %) 36 %
Méprise totale (RI 100 %) 7 %

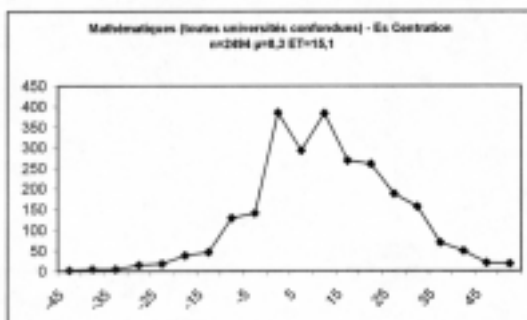
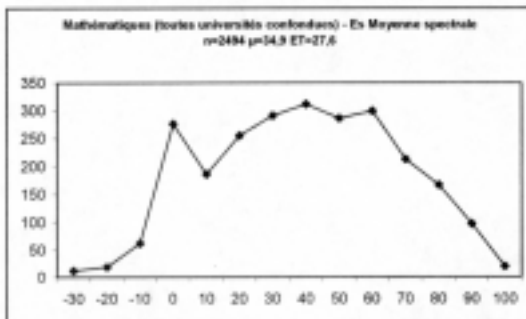
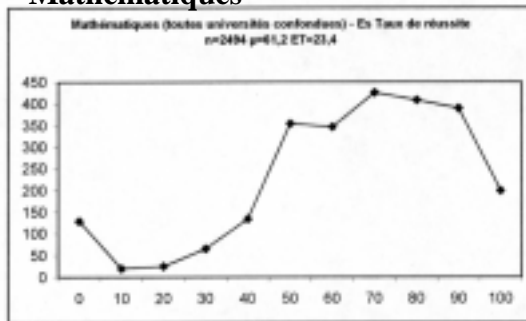
L'hémispectre de droite (les réponses correctes) est nettement en J, comme espéré, mais celui de gauche est « plat », hélas.

56% des connaissances correctes sont utilisables et 18% dangereuses, ce qui est beaucoup. 26 % sont inutilisables, ce qui est la meilleure valeur pour cet indice aux 10 check-up MOHICAN.

Voici les termes classés par ordre (approximatif) décroissant de réussite (dont le % est à gauche du mot) :

100 %			
90	89,2 N→F 87,5 N→G 84,6 F→G	88,7 N→N	
80	77,1 M→N 75,1 G→F	78,2 G→F 75 N→N	79,5 M→M
70	70,5 M→N	70,6 F→M	70,4 M→M
	64,1 G→M 63,2 G→N	63,9 F→N	
60	55,6 M→M 53,2 F→N 51,7 M→M	55,8 G→M	
50			
40	40,0 F→M		
30	27,8 N→M		
20	20,6 G→N		
10			
0%			

Mathématiques



Taux de Réponses Correctes

	n	μ		n	μ
Garçons	1241	66,7*	18-	1576	62,9
Filles	1154	57,1*	19+	819	60,3

L'AE de ces garçons vaut +0,41.
Nette supériorité du taux de réussite moyenne chez ces garçons (#).

Moyenne Spectrale

	n	μ		N	μ
Garçons	1241	41,5*	18-	1576	36,7*
Filles	1154	28,7*	19+	819	32,8*

L'AE de ces garçons vaut +0,46.
L'AE de ces filles 18- vaut +0,14.
Nette supériorité de ces garçons (#) et très légère de ces 18 ans et moins (#).

Erreur Moyenne de Centration (Idéal = 0)

	n	μ		n	μ
Garçons	1241	8,7	18-	1576	7,2*
Filles	1154	7,9	19+	819	10,6*

Tendance générale à la surestimation, avec 2 pics : 0 % et + 10 %, mais une forte proposition de plus de 15 %.
Ces 19 ans et + (#) se surestiment légèrement plus (AE = +0,22) que ces 18 ans et moins.

Réalisme par Calibration (Idéal = 100)

	n	μ		n	μ
Garçons	1241	79,4*	18-	1576	77,6
Filles	1154	77,2*	19+	819	77,9

L'AE de ces garçons vaut +0,17.
Très légère supériorité en réalisme de ces garçons (#).

Corrélation avec	Vocab.	Syntaxe	Compré.	Géo.	Math.	Phys.	Chimie	Biologie	Art	Histoire
En Moyenne spectr.	0,37*	0,29*	0,19*	0,38*	-----	0,54*	0,41*	0,27*	-0,16	0,13
En Taux de Réussite	0,39*	0,34*	0,25*	0,44*	-----	0,67	0,55*	0,44*	-0,04	0,06
En Erreur de Centr.	0,50*	0,51*	0,40*	0,46*	-----	0,49*	0,42*	0,47*	-0,02	0,09
En Réalisme	0,26*	0,26*	0,12*	0,17*	-----	0,46*	0,37*	0,28*	-0,02	-0,47
n	2110	2079	2040	2049	2494	2472	2468	2471	12	12
nombre de questions	45	12	6	10	22	10	8	10	25	25

* corrélations significatives marquées à $p < 0,05$

Physique

(Auteur : P. Chapelle, FPMs) : 8 universités : 2520 étudiants

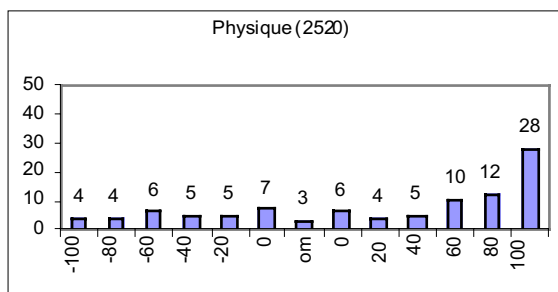
Choix des questions

Les 10 questions de l'épreuve visaient à toucher 10 domaines différents. Il est regrettable que chaque domaine n'aie pu être testé par plusieurs questions (4 par exemple). Le fait que la (seule) question sur la vitesse soit la moins bien réussie n'implique pas que ce soit le concept le moins bien maîtrisé. Pour les questions sur l'électricité, la même remarque (inversée) s'impose.

Indices globaux

Le taux de réussite moyen est **65 %**, avec 32 % d'erreurs et 3 % d'omissions, sur 2520 étudiants testés.

Analyse spectrale



Qualité spectrale : IA-ID-OM-CD-CA = **14-17-3-15-50**

Confiance Collective **71,4 %**
Prudence Collective **44,5 %**
Discriminance Collective **26,9 %**
Maîtrise totale (RC 100 %) **28 %**
Méprise totale (RI 100 %) **4 %**

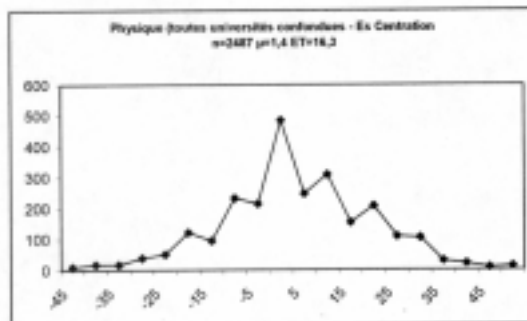
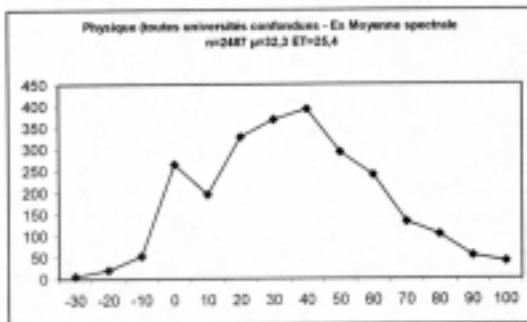
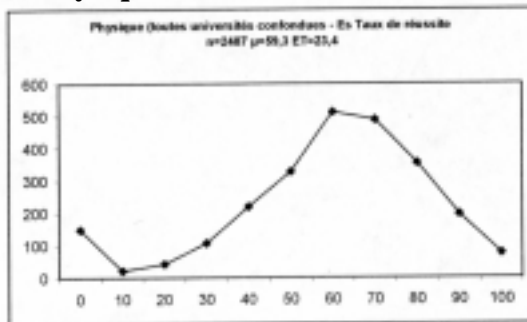
L'hémispectre de droite (les réponses correctes) est nettement en J, comme espéré, mais celui de gauche est l'est beaucoup moins, hélas.

50% des connaissances correctes sont **utilisables** et **14%** **dangereuses, ce qui est très peu, heureusement. 36 %** sont inutilisables, hélas.

Voici les termes classés par ordre (approximatif) décroissant de réussite (dont le % est à gauche du mot) :

100 %
91,1 Electricité (bobine)
90 89,8 Electricité (lumière)
86,5 Propriétés des corps
80 79,1 Accélération
74,5 Electricité (Résistance)
70 70,8 Masse
60
51,9 Forces
50
44,1 Gravitation
40 41,3 Optique
30
24,1 Vitesse
20
10
0%

Physique



Taux de Réponses Correctes

	n	μ
Garçons	1244	65,8*
Filles	1146	53,9*

	n	μ
18-	1574	59,2*
19+	814	61,9*

Moyenne Spectrale

	n	μ
Garçons	1244	39,6*
Filles	1146	25,3*

	n	μ
18-	1574	31,2*
19+	814	35,8*

L'AE de ces garçons
vaut +0,17

L'AE de ces
18+ vaut +0,18

Très légère supériorité de ces garçons (#) sur
les filles (#) et de ces 19 ans et + ((#) sur Ces
18 ans et -.

Erreur Moyenne de Centration (Idéal = 0)

	n	μ
Garçons	1244	8,7
Filles	1146	7,9

	n	μ
18-	1574	0,4*
19+	814	3,0*

Réalisme par Calibration (Idéal = 100)

	n	μ
Garçons	1244	78,9*
Filles	1146	73,7*

	n	μ
18-	1574	76,4*
19+	814	74,2*

Corrélation avec	Vocab.	Syntaxe	Compré.	Géo.	Math.	Phys.	Chimie	Biologie	Art	Histoire
En Moyenne spectr.	0,28*	0,20*	0,14*	0,30*	0,54*	-----	0,39*	0,32*	0,07	0,35
En Taux de Réussite	0,34*	0,26*	0,22*	0,36*	0,67*	-----	0,56*	0,50*	0,21	0,42
En Erreur de Centr.	0,45*	0,44*	0,33*	0,38*	0,49*	-----	0,40*	0,43*	0,23	0,61*
En Réalisme	0,18*	0,20*	0,04*	0,17*	0,46*	-----	0,29*	0,23*	0,16	0,37
n	2114	2082	2043	2052	2472	2487	2474	2475	11	11
nombre de questions	45	12	6	10	22	10	8	10	25	25

* Corrélations significatives marquées à $p < 0,05$

Chimie

(Auteur A. Cornélis, ULg) : 8 universités – 2522 étudiants

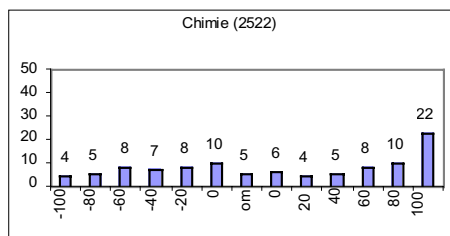
Choix des questions

Les 8 questions de l'épreuve visent à toucher des « points » différents de matière. Les contraintes de temps ont fait « fondre » le check-up initialement prévu par l'auteur (20 questions) à ce nombre quasi dérisoire, si bien que l'analyse spectrale reste le résultat le plus généralisable.

Indices globaux

Le taux de réussite moyen est **55 %**, avec 41 % d'erreurs et 5% d'omissions, sur 2522 étudiants testés.

Analyse spectrale



Qualité spectrale : IA-ID-OM-CD-CA = **17-25-5-15-40**

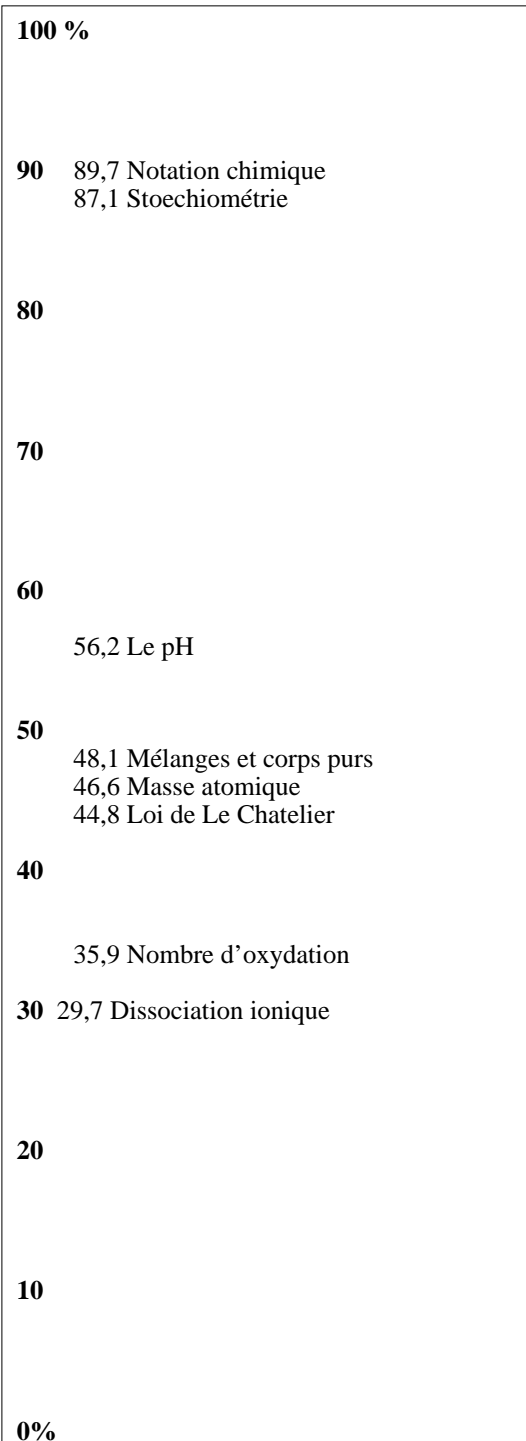
Confiance Collective **68,4 %**
Prudence Collective **41,0 %**
Discriminance Collective **27,4 %**
Maîtrise totale (RC 100 %) **22 %**
Méprise totale (RI 100 %) **4 %**

En comparaison, en Mathématique FSC (Confiance Collective) vaut 80 %, FSI (Prudence Collective) vaut 55 % et la fracture subjective (Discriminance Collective) 25 % !

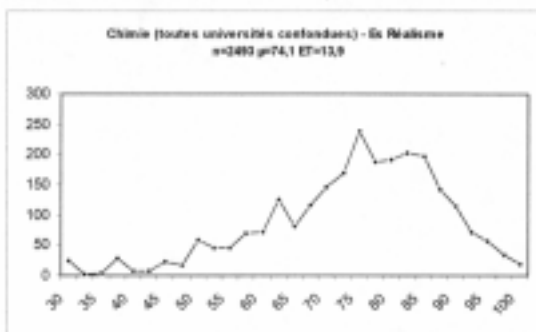
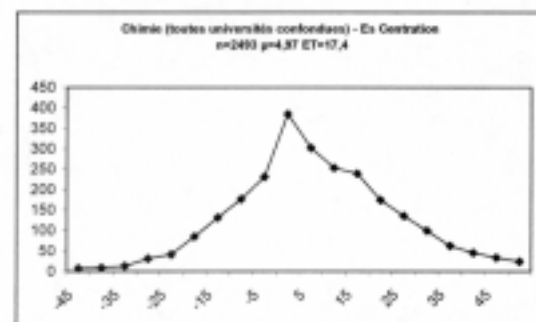
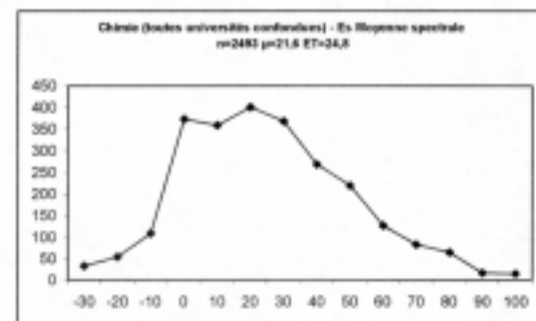
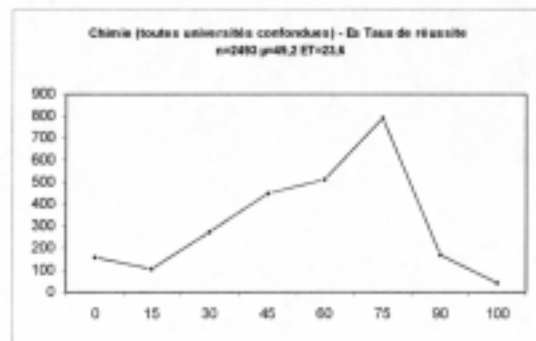
L'hémispectre de droite (les réponses correctes) est nettement en J, comme espéré. Celui de gauche aussi, mais de façon moins prononcée.

40% des connaissances correctes sont **utilisables** et **17% dangereuses**, ce qui est beaucoup. 43% sont, hélas **inutilisables**.

Voici les termes classés par ordre (approximatif) décroissant de réussite (dont le % est à gauche du mot) :



CHIMIE



Taux de Réponses Correctes

	n	μ
Garçons	1246	51,9*
Filles	1153	47,7*

L'AE en faveur de ces garçons vaut +0,18

	N	μ
18-	1582	49,3*
19+	815	50,9*

L'AE de ces 19+ vaut +0,07

Moyenne Spectrale

	n	μ
Garçons	1246	23,4*
Filles	1153	20,4*

L'AE en faveur de ces garçons vaut +0,08

	n	μ
18-	1582	21,3*
19+	815	23,3*

L'AE de ces 19+ vaut +0,12

Erreur Moyenne de Centration (Idéal = 0)

	n	μ
Garçons	1246	6,4*
Filles	1153	3,4*

L'AE en faveur de ces garçons vaut +0,14

	N	μ
18-	1582	4,1*
19+	815	6,6*

L'AE de ces 19+ vaut +0,17

Donc surestimation très légèrement plus forte chez ces garçons et chez ces 19 ans et + .

Réalisme par Calibration (Idéal = 100)

	n	μ
Garçons	1246	74,1
Filles	1153	74,8

	N	μ
18-	1582	74,9*
19+	815	73,7*

L'AE de ces 18- vaut +0,08

Corrélation avec	Vocab.	Syntaxe	Compré.	Géo.	Math.	Phys.	Chimie	Biologie	Art	Histoire
En Moyenne spectr.	0,27*	0,19*	0,17*	0,24*	0,41*	0,39*	-----	0,35*	0,02	-0,54
En Taux de Réussite	0,30*	0,22*	0,22*	0,27*	0,55*	0,56*	-----	0,51*	-0,01	-0,67*
En Erreur de Centr.	0,37*	0,35*	0,28*	0,35*	0,42*	0,40*	-----	0,39*	-0,21	-0,37
En Réalisme	0,14*	0,13*	0,08*	0,09*	0,37*	0,29*	-----	0,23*	0,54	0,03
n	2127	2094	2057	2067	2468	2474	2493	2488	11	11
nombre de questions	45	12	6	10	22	10	8	10	25	25

* Corrélations significatives marquées à $p < 0,05$

Biologie

(Auteur J.-C. Verhaeghe, ULB) : 8 universités – 2526 étudiants

Choix des questions

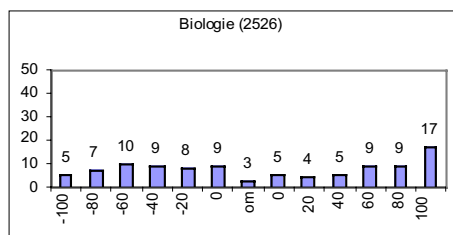
Les 10 questions de l'épreuve visent à toucher des thèmes différents. Hélas, une seule question par thème ne permet pas d'extrapoler le taux de réussite à la compétence sur le thème.

Le résultat exceptionnellement élevé à la question sur le SIDA, résulte probablement de la multiplicité des sources d'information sur le sujet et de l'importance que les jeunes leur ont (probablement) accordé.

Indices globaux

Le taux de réussite moyen est **49 %**, avec 48 % d'erreurs et 2 % d'omissions, sur 2526 étudiants testés.

Analyse spectrale



Qualité spectrale : IA-ID-OM-CD-CA = 22-26-3-14-35

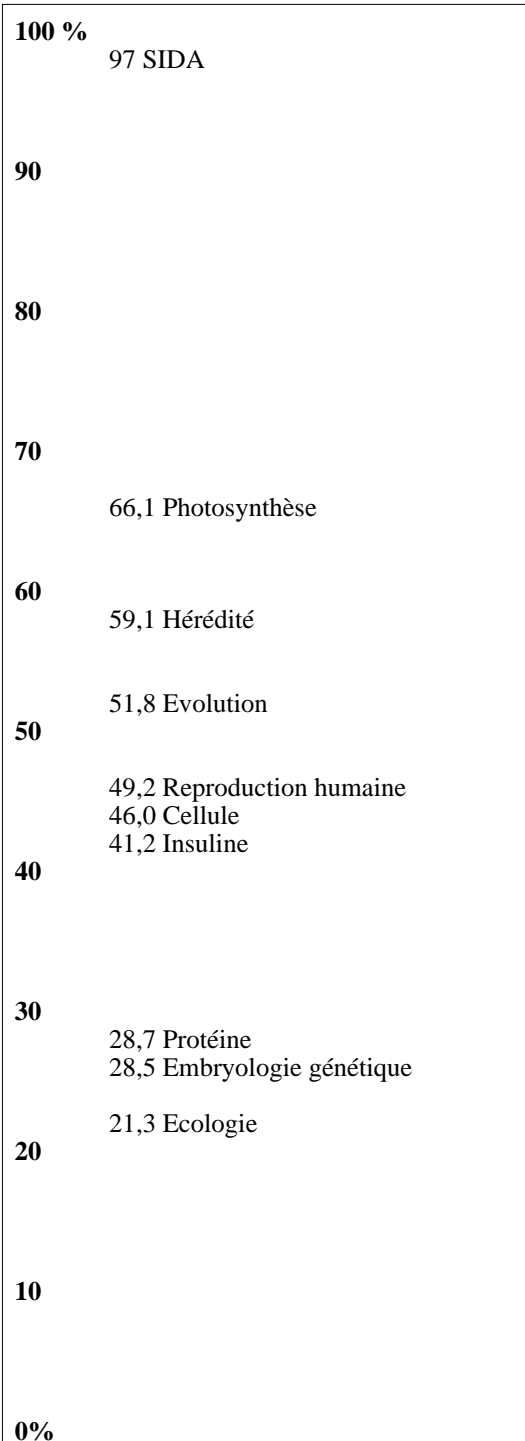
Confiance Collective 66,1 %
Prudence Collective 45,4 %
Discriminance Collective 20,7 %
Maîtrise totale (RC 100 %) 17 %
Méprise totale (RI 100 %) 5 %

En comparaison, en Mathématique FSC (Confiance Collective) vaut 80 %, FSI (Prudence Collective) vaut 51% et la fracture subjective (Discriminance Collective) 25 % !

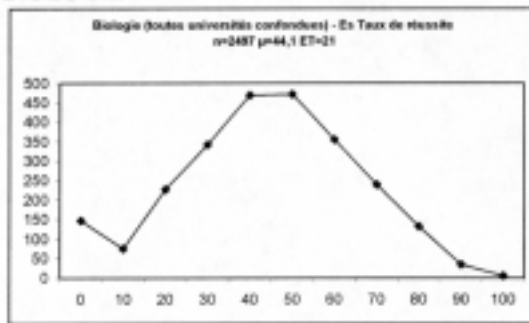
L'hémispectre de droite (les réponses correctes) est en J, comme espéré. Celui de gauche est plutôt plat, hélas.

35% des connaissances correctes sont **utilisables** et **22%** **dangereuses, ce qui est beaucoup**. 43% sont, hélas inutilisables.

Voici les termes classés par ordre (approximatif) décroissant de réussite (dont le % est à gauche du mot) :



BIOLOGIE

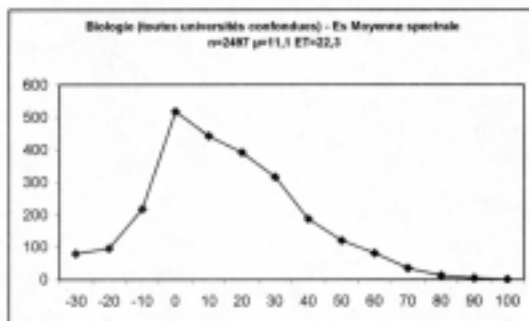


Taux de Réponses Correctes

	n	μ
Garçons	1248	45,5*
Filles	1155	43,7*

	n	μ
18-	1582	42,6
19+	820	48,4

L'AE de ces garçons vaut +0,08



Moyenne Spectrale

	n	μ
Garçons	1248	11,7
Filles	1155	10,7

	n	μ
18-	1582	9,2*
19+	820	15,1*

..... L'AE en faveur
..... de ces 19+
vaut 0,71



Erreur Moyenne de Centration (Idéal = 0)

	N	μ
Garçons	1248	11,4*
Filles	1155	9,2*

	n	μ
18-	1582	9,5*
19+	820	12,1*

L'AE en faveur de ces garçons vaut +0,12
L'AE de ces 19+ vaut +0,14
Ces garçons se surestiment plus que ces filles et ces 19 ans et + plus que ces 18 ans et -.



Réalisme par Calibration (Idéal = 100)

	n	μ
Garçons	1248	72,3
Filles	1155	72,9

	n	μ
18-	1582	72,7
19+	820	72,3

Corrélation avec	Vocab.	Syntaxe	Compré.	Géo.	Math.	Phys.	Chimie	Biologie	Art	Histoire
En Moyenne spectr.	0,28*	0,19*	0,21*	0,15*	0,27*	0,32*	0,35*	-----	0,28	-0,06
En Taux de Réussite	0,30*	0,21*	0,21*	0,20*	0,44*	0,50*	0,51*	-----	0,45	0,00
En Erreur de Centr.	0,46*	0,41*	0,34*	0,36*	0,47*	0,43*	0,39*	-----	-0,18	0,06
En Réalisme	0,20*	0,11*	0,14*	0,08*	0,28*	0,23*	0,23*	-----	0,06	-0,09
n	2129	2097	2058	2067	2471	2475	2488	2497	11	11
nombre de questions	45	12	6	10	22	10	8	10	25	25

* Corrélations significatives marquées à $p < 0,05$

Connaissances historiques, économiques et actualité

(Auteurs D. Leclercq & F. Georges, ULg) : 8 universités – 1418 étudiants

Choix des questions

Les 25 questions de l'épreuve se répartissent comme suit :

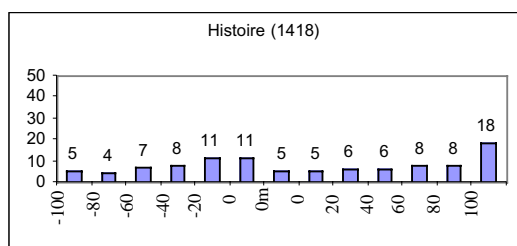
- 4 d'histoire (1789, Hiroshima, Congo, Hégire),
- 4 de terminologie se référant à l'histoire (iconoclaste, spartiate, précolombien, kolkhoze),
- 9 sur les institutions européennes et mondiales (UNICEF, OMS, OTAN, ONU, GATT, BIRD, UNESCO, Président E.U., Euro date),
- 4 sur l'économie (Euro BEF, minimex, % chômage, % inflation),
- 4 sur les médias et l'actualité (Rossel, No Comment, Proche Orient, Euro date).

La question sur l'histoire récente (Hiroshima) est moins bien réussie qu'une plus éloignée (ex : 1789) et les indices économiques wallons (chômage et inflation) sont mal connus.

Indices globaux

Le taux de réussite moyen est 50 %, avec 46 % d'erreurs et 5% d'omissions, sur 1418 étudiants testés.

Analyse spectrale



Qualité spectrale : IA-ID-OM-CD-CA = 16-30-5-17-34

Confiance Collective 64,3 %

Prudence Collective 38,7 %

Discriminance Collective 25,6 %

Maîtrise totale (RC 100 %) 18 %

Méprise totale (RI 100 %) 5 %

En comparaison, en Mathématique FSC (Confiance Collective) vaut 80 %, FSI (Prudence Collective) vaut 51% et la fracture subjective (Discriminance Collective) 25 % !

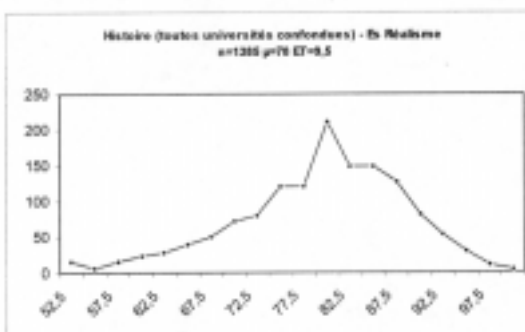
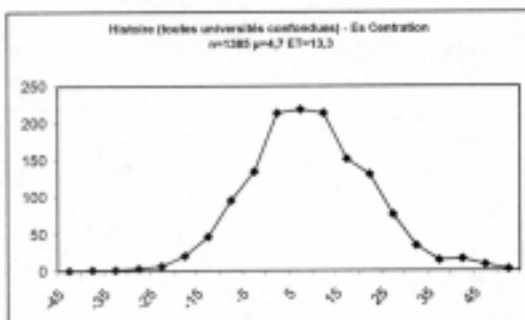
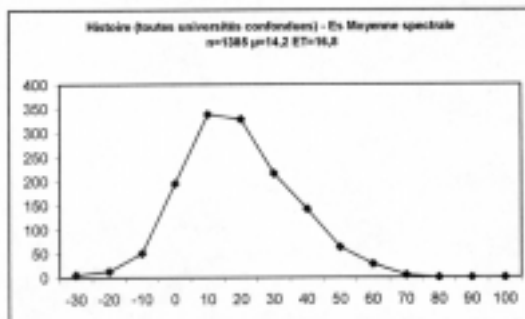
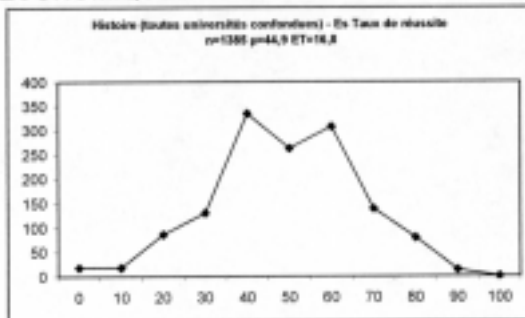
L'hémispectre de droite (les réponses correctes) est en J, comme espéré. Celui de gauche l'est aussi, ce qui est un signe de réalisme.

34% des connaissances correctes sont **utilisables** et **16%** **dangereuses**, ce qui est beaucoup. 50% sont, hélas **inutilisables**.

Voici les termes classés par ordre (approximatif) décroissant de réussite (dont le % est à gauche du mot) :

100 %	
92,8	1 Euro = 40,3399 BEF
91,5	1789
90	
76	UNICEF
74,9	Proche Orient
72,8	OMS
72,8	Spartiate
70	
65,0	EURO (date)
65,3	minimex
60,5	ONU
60	
59,1	GATT
55,5	Kolkhoze
52	No comment (AUCUNE)
50,6	BIRD
51,9	Congo
50	
44,7	Hégire
45	OTAN
43,4	Président U.E.
42,8	précolombien
40	
37,8	UNESCO
34,1	% Chômage wallon
30	
27,2	ONU
25,1	taux d'inflation annuel
23,3	iconoclaste
22,7	Rossel
20,5	Hiroshima (AUCUNE)
20	
10	
0%	

CONNAISSANCES HISTORIQUES, ECONOMIQUES ET ACTUALITE



Taux de Réponses Correctes

	n	µ
Garçons	478	50,9*
Filles	884	41,6*

	n	µ
18-	902	45,6
19+	459	43,6

L'AE en faveur de ces garçons vaut +0,55

Moyenne Spectrale

	n	µ
Garçons	478	19,0*
Filles	884	11,6*

	n	µ
18-	902	14,6
19+	459	13,6

L'AE en faveur de ces garçons vaut +0,44

Erreur Moyenne de Centration (Idéal = 0)

	n	µ
Garçons	478	7,8*
Filles	884	2,8*

	n	µ
18-	902	3,8*
19+	459	6,3*

L'AE en faveur de ces garçons vaut +0,38

L'AE de ces 19+ vaut +0,20

Surestimation plus importante chez ces garçons que chez ces filles et chez ces 19 ans et plus que chez ces 18 ans et moins.

Réalisme par Calibration (Idéal = 100)

	n	µ
Garçons	478	76,9*
Filles	884	78,5*

	n	µ
18-	902	77,4
19+	459	78,2

L'AE de ces filles vaut +0,17

Corrélation avec	Vocab.	Syntaxe	Compré.	Géo.	Math.	Phys.	Chimie	Biologie	Art	Histoire
En Moyenne spectr.	0,45*	0,25*	0,17*	0,37*	0,13	0,35	-0,54	-0,06	0,52*	-----
En Taux de Réussite	0,48*	0,25*	0,16*	0,41*	0,06	0,42	-0,67*	0,00	0,57*	-----
En Erreur de Centr.	0,48*	0,43*	0,33*	0,39*	0,09	0,61*	-0,37	0,06	0,58*	-----
En Réalisme	0,27*	0,23*	0,08*	0,21*	-0,47	0,37	0,03	-0,09	0,40*	-----
n	1241	1250	1081	1248	12	11	11	11	1366	1385
nombre de questions	45	12	6	10	22	10	8	10	25	25

* Corrélations significatives marquées à $p < 0,05$

Connaissances artistiques et culturelles

(Auteurs D. Leclercq-ULg & Lamblin, Insp. C.F.) : 8 universités – 1399 étudiants

Choix des questions

Les 25 questions de l'épreuve sont réparties comme suit :

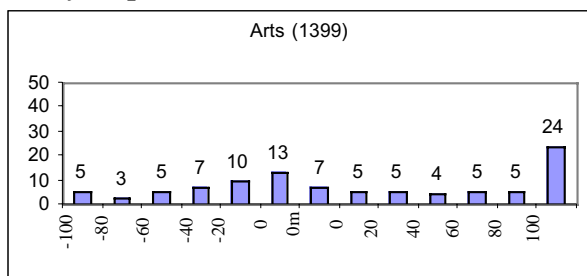
- 5 sur la littérature,
- 2 sur le théâtre,
- 4 sur la peinture/sculpture,
- 4 sur l'architecture,
- 6 sur le cinéma,
- 2 sur la musique,
- 2 sur la danse.

On constate que 4 des 6 questions sur le cinéma sont dans le peloton de tête des taux de réussite et que le théâtre et la danse sont les moins bien connus. L'écart entre la musique pop (Queen) et la musique classique (Beethoven) est flagrant. Il serait probablement inversé chez des adultes de la soixantaine.

Indices globaux

Le taux de réussite moyen est **49 %**, avec 44 % d'erreurs et 7 % d'omissions, sur 1399 étudiants testés.

Analyse spectrale



Qualité spectrale : IA-ID-OM-CD-CA = **13-30-7-14-34**

Confiance Collective **70%**
Prudence Collective **35,3 %**
Discriminance Collective **34,7 %**
Maîtrise totale (RC 100 %) **24 %**
Méprise totale (RI 100 %) **5 %**

FSC (Confiance Collective) vaut 80 %, FSI (Prudence Collective) vaut 51% et la fracture subjective (Discriminance Collective) 25 % !

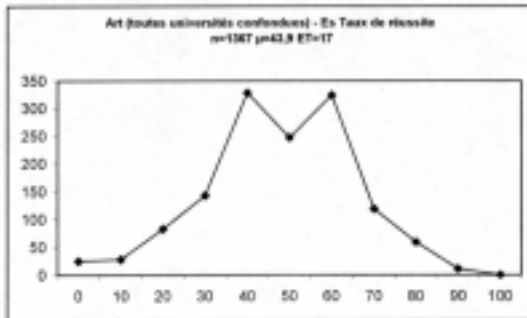
L'hémispectre de droite (les réponses correctes) est très nettement en J, comme espéré. Celui de gauche l'est aussi, ce qui est un signe de réalisme.

34% des connaissances correctes sont **utilisables** et **13%** **dangereuses**. 53% sont, hélas, **inutilisables**.

Voici les termes classés par ordre (approximatif) décroissant de réussite (dont le % est à gauche du mot) :

100 %	93,6 Cinéma (Philadelphia)
	91,6 Cinéma (Schindler/Shoa)
	91,2 Musique (Queen)
90	
	85,2 Cinéma (Rain Man)
80	
	77,0 Littérature (Mein Kampf)
	73,1 Cinéma (Spielberg) (TOUTES)
70	
60	
	58,7 Peinture/Sculpture (Impressionnisme)
	56,8 Architecture (Gothique)
	54,8 Architecture (Gothique)
	51,8 Littérature (Kant)
50	
	49,4 Peinture/sculpture (cubisme)
	48,2 Littérature (A. Huxley)
	47,8 Architecture (Parthénon) (AUCUNE)
	46,8 Danse (Boléro) 41,8 Peint/scul.belge
XXe siècle	
40	
	36,7 Cinéma (Stanley Kubrik) (TOUTES)
	35,0 Architecture (Golden Gate) (AUCUNE)
	35,3 Littérature (Baudelaire)
30	
	30,2 Musique (Beethoven) (AUCUNE)
	26,8 Littérature (Hugo)
	23,4 Peinture/sculpture (Breughel) (AUCUNE)
20	
	21,3 Cinéma (Sentiers Gloire/décimation)
10	
	16,7 Danse (Jorge Donn) (AUCUNE)
	15,3 Théâtre (Dr Knock) (AUCUNE)
0%	
	6,22 Théâtre (Bertold Brecht) (TOUTES)

CONNAISSANCES ARTISTIQUES ET CULTURELLES



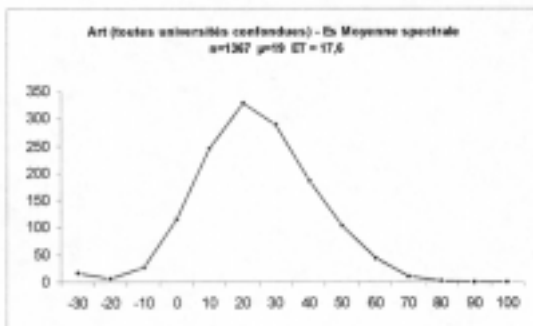
Taux de Réponses Correctes

	n	μ
Garçons	470	46,6*
Filles	875	42,6*

	n	μ
18-	895	44,9*
19+	449	41,9*

L'AE de ces garçons vaut +0,23

L'AE de ces 18- vaut +0,18

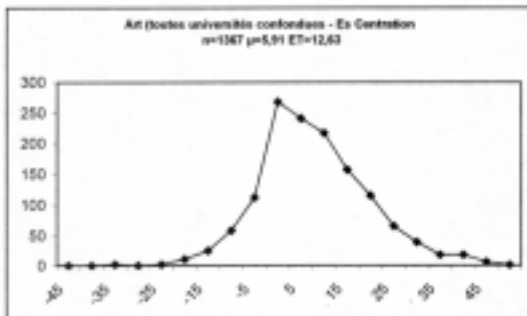


Moyenne Spectrale

	n	μ
Garçons	470	21,5*
Filles	875	17,8*

	n	μ
18-	895	19,7
19+	449	17,9

L'AE de ces garçons vaut +0,21



Erreur Moyenne de Centration (Idéal = 0)

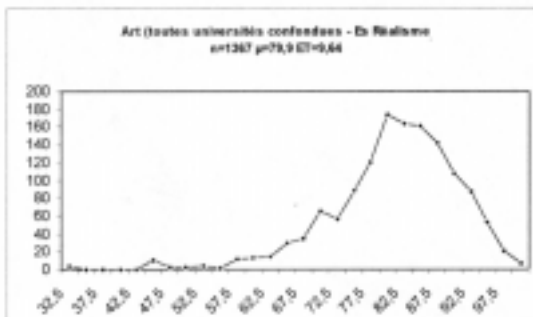
	n	μ
Garçons	470	8,1*
Filles	875	4,7*

	N	μ
18-	895	5,2*
19+	449	7,1*

L'AE de ces garçons vaut +0,27

L'AE de ces 18- vaut +0,18

Ces garçons se surestiment plus que ces filles et ces 19+ se surestiment plus que ces 18-.



Réalisme par Calibration (Idéal = 100)

	n	μ
Garçons	470	79,9
Filles	875	79,7

	n	μ
18-	895	79,9
19+	449	79,7

Corrélation avec	Vocab.	Syntaxe	Compré.	Géo.	Math.	Phys.	Chimie	Biologie	Art	Histoire
En Moyenne spectr.	0,42*	0,34*	0,22	0,30*	-0,16	0,07	0,02	0,28	-----	0,52*
En Taux de Réussite	0,47*	0,38*	0,25*	0,38*	-0,04	0,21	-0,01	0,45	-----	0,57*
En Erreur de Centr.	0,43*	0,40*	0,28*	0,34*	-0,02	0,23	-0,21	-0,18	-----	0,58*
En Réalisme	0,24*	0,23*	0,07*	0,19*	-0,02	0,16	0,54	0,05	-----	0,40*
n	1228	1234	1066	1232	12	11	11	11	1367	1366
nombre de questions	45	12	6	10	22	10	8	10	25	25

(#) de cette population, qui n'est PAS un échantillon aléatoire de la population des étudiants de la Communauté française.

* Corrélations significatives marquées à $p < 0,05$

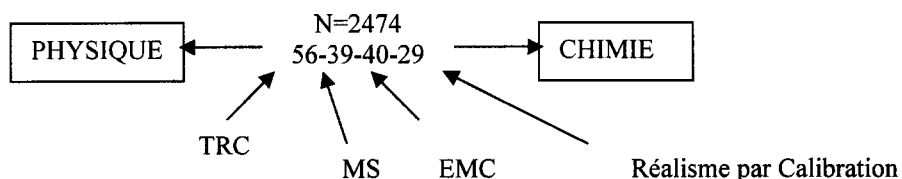
5. La carte de présentation de toutes les intercorrélations (ci-contre)

5.1 Les corrélations sont **TOUTES POSITIVES** et nous nous permettons (pour gagner en place et en lisibilité), d'écrire le centuple de leur valeur. Donc 57 signifie 0,57.

5.2 **Quatre corrélations** figurent sous le nombre d'observations (N) qui a servi à calculer les corrélations. Ces 4 valeurs de corrélations, sont dans l'ordre :

- celle qui concerne les **TRC** (ou **Taux de Réponses Correctes**, en %),
- celle qui concerne les **MS** (ou **Moyennes Spectrales**)
- celle qui concerne les **EMC** (ou **Erreurs Moyennes de Centration** soit la Certitude Moyenne moins l'exactitude moyenne ou le Taux Réponses Correctes)
- celle qui concerne le **Réalisme** par Calibration (ou 100-EMAC, cette dernière valeur constituant l'Erreur Moyenne Absolue de Certitude, ce qui ne permet pas de compenser les surestimations par des sous-estimations et vice versa.

Exemple :



5.3 Rappel sur la **Moyenne Spectrale** ou **MS**

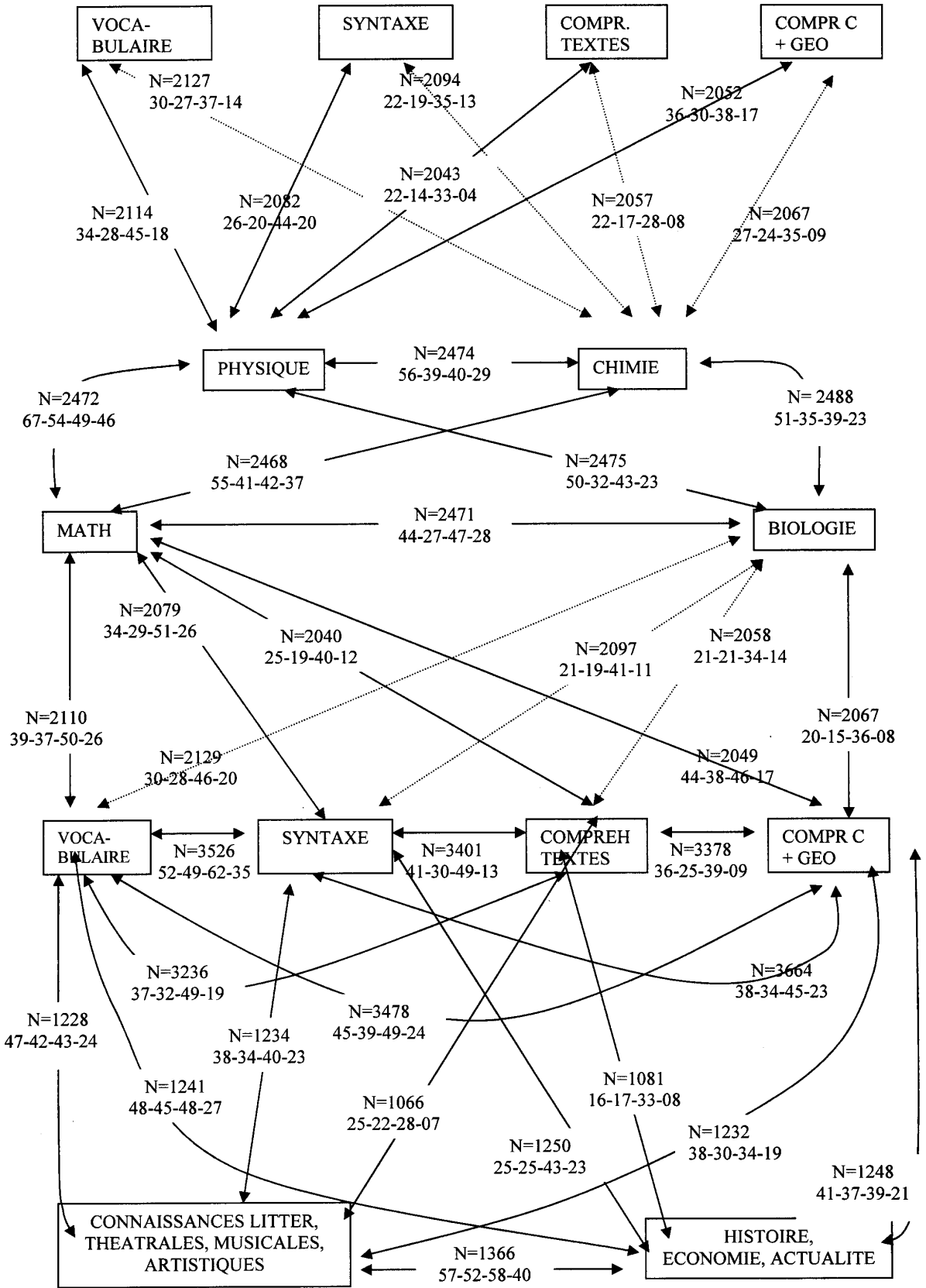
a) **Son calcul** : la MS est obtenue en attribuant 20 points à une réponse correcte avec certitude 20% et -20 points à une réponse incorrecte avec certitude 20%. Le maximum possible est donc 100 et le minimum possible -100.

La moyenne Spectrale d'un étudiant est quasi toujours inférieure à son taux de Réponses Correctes. Elle est égale au Taux de Réponses correctes quand toutes les réponses correctes ont été données avec certitude 100% et que les autres questions ont fait l'objet d'omissions. On peut y voir une sorte de « *correction for guessing* ».

b) **Sa signification** : A **égalité de Réponses Correctes** (par exemple 6 correctes sur 10 réponses), des étudiants différents peuvent avoir des **Moyennes Spectrales très différentes**, comme le montrent les 3 exemples ci-après :

- L'étudiant A a donné ses 6 réponses avec la certitude 100% et omis pour les 4 autres questions ; sa moyenne spectrale est de 60.
- L'étudiant B a répondu 6 fois correctement avec 0% de certitude et s'est trompé 4 fois avec 100% de certitude. Sa moyenne spectrale vaut -40.
- L'étudiant C a donné 6 réponses correctes avec la certitude 40% et 4 incorrectes avec la certitude 60%. Sa Moyenne Spectrale vaut 0.

La carte ci-contre présente l'avantage de l'exhaustivité, mais l'inconvénient de ne plus pouvoir recevoir des informations supplémentaires. Les pages qui la suivent « focalisent » les corrélations pour une seule variable (le %RC) à la fois, avec les moyennes. Ce sont, évidemment les mêmes corrélations de base.



6. Groupements des corrélations concernant les Taux de Réponses Correctes (% RC)

6.1 Les corrélations entre les check-up **sciences de la nature** (math, physique, chimie, biologie) et les 4 épreuves communes sont les suivantes (avec une Moyenne = 0,28)

0,28	Vocabulaire	Syntaxe	Compréh. Texte	Compréh. Géo	Moyenne
Math	0,39	0,34	0,25	0,44	0,355
Physique	0,34	0,26	0,22	0,36	0,295
Chimie	0,30	0,22	0,22	0,27	0,25
Bio	0,30	0,21	0,21	0,20	0,23
	0,33	0,26	0,225	0,32	

6.2 Les corrélations entre les 4 Check-up communs et les 2 Check-up **sciences humaines** (Art et Histoire-éco-actualité) sont les suivants, avec une Moyenne = 0,35 :

0,35	Vocabulaire	Syntaxe	Compréh. Texte	Compréh. Géo	Moyenne
Art	0,47	0,38	0,25	0,38	0,37
Histoire-éco-Actualité	0,48	0,25	0,16	0,41	0,32
	0,475	0,315	0,205	0,395	

6.3 Les corrélations **intra** checkup **communs** ont une Moyenne = 0,42 :

0,42	Vocabulaire	Syntaxe	Compréh. Texte	Compréh. Géo	Moyenne
Vocabulaire	x	0,52	0,37	0,45	0,45
Syntaxe		X	0,41	0,38	0,44
Compréh. Texte			X	0,36	0,38
Compréh. Géo				x	0,40

6.4 Les Corrélations **intra** check-up **sciences de la nature** (math, physique, chimie, biologie) ont une Moyenne = 0,54 :

0,54	Math	Physique	Chimie	Biologie	Moyenne
Math	x	0,67	0,55	0,44	0,55
Physique		X	0,56	0,50	0,58
Chimie			X	0,51	0,54
Biologie				x	0,48

6.5 La corrélation entre les 2 check-up spécifiques aux **sciences humaines** (d'une part, **Connaissances Artistiques** et d'autre part, **Histoire-Economie-Actualité**) vaut : **0,57**.

6.6 Constats pour les **Taux de Réponses Correctes**

1. Les **épreuves spécifiques aux facultés de sciences de la nature** (math, physique, chimie, biologie) sont bien corrélées entre elles (de 0,50 à 0,67 pour 5 corrélations sur les 6 possibles).
2. Les **deux épreuves spécifiques aux sciences humaines** (Art et Histoire-éco-actualité) sont bien corrélées entre elles (0,57).
3. Le **Vocabulaire** est bien corrélé avec la Syntaxe (0,52) puis avec les deux épreuves spécifiquement Sciences Humaines (0,48 et 0,47).
4. La **Compréhension-Géo (de cartes, de graphiques et de tableaux)** est moyennement corrélée avec le vocabulaire (0,45), les math (0,44), l'histoire (0,41), l'art (0,38), la syntaxe (0,38), la compréhension de textes (0,36) et la physique (0,36). Cette épreuve, conçue comme testant des **compétences « transversales »** apparaît donc bien comme mesurant des compétences en intersection à la fois avec les épreuves de sciences humaines et celles de sciences. Elle apparaît dans l'échelle du milieu dans la représentation suivante en 3 échelles où figurent les corrélations (dont les valeurs sont indiquées le long de l'échelle) les plus élevées (supérieures à 0,33). Les corrélations inter-épreuves relevant des sciences de la nature figurent à gauche et celles relevant des sciences humaines à droite.

<p>67 Math -Physique</p> <p>56 Physique-Chimie 55 Math-Chimie</p> <p>51 Chimie- Biologie 50 Physique - Biologie</p>	<p>45 Math Compréh. Gé -----45</p> <p>44 Biologie - Math -----Compréh. Gé</p> <p>41 Math Compréh. Gé -----41</p> <p>39 Math -----Vocabulaire</p> <p>38 Physique Compréh. Gé -----38</p> <p>37 Physique -----</p> <p>36 Physique -----Compréh. Gé-----36</p> <p>34 Physique -----Vocabulaire</p> <p>34 Math -----34</p>	<p>57 Art - Histoire</p> <p>52 Vocabulaire - Syntaxe</p> <p>48 Vocabulaire-Histoire 47 Vocabulaire -Art</p> <p>Vocabulaire</p> <p>41 Syntaxe - Compréh. Texte Histoire</p> <p>Art - Syntaxe 37. Vocabulaire -Compréh. Texte Compréh. Texte</p> <p>Syntaxe</p>
---	--	--

7. Groupements des corrélations concernant les Moyennes Spectrales

7.1 Les 16 corrélations entre les 4 check-up **sciences de la nature** (math, physique, chimie, biologie) et les 4 épreuves communes sont les suivantes (avec une Moyenne = 0,26)

0,26	Vocabulaire	Syntaxe	Compréh. Texte	Compréh. Géo	Moyenne
Math	0,37	0,29	0,40	0,46	0,38
Physique	0,28	0,20	0,14	0,30	0,23
Chimie	0,27	0,19	0,17	0,24	0,22
Bio	0,28	0,19	0,21	0,15	0,21
	0,3	0,22	0,23	0,29	

7.2 Les 8 corrélations entre les 4 Check-up communs et les 2 Checkup **sciences humaines** (Art et Histoire-éco-actualité) sont les suivants (avec une Moyenne = 0,32)

0,32	Vocabulaire	Syntaxe	Compréh. Texte	Compréh. Géo	Moyenne
Art	0,42	0,34	0,22	0,30	0,32
Histoire-éco-Actualité	0,45	0,25	0,17	0,37	0,31
	0,44	0,30	0,20	0,34	

7.3 Les 6 corrélations **intra** check-up **communs** ont une Moyenne = 0,37 :

0,37	Vocabulaire	Syntaxe	Compréh. Texte	Compréh. Géo	Moyenne
Vocabulaire	x	0,49	0,32	0,39	0,40
Syntaxe		X	0,30	0,34	0,38
Compréh. Texte			X	0,39	0,34
Compréh. Géo				x	0,37

7.4 Les 6 Corrélations **intra** check-up **sciences de la nature** (math, physique, chimie, biologie) ont une Moyenne = 0,38

0,38	Math	Physique	Chimie	Biologie	Moyenne
Math	x	0,54	0,41	0,27	0,40
Physique		X	0,39	0,32	0,42
Chimie			X	0,35	0,38
Biologie				x	0,31

7.5 La corrélation entre les 2 checkup spécifiques aux **sciences humaines** (d'une part, **Connaissances Artistiques** et d'autre part, **Histoire-Economie-Actualité**) vaut : **0,52**

8. Groupements des corrélations concernant les Erreurs Moyennes de Centration

8.1 Les 16 corrélations entre les 4 check-up **sciences de la nature** (math, physique, chimie, biologie) et les 4 épreuves communes sont les suivantes (avec une Moyenne = **0,43**)

0,43	Vocabulaire	Syntaxe	Compréh. Texte	Compréh. Géo	Moyenne
Math	0,50	0,51	0,40	0,46	0,47
Physique	0,45	0,44	0,33	0,38	0,40
Chimie	0,37	0,35	0,28	0,35	0,34
Biologie	0,46	0,41	0,34	0,36	0,39
	0,45	0,43	0,34	0,39	

8.2 Les 8 corrélations entre les 4 check-up communs et les 2 Check-up **sciences humaines** (Art et Histoire-éco-actualités) sont les suivants, avec une Moyenne = 0,38

0,38	Vocabulaire	Syntaxe	Compréh. Texte	Compréh. Géo	Moyenne
Art	0,43	0,40	0,28	0,34	0,36
Histoire-éco-Actualité	0,48	0,43	0,33	0,39	0,41
	0,46	0,42	0,31	0,37	

8.3 Les corrélations **intra** check-up **communs** ont une Moyenne = 0,49

0,49	Vocabulaire	Syntaxe	Compréh. Texte	Compréh. Géo	Moyenne
Vocabulaire	x	0,62	0,49	0,49	0,53
Syntaxe		X	0,49	0,45	0,52
Compréh. Texte			X	0,39	0,46
Compréh. Géo				x	0,44

8.4 Les Corrélations **intra** check-up **sciences de la nature** (math, physique, chimie, biologie) ont une Moyenne = 0,43

0,43	Math	Physique	Chimie	Biologie	Moyenne
Math	x	0,49	0,42	0,47	0,46
Physique		X	0,40	0,43	0,44
Chimie			X	0,39	0,40
Biologie				x	0,43

8.5 La corrélation entre les 2 check-up spécifiques aux **sciences humaines** (d'une part, **Connaissances Artistiques** et d'autre part, **Histoire-Economie-Actualité**) vaut : **0,58**

Conclusions sur les EMC

Les corrélations moyennes sont assez élevées (0,43 en moyenne). Les étudiants qui se surestiment (ou se sous-estiment) à une épreuve sont, dans une certaine mesure ($r = 0,43$) les mêmes que dans une autre épreuve. Cette stabilité ouvre des perspectives d'intervention : on peut espérer améliorer ce comportement d'auto-évaluation qui apparaît comme assez systématique.

9. Groupements des corrélations concernant le Réalisme par Calibration

9.1 Les corrélations entre les check-up **sciences de la nature** (math, physique, chimie, biologie) et les 4 épreuves communes sont les suivantes (avec une Moyenne = 0,15)

0,15	Vocabulaire	Syntaxe	Compréh. textes	Compréh. géo	Moyenne
Math	0,26	0,26	0,12	0,17	0,20
Physique	0,18	0,20	0,04	0,17	0,15
Chimie	0,14	0,13	0,08	0,09	0,11
Bio	0,20	0,11	0,14	0,08	0,13
	0,20	0,18	0,10	0,13	

9.2 Les corrélations entre les 4 Check-up communs et les 2 Check-up **sciences humaines** (Art et Histoire-éco-actualité) sont les suivants, avec une Moyenne = 0,19

0,19	Vocabulaire	Syntaxe	Compréh. Texte	Compréh. Géo	Moyenne
Art	0,24	0,23	0,07	0,19	0,18
Histoire-éco-Actualité	0,27	0,23	0,08	0,21	0,20
	0,26	0,23	0,08	0,20	

9.3 Les corrélations **intra** check-up **communs** ont une Moyenne = 0,21

0,21	Vocabulaire	Syntaxe	Compréh. Texte	Compréh. Géo	Moyenne
Vocabulaire	x	0,35	0,19	0,24	0,26
Syntaxe		X	0,13	0,23	0,24
Compréh. Texte			X	0,09	0,14
Compréh. Géo				x	0,19

9.4 Les Corrélations **intra** check-up **sciences de la nature** (math, physique, chimie, biologie) ont une Moyenne = 0,31

0,31	Math	Physique	Chimie	Biologie	Moyenne
Math	x	0,46	0,37	0,28	0,37
Physique		x	0,29	0,23	0,33
Chimie			x	0,23	0,30
Bio				x	0,25

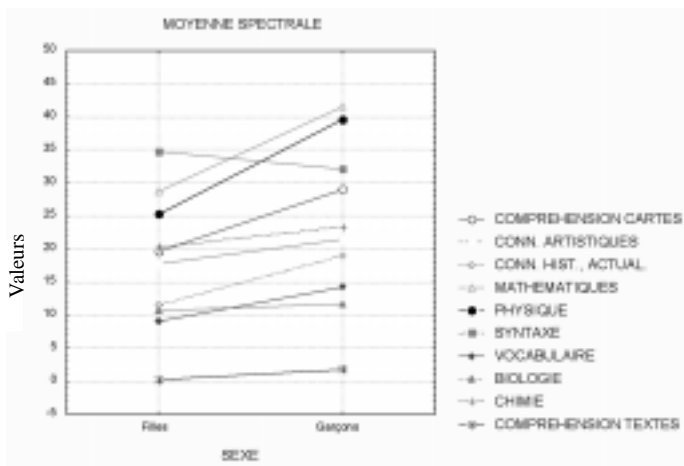
9.5 La corrélation entre les 2 check-up spécifiques aux **sciences humaines** (d'une part, **Connaissances Artistiques** et d'autre part, **Histoire-Economie-Actualité**) vaut : **0,40**

Conclusions sur le réalisme par Calibration

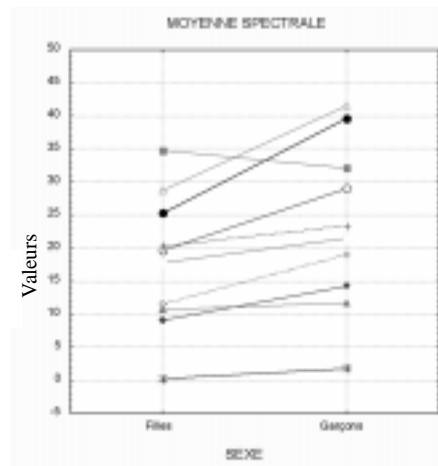
Les corrélations des réalismes intertests sont moins élevées que les EMC (Erreurs Moyennes de Centration) ; c'est donc ce dernier indice qui sera privilégié pour envisager des interventions pédagogiques.

10. Analyse par sexe

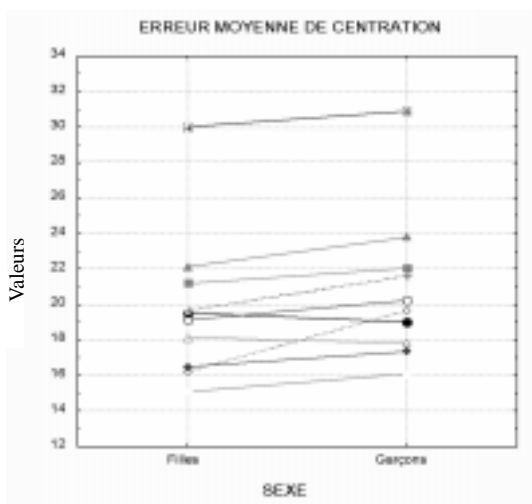
Les nombres d’étudiantes et d’étudiants diffèrent d’un check-up à l’autre



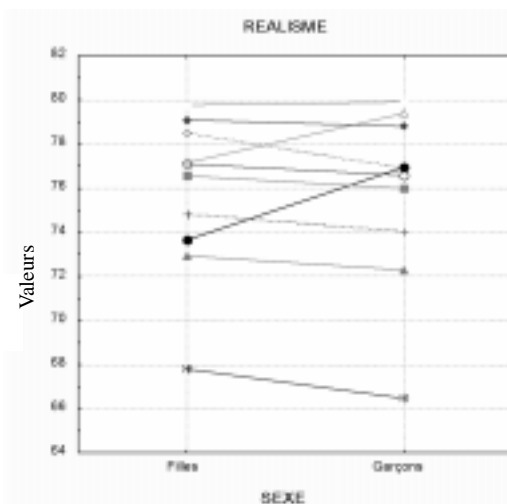
Les garçons sont meilleurs dans tous les tests, sauf pour la syntaxe.



Les garçons sont meilleurs dans tous les tests, sauf pour la syntaxe.



Les garçons se surestiment en moyenne d’1% en plus que les filles. Tous deux se surestiment spécialement en COMPREHENSION de texte.



Le réalisme est généralement inférieur d’un point chez les garçons. Tous deux ont un réalisme faible en COMPREHENSION de texte.

Rappel : Erreur Moyenne de Centration = Certitude Moyenne-Exactitude Moyenne. Si le résultat est positif, il y a surestimation et s’il est négatif, il y a sous-estimation. Attention ! Une certitude moyenne parfaitement égale à l’exactitude moyenne (EMC = 0) peut résulter de la compensation de surestimations par des sous-estimations.

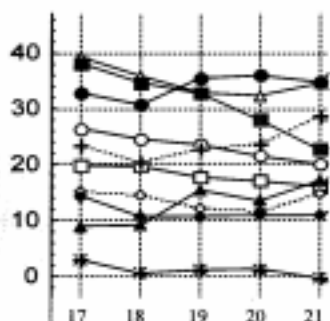
Le réalisme par Calibration se calcule par la formule $100 - EMC$ où EMC est la moyenne des Erreurs de Certitude en valeur Absolue, ce qui exclut les compensations des surestimations par des sous-estimations. EMC est toujours plus grand ou égal à EMC.

11. Analyse par l'âge à l'entrée

Age	17	18	19	20	21	22	23	24	25	26	27
Nombre	582	2135	888	326	119	57	22	16	8	5	2
%	14	51	21	8	3	1	0.5	0.4	0.2	0.1	<0.1

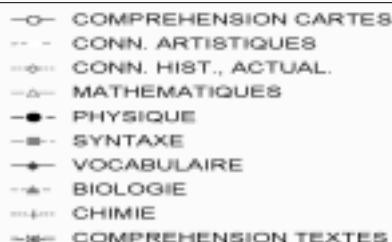
NB : Les nombres de sujets ne sont suffisants que de 17 à 21 ans. Nos commentaires ne portent donc que sur ces valeurs.

Moyenne spectrale

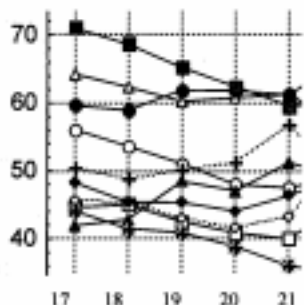


La performance est d'autant meilleure que les sujets sont jeunes à l'entrée. Par rapport aux étudiants « à l'heure » (18 ans), la différence moyenne des « avancés d'un an » (17 ans) est +2 pour la Syntaxe, la Compréhension de cartes, les Connaissances Artistiques, la Compréhension de Textes et -2 pour les « retardés d'un an ». Pour les sciences, l'âge (+ de 18 ans) est au contraire lié à une amélioration.

Entre 17 ans et 21 ans, la différence moyenne est de 5 environ.



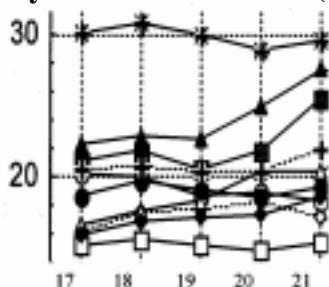
%RC ou Taux de Réussite



Entre 17 ans et 21 ans, la performance moyenne décroît de 10 % pour Syntaxe, compréhension de cartes et de textes, de 5% pour Connaissances artistiques, de 2% pour Connaissances historiques, Math est stationnaire pour Vocabulaire, croît de 2% pour Physique, de 5% pour chimie. de 10% pour Biologie.

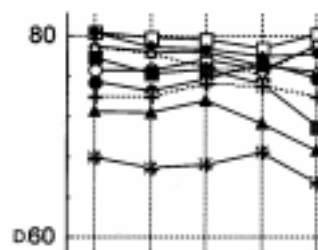
Dans les deux cas, ce phénomène est dû à la performance meilleure des Français, plus âgés, et plus présents en facultés des sciences de la nature qu'en sciences humaines.

Erreur Moyenne de Centration (idéal = 0)



La surestimation est générale pour tous les check-up. Elle culmine pour « compréhension de texte » de la même façon (+30%) pour tous les âges. Elle croît légèrement avec l'âge pour Biologie, Chimie, Syntaxe, Vocabulaire, Connaissances Historiques.

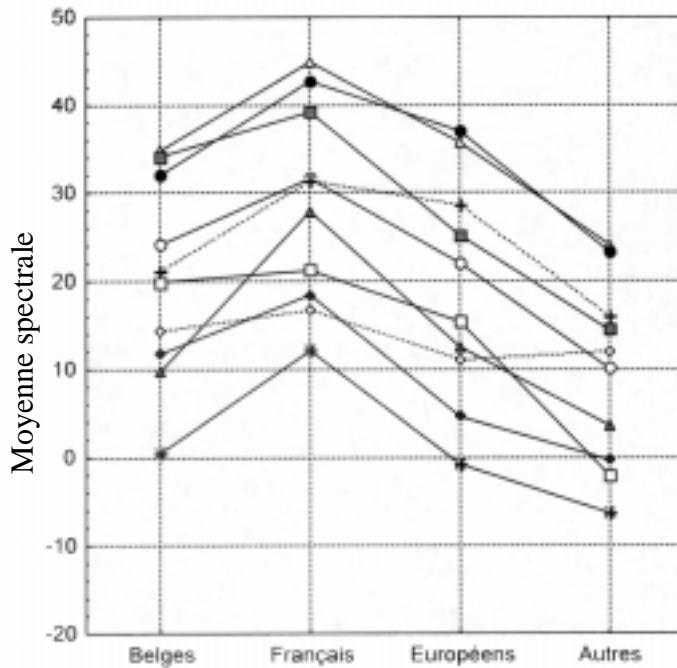
Réalisme par Calibration (idéal = 100)



Le réalisme est en général égal pour tous les âges, sauf pour la Syntaxe et la Biologie.

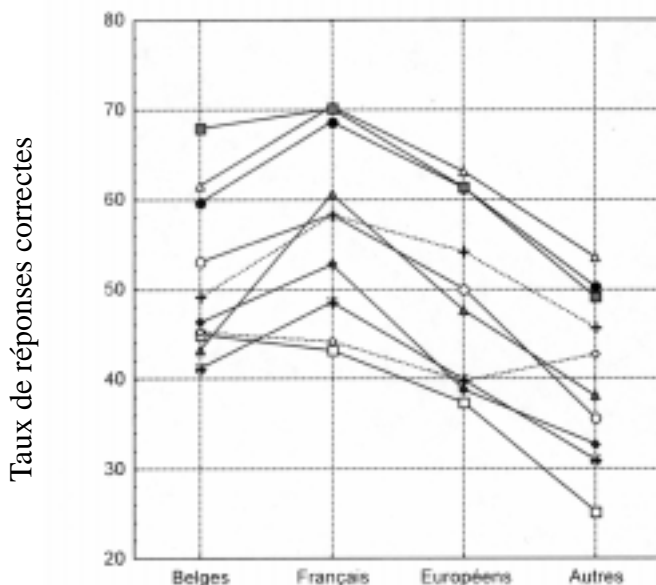
12. Analyse par nationalité

Nationalité	Belges	Français	Européens	Autres
Nombre d'étudiants	3668	225	142	130
%	88	5	4	3

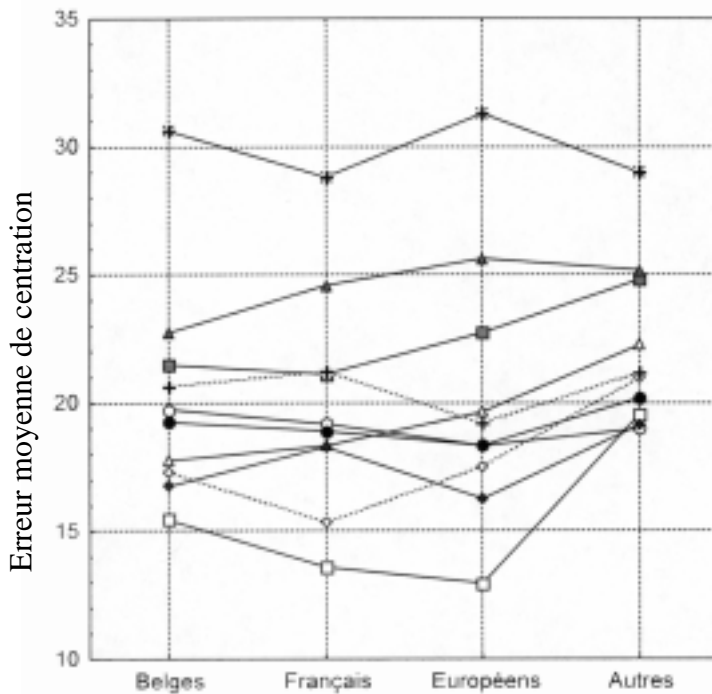


Les **Français** (qui ont effectué une année d'études supérieures supplémentaire en moyenne) ont des résultats supérieurs aux autres dans tous les check-up.

En biologie, chimie et physique, les **Etrangers européens** sont supérieurs aux Belges. En math et compréhension de textes, ils sont égaux.



Même constat : Supériorité des étudiants français. Quoi d'étonnant quand on sait qu'ils entrent dans les universités de la communauté française de Belgique parfois après avoir effectué une ou deux années (DEUG) d'études universitaires en France et s'être préparés aux examens d'entrée aux Grandes Ecoles.



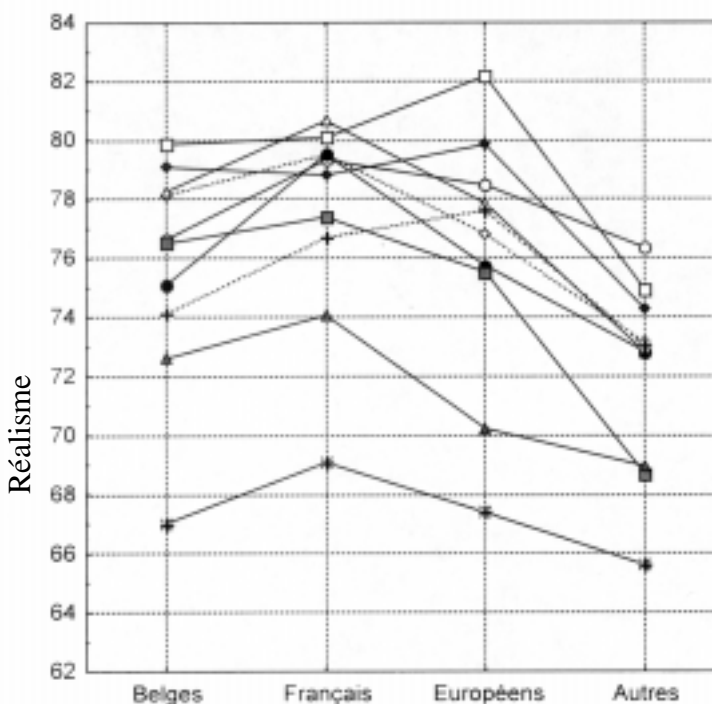
Par rapport aux étudiants d'autres nationalités,

a) les **Belges** ont une Erreur Moyenne de Centration (EMC) *inférieure* (donc meilleure) en math, biologie,

supérieure (donc moins bonne) en compréhension de cartes, connaissances artistiques, physique

b) les **Français** sont meilleurs (EMC plus faible) en Histoire, économie et actualité, compréhension de textes, syntaxe

c) les **Européens** sont meilleurs (EMC la plus faible) en art, physique, chimie.



Excepté pour la compréhension de textes, les **Belges** et les **Français** ont un réalisme par Calibration compris entre 73 et 80. (donc une EMAC ou Erreur Moyenne Absolue de Centration comprise entre 20 et 27)

Les **Européens** ont un réalisme par Calibration plus « dispersé » : de 67 à 82.

Les **Autres** ont un réalisme par Calibration plus faible : compris entre 69 et 77.

NB : Les nombres d'étudiants par nationalité diffèrent d'un test à l'autre.

13. Conclusions

En section F du présent chapitre, nous avons pu mettre en évidence une stabilité moyenne assez forte (corrélations moyennes de 0,54) des Taux de Réussite entre divers check-up reliés par le contenu. Comme on pouvait s'y attendre, les « forts en math » sont aussi, grosso modo, les « forts en chimie, en biologie et en physique », les « forts en Histoire » sont aussi grosso modo les « forts en Connaissances artistiques ».

De telles corrélations étaient attendues et ne surprennent pas.

Les garçons ont en moyenne des scores légèrement supérieurs aux filles en taux d'exactitude et en Moyenne Spectrale.

Par contre, on en savait moins sur la stabilité inter check-up chez un même étudiant d'indices métacognitifs comme le Réalisme par Calibration et l'**Erreur de Centration** (EMC), ce dernier étant en outre un indice de sur et sous-estimation. La corrélation moyenne des EMC entre les check-up de 0,43 mise en évidence en section H du présent chapitre est lourde de conséquences. Elle signifie qu'il existe des tendances assez stables chez les étudiants à se surestimer ou se sous-estimer et que l'on peut imaginer, le sachant, avoir un effet sur le comportement auto-évaluatif par un entraînement approprié.

On constate, en moyenne, une **surestimation, plus forte chez les garçons que chez les filles**, avec des variations de check-up à check-up.

On a pu constater aussi, et de manière répétée pour les 10 check-up, l'**ampleur des différences inter-individus** sur chacun de ces indices, dont certains sont nouveaux, comme la Moyenne Spectrale (**MS**), par exemple.

Comme annoncé en section H du chapitre 1, ce sont surtout les **indices métacognitifs** (indices de Réalisme par Calibration, d'Erreur Moyenne de Centration, de Confiance, de Prudence et de Discriminance qui pourraient servir de références à des expériences ultérieures même si on utilise d'autres questions.

Bien entendu, si les **mêmes questions** sont réutilisées, la comparaison sera possible, mais en tenant compte du public-cible, du moment de l'année, des conditions de passation, des consignes, de la préparation des étudiants testés.

Quant à l'**âge à l'entrée** : On ne constate pas de supériorité systématique des résultats en faveur des « -de 18 ans » par rapport aux « 19 ans et plus », alors que, nous le savons, les taux de réussite de la première année de candidature sont d'habitude forts différents en moyenne (voir Introduction section A).

Enfin, le phénomène « **supériorité moyenne des Etudiants français** » a été expliqué en termes d'âge plus élevé à l'entrée (un an de plus en moyenne) dû à une expérience préalable de l'enseignement supérieur et à la préparation de concours et d'examens.

Ces données « brutes » sont à mettre en correspondance avec les données corrélationnelles avec le succès en première année (voir chapitre 7).

Chapitre 4

Incorporation des feedbacks MOHICAN dans les interviews individuelles

*B. Noël
S. Goemaere,
FUCaM*

Les lignes qui suivent traitent du feedback personnalisé du check-up MOHICAN donné aux étudiants inscrits en première année aux Facultés universitaires catholiques de Mons (FUCaM).

1. Contexte

Les FUCaM regroupent deux facultés, une faculté de sciences de gestion et une autre de sciences politiques. En première année, les effectifs de première génération oscillent entre 150 et 200 étudiants, les filles et les garçons étant représentés quasiment dans les mêmes proportions. Dans la région montoise, seulement 15% des jeunes en âge d'entrer à l'université choisissent l'enseignement universitaire. Nos étudiants proviennent de plus en plus d'un milieu socioculturel défavorisé et les enseignants constatent une maîtrise imparfaite de la langue maternelle pratiquée par les nouveaux étudiants.

Avant d'entrer dans le vif du sujet, nous resituerons le cadre historique du test de français aux FUCaM. Depuis 93-94, nous proposons aux nouveaux étudiants, dès la première semaine de cours, un test informatif de français composé de plusieurs sous-épreuves : vocabulaire, articulations logiques, prise et traitement de l'information et compréhension de texte. A l'instar du check-up MOHICAN, ce test n'est d'aucune manière un examen d'entrée aux Facultés. L'objectif essentiel de sa passation est de faire prendre conscience à l'étudiant de ses lacunes éventuelles et de le rendre attentif aux conseils prodigués en la matière. A l'issue de la correction du test, c'est pour l'équipe d'accompagnement l'occasion d'avoir un échange individuel avec la plupart d'entre eux. De 93-94 à 01-02, le nombre d'étudiants qui se sont présentés à l'entretien n'a fait qu'augmenter : de 76% la première année, nous sommes passés à 93% en octobre 2001.

2. Préparation des rencontres individuelles

En octobre 1999, les FUCaM ont fait partie des institutions universitaires qui ont participé au projet MOHICAN. Le test habituel fut donc remplacé par le check-up MOHICAN ; 93% des étudiants ont répondu à l'invitation et ont reçu un feedback individuel sur leurs résultats.

A partir des documents « Comment lire les résultats » (voir chapitre 2), nous avons conçu une feuille récapitulative simplifiée des données qui nous apparaissaient les plus pertinentes, les plus lisibles pour notre public étudiant et qui, de plus, nous permettaient de donner un feedback individualisé endéans un laps de temps relativement bref. Rappelons que ce rendez-vous habituel de début d'année sert avant tout de prétexte à repérer les difficultés d'adaptation éventuelles des étudiants et que nous disposons d'une vingtaine de minutes pour mener à bien l'entièreté de l'entretien. Nous avons donc repris chacune des sous-épreuves constituant les tests 1 et 2 (voir feuille annexée) en notant, pour chacune d'elles, le nombre de questions posées, le nombre de réponses données par l'étudiant, le nombre de réponses correctes, le pourcentage de réponses correctes ou la facilité objective, le pourcentage d'erreur globale de certitude (surestimation ou sous-estimation) et enfin le réalisme par Calibration. En fin d'entretien, nous remettons à chaque étudiant l'enveloppe MOHICAN renfermant ses résultats confidentiels, personnalisés et détaillés qu'il pourrait découvrir par lui-même.

3. Le déroulement des rencontres individuelles

Lors de l'entretien, nous mettions en évidence les sous-épreuves les moins bien réussies, les erreurs commises le plus fréquemment par l'étudiant, la surestimation systématique ou, plus rare, la sous-estimation de ses résultats, en sollicitant de sa part une justification éventuelle, une reconnaissance de ses points faibles et de ses points forts. Par exemple, en vocabulaire, de nombreux étudiants invoquaient une maîtrise imparfaite et habituelle de ce type de savoirs (« je n'ai jamais été fort en vocabulaire », « cela a toujours été mon point faible en langue maternelle », etc.). En compréhension, la plupart des étudiants ont prétexté la longueur exagérée du texte à lire et le nombre trop élevé de questions posées à son sujet (« je n'ai pas eu le temps de (re) lire le texte », « je n'ai pas eu le temps de répondre à toutes les questions », etc.). Face à des résultats médiocres, nous avons été amenées à rassurer les étudiants en invoquant la difficulté de l'épreuve, son caractère expérimental (ce pretesting du check-up devrait servir à la construction ultérieure d'une épreuve mieux adaptée, plus spécifique à leur orientation d'étude, probablement plus courte aussi). Ensuite, nous donnions quelques conseils, formulions quelques recommandations remédiatives mais sans pouvoir proposer une aide effective dans la maîtrise de la langue maternelle, faute de ressources humaines pour prendre en charge ce type de difficultés. Par exemple, en cas de résultats insuffisants en vocabulaire, de surestimation de leurs résultats, nous préconisons de recourir de façon plus systématique au dictionnaire, de faire appel au professeur ou aux assistants qui dispensent les cours non quantitatifs en première année ; en ce qui concerne la syntaxe et les articulations logiques, nous proposons de relire leurs résumés (en philosophie, sociologie ou histoire) ou de les comparer à ceux de leurs pairs de façon à être sûrs de leur support essentiel d'étude en vue de l'examen. Ajoutons toutefois que, face à quelques étudiants demandeurs, nous avons entrepris, via les collègues chargés de l'encadrement des travaux dirigés en philosophie, des opérations formatives destinées à aider les plus démunis sur le plan formel.

4. Quelques éléments révélateurs

A l'occasion de l'entretien, nous avons également relevé quelques commentaires que nous livrons tels quels sans aucune analyse statistique des données :

- De nombreux étudiants nous disent être venus au rendez-vous pour connaître leur résultat aux différentes sous-épreuves : c'est d'ailleurs la raison invoquée dans notre invitation à nous rencontrer (assortie d'une date et d'un moment de rendez-vous).
- Malgré des résultats peu brillants, nous n'avons pas enregistré de découragement et nous avons constaté en général une écoute attentive de nos conseils.
- De nombreux étudiants ont avoué avoir laissé tomber la résolution de la (des) sous-épreuve (s) de compréhension et/ou de lecture des documents géographiques, faute de temps. Nous enregistrons en effet des résultats faibles à ces deux parties d'épreuve, respectivement 59% de réponses incorrectes pour l'une, 44% pour l'autre.
- Plusieurs étudiants nous ont également confié, qu'en ce qui concerne les items relatifs au degré de certitude de leurs réponses, ils n'arrivaient plus à faire l'effort de discriminer finement entre 20, 40, 60 voire 80% de certitude, surtout en fin d'épreuve.
- Plusieurs étudiants enfin, nous ont questionnées sur la position relative des FUCaM par rapport aux autres universités participantes. N'ayant pas encore reçu cette information plus globale, nous répondions ne pas être en possession, à ce moment-là, des données y afférentes.

5. L'après-MOHICAN

Suite à l'opération MOHICAN suspendue après 99, nous avons réintroduit notre test habituel mais en apportant quelques modifications importantes, tant du point de vue formel que du contenu, dont plusieurs furent empruntées au check-up. Avec l'accord de M. Monballin, nous y avons injecté quelques propositions de définitions qui s'étaient révélées mal maîtrisées par les étudiants et ayant recueilli une certitude moyenne élevée à très élevée, ce que Leclercq et Poumay (2002, chap. 10) désignent par « imprudence ». Il s'agit des mots suivants : arbitraire, empirique et réactionnaire. De même, après demande auprès de Ph. Hougardy, nous avons choisi de remplacer la question de synthèse présente dans notre test depuis ses débuts, en sélectionnant cinq questions parmi celles de l'épreuve de compréhension qu'il avait construite pour le check-up MOHICAN. Dans l'avenir, une analyse comparée de quelques résultats partiels entre institutions universitaires pourrait être envisagée.

A l'issue de nos entretiens, qu'ils soient consécutifs à la passation de notre test ou à celle du check-up MOHICAN, notre grand regret est de ne pouvoir proposer aux étudiants concernés les moyens susceptibles de combler leurs manques.

Chapitre 5

La consultation du feedback MOHICAN : Fréquences, intérêts et implications

*M. Delforge et B. Diricq, Faculté Polytechnique de Mons
P. Dupont, Université de Mons–Hainaut*

Le présent chapitre traite des déclarations des étudiant(e)s sur la manière dont ils (elles) ont accueilli le feed-back des check-up MOHICAN qui leur était proposé. Deux institutions, l'université de Mons – Hainaut et la Faculté Polytechnique de Mons, se sont associées dans cette démarche et ont soumis leurs étudiants, quelques jours après la remise des résultats, à un questionnaire dont l'objectif était d'une part d'estimer le taux de consultation et, d'autre part, d'investiguer les sources du comportement étudiant à cet égard.

Le questionnaire tente d'aborder le feed-back à la fois sous l'angle de son contenu et de l'intérêt qu'il éveille chez les étudiants et sous l'angle de sa présentation, jugée par trop complexe par plusieurs des membres du groupe de travail ayant présidé à sa création.

Enfin, la littérature en pédagogie universitaire l'a montré à souhait, les résultats des recherches qui sont menées dans son champ sont toujours hautement contextualisés : une même pratique, une même variable mises en œuvre dans des environnements différents voient leur impact modifié du tout au tout. Il nous a donc paru intéressant de soumettre à ce même questionnaire des populations, différenciées s'il en est. Notre premier point s'attache à les décrire.

1. Deux groupes contrastés

Nous l'avons dit, les deux facultés investiguées sont très différentes, tant dans leurs caractéristiques institutionnelles que dans leur recrutement.

La Faculté de Psychologie et des Sciences de l'Éducation est relativement jeune puisque sa création remonte à une trentaine d'années. Elle est l'une des cinq facultés de l'Université de Mons – Hainaut. Bon an, mal an, toutes années et tous programmes confondus, elle avoisine le millier d'étudiants (Moreau, 1999 : 22) ; en première année, les effectifs varient entre 400 et 500 étudiants et ce sont majoritairement des filles qui choisissent cette orientation : elles représentent environ trois quarts des inscrits (Moreau, 1999 : 19). Du point de vue de l'origine sociale, c'est principalement parmi les couches populaires que la faculté recrute : les chiffres et les typologies utilisées varient selon les sources mais on croit pouvoir affirmer sans trop de risque d'erreur que 50 % des inscrits proviennent de familles où le père est inactif ou « col bleu » (Moreau, 1999 : 16 et 17). En ce sens, elle constitue une réelle force d'ascension sociale dans une région très défavorisée.

La Faculté Polytechnique de Mons est quant à elle la plus ancienne école d'ingénieurs de Belgique. Il s'agit d'un établissement monofacultaire en Sciences appliquées qui compte entre 1000 et 1200 étudiants. La première candidature rassemble, en moyenne de 250 à 300 inscrits. Comme toute faculté de Sciences appliquées à l'heure actuelle, sa population féminine oscille entre 15 et 20 %. L'aspect socialement élitiste du recrutement de ce type de faculté est perceptible. Toutefois, l'examen des chiffres de la densité d'ingénieurs diplômés par arrondissement belge couplé à la situation économiquement défavorisée du Hainaut occidental en font également un réel tremplin dans l'ascension sociale des habitants de la région montoise. L'obligation légale de réussir des épreuves d'admission qui est imposée à toutes les facultés de Sciences appliquées limite son recrutement. Ainsi, tous les étudiants inscrits en première candidature à la FPMs ont présenté au moins 5 épreuves mathématiques et, selon le cas, des épreuves complémentaires.

2. Intérêt pour les résultats des tests précoces

Les questionnaires d'enquête ont été diffusés lors d'une heure de cours. Ceci induit donc une spécificité dans les caractéristiques des répondants : les décrocheurs précoces et les peu assidus aux séances en sont absents. Au total, les réponses de 260 étudiants sont exploitables.

Taux de consultation et caractéristiques des étudiants

Nombre d'étudiants (total = 260)	76	184
	Psycho péda	Polytechnique
% de consultation parmi		
La section (G+F)	72,4 %	83,7 %
Les Filles	71,9 %	69,0 %
Les Garçons	73,7 %	86,5 %

Tableau I : Pourcentage de consultation selon les facultés et le sexe.

Au total des deux facultés, **80,4 %** des étudiants (209 sur 260) disent avoir examiné leurs résultats. Toutefois, les futurs polytechniciens sont significativement plus nombreux à l’avoir fait¹⁸. Si en psychologie et sciences de l’éducation le sexe n’apparaît pas comme une variable discriminante¹⁹, elle crée un croisement significatif pour le sous-échantillon de la faculté polytechnique²⁰ : en effet, un tiers des filles dit n’avoir pas pris connaissance du feed-back contre seulement 13,5 % des garçons.

Autre variable indépendante, l’appartenance sociale peut-elle se révéler déterminante dans la consultation des résultats ?

Diplôme le plus élevé obtenu par le père	Psycho péda	Polytechnique
Pas ou peu diplômés	68	84,4
Diplômés ens. Secondaire	78.3	73,7
Diplômés ens. supérieur	73.1	90,6

Tableau II : Pourcentage de consultation selon les facultés et l’appartenance sociale

Nous retenons comme critère de niveau social le diplôme le plus élevé obtenu par le père. Si, comme c’était le cas précédemment pour la variable sexe, la ventilation des répondants entre « consulté » et « pas consulté » ne crée pas de différence significative parmi les étudiants de psychopéda²¹, c’est en revanche le cas chez les futurs ingénieurs²². Les étudiants dont le père possède un diplôme de l’enseignement secondaire ont consulté leurs résultats dans une moindre proportion que les autres et ceux issus de familles où le père est détenteur d’un diplôme de l’enseignement supérieur ont consulté le bilan des épreuves Mohican dans une proportion plus élevée.

Enfin, dernière piste investiguée pour rendre compte du taux de consultation : le fait d’avoir connu l’échec dans la scolarité antérieure sensibilise-t-il à l’intérêt de tests informatifs ? Il est difficile d’apporter une réponse fiable à cette question pour le groupe des polytechniciens. En effet, dans notre échantillon, seuls 7 étudiants ont connu le redoublement. Il se fait qu’ils ont tous consulté mais il serait abusif de tableur sur un effectif si restreint. Chez les futurs psychologues et licenciés en sciences de l’éducation, le croisement est non significatif²³.

¹⁸ ($\chi^2 = 4.377$, d.l. = 1, P = .036)

¹⁹ ($\chi^2 = .022$, d.l. = 1, P = .882)

²⁰ ($\chi^2 = 5.474$, d.l. = 1, P = .019)

²¹ ($\chi^2 = .640$, d.l. = 2, P = .726)

²² ($\chi^2 = 7.227$, d.l. = 2, P = .027)

²³ ($\chi^2 = 1.430$, d.l. = 1, P = .232)

2.1 Pourquoi ils n'ont pas consulté leurs résultats...

Interrogés sur les raisons de la non consultation, les étudiants évoquent les motifs suivants :

Raisons invoquées	Total	Psy	Polytec
Les questions ne me semblent pas représentatives, dans leur contenu , de ce qui est attendu de moi à l'université.	58,3	27,8	76,7
Les questions ne me semblent pas représentatives, dans leur niveau de difficulté de ce qui est attendu de moi à l'université.	51,0	10,5	76,7
Quels que soient mes résultats, ils ne changeront pas ma manière de travailler.	37,5	55,6	26,7
J'ai oublié.	32,7	42,1	26,7
Je n'ai pas répondu avec tout le sérieux nécessaire et les résultats ne sont pas le reflet de mes capacités.	20,4	21,1	20,0
Le simple fait de répondre aux questions me donne une indication sur mon niveau.	18,4	31,6	10,0
J'ai perdu mon numéro check-up.	16,3	10,5	20,0
Je ne souhaite pas cautionner un système de tests, quel qu'il soit, en début de cursus universitaire.	8,2	21,1	0

Tableau III : raisons de la non consultation selon la faculté en pourcentage

Nos réponses sont de l'ordre du déclaratif et dans une certaine mesure, elles sont le reflet de désirabilité sociale et de mécanismes de défense. Une attribution causale externe mettant en question la qualité du test est une attitude plus « confortable » à vivre qu'une réelle interrogation sur les mobiles profonds qui ont amené à ne pas faire la démarche de consultation.

À l'exception de la proposition « *Je n'ai pas répondu avec tout le sérieux nécessaire et les résultats ne sont pas le reflet de mes capacités* », les raisons qui ont poussé les étudiants à ne pas prendre connaissance des résultats de l'évaluation opposent fortement les deux facultés. Les futurs polytechniciens concentrent leurs critiques sur le test lui-même : ni le contenu, ni le niveau de difficulté ne leur semblent adaptés. De leur côté, les étudiants de l'U.M.H. mettent en avant l'absence d'impact qu'auraient les résultats sur leur manière de travailler. Dès lors, à quoi bon les lire ?

Seuls les futurs psychopédagogues entrevoient la dimension politique de l'opération : il s'en trouve un sur cinq pour mettre en cause l'existence de tels tests en début de cursus (même s'ils étaient en principe informés de leur visée purement expérimentale et informative). Aucun futur ingénieur (les seuls, dans l'enseignement supérieur en Communauté française, à devoir franchir le barrage de l'épreuve d'admission) ne critique ce système, sans doute parce que tous ont surmonté l'obstacle de la sélection avec succès.

2.2 Pourquoi ils ont consulté leurs résultats...

Les 80,4 % d'étudiants (209 sur les 260, dont 55 en psycho pédagogie et 154 en Polytechnique) qui ont consulté leurs feedbacks se répartissent de la manière suivante :

	Psychopéda	Poly-tech	Psychopéda	Polytech
Par curiosité pour les solutions correctes	15	10	27,3%	6,4%
Par intérêt pour les résultats personnels	35	133	63,6%	86,4%
Parce que cela a semblé utile	5	11	9,1%	7,1%
Total de ceux qui ont consulté	55	154	100%	100%

Pour 80 % d’entre eux, le principal intérêt de la lecture du feed-back réside dans le résultat. Peu d’étudiants (7,7 %) voient une utilité dans l’information qui leur est proposée. De ce point de vue, les étudiants des deux facultés ne se différencient pas. Par contre, bien qu’on retrouve une large majorité des deux groupes dans la catégorie « curieux des résultats », les étudiants de la faculté polytechnique s’y retrouvent préférentiellement²⁴. C’est ainsi que parmi les étudiants en psychologie et sciences de l’éducation, plus d’un quart se dit préférentiellement intéressé par les réponses aux questions.

Les feed-back remis aux étudiants rassemblaient une masse importante d’informations. Une majorité d’étudiants dit avoir éprouvé des **difficultés dans la lecture du feed-back** (58,6 %). L’aspect « avalanche de chiffres » a particulièrement rebuté les étudiants en psychopédagogie : seul un étudiant sur 5 dit ne pas avoir rencontré de difficulté, alors qu’en Polytechnique la moitié des étudiants (49 %) n’a pas eu de difficulté à lire les feedbacks.²⁵

Quel type d’information vous a intéressé(e) ?	Total	Psycho pédas	Polytech niques
Le taux de réponses correctes pour les différentes parties	39,4%	32,7	42,0
La position à l’intérieur du groupe	22,3%	20,0	23,2
Le taux global de réponses correctes	20,2%	25,5	18,1
Les informations relatives aux degrés de certitude	9,8	5,5	11,6
Les réponses aux questions	8,3	16,4	5,1

Tableau IV : intérêt porté aux différents types d’information selon la faculté en pourcentage.

Le désintérêt pour les solutions constaté à la question précédente se retrouve avec, comme précédemment, une moindre fréquence pour les étudiants en psychopédagogie de l’UMH.

D’une manière générale, les étudiants accordent une plus grande attention aux **résultats spécifiques** à une partie donnée. Ceci est surtout vrai pour les futurs polytechniciens. Il ne serait pas étonnant que nos sujets privilégient les informations relatives à ce qui leur semble plus directement en phase avec leur filière d’études. Par exemple, les futurs ingénieurs s’intéresseraient plus volontiers aux résultats de mathématiques ou de physique qu’à ceux de biologie ou de langue française.

Les **degrés de certitude** recueillent peu de suffrages. Peu employés dans les deux institutions qui font l’objet de notre étude, il n’est pas étonnant que les étudiants y soient peu sensibilisés.

Enfin, dans l’ensemble, une proportion non négligeable d’étudiants s’intéresse à sa **position dans le groupe**. Cette information peut sembler fort écartée de l’idée de maîtrise de pré requis, mais il n’est pas étonnant de voir les nouvelles recrues s’y intéresser. Le moment de la transition secondaire – université, notamment par le changement de groupe-classe qu’elle suppose, suscite chez le nouvel étudiant, une interrogation quant à sa place parmi les pairs.

Nous avons également voulu savoir (et cette interrogation n’est pas détachée de préoccupations éthiques) quel était l’**impact de la lecture des résultats sur les étudiants**, tant d’un point de vue de la représentation de soi qu’en ce qui concerne leurs intentions déclarées.

Un tiers des étudiants se dit *rassurés* quant à sa capacité à entreprendre des études universitaires. Cette moyenne de 35,6 % cache une forte disparité entre les deux facultés : ce sont surtout les futurs polytechniciens qui sont dans ce cas (42,9 %) ; les étudiants de l’U.M.H. ne sont que 14,8 % à avoir coché cette proposition (cette différence est effectivement significative : $\chi^2 = 13,717$, d.l. = 1, P = .000).

²⁴ ($\chi^2 = 16,698$, d.l. = 2, P = .000)

²⁵ ($\chi^2 = 14,187$, d.l. = 2, P = .001).

Pour une *large majorité*, les résultats sont conformes à leurs prévisions (63% en psycho, 72,5 % en polytechnique).

Tout comme les étudiants qui n'avaient pas consulté leurs résultats, ceux qui l'ont fait sont encore plus nombreux (55,8%) à **affirmer qu'ils ne changeront pas leur manière de travailler**. Ici aussi les deux sous-groupes diffèrent : si les futurs ingénieurs se répartissent presque équitablement entre les deux positions (48,1 % / 51,9 %), les futurs licenciés en psychologie et sciences de l'éducation sont beaucoup plus tranchés dans leur opinion : 8 étudiants sur 10 *ne modifieront pas leur comportement d'étude* (différence significative : $\chi^2 = 14.322$, d.l. = 1, P = .000).

Il semble que *les aides méthodologiques et les « soutiens matières » ne gagneront pas en fréquentation* suite à la lecture des résultats : 88,9 % (aide méthodologique) et 91,8 % (soutiens matières) ne songent pas à s'y investir. Les futurs psychologues et licenciés en sciences de l'éducation se montrent particulièrement réfractaires : seuls 3,7 % imaginent fréquenter les aides qui leur sont offertes (la différence entre facultés est significative : $\chi^2 = 4.011$, d.l. = 1, P = .045). Travailler plus ne mobilise ni les psy, ni les autres : 74,5 % des étudiants *ne pensent pas à augmenter leur quantité de travail* (différence non significative entre les deux sous-groupes). Chose heureuse : *aucun ne se dit découragé* à la lecture des résultats.

Nous l'avons vu du côté des personnes qui n'avaient pas consulté leurs résultats, les étudiants dans leur ensemble et les futurs ingénieurs particulièrement se montrent très critiques par rapport au test lui-même. Si le niveau de difficulté semble adapté à un quart des interrogés (tant polytechniciens que psychopédagogues), l'absence d'adéquation entre son contenu et les études entreprises est dénoncée par 94,5 % des étudiants de l'U.M.H. Logiquement, les recrues de la Faculté Polytechnique sont moins sévères sur ce point : ils ont sans doute moins de mal à envisager la continuité disciplinaire entre les tests et leur orientation (différence significative donc : $\chi^2 = 8.668$, d.l. = 1, P = .003).

3. En guise de conclusion, quelques interrogations...

L'abondante littérature que la dernière décennie a consacrée aux nouvelles conditions d'accueil des étudiants universitaires et aux déterminants de la réussite académique nous engage à beaucoup de modestie dans la généralisation des résultats qui précèdent. Les importantes différences constatées entre les deux sous-échantillons nous montrent, si besoin en était, que l'exploitation des résultats de recherche passe par une nécessaire contextualisation des données. La suite de nos conclusions prendra la forme d'interrogations.

En ce qui concerne l'existence de tests, même à visée purement informative en début de cursus universitaire, s'ils sont d'un intérêt indéniable pour les chercheurs, remplissent-ils vraiment auprès des étudiants, les bénéficiaires ultimes, leur rôle s'ils constituent une opération « isolée » ? Ne conviendrait-il pas, dans le but de sensibiliser les étudiants à l'apport que peut constituer ce genre d'information, d'insérer ce type d'opération dans un contexte plus large d'accompagnement des étudiants fraîchement inscrits ?

Si la confiance en soi s'avère un facteur déterminant dans la réussite universitaire, il est patent que cette confiance ne doit pas être irréaliste mais plutôt reposer sur une analyse lucide de ses propres possibilités (De Ketele, 1990 : 6). En ce sens, une information valide ne peut que constituer un atout indéniable pour l'étudiant. La validité prédictive et formative des check-up MOHICAN est

étudiée ailleurs. Il était impossible de leur conférer un statut de validité la première année, au moment de la remise du feed-back.

Mener sa vie d'étudiant dans un environnement pour lequel on possède des clefs de lecture efficaces est indéniablement un avantage. Ainsi, se penchant sur les sources d'informations utilisées par les premières candidatures, Donnay (1995 : 109) montre que les étudiants qui réussissent en fin d'année mettent davantage en avant le caractère fiable des sources qu'ils consultent. Nous l'avons vu, de ce point de vue, le feed-back MOHICAN rencontre plutôt la méfiance.

Tant dans le cas de la confiance en soi que dans celui de la fiabilité de l'information, la contextualisation des épreuves nous semble être la meilleure voie. Le check-up ne doit-il pas dès lors évoluer vers l'adoption de modalités spécifiques à chaque cursus ?

Il nous paraîtrait essentiel d'insister sur les indicateurs qualitatifs de l'insertion universitaire (Dupont et Ossandon, 1994), d'autant qu'il existerait un dénominateur commun au décrochage : la difficulté d'évoluer et de s'auto-organiser dans la mécanique universitaire.

Par ailleurs, si la comparaison entre prédiction, actions correctrices et démentis de prédiction constitue le stade épistémologique ultime, nous en sommes encore loin (Romainville, 1997 : 82, Duru – Bellat, 1995 : 414), quoiqu'on en trouvera une ébauche dans l'opération RESSAC. La prédiction parfaite, celle qui dans tous les cas se confondrait avec la réalité à venir n'est ni possible ni souhaitable. Du point de vue de l'enseignant, il importe de contrer le phénomène de translation simple par lequel les situations de départ, individuelles et sociales, sont transposées sans modifications réelles. Du point de vue de l'étudiant, l'attitude la plus souhaitable est de s'appuyer, au besoin, sur la résilience et de s'inscrire dans une perspective sartrienne : « L'important n'est pas tant ce qu'on m'a fait que ce que je fais de ce qu'on m'a fait » ?

Autre crainte, ces « bilans » de culture générale à l'entrée à l'université peuvent – bien que ce ne soit pas l'intention – apparaître comme un bilan de l'enseignement secondaire. Ne serait-il pas opportun d'envisager des épreuves à destination exclusivement des étudiants du secondaire afin de les sensibiliser progressivement à cette approche autoévaluative à laquelle ils sont si peu préparés ?

Plus généralement, sans doute faudrait-il s'interroger sur les profils de formation des étudiants que nous recevons et que nous avons pour mission de former : qu'attendons-nous qu'ils maîtrisent à la base ? Quel domaine disciplinaire ou quelle compétence plus vaste sont indispensables ?

Enfin, d'un point de vue déontologique, relever des lacunes ne peut être fait que si l'on peut proposer aux étudiants concernés les moyens de les combler. Ce qui plaide une fois de plus pour une intégration plus complète des dispositifs destinés aux étudiants.

4. Bibliographie

- De Ketele, J.-M., Le passage de l'enseignement secondaire à l'enseignement supérieur : les facteurs de réussite, *Vie pédagogique*, 66, 1990.
- Donnay, J., Quelques images que se font les étudiants à propos des examens de première année universitaire, *Actes du Colloque de l'Association Internationale de Pédagogie Universitaire*, Hull, 1995.
- Dupont, P., Ossandon, M., *La pédagogie universitaire*, Paris, Presses Universitaires de France, Collection Que sais-je ?, 1994.
- Duru – Bellat, M., Des tentatives de prédiction aux écueils de la prévention en matière d'échec en première année d'université, *Savoir*, 7 (3), 1995.
- Boxus et al., Groupe de travail « Réussite en candidatures », *Rapport de travail*, Bruxelles : Conseil interuniversitaire francophone, 1993.
- Leclercq et al., Groupe de travail « Réussite en candidatures », *Franchir le cap des candis*, Bruxelles : Conseil interuniversitaire francophone, 1997.
- Moreau, M.-L., *Rapport sur la faculté de psychologie et des sciences de l'éducation*, Mons : U.M.H., 1999.
- Romainville, M., Peut-on prédire la réussite d'une première année universitaire, *Revue française de Pédagogie*, 119, 1997.

Chapitre 6

L'information des feedbacks pour les professeurs L'opération VOCA PROFS à l'ULg ou Estimation par les professeurs des résultats des étudiants au check-up de Vocabulaire ²⁶

D. Leclercq, Président du Groupe de Travail CIUF « Réussite en candidatures »

J.-L. Gilles, Directeur du SMART, Université de Liège

C. Dupont, Chercheur au LabSET, Université de Liège

²⁶ Cette recherche-action a été possible grâce au Conseil des Etudes de la FAPSE, ULg, au secrétariat facultaire qui a encodé une partie des données, grâce au CAFEIM-FAPSE et au secrétariat du STE – ULg qui ont pris en charge la plus grande partie des encodages.

1. Quel est l'intérêt du feedback des résultats aux check-up « Vocabulaire » pour les encadrants ?

On entend souvent dire que les enseignants connaissent bien les mots qui font difficulté aux étudiants. C'est à la fois vrai et faux. Sur la matière de leur enseignement, ils savent évidemment, preuves orales et écrites à l'appui, quels sont les termes qui posent problème. Pour les termes généraux du français, il faut être plus nuancé. Dans l'expérience qui suit (et dont nous n'avons aucune garantie qu'elle soit généralisable à d'autres sections et d'autres encadrants), on verra que les professeurs ont une **bonne estimation moyenne des termes plus difficiles** que d'autres (avec un optimisme certain, bref ils seront déçus par les résultats objectifs) mais sans fiabilité prédictive sur chaque terme précis. Ce qui indique l'intérêt de fournir cette information que constitue le % de réussite du groupe pour chaque question. Voici comment ces données ont été mises en évidence.

Les sections et facultés ont reçu dans le courant du mois de novembre 1999 les statistiques relatives à la distribution des résultats de leurs étudiants, sans qu'il soit possible d'identifier les résultats individuels. Nous avons fait une exception pour un check-up dans une section d'une université : les résultats au check-up de Vocabulaire des étudiants en psychologie et Sciences de l'éducation de l'université de Liège. Les étudiants ont reçu leurs résultats quasi immédiatement (dans le courant du mois d'octobre 1999), mais nous avons **attendu le mois de mai 2000** (la fin de l'année) pour demander à 15 enseignants (et assistants) concernés par cette section

« A votre avis, pour chacun des termes ci-dessous qui ont été présentés en QCM à 5 solutions proposées (plus AUCUNE plus TOUTES), quel a été le taux de réussite des 200 étudiants de première année en psychologie et sciences de l'éducation qui ont répondu à ces questions le 10 octobre 1999 ? ».

Mentionnons de suite le **biais** le plus important de ce questionnement : ces enseignants n'avaient pas les questions elles-mêmes sous les yeux, mais les seuls mots de vocabulaire (voir la feuille de consigne ci-après). Or la formulation des questions (à choix multiple, rappelons-le) a indéniablement pesé sur les taux de réussite, et ne pas en disposer ne peut avoir que handicapé le travail d'estimation des collègues.

Les résultats sont présentés par ordre croissant de TRP (Taux de Réussite Prédit) pour ces étudiants **estimés** par leurs encadrants (professeurs et assistants). Il s'agit pour chaque mot de la Moyenne des TRP pour les 15 encadrants.

Ici c'est le terme « **allégation** » (TRP de 38,1%, alors que sa Facilité Objective est de 35%) qui vient en pire position (alors qu'en réalité, pour ce groupe d'étudiants, il y a 12 autres mots dont la Facilité Objective est pire), et le terme « **légal** » (TRP de 95,6%) qui vient en meilleure position (alors qu'en réalité pour ce groupe d'étudiants 12 autres mots sont mieux réussis).

Les discordances entre les FO et les TRP constituent des surprises, tantôt **bonnes** (les étudiants réussissent mieux qu'anticipé), tantôt **mauvaises** (les étudiants réussissent moins bien que prévu).

Les professeurs ont été aussi invités à donner un avis sur l'importance pour leur discipline de chacun de ces mots (non techniques) de la langue française, sur une échelle de 0 à 3. Nous avons calculé la relation entre ce « poids » et la réussite.

2. Les jugements des professeurs

Estimation par les 15 enseignants des Taux de Réussite de chacune des 45 questions sur l'épreuve de Vocabulaire MOHICAN passée le 10 octobre 1999 par les étudiants de la FAPSE.

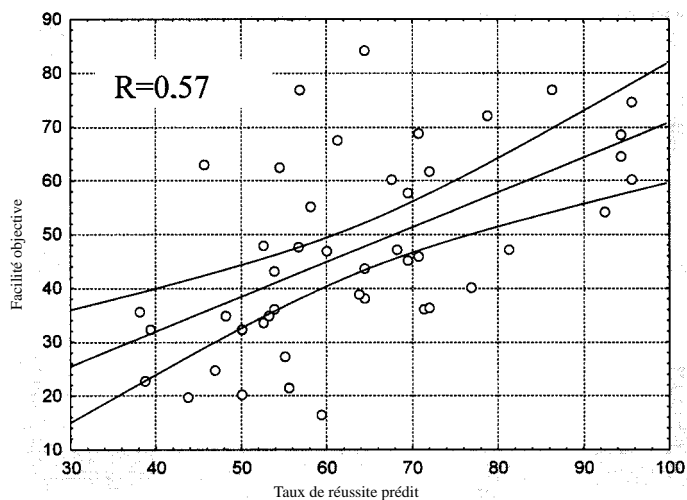
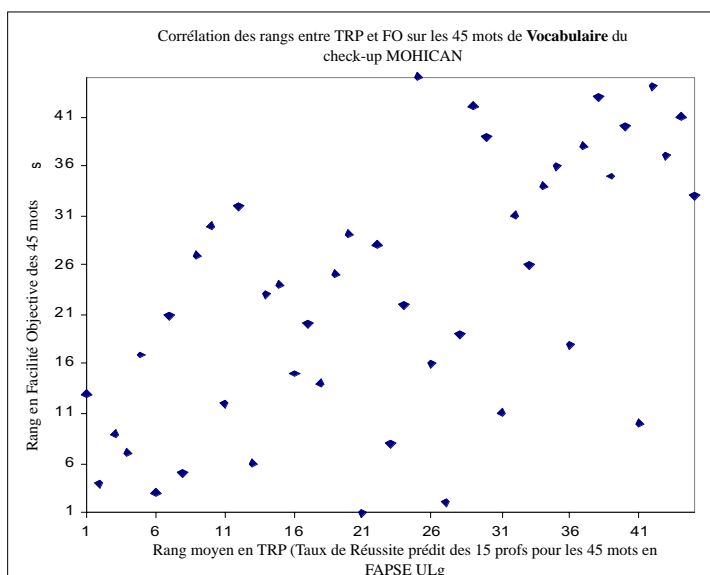
N.B. Tous les mots étaient en contexte et le QCM était à 5 solutions plus 6 « AUCUNE » et 7 « TOUTES ».

	JUGEZ CES IMPORTANCES				Mot (verbe conjugué ou nom ou adjectif)	ESTIMEZ le % de réussite										
	1	2	3			0	10	20	30	40	50	60	70	80	90	100
1	1	2	3		corrobore	0	10	20	30	40	50	60	70	80	90	100
2	1	2	3		prérogatives	0	10	20	30	40	50	60	70	80	90	100
3	1	2	3		stigmatisé	0	10	20	30	40	50	60	70	80	90	100
4	1	2	3		sous-jacentes	0	10	20	30	40	50	60	70	80	90	100
5	1	2	3		latentes	0	10	20	30	40	50	60	70	80	90	100
6	1	2	3		divergent	0	10	20	30	40	50	60	70	80	90	100
7	1	2	3		dénégations	0	10	20	30	40	50	60	70	80	90	100
8	1	2	3		discrédit	0	10	20	30	40	50	60	70	80	90	100
9	1	2	3		a inféré	0	10	20	30	40	50	60	70	80	90	100
10	1	2	3		inhérentes	0	10	20	30	40	50	60	70	80	90	100
11	1	2	3		clivages	0	10	20	30	40	50	60	70	80	90	100
12	1	2	3		à l'instar de	0	10	20	30	40	50	60	70	80	90	100
13	1	2	3		éludé	0	10	20	30	40	50	60	70	80	90	100
14	1	2	3		judicieuse	0	10	20	30	40	50	60	70	80	90	100
15	1	2	3		réfute	0	10	20	30	40	50	60	70	80	90	100
16	1	2	3		précarité	0	10	20	30	40	50	60	70	80	90	100
17	1	2	3		présomptions	0	10	20	30	40	50	60	70	80	90	100
18	1	2	3		partiale	0	10	20	30	40	50	60	70	80	90	100
19	1	2	3		habileté	0	10	20	30	40	50	60	70	80	90	100
20	1	2	3		exhaustif	0	10	20	30	40	50	60	70	80	90	100
21	1	2	3		griefs	0	10	20	30	40	50	60	70	80	90	100
22	1	2	3		émanciper	0	10	20	30	40	50	60	70	80	90	100
23	1	2	3		subordonner	0	10	20	30	40	50	60	70	80	90	100
24	1	2	3		asservir	0	10	20	30	40	50	60	70	80	90	100
25	1	2	3		affranchir	0	10	20	30	40	50	60	70	80	90	100
26	1	2	3		sommation	0	10	20	30	40	50	60	70	80	90	100
27	1	2	3		invective	0	10	20	30	40	50	60	70	80	90	100
28	1	2	3		ordre	0	10	20	30	40	50	60	70	80	90	100
29	1	2	3		addition	0	10	20	30	40	50	60	70	80	90	100
30	1	2	3		permis	0	10	20	30	40	50	60	70	80	90	100
31	1	2	3		inhibé	0	10	20	30	40	50	60	70	80	90	100
32	1	2	3		proscrit	0	10	20	30	40	50	60	70	80	90	100
33	1	2	3		égal	0	10	20	30	40	50	60	70	80	90	100
34	1	2	3		retarder	0	10	20	30	40	50	60	70	80	90	100
35	1	2	3		contester	0	10	20	30	40	50	60	70	80	90	100
36	1	2	3		ajourner	0	10	20	30	40	50	60	70	80	90	100
37	1	2	3		anticiper	0	10	20	30	40	50	60	70	80	90	100
38	1	2	3		probité	0	10	20	30	40	50	60	70	80	90	100
39	1	2	3		réactionnaire	0	10	20	30	40	50	60	70	80	90	100
40	1	2	3		exhorter	0	10	20	30	40	50	60	70	80	90	100
41	1	2	3		arbitraire	0	10	20	30	40	50	60	70	80	90	100
42	1	2	3		empirique	0	10	20	30	40	50	60	70	80	90	100
43	1	2	3		déperdition	0	10	20	30	40	50	60	70	80	90	100
44	1	2	3		allégation	0	10	20	30	40	50	60	70	80	90	100
45	1	2	3		subversion	0	10	20	30	40	50	60	70	80	90	100

2.2 TRC

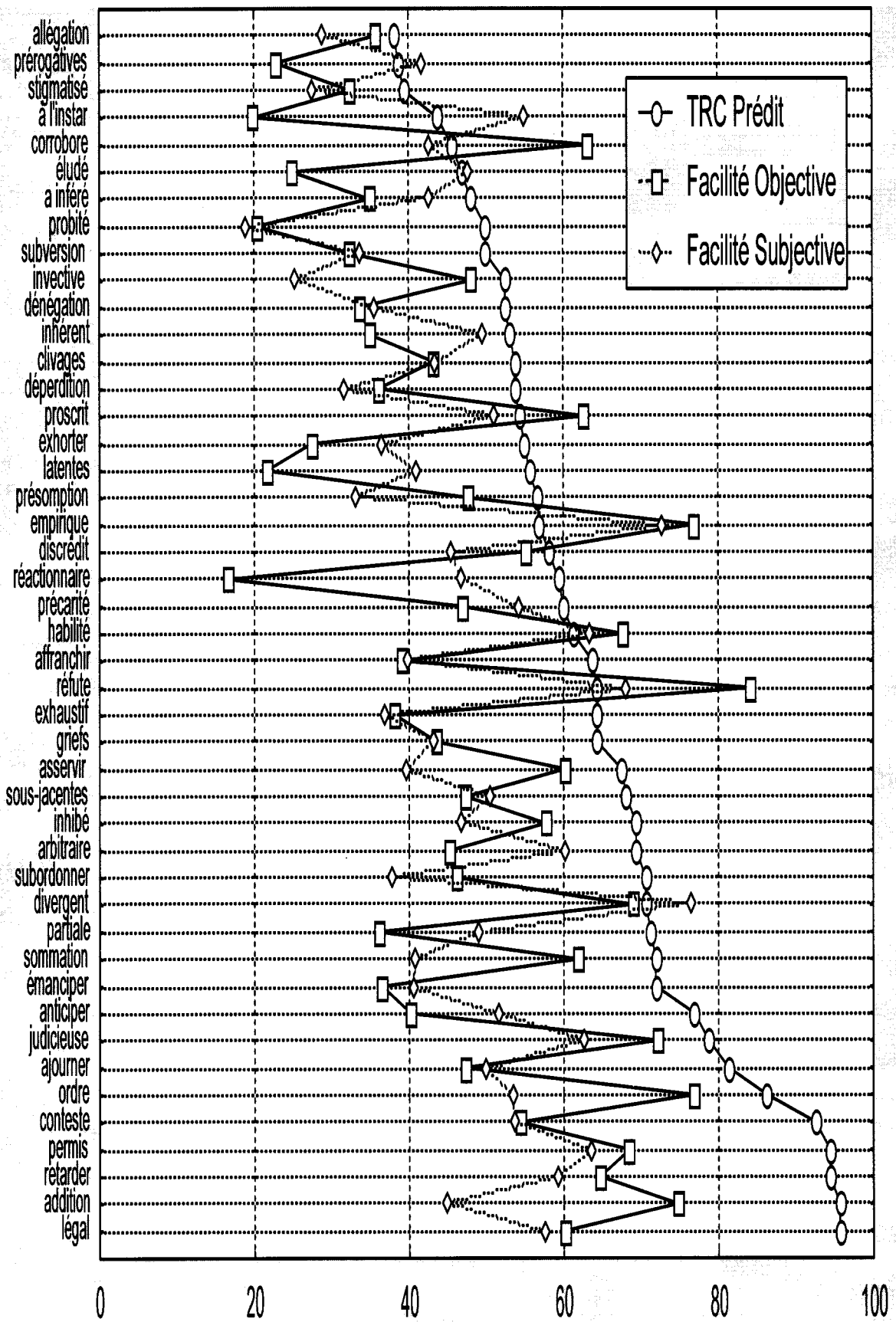
rangs	TRP	FO
légal	1	13
addition	2	4
retarder	3	9
permis	4	7
conteste	5	17
ordre	6	3
ajourner	7	21
judicieux	8	5
anticipe	9	27
émancipe	10	30
sommation	11	12
partiale	12	32
divergent	13	6
subordonner	14	23
arbitraire	15	24
inhibé	16	15
sous-jacentes	17	20
asservir	18	14
griefs	19	25
exhaustif	20	29
réfute	21	1
affranchir	22	28
habilité	23	8
précarité	24	22
réaction	25	45
discrédit	26	16
empirique	27	2
présomption	28	19
latentes	29	42
exhorter	30	39
proscrit	31	11
déperdition	32	31
clivages	33	26
inhérent	34	34
dénégation	35	36
invective	36	18
subversif	37	38
probité	38	43
a inféré	39	35
éludé	40	40
corroborer	41	10
à l'instar	42	44
stigmatisé	43	37
prérogatives	44	41
allégation	45	33

La corrélation de rang vaut 0,596 alors qu'en valeur absolue elle vaut 0,576. Ces professeurs, cette fois là, pour ces mots et ces étudiants (nous nous garderons de généraliser à tous les moments, à tous les mots, à tous les étudiants et à tous les professeurs) se sont montrés très légèrement plus capables d'estimer les **rangs** des mots dans l'échelle de difficultés que les **ampleurs de différences** relatives entre eux. Quant aux valeurs absolues, ils ont sous-estimé les difficultés des étudiants d'au moins 20% en moyenne.

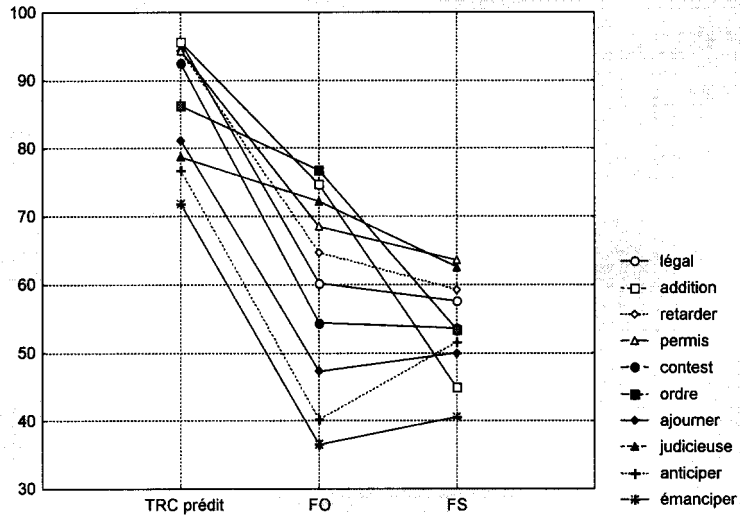


La corrélation entre les facilités subjectives et objective des étudiants est à peine meilleure : 0,63.

C'est entre le TRP et la facilité subjective des étudiants que la corrélation est la plus faible : 0,47.

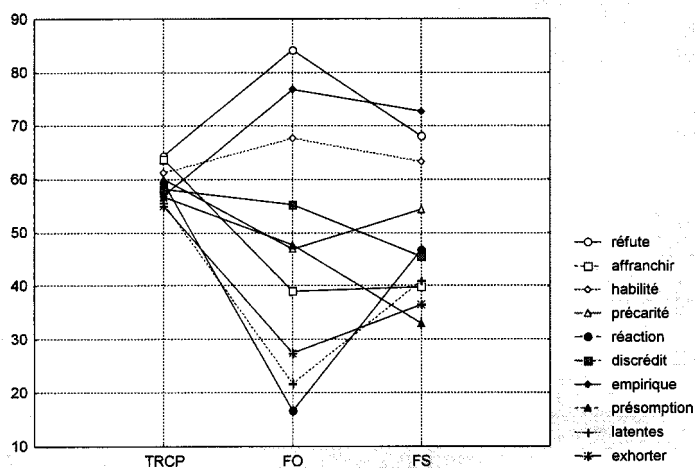


La difficulté qu'ont les encadrants à prédire en valeur absolue, mais le caractère fondé de certaines intuitions sont illustrées par les deux graphiques suivants. Si l'on range les mots depuis celui qui a été prédit (TRP) comme le plus facile jusqu'au plus difficile, les dix premiers de ce classement sont les suivants :



La colonne TRC Prédit indique le Taux de Réponses Correctes Prédit ; FO indique les Taux de réussite ou facilités Objectives et FS les Facilités Subjectives ou degré moyen de certitude (que la réponse soit correcte ou non). Le classement prédit est excellent (l'ordre est quasi le même que la réalité) mais la surestimation est flagrante, et supérieure en moyenne à 20%. Par contre, les étudiants ont été plus « prudents » encore dans leurs degrés de certitude.

A un autre endroit du classement, ce qui vient d'être dit ne tient plus. Du plus 21° (réfute) au 30° (émanciper), il en va autrement :



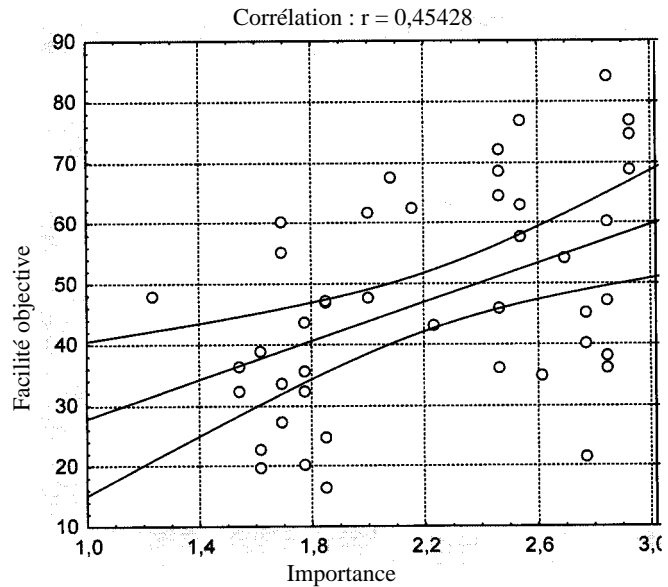
On voit que ce sont les TRP qui sont les plus regroupés, les FO étant très dispersés. L'ordre, par contre, reste très parallèle à la réalité.

2.3 Les importances

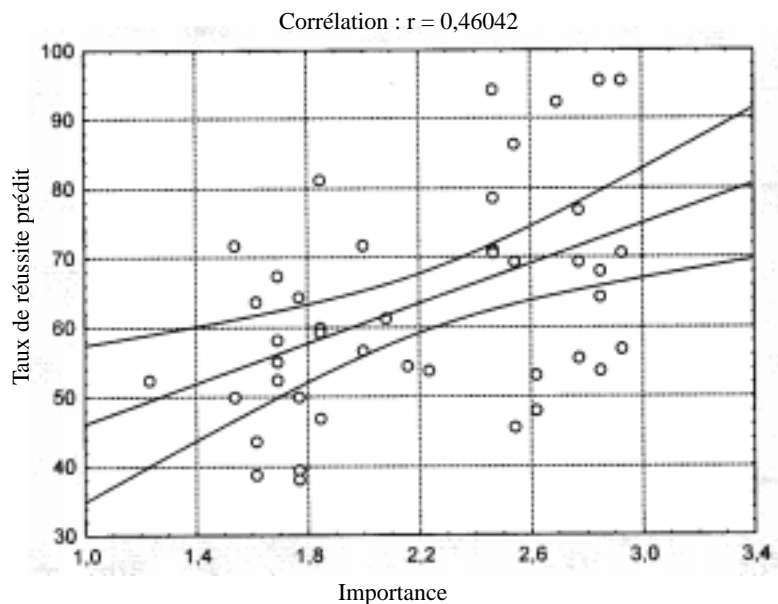
13 professeurs ont rempli ce volet, d'où un score maximal de 39 points par mot, un mot très important étant noté 3.

divergent	38
empirique	38
addition	38
exhaustif	37
réfute	37
légal	37
sous-jacentes	37
arbitraire	36
latentes	36
anticiper	36
conteste	35
a inféré	34
inhérent	34
ordre	33
inhibé	33
corrobores	33
judicieuse	32
permis	32
subordonner	32
retarder	32
partiale	32
clivages	29
proscrit	28
déperdition	28
habilité	27
sommaton	26
précarité	24
présomption	24
réaction	24
éludé	24
ajourner	24
probité	23
allégation	23
griefs	23
stigmatisés	23
asservir	22
discrédit	22
exhorter	22
dénégations	22
affranchir	21
à l'instar	21
prérogatives	21
émanciper	20
subversion	20
invective	16

Heureusement, plus les mots sont jugés importants, plus ils sont réussis par les étudiants... mais avec de très grandes différences d'un mot à l'autre.



La même corrélation s'observe...



3. Conclusion

Sur cette matière (Vocabulaire), des sondages avec feedbacks aux étudiants sont justifiés. C'est évidemment sur des mots utilisés dans son cours que ce genre d'information intéresse le professeur. C'est pourquoi chaque professeur pourrait envisager de développer son propre check-up de vocabulaire, non pas pour assurer la maîtrise de ces quelques termes de la langue française, mais pour instaurer le réflexe de recourir au dictionnaire quand ce réflexe n'est pas encore suffisamment présent. Or il ne l'est pas comme ont pu le constater Leclercq et al. (2002).

4. Bibliographie

Leclercq, D., Simon, F., Marotte, P., Verschueren A. et Lacaille C. (2002). Former des étudiants de première candidature universitaire à des compétences transversales : lesquelles et comment ?, *2^o Congrès des chercheurs en Education*, Louvain-La-Neuve.

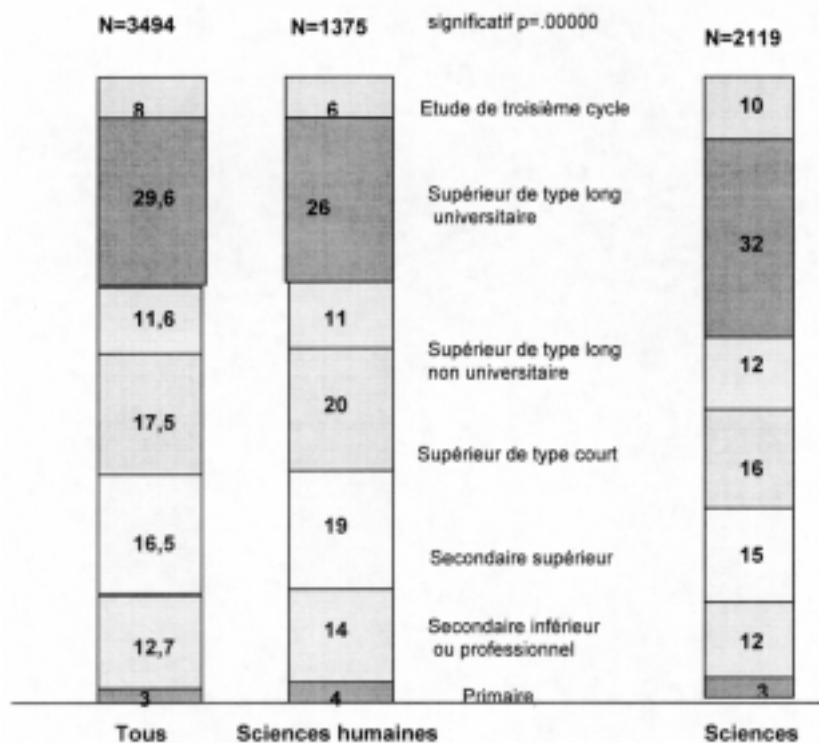
Chapitre 7

Liens entre caractéristiques de départ (dont les résultats aux check-up) et les réussites en première candidature

*D. Leclercq, Président du Groupe de Travail CIUF « Réussite en candidatures »
P. Detroz, chercheur au SMART ULg*

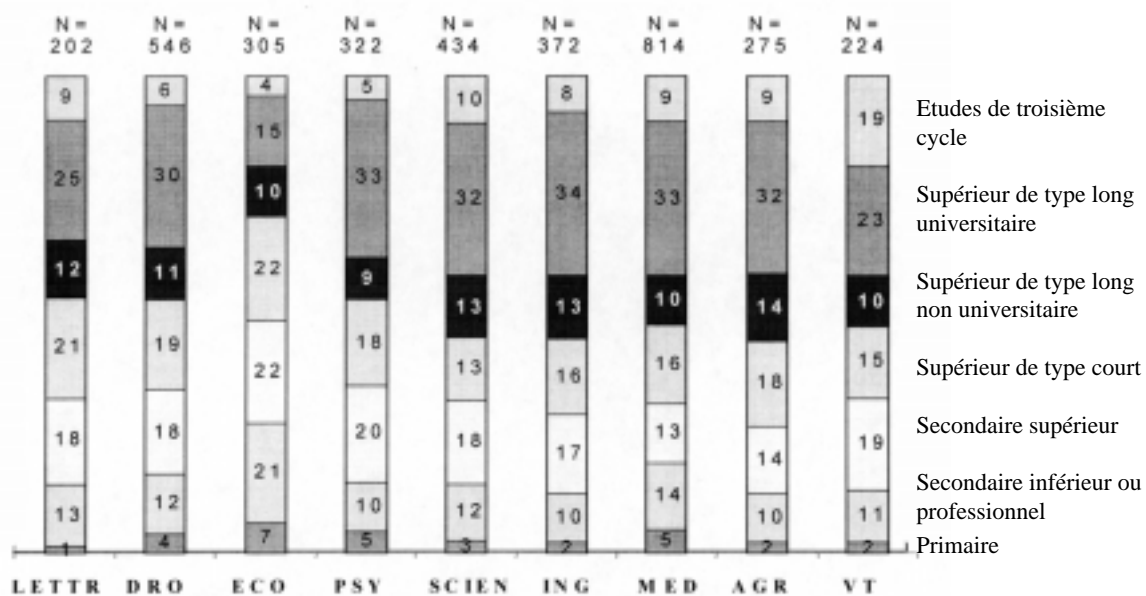
1. Liaisons entre origine sociale et choix d'études

Orientation vers les sciences ou sciences humaines et diplôme du père



Parmi les étudiants inscrits en sciences de la nature, 42% ont un père universitaire. Parmi les inscrits en sciences humaines, 32% sont dans ce cas.

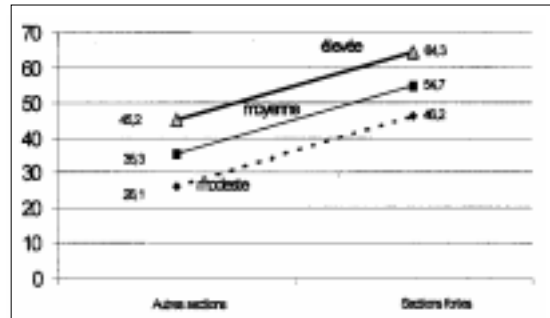
Choix de faculté et plus haut diplôme du père



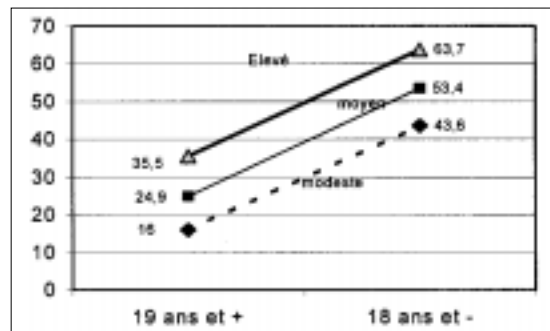
On comparera ces données à celles de Leclercq et al. 1998, p. 42.

2. L'origine socio-culturelle et la réussite

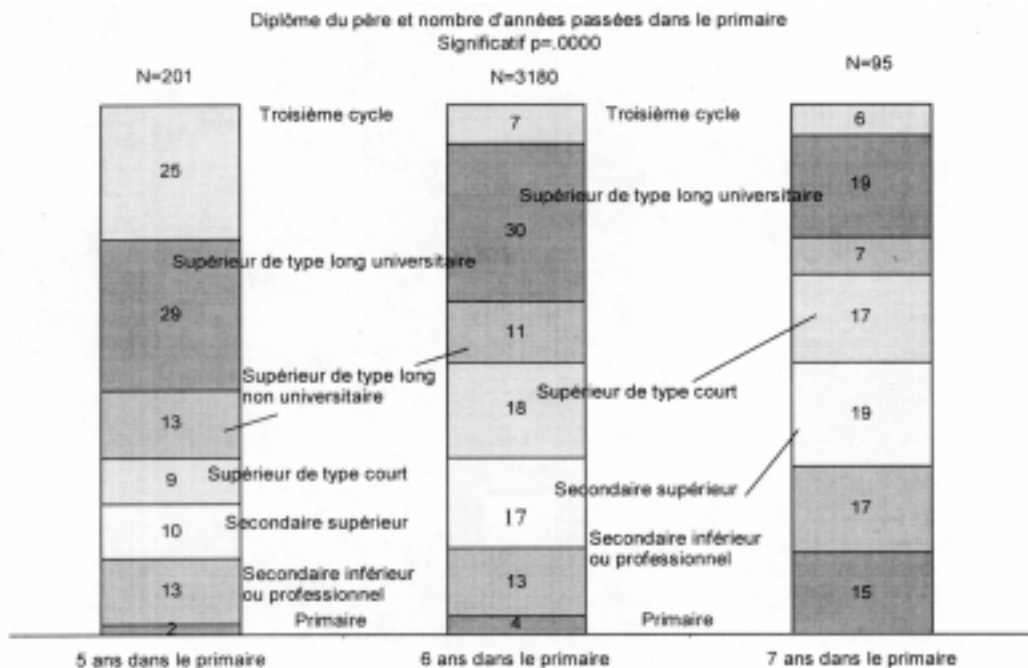
Beguin a montré que pour l'année 1989-90 à l'UCL, les étudiants d'origine sociale élevée ont un taux moyen de réussite plus élevé de 20% que les étudiants d'origine modeste, ceux issus d'une section forte dans le secondaire ont un taux moyen supérieur de 20% à ceux qui sont d'une filière faible, (Leclercq, Beguin Lambert, De Kerchove, Pestiau, 1998, p. 38).



Dans la même étude de Beguin (1991) sur la même cohorte, ceux qui n'ont jamais redoublé à l'entrée ont un taux moyen de réussite supérieur de 30% sur ceux qui arrivent avec un an (ou plus) de « retard » (Leclercq, Beguin Lambert, De Kerchove, Pestiau, 1998, p. 39).



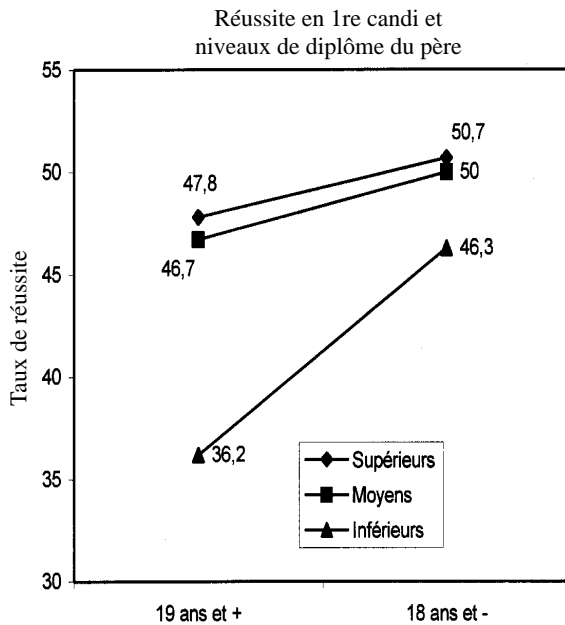
Voici, pour la cohorte MOHICAN, la liaison entre diplôme du père et temps passé dans le primaire :



On voit que le lien est très fort (54% des étudiants ayant passé seulement 5 ans dans le primaire ont un père universitaire ; ce taux est de 25% pour les étudiants qui y ont passé 7 ans). On verra (en section D ci-après) que le nombre d'années passées dans le primaire est, pour la cohorte MOHICAN, très lié à la réussite en première candidature.

Avant de présenter les liaisons entre variables « prédictives » et réussite, il importe de préciser (en p. 133) comment se lit la représentation « tétramétrique » des résultats.

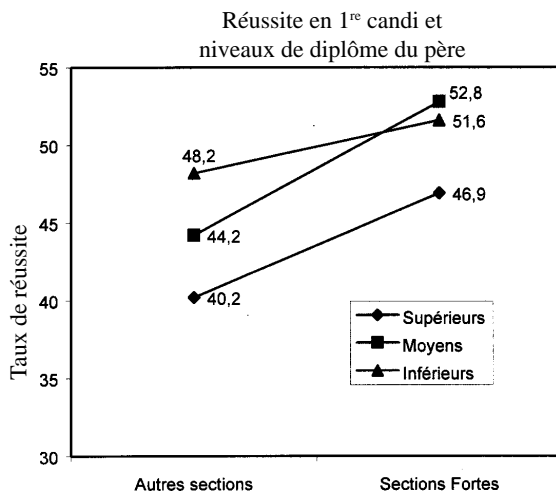
Trois groupes ont été constitués sur base du score « diplôme du père » (voir section A ci-avant). Le groupe « inférieur (1) » a un score diplôme compris entre 1 et 3, le groupe « moyen » (2) 4 ou 5 et le groupe (3), de 6 ou 7.



Pour les étudiants entrés « à l'heure ou en avance » (18 ans et moins), les étudiants dont le diplôme du père est dans une des catégories inférieures n'ont un taux de réussite inférieur que de 4% environ.

Par contre, les étudiants entrant avec une année de retard ou plus, ont un taux de réussite moyen inférieur de plus de 10% par rapport aux étudiants dont le père a un diplôme d'une catégorie moyenne ou élevée.

Comme on le sait par ailleurs, les étudiants entrant avec un retard réussissent moins bien que ceux qui entrent « à l'heure », mais, comme on le voit, la variable « diplôme du père » a une interaction très importante avec cet âge à l'entrée dans le lien avec la réussite, surtout pour les étudiants dont le diplôme du père est d'un niveau inférieur.



Pour les étudiants issus d'une section forte (latin math, latin sciences, scientifiques fortes), les étudiants dont le père a un diplôme de catégorie élevée ou moyenne ont des résultats assez équivalents, ceux dont le père a un diplôme de catégorie inférieure ayant un taux de réussite moyen inférieur de 6% environ.

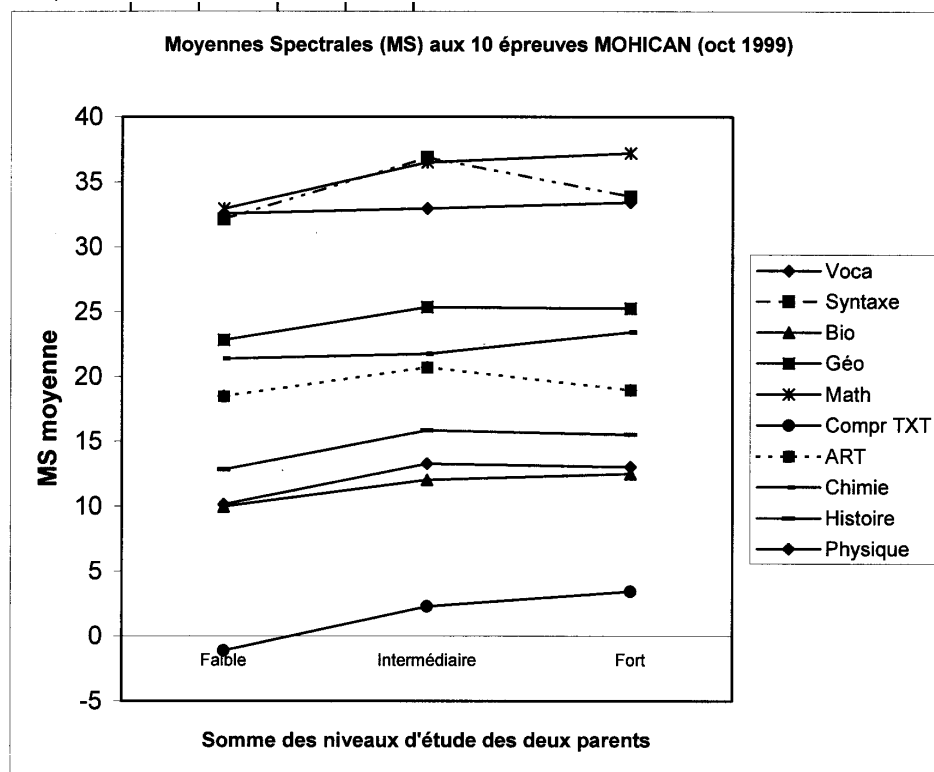
Pour les étudiants issus d'autres sections, les taux de réussite s'échelonnent strictement selon l'ordre des catégories de diplôme du père, avec 4% de différence d'une catégorie à l'autre.

On rapprochera ces données de celles qu'a exposées le recteur de l'UCL ; le Professeur M. Crochet lors de la rentrée académique 2002 – 2003 dans son discours intitulé « L'inégalité des chances à l'entrée à l'université ».

Trois autres groupes ont été constitués sur base de la somme des scores « diplôme le plus élevé des parents » (voir section A ci-avant). Le **groupe 1** a une somme comprise entre 1 et 8. Le **groupe 2** entre 9 et 11 et le **groupe 3**, de 12 et plus. Quand le diplôme d'un seul parent est connu, il est multiplié par deux. Les nombres sur lesquels les statistiques qui suivent ont été calculées sont inférieurs à ceux du chapitre 4, car pour certains étudiants, l'information sur les diplômes des parents manquait.

		Moyennes spectrales (MS)									
MS	Voca	Syntaxe	Bio	Géo	Math	Compr	ART	Chimie	Histoire	Physique	
Faible	10,13	32,14	9,99	22,81	32,94	-1,11	18,49	21,37	12,84	32,56	
Intermédiaire	13,29	36,87	12,03	25,34	36,49	2,27	20,69	21,74	15,84	32,96	
Fort	13,04	33,88	12,50	25,25	37,20	3,42	18,96	23,41	15,52	33,43	
p =	0,00	0,00	0,05	0,01	0,00	0,00	0,17	0,22	0,01	0,79	

Nombres	Voca	Syntaxe	Bio	Géo	Math	Compr	ART	Chimie	Hist	Phys
Faible	1442	1391	829	1373	826	1261	609	828	618	824
Intermédiaire	962	967	634	957	629	887	341	630	341	628
Fort	1044	1034	759	1018	759	962	321	759	325	757
TsGrpes	3448	3392	2222	3348	2214	3110	1271	2217	1284	2209



Pour les Moyennes Spectrales (MS),

le groupe 1 a une moyenne inférieure au groupe 3 ET au groupe 2 pour les 10 check-up.

le groupe 2 a une moyenne inférieure au 3 pour bio, math, compréhension de textes, chimie et physique.

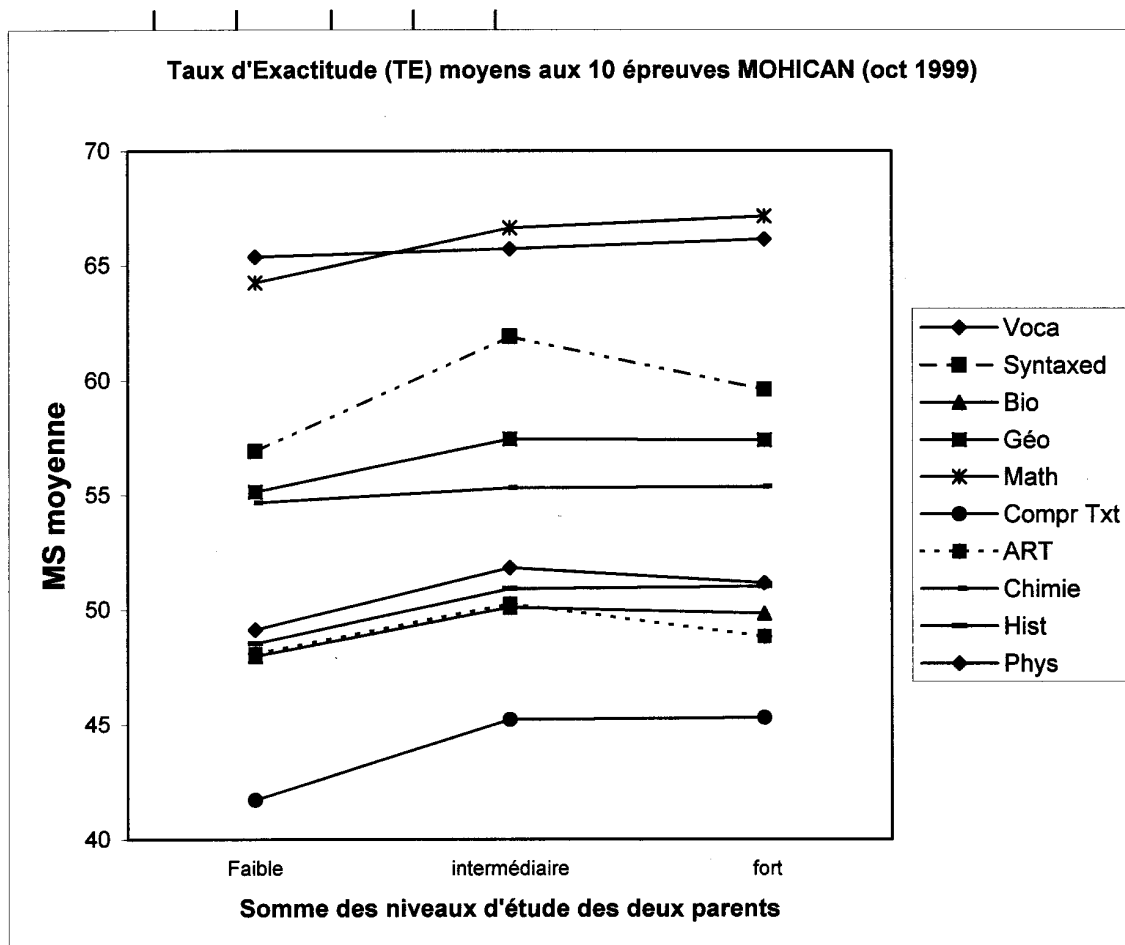
et une moyenne supérieure au groupe 3 dans les 5 autres check-up.

La significativité des différences (analyse de variance pour trichotomie) atteint p. 05 sauf pour art, chimie, et physique.

Pour les Taux d'Exactitude (TE),

TE	Voca	Syntaxe	Bio	Géo	Math	Compr	ART	Chimie	Histoire	Physique
Faible	49,13	56,91	47,99	55,15	64,25	41,73	48,09	54,66	48,53	65,38
intermédiaire	51,84	61,92	50,11	57,46	66,84	45,23	50,26	55,32	50,92	65,72
fort	51,18	59,62	49,84	57,42	67,15	45,30	48,83	55,37	51,03	66,14
p =	0,00	0,00	0,04	0,00	0,00	0,00	0,11	0,74	0,01	0,68

Nombres	Voca	Syntaxe	Bio	Géo	Math	Compr	ART	Chimie	Hist	Phys
Faible	1442	1674	829	1373	826	1261	609	828	618	824
intermédiaire	962	1119	634	957	629	887	341	630	341	628
fort	1044	1209	759	1018	759	962	321	759	325	757
TsGrpes	3448	4002	2222	3348	2214	3110	1271	2217	1284	2209



Le groupe 1 a une moyenne inférieure au groupe 3 ET au groupe 2 pour les 10 check-up.
 Le groupe 2 a une moyenne inférieure au 3 pour math, compréhension de Textes, chimie et histoire et une moyenne supérieure au groupe 3 dans les 6 autres check-up.

La significativité des différences (analyse de variance pour trichotomie) atteint p. 05 sauf pour art, chimie et physique.

3. La tétramétrie des résultats d'une cohorte d'étudiants universitaires

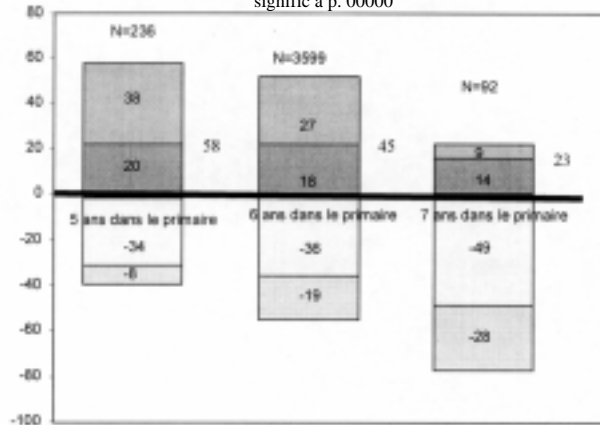
La tétramétrie (ou mesure par décomposition en quatre parties) consiste à présenter les résultats de la cohorte aux examens :

- arithmétiquement en répartissant les 100% de l'effectif des étudiants de la cohorte en **4 groupes (donc 4 pourcentages dont la somme arithmétique fait 100%)**
- et graphiquement, sous forme d'un **ludion** ou « flotteur » positionné par rapport à un niveau (celui de la « barre » de la réussite) et composé de 4 segments rectangulaires empilés, du meilleur (en haut) au pire (en bas) :
 - R1 = **Réussite en première session** (surface au-dessus de toutes les autres)
 - R2 = **Réussite en seconde session** (surface juste au-dessus de la barre)
 - Eap = **Echecs après persévération** : ces étudiants ont échoué après s'être présentés à tous les examens, donc en ayant fait preuve de persévération (surface juste sous la barre)
 - A = **Abandons** : ces étudiants échouent parce qu'ils ne se sont pas présentés aux examens ou à TOUS les examens (surface la plus basse).
 - Avec **R** = Réussites ($R = R1+R2$) et **E** = Echecs ($E = Eap + A$).

Dans l'exemple ci-dessous, les R sont rappelés (respectivement 58%, 45% et 23%) dans les 3 sous-populations contrastées.

Nous parlons de **ludion** car il s'agit d'un « *appareil de démonstration de physique... qui monte et qui descend... quand on modifie la pression* »²⁷. C'est en effet notre conviction qu'il est possible, par des efforts (pression) tantôt des apprenants, tantôt des enseignants, tantôt d'instances organisatrices (universités, ministères) de changer la position du ludion par rapport à la barre et la tétramétrie des résultats d'une cohorte universitaire. La « **barre de la réussite** » est représentée dans le graphique ci-après par une grosse barre noire horizontale; on pourrait la comparer au niveau d'un fleuve: ce qui est au-dessus indique la réussite; ce qui est en-dessous l'échec. Plus le ludion émerge, ressort au-dessus de la barre de la réussite, plus le pourcentage de réussites est élevé. Voici un exemple de tétramétrie :

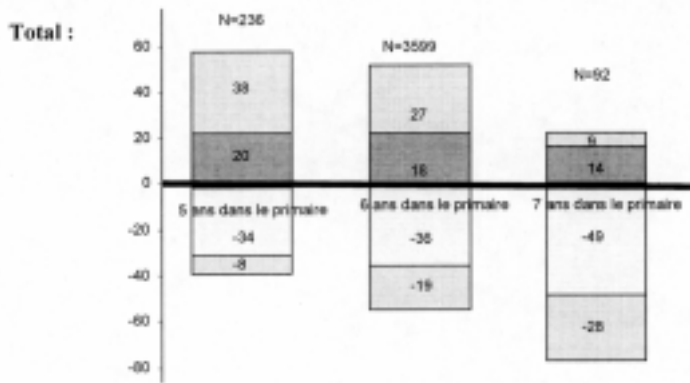
Pourcentage de Réussite (R1-R2) – d'Echec et d'Abandon en fonction du nombre d'années passées dans le primaire signific à p. 00000



²⁷ La définition (du dictionnaire Robert, 1995) est volontairement tronquée, car la comparaison n'est que partielle. Voici la définition complète de « ludion » : « Appareil de démonstration de physique, formé d'une sphère creuse percée d'un trou à sa partie inférieure... qui monte et qui descend dans un bocal fermé par une membrane, quand on y modifie la pression ».

4. Liens entre caractéristiques individuelles d'entrée et la tétramétrie des résultats

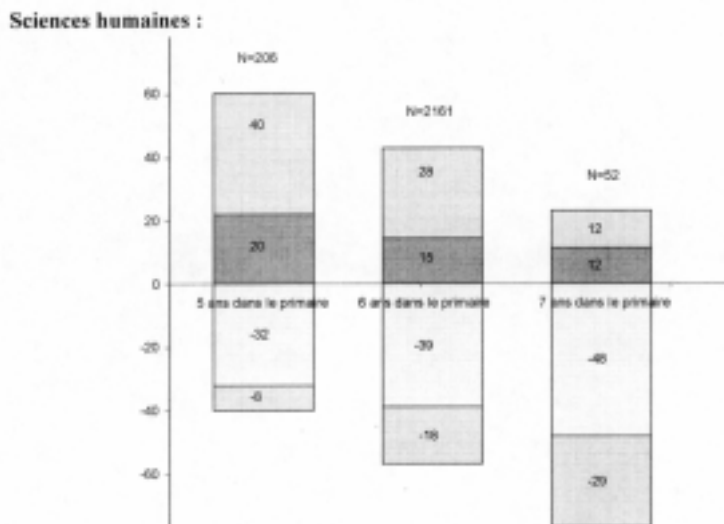
4.1 Pourcentages de R1, R2, Echecs et Abandons en fonction du nombre d'années passées dans le primaire



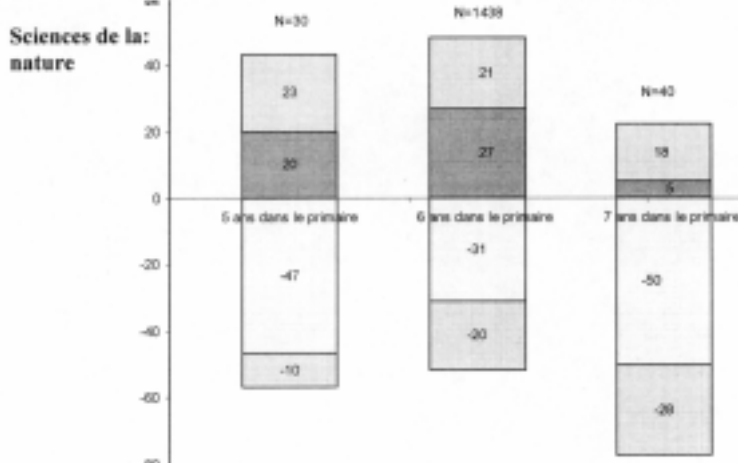
La liaison entre cette variable de « Présage » et la réussite est élevée, comme on pourrait s'y attendre, tout spécialement sur les deux situations extrêmes : les taux d'abandon et les taux de réussite en première session.

Par rapport à la durée normale (6 ans) où le taux de R est de 45 %, le taux de **R** est 58 % (**13 % de plus**) pour 5 ans et 23 % (**22 % de moins**) pour 7 ans.

Soit 35% de différence entre les deux groupes extrêmes.

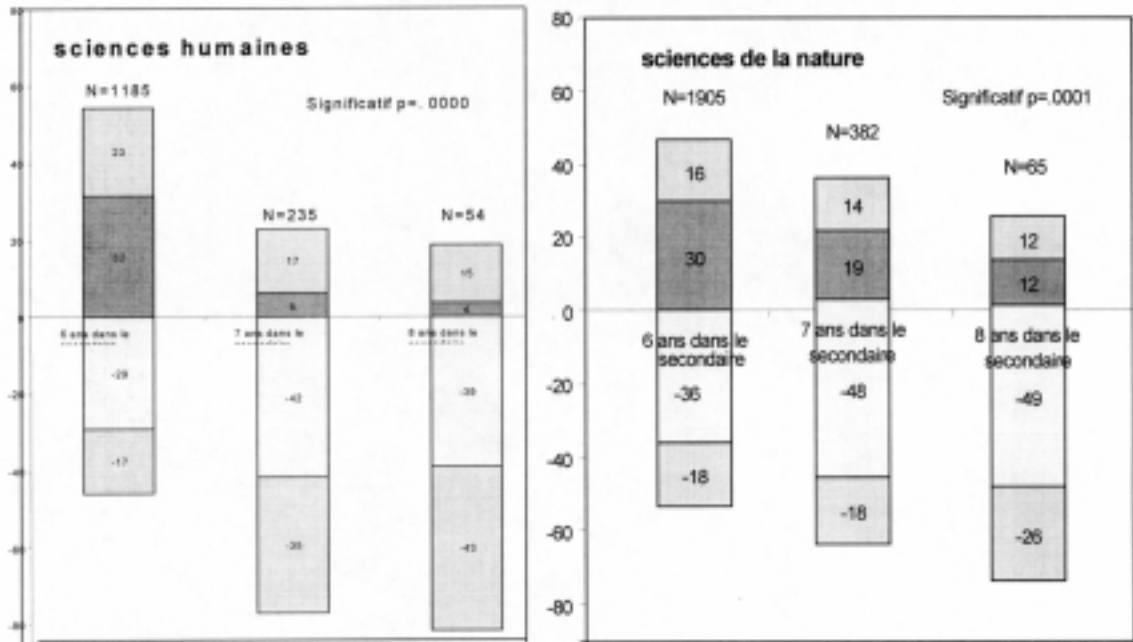


Le phénomène est plus accentué en Sciences humaines où se retrouvent le plus grand nombre des étudiants n'ayant passé que 5 ans dans le primaire.



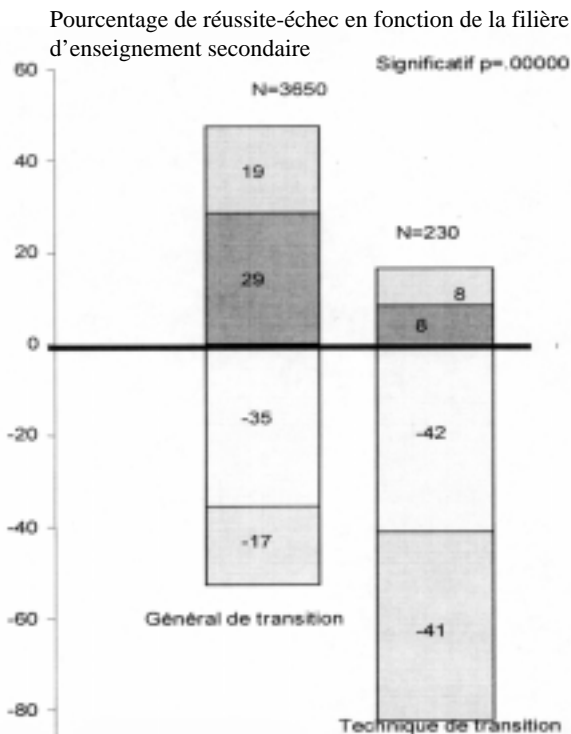
Le phénomène est plus rare en Sciences de la nature (seulement 30 étudiants ayant passé 5 ans dans le primaire), ce qui rend les valeurs statistiques moins fiables qu'en sciences humaines.

4.2 Pourcentage de réussite-échec en fonction du nombre d'années passées dans le secondaire



La liaison est très forte, plus encore pour les Sciences humaines que pour les Sciences de la nature.

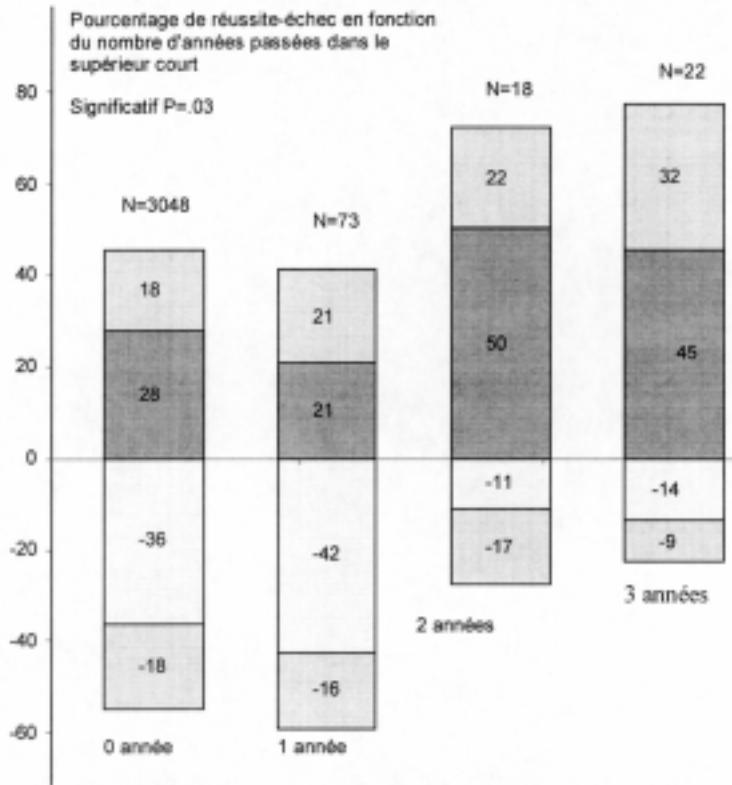
4.3 Pourcentage de réussite-échec en fonction de la filière d'études du secondaire



Les étudiants issus de l'enseignement secondaire Technique de Transition (TT), qui ne représentent que 6% de la cohorte MOHICAN (230/(3650+230)) réussissent à 16 % alors que ceux qui sont issus de l'enseignement secondaire Général de Transition réussissent à 48 %.

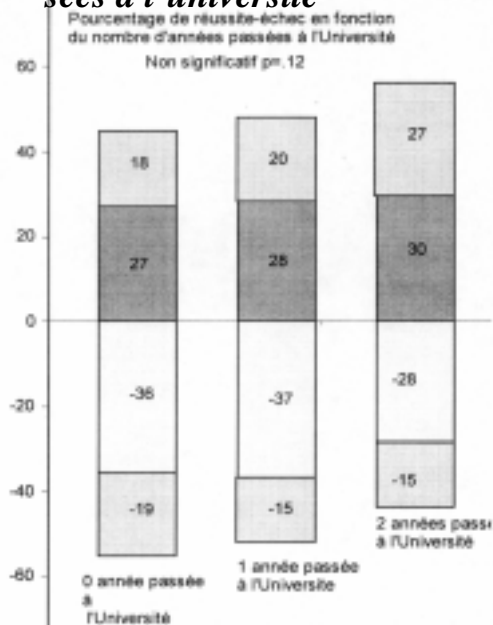
Le taux d'abandons (41 %) des TT est particulièrement élevé.

4.4 Pourcentage de réussite-échec en fonction du nombre d'années passées dans le supérieur court



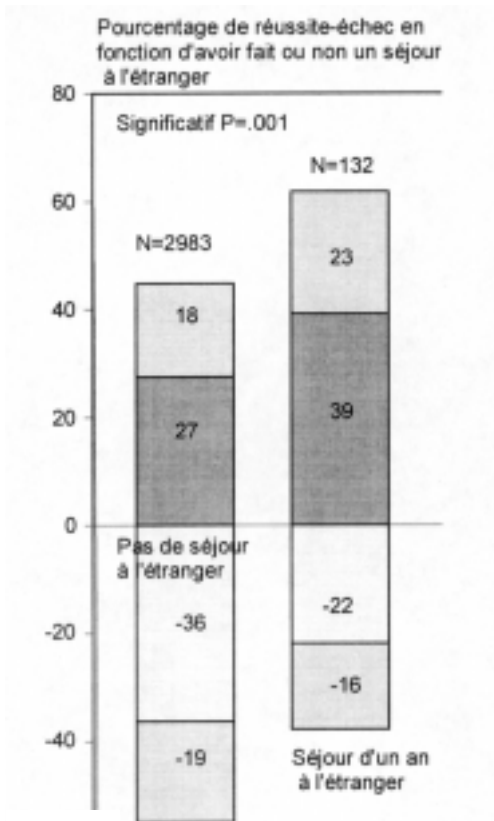
Ceux qui ont passé 1 an dans le supérieur court réussissent légèrement moins bien (42 % au lieu de 46 %) mais ceux qui y ont passé 2 ou 3 ans réussissent mieux (72 % et 77 % au lieu de 46%).

4.5 Pourcentage de réussite-échec en fonction du nombre d'années passées à l'université



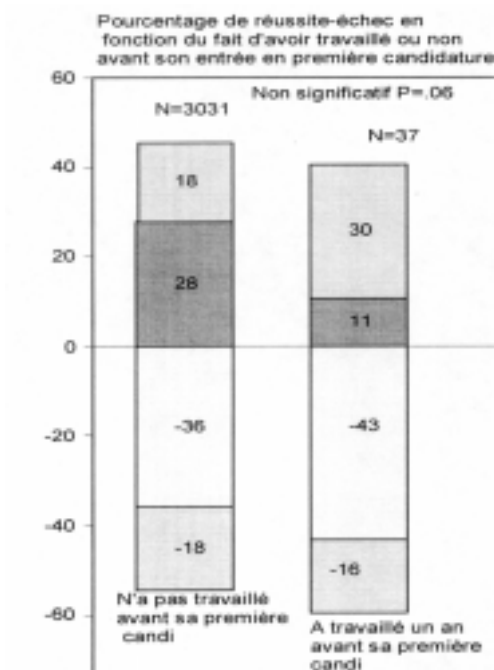
Les doublants (réorientés ou non) réussissent légèrement mieux (48 % au lieu de 45 %) et les « triplants » mieux encore (57 %).

4.6 Pourcentage de réussite-échec en fonction d'un séjour ou non à l'étranger



Ce n'est probablement pas le séjour à l'étranger en lui-même ou à lui seul qui a favorisé la réussite, mais probablement une combinaison des effets de l'appartenance à un milieu socio-culturel supérieur à la moyenne, à une maturité plus grande (entrée à l'université avec un an de plus), et bien sûr le séjour lui-même.

4.7 Pourcentage de réussite-échec en fonction du fait d'avoir travaillé ou non avant son entrée en 1^{re} candidature

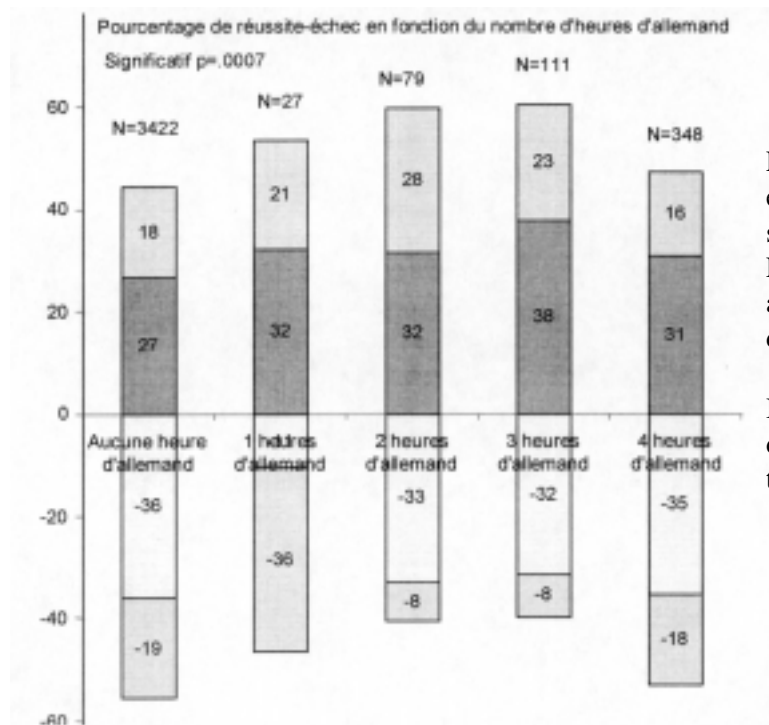


Le nombre d'étudiants concernés est trop faible pour tirer des conclusions.

On peut se demander si le nombre d'étudiants devant travailler PENDANT leurs études est en augmentation ou non ces dernières années. Nous ne disposons pas des données sur ce sujet.

L'étude, en 2001, de Creyf's-Solvus (L'Echo, 17-4-2002) montre que pour ce qui concerne LEURS observations (donc dans l'intérim) « si les étudiants font des petits boulots, c'est surtout pour arrondir leurs fins de mois. Il s'agit en général d'un revenu d'appoint. Chez Creyf's, le plus grand nombre des étudiants travaillent entre 5 et 13 heures par semaine (hors vacances). Peu dépassent le quota de gain qui permet de garder une indépendance fiscale vis-à-vis des parents. Ils sont souvent en 2^o cycle. »

4.8 Pourcentage de réussite-échec en fonction du nombre d'heures d'allemand dans le secondaire

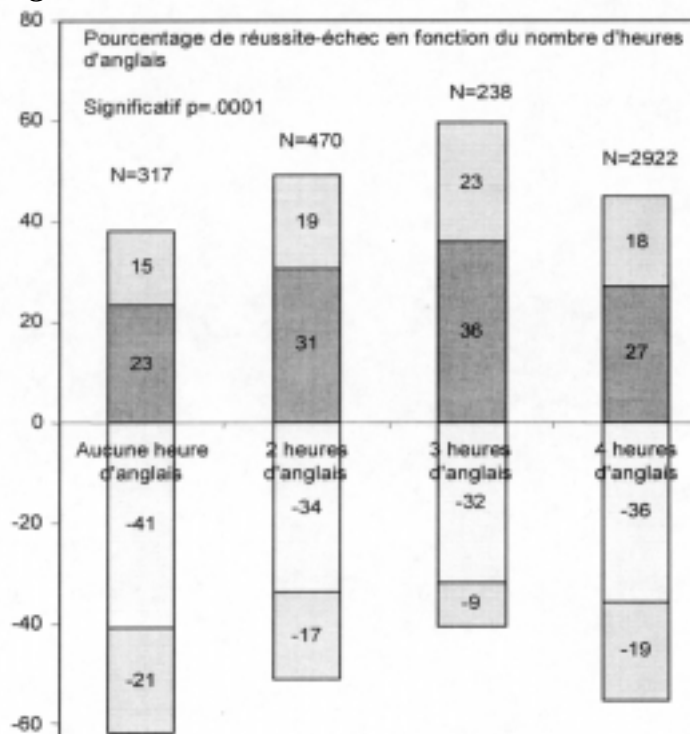


Le nombre d'heures d'allemand est lié positivement avec la réussite.

Le groupe des 348 étudiants qui avaient 4 heures ne s'inscrit pas dans cette tendance.

Il se pourrait qu'il s'agisse d'étudiants dont la localisation est particulière.

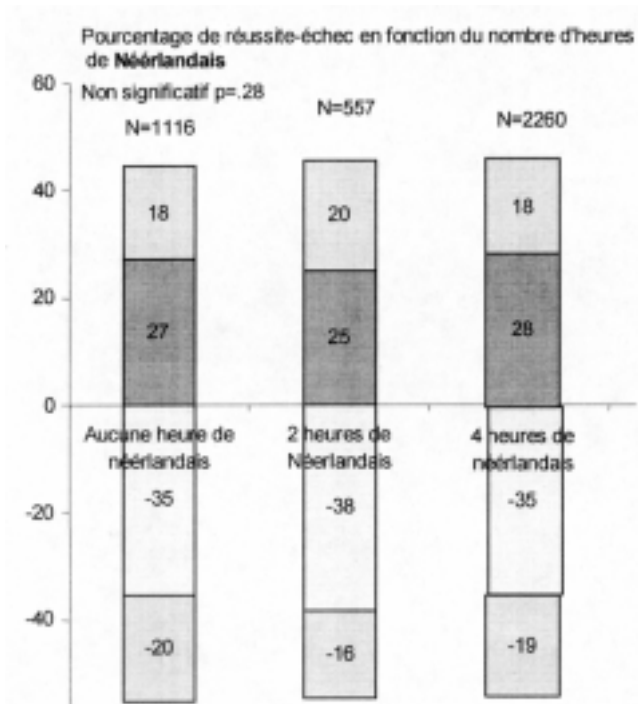
4.9 Pourcentage de réussite-échec en fonction du nombre d'heures d'anglais dans le secondaire



Ici la norme (le plus grand nombre est « 4 heures d'anglais »).

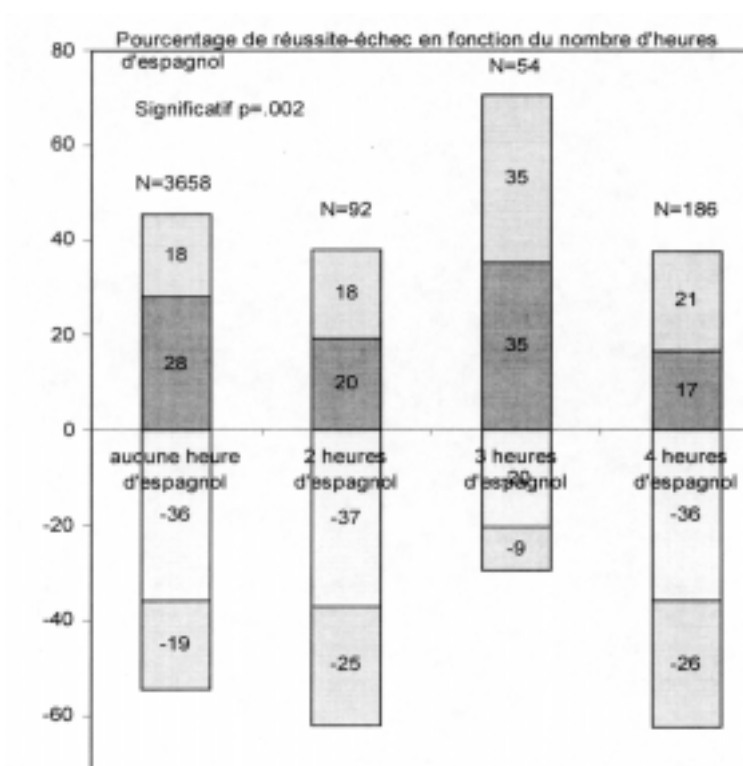
Le groupe « 3 heures d'anglais » a de meilleurs résultats. S'agit-il d'un groupe qui a plus d'allemand ou de néerlandais ?

4.10 Pourcentage de réussite-échec en fonction du nombre d'heures de néerlandais dans le secondaire



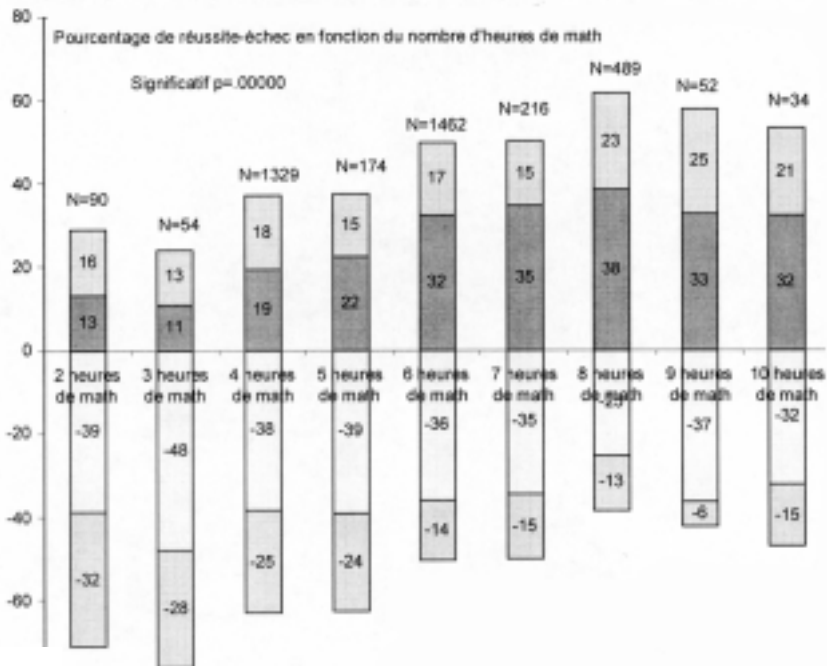
Pas de différence significative de réussite / échec en fonction du nombre d'heures de néerlandais.

4.11 Pourcentage de réussite-échec en fonction du nombre d'heures d'espagnol dans le secondaire



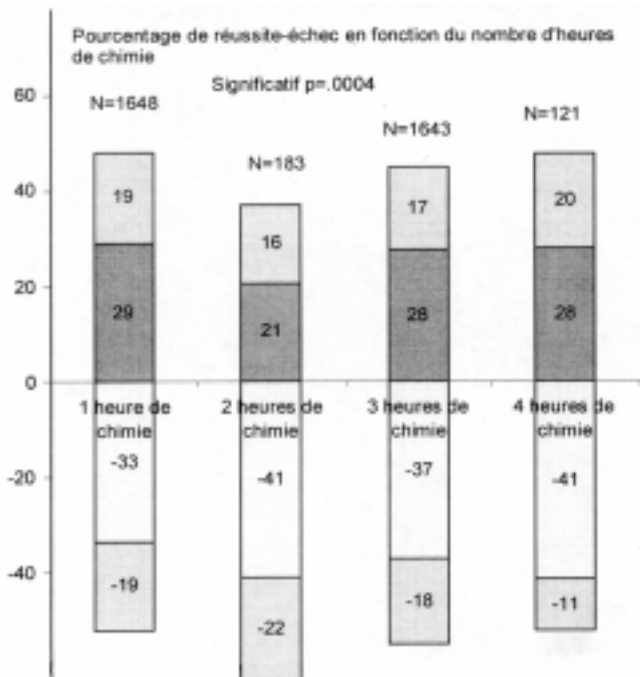
Ces résultats sont difficiles à interpréter. Il faudrait considérer dans quelle (s) section (s) universitaires sont inscrits les étudiants des divers groupes.

4.12 Pourcentage de réussite-échec en fonction du nombre d'heures de math dans le secondaire



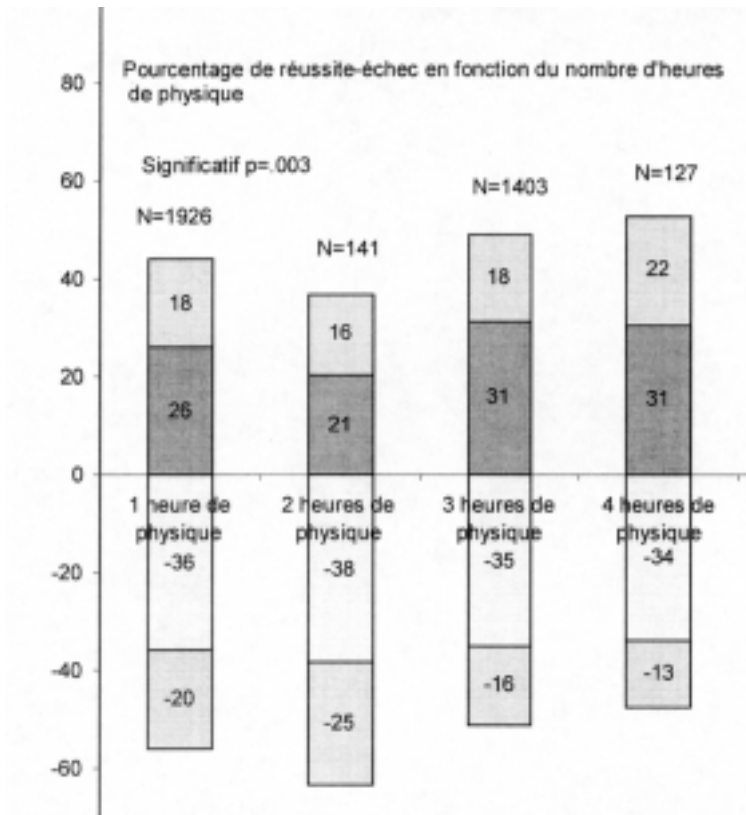
Le Taux de R augmente parallèlement au nombre d'heures de math, et ce jusqu'à 8 heures par semaine. Il n'en est pas de même pour les « 9 et 10 heures / semaine ». S'agirait-il d'étudiants qui se sont engagés dans une 7^e année secondaire de « math spéciale » ?

4.13 Pourcentage de réussite-échec en fonction du nombre d'heures de chimie dans le secondaire



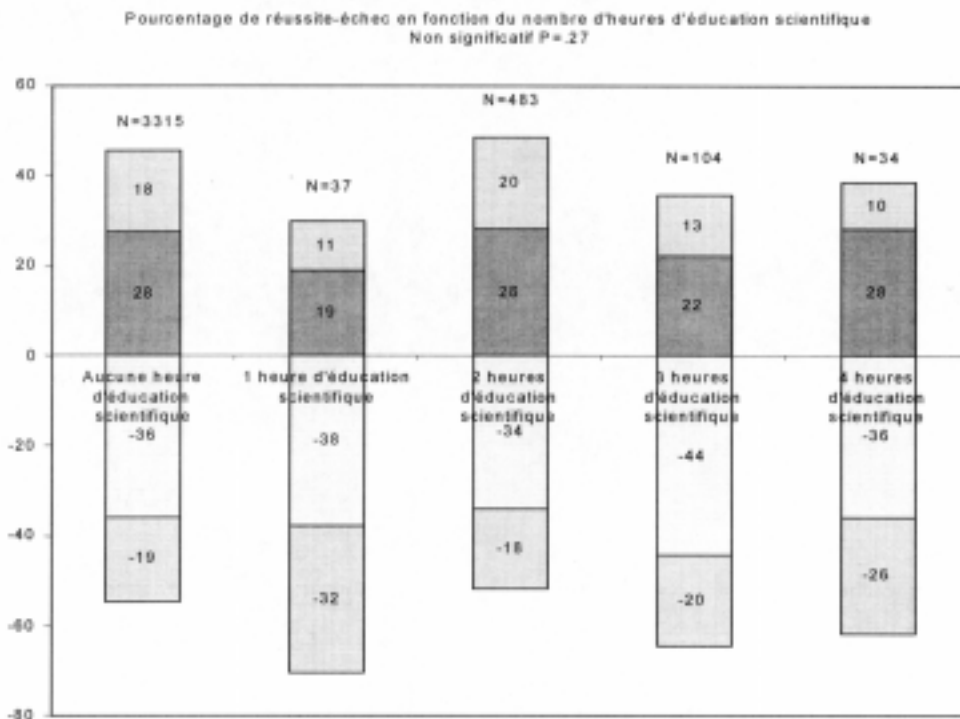
La répartition de ces étudiants dans les facultés explique probablement ces résultats complexes (réussite meilleure avec 1 heure seulement).

4.14 Pourcentage de réussite-échec en fonction du nombre d'heures de physique dans le secondaire



Même interprétation que pour la chimie.

4.15 Pourcentage de réussite-échec en fonction du nombre d'heures d'éducation scientifique dans le secondaire



Les différences n'étant pas significatives, nous ne les interprétons pas.

5. Corrélations avec la réussite/échec (1/0) en 1^{re} candidature

5.1 Vue d'ensemble

Les nombres d'étudiants sur lesquels les corrélations ont été calculées (voir pages qui suivent) ne sont pas les mêmes pour tous les check-up, car certains résultats peuvent manquer pour des étudiants à certains check-up.

% de Rép. Correctes au check up de.....	Voca	Synt	CompT	Géo	Math	Phys	Chim	Bio	Art	Hist	MOY
Total	0,210	0,189	0,100	0,161	0,321	0,204	0,219	0,203	0,163	0,188	0,195
Philo et Lettres	0,123	0,181	0,075	0,041					0,191	0,190	
Droit	0,276	0,320	0,097	0,211					0,231	0,230	
Psycho-Education	0,192	0,211	0,093	0,266					0,103	0,192	
Sc. Economiques	0,215	0,217	0,194	0,233					0,156	0,209	
Sciences	0,160	0,105	0,070	0,134	0,321	0,214	0,175	0,218			
Sciences Appliquées	0,198	0,102	0,045	0,119	0,115	0,067	0,145	0,054			
Médecine	0,206	0,189	0,068	0,173	0,356	0,196	0,263	0,249			
Médecine vété	0,287	0,128	0,332	0,178	0,448	0,225	0,31	0,305			
Sc. Agronomiques	0,204	0,205	0,120	0,168	0,381	0,229	0,170	0,169			

Moyenne Spectrale (-100 à +100)	Voca	Synt	CompT	Géo	Math	Phys	Chim	Bio	Art	Hist	MOY
Total	0,221	0,198	0,114	0,160	0,29	0,196	0,23	0,209	0,154	0,203	0,197
Philo et Lettres	0,193	0,217	0,121	0,076					0,181	0,232	
Droit	0,297	0,338	0,071	0,193					0,214	0,252	
Psycho-Education	0,176	0,240	0,132	0,216					0,079	0,155	
Sc. Economiques	0,241	0,188	0,128	0,233					0,140	0,247	
Sciences	0,190	0,077	0,071	0,136	0,23	0,148	0,16	0,220			
Sciences Appliquées	0,199	0,146	0,058	0,109	0,12	0,108	0,18	0,049			
Médecine	0,209	0,197	0,106	0,186	0,33	0,214	0,25	0,257			
Médecine vété	0,311	0,186	0,312	0,252	0,46	0,268	0,38	0,325			
Sc. Agronomiques	0,198	0,186	0,129	0,161	0,33	0,171 ²	0,18	0,150			

5.1.1 Données de base pour les calculs

Les corrélations ont été calculées entre les réussites et les résultats à 8 check-up pour 5 facultés (40 corrélations) et les résultats à 6 check-up pour 4 facultés (24 corrélations), soit un total de **64 corrélations**. En outre, chaque résultat a été calculé **de deux façons différentes**: d'une part le Taux de Réponses Correctes et d'autre part la Moyenne Spectrale, faisant intervenir les degrés de certitude, la connaissance subjective. Dans 26 cas sur 64 (42%), ce sont les scores « objectifs » les plus prédictifs que les subjectifs, dans 37 cas (58%) ce sont les subjectifs les plus prédictifs et que dans un cas, il y a parfaite égalité. Ce constat plaide pour le maintien des degrés de certitude dans les recherches et les pratiques.

5.1.2 Toutes les corrélations sont positives !

Il y a donc un lien entre ces variables et la réussite, même si ce lien est extrêmement faible pour certaines et faible en moyenne : 0,195 (en pratique 0,2), ce qui n'explique que 4% de la variance des réussites, ce qui exclut d'envisager de sélectionner sur la base d'un résultat à un seul check-up. Et de plusieurs check-up ? Ce sera traité en section F ci-après.

5.1.3 Différences inter-facultaires

La prédictivité des résultats à un check-up n'est pas la même selon les facultés, comme on pouvait s'y attendre.

5.1.4 Effet homogénéisant de l'examen d'entrée

La prédictivité des résultats (% RC) aux check-up est la plus faible pour les **Sciences Appliquées** (entre 0,045 et 0,198). Ce phénomène s'explique par la sélection préalable qui a été opérée. La présence des prérequis étant « assurée », ce sont d'autres facteurs, non testés dans MOHICAN, qui influent sur la réussite.

5.1.5 Les étudiants français « gonflent » la prédictivité en faculté de Vétérinaires

La prédictivité des résultats (% RC) aux check-up est la plus forte pour les étudiants en **Médecine Vétérinaire** (le maximum étant de 0,128 et 0,448, soit 20 % de variance expliquée au maximum) de l'échantillon MOHICAN. Ce phénomène s'explique dans une bonne mesure par la présence croissante, ces dernières années, d'étudiants français dans cette faculté. Ainsi, en 1999-2000, pour l'échantillon MOHICAN, dans les autres facultés, les étudiants français constituaient 3% de l'effectif (123 sur 4048) et 40% de l'effectif (102 sur 252) en Médecine vétérinaire. En outre, alors que les Belges de l'échantillon MOHICAN ne sont que 20 % à avoir déjà effectué une année universitaire ou plus (697 sur 3452), les Français dans ce cas sont 46% (83 sur 180).

Dans la faculté de médecine vétérinaire de l'échantillon MOHICAN, 32% des Belges ont déjà passé une année au moins à l'université (doublements) et 50% des Français, mais eux, ce n'est pas par « doublement » généralement. Dans leur pays, la plupart de ces étudiants français se sont préparés intensément non seulement au Baccalauréat, mais à des épreuves universitaires ou à des concours d'entrée. Le check-up de mathématique est réussi à 59% par les (142) Belges et à 73% par les (102) Français. La maturité pèse également : l'âge moyen des étudiants français est, évidemment plus élevé. Ainsi, pour les Belges de l'échantillon MOHICAN, la moyenne d'âge est 18,4 ans et pour les Français 19,4 ans.

5.2 Les trois épreuves les plus prédictives par faculté

L'épreuve check-up la plus prédictive n'est pas la même pour les diverses facultés. Voici, pour chaque faculté, les trois épreuves check-up les plus prédictives. Celles qui sont marquées d'un * indiquent qu'il s'agit non pas du pourcentage de réponses correctes, mais du score spectral, qui incorpore les degrés de certitude.

Philo-Lett	Droit	Psycho-Ed	Eco	Sciences	Sc.Appl	Médecine	Vété Math* (0,46)	Agro
						Math (0,36)		Math (0,38)
	Syntaxe* (0,33)			Math (0,32)			Comp.T (0,33) Bio* (0,32)	
	Vocab* (0,30)							
		Comp.G (0,26)				Chimie (0,26) Biologie* (0,25)		
	Hist-Eco* (0,25)		Hist-Eco* (0,25) Vocab* (0,24) Com.G(*) (0,23)					
Hist-Eco* (0,23)		Syntaxe* (0,24)		Biologie* (0,22) Physique (0,21)				Physique (0,23)
Syntax* (0,21) Vocab* (0,19)		Voca – hist (0,19)			Vocab.* (0,19)			Syntaxe (0,20)
					Syntaxe* (0,14) Chimie (0,14)			

Sur les 27 prédicteurs des « trios de tête » pour chaque faculté, 11 sont des scores SANS certitude et 15 AVEC certitude, et 1 obtient le même score avec et sans certitude, ce qui confirme l'intérêt de continuer à incorporer les degrés de certitude dans les recherches et les pratiques, même si l'on ne comprend pas encore (quoique : 10 sur 12 * pour les sciences humaines, 5 sur 15 pour les sciences dures) pour quelles disciplines ou quelles facultés ou quelles interactions des deux les indices « incorporant la subjectivité » sont plus prédictifs et pourquoi.

5.3 Les facultés pour lesquelles chaque épreuve prédit le mieux

Vocab	Syntax	Comp.T	Comp. G	Math Vété* (0,46)	Phys	Chim	Bio	Art	Hist-Eco Actualité
				Agro (0,38)	Vété* (0,38)	Vété (0,38)			
	Droit* (0,34)	Vété (0,33)		Médecine (0,35)					
Vété* (0,31) Droit* (0,30)				Sciences (0,32)			Vété* (0,32)		
Sc. Eco* (0,27)			Psycho* (0,26) Vété (0,25) Eco* (0,23)			Médecine (0,26)	Médecine (0,25)		Droit*- Eco* (0,25)
	Psycho* (0,24) PhiloL- sc.eco* (0,21)				Agro (0,23) Médecine - sciences (021)		Sciences* (0,22)	Droit (0,23)	PhiloL* (0,23)
		Eco (0,19)				Sciences appl* Agro* (0,17)	Agro (0,16)	PhiloLett (0,19)	

Colonnes 1 et 3 : Vocabulaire et Compréhension de textes : il pourrait apparaître étonnant que ce soit pour les Vétérinaires que la prédictivité soit la plus forte. Rappelons notre **hypothèse** selon laquelle la supériorité des Français, **nombreux dans cette faculté**, en réussite de fin d'année n'est pas due seulement à leur supériorité aux check-up (ici « Vocabulaire » et « Compréhension de textes »).

Colonne 2 : Syntaxe : il est assez logique que la meilleure prédictivité s'observe pour le Droit, étant donné la nature des performances (examiner la logique et donc la syntaxe de textes)

Colonne 4 : Compréhension de graphiques, schémas, cartes : la lecture de graphique (et la compréhension des statistiques) est le check-up isolé le plus prédictif dans la réussite des cours de candidature en psychologie.

Colonnes 5 à 8 : Math-Physique-Chimie-Biologie : la supériorité prédictive des mathématiques ne surprend pas.

Colonnes 9 et 10 : Art et Histoire-Economie-Actualité : que ces deux check-up de culture générale soient liés à la réussite en Droit et en Philo et lettres ne surprend pas.

6. Discriminativité Rétrospective (DR = MR – ME) de chaque check-up quant à la réussite en candidature

N = Nombre d'étudiants sur lesquels ont été calculés les indices (N = NR + NE).

% R1C = Pourcentage de Réussites en 1^{er} candi

NR = Nombre d'étudiants ayant réussi et NE ayant échoué en 1^{er} candidature

r = Corrélation entre le check-up d'octobre 1999 et l'échec (0) ou la réussite (1) en juin ou septembre 2000

MR = Moyenne à l'épreuve des étudiants ayant Réussi ; ME = ayant Echoué en première candidature

ET R = Ecart-Type à l'épreuve des étudiants ayant Réussi ; ET E = ayant Echoué en 1^{er} candidature

Toutes les corrélations et toutes les différences de moyennes (MR-ME) sont significatives au-delà de P. 05.

Vocabulaire : lien avec la réussite 0/1 en 1^{er} candidature et DR (MR – ME)

N	% R1C	NR	NE		% RC (0 à 100)					Moyenne Spectrale (-100 à +100)				
					r	MR	ME	ET R	ET E	r	MR	ME	ET R	ET E
3327	45,18%	1503	1824	Tous	0,210	53,99	47,13	15,95	15,71	0,221	16,44	7,83	19,45	18,35
219	46,6%	102	117	Philo et Lettres	0,123	53,79	49,81	15,70	16,18	0,193	19,76	12,36	19,42	18,36
422	47,2%	199	223	Droit	0,276	56,89	47,10	17,68	16,37	0,297	20,66	8,88	21,37	16,49
360	49,2%	177	183	Psycho-Education	0,192	49,27	43,21	15,90	15,10	0,176	10,50	3,81	19,36	17,98
176	46,6%	82	94	Sc. Economiques	0,215	47,96	41,80	14,77	13,28	0,241	8,9	1,21	16,33	14,87
309	39,8%	123	186	Sciences	0,160	52,84	47,43	15,77	16,68	0,190	16,37	8,30	20,80	20,31
420	51,9%	218	202	Sc. Appliquées	0,198	59,20	52,98	15,78	15,00	0,199	20,68	13,07	19,57	17,73
790	42,5%	336	454	Médecine	0,206	51,37	44,80	15,23	15,56	0,209	12,92	5,18	17,94	17,90
252	48,4%	122	130	Médecine vété	0,287	31,34	19,72	22,75	25,27	0,311	21,96	10,25	17,60	18,14
379	38,0%	144	235	Sc. Agronomiques	0,204	55,35	49,09	14,03	14,94	0,198	17,04	9,30	17,36	19,37

Syntaxe : lien avec la réussite et DR (MR – ME)

NT	% R1C	NR	NE	% RC	% RC (0 à 100)					Moyenne spectrale (-100 à +100)				
					r	MR	ME	ET R	ET E	r	MR	ME	ET R	ET E
3281	45,0%	1478	1803	Tous	0,189	88,19	80,06	20,16	21,72	0,198	40,01	29,25	25,99	26,70
209	44,5%	93	116	Philo et Lettres	0,181	88,81	81,90	17,31	19,82	0,217	44,55	33,73	22,20	25,05
445	48,5%	216	229	Droit	0,32	86,13	75,93	18,51	20,49	0,338	40,07	22,76	23,99	24,12
346	49,4%	171	175	Psycho-Education	0,211	87,19	78,57	18,95	20,92	0,24	39,97	25,61	24,37	25,60
173	45,1%	78	95	Sc. Economiques	0,217	86,15	77,47	19,01	19,95	0,188	35,28	26,00	25,01	23,53
321	37,7%	121	200	Sciences	0,105	85,53	82,10	18,29	21,37	0,077	39,30	35,02	24,29	27,82
413	52,5%	217	196	Sc. Appliquées	0,102	89,40	84,69	23,09	22,49	0,146	38,91	31,07	27,39	25,68
755	42,6%	322	433	Médecine	0,189	88,20	79,53	21,54	22,66	0,197	40,43	29,05	27,36	28,29
250	48,8%	122	128	Médecine vété	0,128	87,45	82,34	20,71	19,05	0,186	41,83	31,66	29,16	24,61
369	37,4%	138	231	Sc. Agronomiques	0,205	88,77	79,43	19,54	22,73	0,186	39,33	28,51	26,27	28,50

Compréhension de texte : lien avec la réussite et DR (MR – ME)

NT	% RIC	NR	NE		% RC					MS				
					r	MR	ME	ET R	ET E	r	MR	ME	ET R	ET E
3037	45,1%	1370	1667	Tous	0,1	46,53	42,20	21,86	20,73	0,114	4,67	-1,31	27,13	24,77
185	44,3%	82	103	Philo et Lettres	0,075	37,60	34,79	17,53	19,26	0,121	0,81	-4,66	18,66	19,83
255	51,8%	132	123	Droit	0,097	37,50	33,20	22,69	21,39	0,071	34,47	23,15	28,62	26,40
358	49,2%	176	182	Psycho-Education	0,093	46,40	42,67	21,52	17,99	0,132	3,2	-3,97	28,34	25,39
168	45,2%	76	92	Sc. Economiques	0,194	45,39	37,50	20,29	19,71	0,128	-1,31	-7,93	27,55	23,73
320	37,8%	121	199	Sciences	0,07	47,52	44,72	19,08	19,28	0,071	4,76	0,67	28,30	27,16
405	52,3%	212	193	Sc. Appliquées	0,045	51,41	49,22	24,62	24,01	0,058	7,02	3,85	28,20	26,44
741	42,9%	318	423	Médecine	0,068	45,49	42,75	20,90	18,68	0,106	4,31	-1,09	26,94	23,50
240	47,9%	115	125	Médecine vété	0,332	50,86	35,33	22,27	21,96	0,312	14,34	-3,57	30,41	24,06
365	37,8%	138	227	Sc. Agronomiques	0,12	51,69	46,55	19,96	20,84	0,129	9,64	2,73	24,86	26,35

Compréhension Géo : lien avec la réussite et DR (MR – ME)

NT	% RIC	NR	NE		% RC					MS				
					r	MR	ME	ET R	ET E	r	MR	ME	ET R	ET E
3239	45,0%	1459	1780	Tous	0,161	60,00	53,56	19,89	19,38	0,16	28,35	21,24	22,52	21,21
195	43,6%	85	110	Philo et Lettres	0,041	46,47	44,81	21,14	18,71	0,076	21,81	18,93	20,33	17,64
439	49,0%	215	224	Droit	0,211	56,60	48,48	19,62	17,98	0,193	22,64	14,63	22,32	18,72
358	49,2%	176	182	Psycho-Education	0,266	54,88	45,16	18,39	16,90	0,216	21,38	12,64	20,63	18,82
173	45,1%	78	95	Sc. Economiques	0,233	59,35	51,79	16,62	17,00	0,233	27,79	18,80	19,35	18,21
320	37,5%	120	200	Sciences	0,134	65,75	60,90	15,96	18,13	0,136	36,70	30,24	23,33	22,40
401	52,4%	210	191	Sc. Appliquées	0,119	71,40	66,75	19,40	19,25	0,109	37,20	32,06	24,27	22,75
742	42,9%	318	424	Médecine	0,173	57,01	50,23	20,08	18,28	0,186	25,55	17,52	21,86	20,30
244	48,8%	119	125	Médecine vété	0,178	59,41	52,80	18,10	18,56	0,252	30,85	21,29	18,04	18,79
367	37,6%	138	229	Sc. Agronomiques	0,168	65,58	59,30	17,09	18,19	0,161	33,94	26,75	21,85	20,99

Math : lien avec la réussite et DR (MR – ME)

NT	% RIC	NR	NE		% RC					MS				
					r	MR	ME	ET R	ET E	r	MR	ME	ET R	ET E
2275	44,6%	1015	1266	Tous	0,321	72,61	60,86	15,00	18,81	0,29	44,62	28,62	25,18	27,27
497	45,1%	224	273	Sciences	0,321	71,57	60,56	14,72	17,77	0,23	40,47	28,31	27,15	25,49
388	51,3%	199	189	Sciences Appliquées	0,115	84,51	82,01	10,08	11,41	0,12	63,37	58,65	18,21	20,50
787	42,6%	335	452	Médecine	0,356	67,19	54,72	15,41	16,79	0,33	36,75	19,20	24,96	24,79
242	48,8%	118	124	Médecine vété	0,448	72,96	57,11	12,58	18,46	0,46	47,35	24,65	20,41	23,59
361	37,1%	134	227	Sc. Agronomiques	0,381	70,62	57,83	11,88	16,62	0,33	40,94	24,88	20,42	23,33

Physique : lien avec la réussite et DR (MR – ME)

NT	% RIC	NR	NE		% RC					MS				
					r	MR	ME	ET R	ET E	r	MR	ME	ET R	ET E
2261	44,5%	1006	1260	Tous	0,204	69,73	62,48	16,36	17,88	0,196	38,70	28,70	25,03	24,71
488	45,5%	222	266	Sciences	0,214	68,29	60,56	16,90	18,10	0,148	35,42	27,64	26,91	24,96
388	51,3%	199	189	Sciences Appliquées	0,067	78,29	76,29	14,64	15,02	0,108	51,73	46,30	24,59	25,42
785	42,2%	331	454	Médecine	0,196	66,85	60,29	15,64	16,62	0,214	34,31	24,42	22,38	22,23
241	48,5%	117	124	Médecine vété	0,225	71,79	64,76	14,89	15,59	0,268	44,61	31,84	22,64	23,34
359	37,0%	133	226	Sc. Agronomiques	0,229	64,51	56,23	16,07	17,49	0,171	30,19	22,15	22,27	22,42

Chimie : lien avec la réussite et DR (MR – ME)

NT	% R IC	NR	NE		% RC					MS				
					r	MR	ME	ETR	ETE	r	MR	ME	ETR	ETE
2261	44,6%	1008	1258	Tous	0,219	59,98	51,17	19,47	19,47	0,23	28,61	17,13	28,81	22,93
485	45,4%	220	265	Sciences	0,175	56,64	49,15	21,71	20,28	0,16	25,03	17,01	27,40	22,83
388	51,3%	199	189	Sciences Appliquées	0,145	64,70	49,13	18,31	19,62	0,18	32,99	23,64	25,44	24,56
786	42,4%	333	453	Médecine	0,263	58,14	47,84	18,86	18,47	0,25	25,54	13,17	24,37	22,49
242	48,8%	118	124	Médecine vété	0,17	65,47	54,03	16,96	18,00	0,38	38,30	19,5	25,83	20,21
360	37,2%	134	226	Sc. Agronomiques	0,17	58,67	51,93	18,47	19,14	0,18	27,15	18,49	24,41	22,54

Biologie : lien avec la réussite et DR (MR – ME)

NT	% R IC	NR	NE		% RC					MS				
					r	MR	ME	ETR	ETE	r	MR	ME	ETR	ETE
2275	44,3%	1007	1263	Tous	0,203	53,27	45,90	18,00	17,32	0,209	16,56	7,10	22,88	21,25
487	45,0%	219	268	Sciences	0,218	50,73	42,57	18,75	17,62	0,22	14,64	4,63	23,73	20,56
387	51,2%	198	189	Sciences Appliquées	0,054	49,34	47,51	17,60	15,97	0,049	11,72	9,79	21,52	17,50
788	42,4%	334	454	Médecine	0,249	53,89	44,87	17,32	17,30	0,257	16,65	4,60	22,34	22,38
252	48,4%	122	130	Médecine vété	0,305	57,72	49,26	13,98	14,33	0,325	31,34	19,72	22,75	25,27
361	37,1%	134	227	Sc. Agronomiques	0,169	53,51	47,71	15,95	16,61	0,15	15,55	9,23	20,30	20,08

Art : lien avec la réussite et DR (MR – ME)

NT	% R IC	NR	NE		% RC					MS				
					r	MR	ME	ETR	ETE	r	MR	ME	ETR	ETE
1104	49,0%	541	569	Tous	0,163	51,08	46,12	15,37	14,54	0,154	21,93	16,64	18,01	15,91
192	47,4%	91	101	Philo et Lettres	0,191	55,38	50,53	12,19	12,71	0,181	29,03	22,98	15,88	16,92
393	49,4%	194	199	Droit	0,231	54,97	48,12	16,84	14,58	0,214	25,29	17,27	20,54	15,74
348	50,3%	175	173	Psycho-Education	0,103	46,10	43,19	14,10	13,93	0,079	16,13	13,70	15,27	15,41
171	45,6%	78	93	Sc. Economiques	0,156	47,18	42,71	12,99	15,09	0,14	17,90	13,92	13,81	14,19

Histoire-Eco-Actualité : lien avec la réussite et DR (MR – ME)

NT	% R IC	NR	NE		% RC					MS				
					r	MR	ME	ETR	ETE	r	MR	ME	ETR	ETE
1119	48,8%	546	579	Tous	0,188	51,6	46,07	14,67	14,19	0,203	16,99	10,45	16,65	14,87
193	47,2%	91	102	Philo et Lettres	0,19	52,39	47,06	12,91	14,60	0,232	22,13	14,78	14,12	16,53
399	49,1%	196	203	Droit	0,23	57,28	50,68	14,57	13,42	0,252	21,36	12,55	18,01	15,85
355	49,9%	177	178	Psycho-Education	0,192	44,36	39,26	12,72	13,33	0,155	9,20	5,11	13,82	12,15
172	45,9%	79	93	Sc. Economiques	0,209	52,86	47,40	14,03	11,68	0,247	17,78	10,92	14,95	12,02

Rappelons cependant que ces check-up n'ont pas été conçus dans un but prédictif. Ainsi, l'épreuve de vocabulaire appliquée aux étudiants en médecine n'a pas été élaborée par des enseignants de candidature en médecine.

7. Régressions multiples ²⁸ sur réussite 0/1

Dans la présente analyse, dite de régression **multiple**, seuls ont été gardés les étudiants pour lesquels nous disposons de TOUTES les données (scores à tous les check-up d'octobre 1999), ce qui n'est pas le cas pour calculer les corrélations **simples** où sont pris en considération tous les étudiants dont nous avons les données pour CETTE variable prédictive (que nous disposions ou non de ses données pour les AUTRES variables prédictives).

Dans les facultés de Sciences humaines, chaque étudiant (sauf données manquantes) a 24 scores prédictifs : ses scores à 6 épreuves, sous 4 formes :

TE = Taux d'Exactitude

MS* = Moyenne Spectrale (le * signifie « fait intervenir les degrés de certitude »)

ERCENTR* = Erreur de centration

REAL* = Réalisme par Calibration

Dans une analyse de régression multiple, on calcule les 24 différentes corrélations entre chacun des 24 différents prédicteurs et la réussite échec (en termes de 0/1). La corrélation la plus élevée détermine la première variable prédictive. On calcule ensuite quelle est la variable qui, combinée à cette première, augmente le plus la première corrélation, et ainsi de suite pour les variables prédictives successives, jusqu'à ce que l'analyse s'arrête par « manque de signification » dans l'amélioration, valeur qui dépend entre autres du nombre de sujets.

% VE représente le % de Variance Expliquée, c'est-à-dire la part de la variance entre les étudiants qui s'explique par ces prédicteurs. On verra que, avec l'augmentation du nombre de prédicteurs, la variance expliquée augmente, mais de moins en moins.

Philo et lettres (N = 154)

NVP	r	%VE	Prédicteurs
1	0,28	8%	Syntaxe MS*
2	0,36	13%	Syntaxe MS* + Histoire MS*
3	0,38	15%	Syntaxe MS* + Histoire MS* + CompGéo TE

Que la Syntaxe soit la plus prédictive et l'histoire la seconde n'a rien d'étonnant en Philo et Lettres.

Droit (N=174)

NVP	Rm	%VE	Prédicteurs
1	0,35	12%	Syntaxe MS*
2	0,40	16%	Syntaxe MS* + CompGéo REAL*
3	0,44	19%	Syntaxe MS* + CompGéo REAL* + Histoire MS*

On observe pratiquement la même combinaison (sauf l'interversion des prédicteurs 2 et 3) qu'en Philo et lettres.

Psycho (N=313)

NVP	Rm	%VE	Prédicteurs
1	0,26	7%	CompGéoTE
2	0,31	9%	CompGéoTE + Syntaxe MS*
3	0,32	11%	CompGéoTE + Syntaxe MS* + Vocab ERCENTR*
4	0,34	12%	CompGéoTE + Syntaxe MS* + Vocab ERCENTR* + CompGéo ERCENTR*

Que la Compréhension Géo (c'est-à-dire de graphiques, de tableaux et de cartes) soit la plus prédictive n'est pas étonnant dans une section où la statistique joue un rôle sélectif important.

²⁸ Sans données manquantes dans aucun des prédicteurs, ce qui explique la différence de nombres quand il s'agit de corréler une seule variable avec la réussite / échec. Les variables socio-économiques n'ont pas été prises en compte.

Eco (N=160)

NVP	Rm	%VE	Prédicteurs
1	0,24	5%	vocabulaire MS*
2	0,32	10%	vocabulaire MS* + CompGéo MS*
3	0,34	12%	vocabulaire MS* + CompGéo MS* + Histoire MS*

Le caractère social de la section nous paraît lié à la variable la plus prédictive (vocabulaire).

Sciences (N=266)

NVP	Rm	%VE	Prédicteurs
1	0,35	12%	MathTE
2	0,40	16%	Math TE + Bio MS*

Sc. Appliquées (N=348)

NVP	r	%VE	Prédicteurs
1	0,20	4%	Vocabulaire TE
2	0,25	6%	Vocabulaire TE + Chim MS*
3	0,28	8%	Vocabulaire TE + Chim MS* + Bio Real*

Médecine (N=669)

NVP	r	%VE	Prédicteurs
1	0,36	13%	MathTE +
2	0,39	15%	MathTE + Bio TE
3	0,41	16%	MathTE + Bio TE + Chim TE
4	0,42	17%	MathTE + Bio TE + Chim TE + Vocab ERCENTR*
5	0,43	18%	MathTE + Bio TE + Chim TE + Vocab ERCENTR* + MathERCENTR*
6	0,44	19%	MathTE + Bio TE + Chim TE + Vocab ERCENTR* + MathERCENTR* + BioERCENTR*
7	0,45	20%	MathTE + Bio TE + Chim TE + Vocab ERCENTR* + MathERCENTR* + BioERCENTR* + Chim ERCENTR*

Cette multitude de prédicteurs reflète-t-elle la multiplicité des talents requis pour réussir les études de première année en médecine ?

Agronomie (N=340)

NVP	r	%VE	Prédicteurs
1	0,38	15%	Math TE
2	0,39	16%	Math TE+ Bio TE

Vété (N=211)

NVP	r	%VE	Prédicteurs
1	0,45	20%	Math MS*
2	0,50	25%	Math MS* + Chimie MS*
3	0,53	28%	Math MS* + Chimie MS* + ComprText TE
4	0,55	32%	Math MS* + Chimie MS* + ComprText TE + ComprText ERCENTR

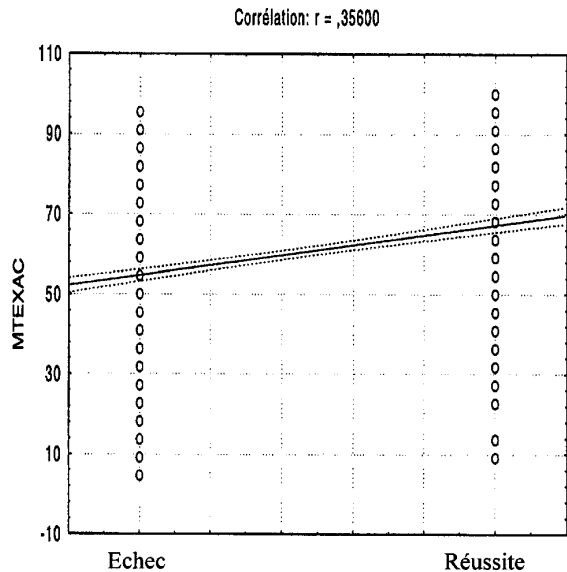
Il faut rappeler ici encore la même remarque qu'en page 22 : les check-up utilisés ici dans un raisonnement prédictif n'ont pas été conçus dans cette perspective.

8. Peut-on sélectionner à partir des liens entre check-up et réussite en candidature ?

Pour introduire ce débat, partons de la corrélation entre le score (en %RC) au check-up de mathématique et la réussite (0/1) en médecine. C'est une des plus élevées (si l'on excepte les vétérinaires) : 0,356.

La moyenne de la distribution des résultats à ce check-up des étudiants ayant ultérieurement réussi leur année est de 67,19 %. La moyenne de la distribution des résultats des étudiants ayant ultérieurement échoué en fin d'année est 54,72 %.

La différence est de 12,47 %, ce qui est important. Cependant, les écarts-types de ces distributions sont importants eux aussi : respectivement 15,41 % et 16,79 %, ce qui donne les deux distributions suivantes, très étalées :

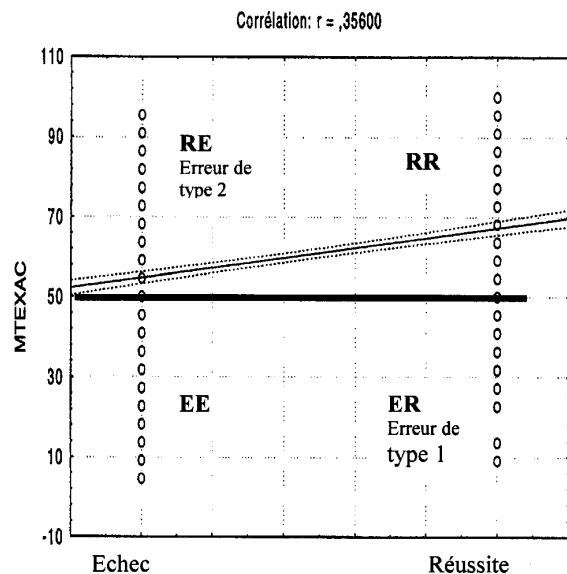


La difficulté de sélectionner sur la base de prédictions est illustrée par les 4 situations déterminées par une « barre de sélection fictive » qui serait fixée arbitrairement à 50% par exemple et qui déterminerait l'admission ou non aux études, une non-admission étant assimilée à un « Echec prédit » :

EE (Echec prédit, Echec observé),
RR (Réussite prédite, Réussite observée),
ER (Echec prédit, Réussite observée) et
RE (Réussite prédite, Echec observé)

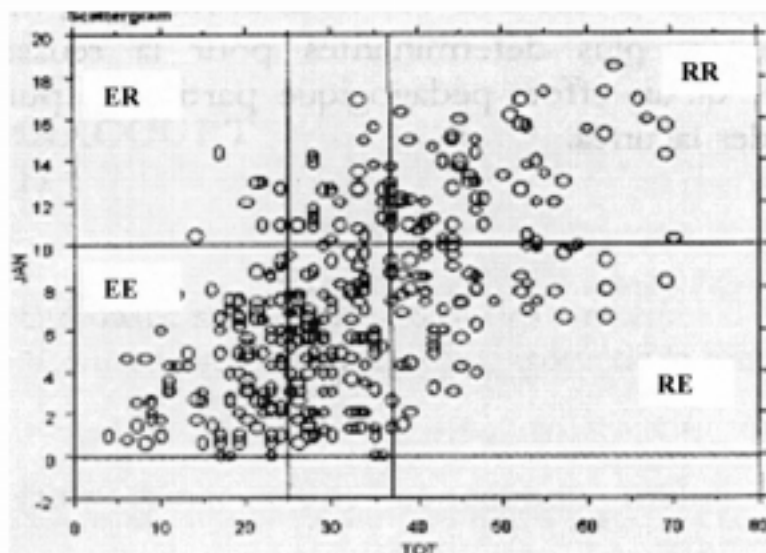
Les ER sont des **erreurs de type 1** (empêcher d'entrer un étudiant qui réussirait, donc en défaveur des étudiants) ou ER seraient très importantes en quantité et, évidemment, qualitativement.

Les RE sont des **erreurs de type 2** (laisser entrer des étudiants qui échoueraient). Elles seraient réduites par rapport à la situation actuelle (où elles sont maximales), mais elles restent très importantes (voir autre démonstration in Leclercq, 1998, 58-59).



Dans les tableaux de corrélations ci-avant, on constate que même en combinant plusieurs variables, les corrélations multiples « plafonnent » à 0,55, ce qui explique 30% de la variance inter-étudiants en réussite.

On retrouve des valeurs familières : 0,56 et 0,60 chez Leclercq (1998, p. 58) et 0,59 chez Lebrun et Lega (2001, 293), dont voici le graphique, assorti par nous des indications EE, ER, RE, RR :



Devant les risques considérables d'erreurs de type 1 (ER) et la démonstration (au chapitre 8) qu'il est possible d'infléchir le cours des choses au lieu d'en observer passivement l'occurrence fatale, nous pensons qu'il n'existe pas de validité prédictive suffisante dans ces prédicteurs check-up qui permettrait à elle-seule de fonder le principe d'une sélection équitable à l'entrée. Ceci rejoint les conclusions de Romainville (1997, 119).

9. Conclusions

Si on ne peut prédire, c'est que bien d'autres facteurs que les prérequis cognitifs interviennent dans la réussite.

De Ketele (1990), dans sa revue critique des recherches sur la réussite dans l'enseignement supérieur distingue trois groupes de facteurs :

- **Les facteurs cognitifs**, parmi lesquels :
 - *Les aptitudes intellectuelles*, à mesurer (par des tests appropriés) selon 4 axes : l'intelligence générale, les intelligences spécifiques (numérique, verbale, spatiale, etc.), le développement moral et social, et le développement de la pensée abstraite (ce que l'on oublie souvent de mesurer).
 - *La réussite scolaire antérieure*, à mesurer par le nombre d'années mises pour terminer le secondaire, la section d'origine et l'école d'origine des étudiants, les résultats obtenus dans des branches particulières prérequis aux orientations du supérieur (les check-up MOHICAN ne mesurant qu'une partie de ces prérequis, et, pour certains, de façon insuffisamment approfondie (trop peu de questions par exemple)).

- *Les capacités cognitives de base* (non pas les notions, mais les capacités exercées sur ces notions). De Ketele en liste 16 :
 1. dire la même chose en d'autres mots,
 2. à l'aide d'une représentation tabulaire ou graphique ou schématique
 3. en langage technique ou symbolique (et réciproquement),
 4. illustrer par des exemples et des contre-exemples différents mais corrects,
 5. trouver ou énoncer le problème dans une situation,
 6. distinguer les informations essentielles, accessoires et parasites,
 7. tenir compte de tous les éléments de la question,
 8. prendre des notes structurées (pour soi),
 9. déceler et énoncer la structuration d'un texte écrit que l'on vient de lire,
 10. d'un exposé que l'on vient d'écouter,
 11. résumer un message de façon succincte et précise,
 12. dans une situation donnée, appliquer un principe ou en tirer les implications (loi, règle,...),
 13. éviter les généralisations abusives,
 14. une question d'évaluation étant posée, se rappeler ou énoncer soi-même (selon le cas) les critères qui vont servir à l'évaluation et les utiliser judicieusement,
 15. rédiger une synthèse écrite structurée,
 16. présenter une synthèse orale structurée.

- *Les styles cognitifs...*

- **L'intérêt pour les études**
 1. La valeur reconnue par l'étudiant au diplôme.
 2. La valeur reconnue à la discipline de base.
 3. La valeur reconnue aux études.
 4. La valeur reconnue par l'étudiant aux animateurs de la formation.
 5. La valeur reconnue par l'étudiant à sa propre personne comme étant capable de réussir dans la filière choisie.
 6. La représentation qu'a l'étudiant de la valeur que les autres lui reconnaissent (surtout les enseignants et les pairs qui le marquent) quant à sa possibilité de bien réussir ses études.

- **La motivation** (intrinsèque ou résultant de pressions externes) **d'entrer dans une filière d'études**

- **La capacité de se fixer des objectifs pertinents, à court et à long terme**
 - Le programme de travail : le fixer, s'y maintenir, l'adapter.
 - Les stratégies d'étude : en surface ou en profondeur.
 - Les activités de dépassement (lectures, conférences).
 - La capacité d'intégration et de structuration.

Tous ces facteurs pourraient faire l'objet d'un deuxième MOHICAN, le premier ayant été fort « disciplinaire ». On peut anticiper que la prédictivité de ces variables combinées les unes aux autres sera plus élevée que celle des épreuves du premier des MOHICAN. Cela ne changera cependant pas l'impossibilité de sélectionner de façon totalement équitable sur base de prédictions. Les raisons sont essentiellement d'ordre éthique. En effet, **quelle que soit la prédiction, on peut la faire mentir... si on le veut**, le « on » étant aussi bien l'institution, l'enseignant que l'étudiant.

C'est une expérience de ce genre de volonté d'infléchir le cours des choses, d'intervenir pour faire mentir les prédictions qui est décrite dans le chapitre 8 (RESSAC) qui suit.

Il se peut cependant qu'un pays décide, et c'est ce que font la plupart des pays qui nous entourent (cf. chapitre 21, section A2) d'instaurer un *numerus clausus* sur base d'un concours à l'entrée. Ce problème est différent : ce n'est pas parce qu'il n'aurait pas assez de chances de réussir que certains sont alors refusés, mais parce qu'ils ne se classent pas en ordre utile pour occuper le nombre restreint de places disponibles.

10. Bibliographie

- Beguin, A., 1968 : une révolution inutile, Quelques chiffres et quelques réflexions à propos de la démocratisation du recrutement des étudiants à l'université, *Socio*, Périodique d'information et d'échange des professeurs de sciences sociales, Louvain-la-Neuve, 1991, 20-26.
- Crochet M., L'inégalité des chances à l'entrée à l'université. Discours de Rentrée académique, Université Catholique de Louvain, 16 septembre 2002.
- De Ketele, J.-M. (1990). Le passage de l'enseignement secondaire à l'enseignement supérieur : les facteurs de réussite. *Vie pédagogique*.
- Debry, M., Leclercq, D. & Boxus, E. (1998). De nouveaux défis pour la pédagogie universitaire, in D. Leclercq (Ed.), *Pour une pédagogie universitaire de qualité*, Sprimont : Mardaga, pp. 55-80.
- Lebrun et Lega (2001). Comment mettre en évidence et développer chez l'apprenant les compétences transversales requises pour le préparer à l'enseignement supérieur ? Actes du 1^o Congrès des Chercheurs en Education (291-294).
- Leclercq, D., Beguin, A., De Kerchove, A.-M., Lambert, J.-P., Pestieau, P. (1998). L'enseignement universitaire est-il démocratique ? in *Pour une Pédagogie universitaire de qualité*. Liège : Mardaga.
- Leclercq, D. (1998). *Pour une pédagogie universitaire de qualité*. Sprimont : Mardaga.
- Romainville, M. (1997). Peut-on prédire la réussite d'une première année universitaire. *Revue française de Pédagogie*, 119.

Chapitre 8

Changer de méthode d'étude suite aux feedbacks : L'opération RESSAC (Résultats d'Épreuves Standardisées au Service de l'Apprentissage en Candidature) ²⁹

D. Leclercq, Président du Groupe de Travail CIUF « Réussite en candidatures »

P. Detroz, Directeur du CAFEIM

*C. Dupont, chercheur au LabSET, Université de Liège
et J.-L. Gilles, Directeur du SMART, Université de Liège*

*et le précieux concours
des Professeurs S. Brédart, M. Crahay et Ch. Mormont
de la Faculté de Psychologie et des Sciences de l'Éducation
de l'Université de Liège*

²⁹ Cette recherche-action a été possible grâce au Conseil des Etudes de la FAPSE, au secrétariat facultaire qui a encodé une partie des données, grâce au CAFEIM-FAPSE et au secrétariat du STE qui ont pris en charge la plus grande partie des encodages.

La présente recherche-action vise à fournir, aux étudiants de la 1^{re} candidature en Faculté de Psychologie et des Sciences de l'Education de l'ULg, des informations diagnostiques de manière à ce qu'ils puissent, au vu des résultats, changer leurs méthodes d'études avant les examens de juin.

1. Les étapes de l'opération RESSAC menée par CAFEIM³⁰

Octobre 1999	Etape 1 : le 10 octobre 1999 , les étudiants de la FAPSE ULg entrant en première candidature sont testés dans le cadre du projet CIUF-MOHICAN au moyen de 6 épreuves «Check-up», dont une donnera deux scores séparés : Lecture de cartes, de graphiques, de tableaux d'une part et Connaissances géographiques d'autre part.
--------------	--

	Nombre de questions	Nombre d'étudiants
Vocabulaire	45	241
Syntaxe	12	229
Compréhension de textes	6	229
Compréhension de cartes, graphiques (Q1 à Q6)	6	229

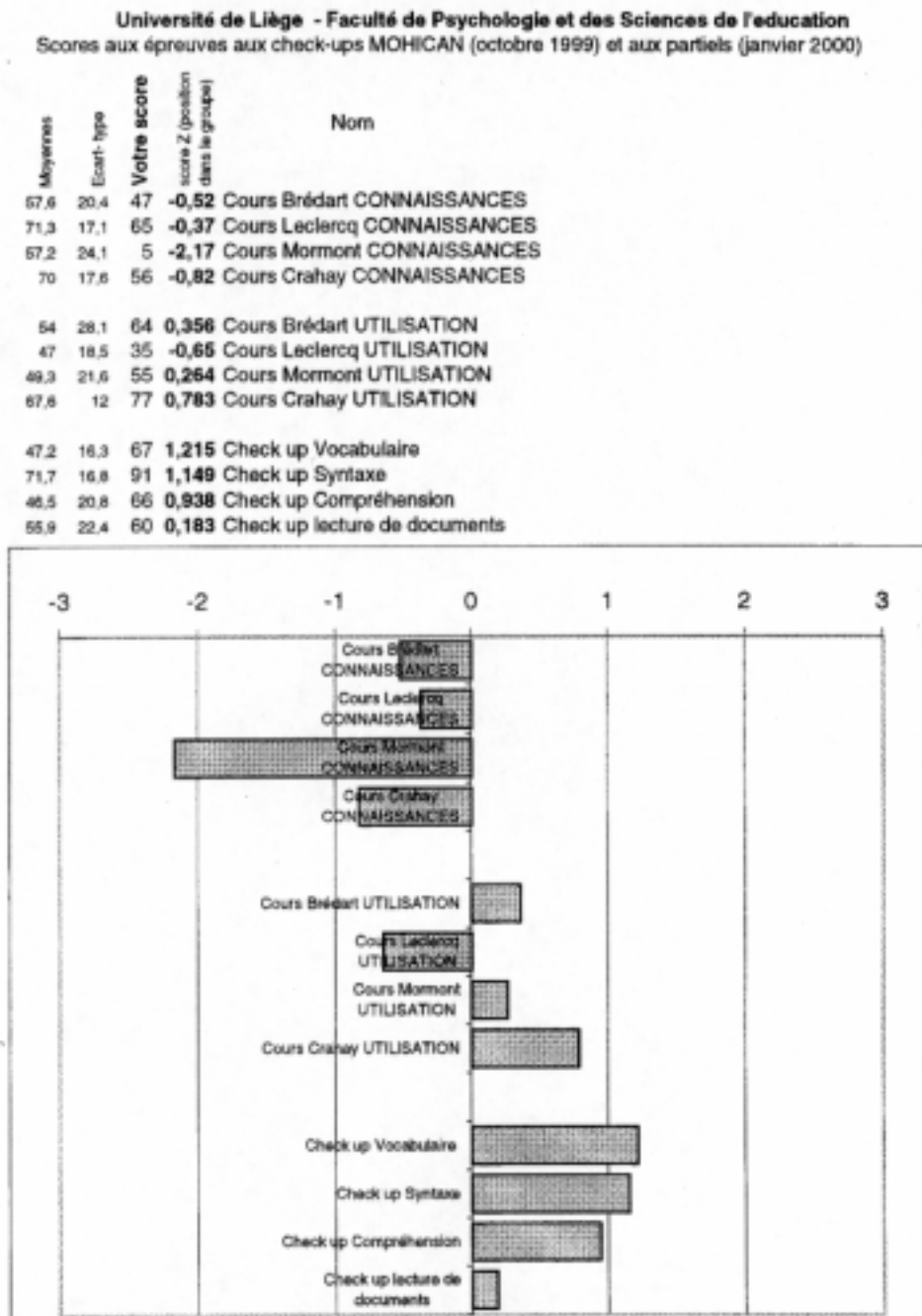
Mi octobre 1999	Etape 2 : Les étudiants reçoivent, chacun dans une enveloppe confidentielle, outre leurs 6 questionnaires (avec les questions, leurs réponses et leurs certitudes) leurs résultats sur 6 feuilles une par check-up (voir chapitre 2).
Novembre 1999	Etape 3 : Le Conseil des Etudes de la FAPSE est informé des résultats statistiques (moyenne de réussite par question, distribution spectrale des réponses des 213 étudiants), aux check-up, sauf celui de vocabulaire.
Janvier 2000	Etape 4 : Des Interrogations Partielles de Janvier (IPJ) ont eu lieu pour 4 cours, ceux donnés par Serge Brédart, Marcel Crahay, Christian Mormont et Dieudonné Leclercq. Elles ont donné lieu à un feedback sanctionnant (une note sur 20 pour chacun, les notes égales ou supérieures à 12 dispensant de la matière).
Avril 2000	Etape 5 : Ces quatre professeurs ont calculé, pour leurs 4 <i>interrogations dispensatoires de janvier 2000</i> , deux scores (en %), l'un portant sur les questions de « <i>connaissances</i> » et l'autre portant sur les questions « <i>d'utilisation des connaissances</i> » de manière à produire 8 notes (exprimées en % de réussite et en notes Z) : – Connaissances Brédart / C Crahay / C Leclercq / C Mormont – Utilisation des Connaissances Brédart / UC Crahay / UC Leclercq / UC Mormont
Mai 2000	Etape 6 : Au tout début de la « <i>bloque</i> », B. Dardenne et D. Leclercq (au nom du Conseil des Etudes) ont envoyé aux étudiants de 1 ^{re} candidature de la FAPSE un Dossier z personnel (voir ci-après) établi par le CAFEIM et leur rappelant leur position (ou note z) par rapport à la moyenne de leur cohorte des 200 testés) dans les mesures obtenues à partir de 4 épreuves MOHICAN très liées à la <i>compréhension</i> : Vocabulaire, Syntaxe, Compréhension de texte, Lecture de cartes, tableaux, graphiques ³¹ et dans les 8 scores de l'étape 5
Juin 2000	Etape 7 : A la fin des examens de 1 ^{re} session, D. Leclercq et J.-L. Gilles ont demandé aux étudiants de 1 ^{re} candidature de remplir un questionnaire d'avis (voir ci-après) sur l'utilité du dossier personnel de mai ; 209 questionnaires remplis ont été recueillis.

³⁰ Le Centre d'Auto-Formation et d'Evaluation Interactive Multimédias de la Faculté de Psychologie et des Sciences de l'Education de l'Université de Liège. Ces opérations ont été menées par D. Leclercq, J.L. Gilles, P. Detroz et M. Pirson

³¹ Ce sont les questions 1 à 6 du check-up MOHICAN de Lecture de documents *géographiques*.

2. La radiographie préparatoire à la bloque (mai 2000)

Au tout début de la période de « bloque » (mai 2000), chaque étudiant a reçu sa « radiographie z », c'est-à-dire les valeurs Z de ses scores aux check-up MOHICAN et aux scores « Mémorisation » et « Compréhension -Utilisation des connaissances » de 4 de leurs examens partiels de janvier. En voici un exemple :



On constate que pour cette étudiante, la compréhension aux check-up est assez bonne, la compréhension aux partiels moyenne et la Mémorisation aux partiels faible.

3. En juin 2000, après les examens

Un questionnaire d'avis leur demandait (entre autres) si ces renseignements avaient été utiles et avaient **influencé leur manière d'étudier**. Voici un exemple de questionnaire rempli par un étudiant différent de celui de la section 2. On notera en passant une circonstance de la vie (accouchement) qui est peut-être liée à la non réussite. Dans bien d'autres cas, ces circonstances défavorables à la réussite sont tout aussi importantes, mais nous restent inconnues.

Université de Liège
Faculté de Psychologie et des Sciences de l'Éducation
 Juin 2000
Enquête sur l'information retirée des feedbacks
 sur les check-up d'octobre 1999 et les partiels de janvier 2000

En octobre, vous avez reçu, dans une enveloppe confidentielle, 6 pages décrivant vos résultats à 6 check-up.

1. Le(s)quel(s) de ces résultats a (ont) été informatif(s) pour vous (entourez) et en quoi ? (expliquez)

Vocabulaire *les résultats ont réduit mon incertitude, maintenant je sais où est mon vocabulaire (il est très bon)*

Syntaxe

Compréhension de texte

Lecture de cartes, tableaux, graphiques géographiques

Histoire, économie, actualité *connaissances faibles*

Connaissances artistiques

2. Suite à ces constatats, avez-vous fréquenté une guidance, remédiation ? *non*
 Laquelle ?

Avez-vous fait autre chose ? *non, à part prendre le dictionnaire quand je ne connais pas un mot*

3. Au début de la période de « bloque », vous avez reçu un graphique mettant en évidence vos points forts et vos points faibles (du type ci-contre). Qu'avez-vous appris sur vos performances grâce à ce tableau ?

je comprends mieux que je m'étudie!

4. Cela a-t-il influencé votre manière d'étudier pour les examens de 1^{re} session ? Comment ?

pas vraiment étant donné que j'ai eu un bébé je m'ai pas beaucoup de temps pour étudier pourtant j'ai vraiment envie de réussir en seconde session

The form includes a bar chart on the right side with the following data (approximate values):

Item	Value
101	1.5
102	1.5
103	1.5
104	1.5
105	1.5
106	1.5
107	1.5
108	1.5
109	1.5
110	1.5
111	1.5
112	1.5
113	1.5
114	1.5
115	1.5
116	1.5
117	1.5
118	1.5
119	1.5
120	1.5
121	1.5
122	1.5
123	1.5
124	1.5
125	1.5
126	1.5
127	1.5
128	1.5
129	1.5
130	1.5
131	1.5
132	1.5
133	1.5
134	1.5
135	1.5
136	1.5
137	1.5
138	1.5
139	1.5
140	1.5
141	1.5
142	1.5
143	1.5
144	1.5
145	1.5
146	1.5
147	1.5
148	1.5
149	1.5
150	1.5
151	1.5
152	1.5
153	1.5
154	1.5
155	1.5
156	1.5
157	1.5
158	1.5
159	1.5
160	1.5
161	1.5
162	1.5
163	1.5
164	1.5
165	1.5
166	1.5
167	1.5
168	1.5
169	1.5
170	1.5
171	1.5
172	1.5
173	1.5
174	1.5
175	1.5
176	1.5
177	1.5
178	1.5
179	1.5
180	1.5
181	1.5
182	1.5
183	1.5
184	1.5
185	1.5
186	1.5
187	1.5
188	1.5
189	1.5
190	1.5
191	1.5
192	1.5
193	1.5
194	1.5
195	1.5
196	1.5
197	1.5
198	1.5
199	1.5
200	1.5

4. Analyses de 8 cas

Nous présentons ci-après 8 cas (sur plus de 200) rebaptisés A, B, C, D, E, F, G, H pour des raisons de confidentialité.

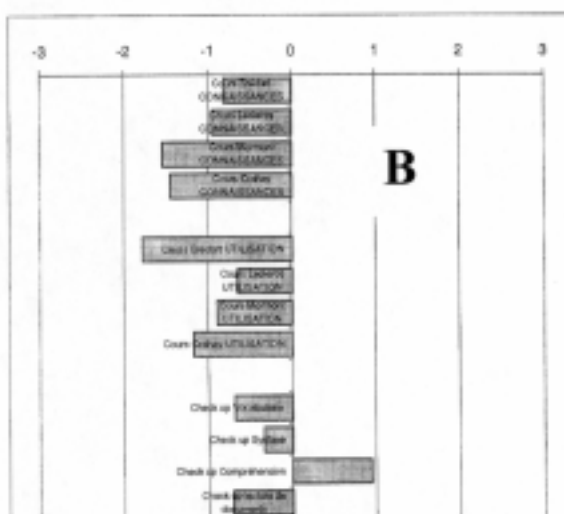
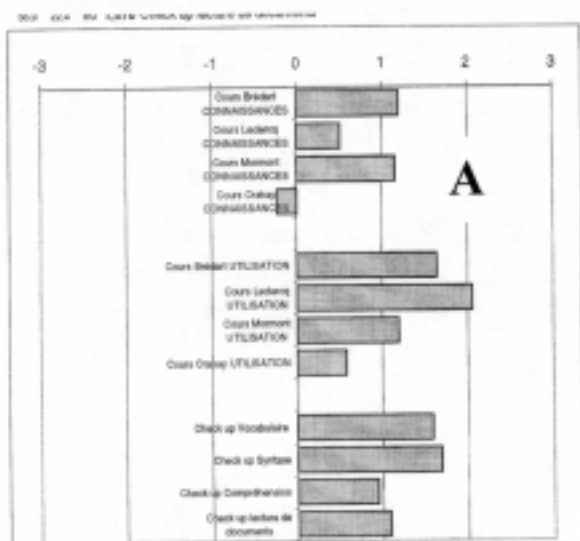
4.1 Les deux extrêmes opposés : A et B

La première (A) n'a pas de raisons de changer sa méthode « gagnante » et réussit. La deuxième (B) ne change pas non plus, mais elle, elle le devrait, comme le montre sa radiographie.

A	Qualité de la radiographie z	B
Excellent partout	Décision de changement de méthode	Faible partout
Pas de changement	Résultat Réussite/ Echec de l'année	Pas de changement
GD en 1 ^{re} session		Ajournée aux 2 sessions

A : Excellente partout, Moyenne en "Mémorisation des connaissances pour le cours du professeur Crahay" (en janvier !). Grande Distinction en 1^{re} session (**GD**)

B : Faible partout, sauf dans le check-up MOHICAN "Compréhension de texte". Elle est Ajournée en 1^{re} ET en 2e session (**AA**)



Ma connaissance en *Histoire-économie-actualité* n'est pas très grande, mais pourtant égale à la moyenne.

Qu'avez-vous appris par ce tableau et ce graphique ? Rien de nouveau, je connaissais déjà mes capacités, mais c'est quand même intéressant de se situer par rapport aux autres étudiants.

Cela a-t-il influencé votre manière d'étudier ? Non je n'ai pas changé ma manière d'étudier car elle semblait convenir.

Commentaire

Absence de modification de méthode apparemment justifiée !

J'ai pu remarquer que j'avais des problèmes en *vocabulaire* et besoin d'aide en *syntaxe*.

Qu'avez-vous appris par ce tableau et ce graphique. Je ne sais plus.

Avez-vous fréquenté une guidance ? Non : j'ai essayé par moi-même de corriger mes erreurs.

Cela a-t-il influencé votre manière d'étudier ? En général, non.

Commentaire

Absence de prise de conscience ou d'action apparemment fatale pour la réussite.

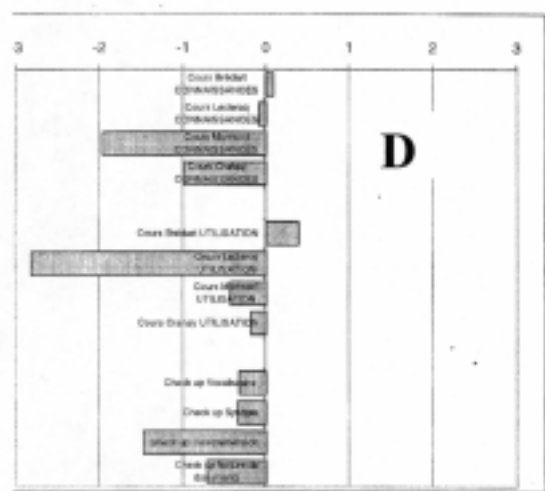
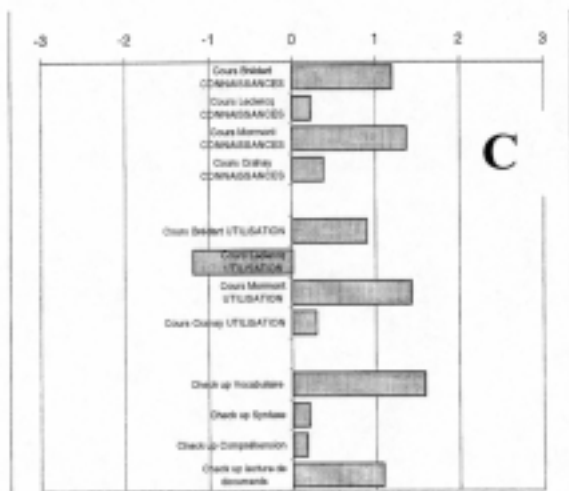
4.2 Deux cas contrastés qui finissent bien, mais pas en même temps

C a eu des résultats MOHICAN et des partiels rassurants. Sa radiographie z est positive, à une seule exception près. Elle a cru que cela était arrivé, elle a relâché ses efforts et a failli rater.

D a une radiographie z négative, à une seule exception près. Il décide de faire très attention à la signification des mots de la langue française (il est d'origine étrangère) et sur la compréhension. Il échoue en première session, mais réussit en seconde.

C : Bon en général (une exception : très faible en "Compréhension" pour Leclercq). Satisfait dès la 1^o session. (S)

D : Faible partout et Moyen en "Mémorisation" et "Utilisation" Brédart. Très faible en "Compréhension" pour Leclercq. Ajourné en 1^o session et Satisfaction en 2^o (AS)



Vocabulaire : je me suis rendue compte que j'avais une bonne base.

Histoire-Economie-Actualité : Faiblesses dans ces domaines, j'ai du mal à m'y intéresser.

Qu'avez-vous appris par ce tableau et ce graphique ?

En rapport à d'autres étudiants de mon entourage, je m'en sors plutôt bien. Cela m'a donné encore plus confiance en moi. Je me suis dit "l'unif, c'est facile !"

Cela a-t-il influencé votre manière d'étudier ?

Oui. Je me suis dit que, vu mes résultats, je pourrais me permettre quelques séances de bronzage ; j'ai passé deux mois "cool" ! Résultat : je suis tombée à court de temps. C'est MA faute...

Mais je reste confiante malgré cela.

Commentaire

Ces avis ont été recueillis après les examens de 1^o session, mais avant que soient connus les résultats de la délibération de 1^o session.

Cette étudiante s'est rendu compte que sa réussite peut être compromise par un excès de confiance. Heureusement, son imprudence ne lui a pas été fatale.

Vocabulaire : J'ai pris conscience du fait que le français, la bonne manipulation de la langue française était extra importante.

Qu'avez-vous appris par ce tableau et ce graphique ?

Que j'avais de grosses lacunes au niveau de la compréhension (notamment des questions)...

Cela a-t-il influencé votre manière d'étudier ?

Oui : au lieu d'étudier bêtement par cœur parce que je n'intégrais pas parfaitement certains concepts, je me suis efforcé de "traduire" les cours théoriques avec mes mots à moi, ce qui eut une meilleure compréhension des concepts, une meilleure restitution.

Commentaire

Cet étudiant, d'origine étrangère, s'est battu sur les deux fronts : celui de la compréhension en profondeur (notamment du vocabulaire et de la langue française) et celui de la compréhension en profondeur en "traduisant dans ses mots", et a fini par gagner au 2^o round (en 2^o session).

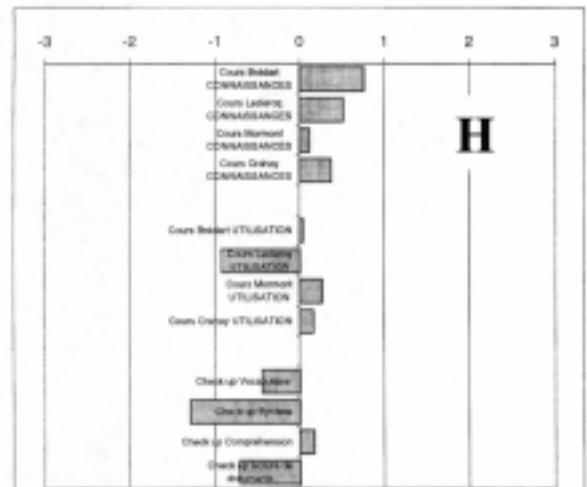
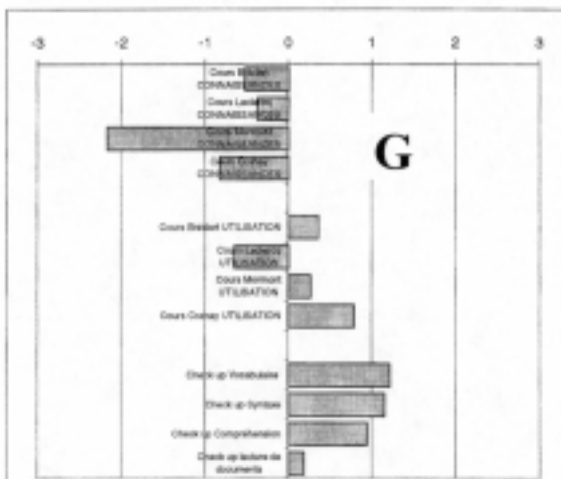
4.4 Deux cas paradoxaux : l'issue surprend

G a de bons scores aux check-up Compréhension, des scores moyens aux partiels (sauf dans un cours) en compréhension et faibles en mémorisation. Elle dit améliorer sa méthode, et pourtant échoue.

H dont les résultats sont « moyens » ne change pas de méthode et réussit.

G : Faible en "Mémorisation" et Bonne en "Compréhension" MOHICAN et Moyenne en "Compréhension - Utilisation" en janvier, sauf pour Leclercq. Ajournée en 1^o et en 2^o session. (AA).

H : Bonne en "Mémorisation" et Faible en "Compréhension - Utilisation" au cours Leclercq et aux check-up MOHICAN sauf "Compréhension de Texte". Moyenne ailleurs. Ajournée en 1^o session et Satisfaction en 2^o (AS)



En lecture de cartes, graphiques : j'ai pu voir un manque de connaissance
Qu'avez-vous appris par ce tableau et ce graphique ? *Que j'avais une mauvaise connaissance et une bonne utilisation, que je ne comprenais pas assez bien la matière.*
Cela a-t-il influencé votre manière d'étudier ?
Oui : j'ai rédigé des résumés et insisté sur les points importants.

En compréhension : Je ne savais pas ce que je valais, et la compréhension est très importante.
Qu'avez-vous appris par ce tableau et ce graphique ?
J'ai une moins bonne compréhension et une meilleure mémoire.
Cela a-t-il influencé votre manière d'étudier ?
Non.

Commentaire :
 Que la performance en compréhension soit plus faible pour le partiel de Leclercq pourrait s'expliquer par le fait que ce professeur utilise des "Solutions Générales Implicites" (des pièges) sollicitant la vigilance, formule à laquelle cette étudiante a peut-être du mal à s'adapter. Les étudiants H, C et D présentent un profil comparable sur ce point, mais eux, réussissent, C dès la première session, H et D lors de la seconde. Les changements de méthode de cette étudiante n'ont pas été suffisants pour réussir. Les raisons, potentiellement multiples, nous sont inconnues.

Commentaire
 Son profil faible en "Compréhension" pour Leclercq pourrait s'expliquer de la même façon que pour l'étudiant G, et le diagnostic a été insuffisant pour changer ses manières d'étudier...mais cette étudiante a réussi quand même, en seconde session. Nous faisons l'hypothèse qu'il s'est passé quelque chose lors de la préparation à la seconde session (changement de méthode, intensité de l'effort), non présent, forcément, dans l'avis émis en juin.

5. Vue d'ensemble des résultats de RESSAC 2000

5.1 Population étudiée

N'ont été gardés, **pour chacun des 4 cours**, que les étudiants qui **A LA FOIS**

- y avaient échoué au partiel de janvier 2000 (mais avaient tenté, et pour lesquels on dispose d'une note, ce qui constitue un « prétest »),
- ont une note pour juin ou septembre 2000,
- ont reçu le feedback de mai 2000
- ont rentré leur questionnaire d'avis de fin juin 2000
- ont donné au questionnaire d'avis de fin juin 2000 une réponse que l'on peut classer SANS AMBIGUÏTE dans l'une des catégories SANS et AVEC ci-dessous

Un grand **nombre d'étudiants** ont ainsi dû être éliminés de l'analyse pour « données manquantes » notamment ceux qui ont abandonné, ceux dont les réponses étaient ininterprétables (ex : « j'ai changé ma méthode de travail »). Par contre, pour ceux qui restent, on peut faire une analyse assez fine.

Parmi ces étudiants (en partie les mêmes et en partie différents pour chaque cours), deux *groupes ont été contrastés* : ceux

- **AVEC déclaration** (en juin après les examens) **de changement de méthode d'étude, soit d'étudier en profondeur**, de ne plus mémoriser sans comprendre, **soit de mémoriser plus** (là où le diagnostic était « faiblesse en connaissances »).
- **SANS une telle décision**, c'est-à-dire inutilisables, ambiguës, etc. ou « cela ne m'a rien appris », ou « pas de réponse » ou « Cela ne m'a PAS influencé(e) ».

5.2 Le tableau résumant les résultats (surtout colonnes 8, 9 et 10) obtenus sur 88 étudiants

1	2	3	4	5	6	7	8	9	-10	11	12
	N sans	N avec	Total	R sans	R avec	R Total	%R sans	%R avec	Diff	r sans	r avec
Total	99	105	204	28	45	73	28%	43%	15%	0,48	0,285
Leclercq	28	34	62	9	19	28	32%	56%	24%	0,44	-0,21
Mormont	23	27	50	2	6	8	9%	22%	14%	0,32	0,53
Brédart	33	31	64	16	17	33	48%	55%	6%	0,50	0,49
Crahay	15	13	28	1	3	4	7%	23%	16%	0,64	0,33

La colonne 2 contient les nombres de scores (à 4 cours) d'étudiants ayant échoué le partiel de janvier de ce cours, qui ont reçu les « radiographies Z » mais les ont interprété SANS prendre la décision d'étudier en PROFONDEUR, de creuser la COMPREHENSION avant la mémorisation.

La colonne 3 contient les nombres de scores d'étudiants dans le même cas mais AVEC cette décision.

La colonne 4 contient la somme des deux.

La colonne 5 contient le nombre de **réussites SANS (exprimé en % en colonne 8)** pour ces 4 cours.

La colonne 6 contient le nombre de **réussites AVEC (en % en colonne 9)** pour ces 4 cours correspondant aux situations de la colonne 2.

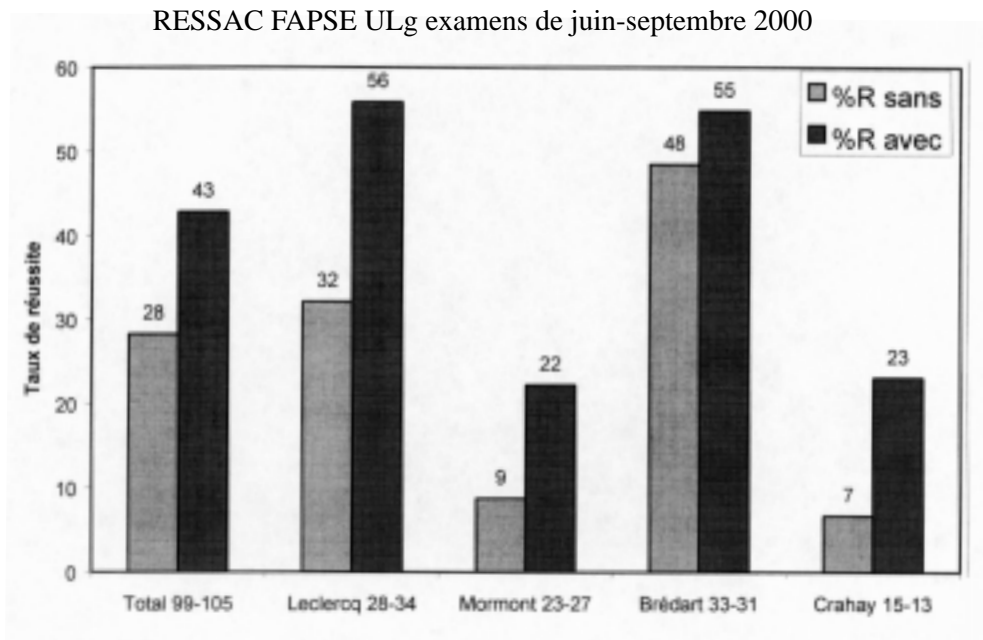
La colonne 7 contient le total des réussites (col 5 + col 6).

Les colonnes 8,9 et 10 contiennent les pourcentages de réussite AVEC et SANS changement de méthode.

Les colonnes 11 et 12 contiennent les corrélations entre la note en janvier et celle en juin-septembre 2000.

5.3 La représentation graphique des résultats

Voici la version graphique des mêmes données :



Rappel: AVEC signifie « Avec déclaration (lors de la prise d'avis de juin 200) de modification de méthode d'étude ». SANS signifie « Sans une telle déclaration ».

5.4 Interprétations

1. Sur ces 204 examens partiels ratés en janvier 2000, 73 se sont soldés par une réussite en juin-septembre 2000, soit 35%, mais avec 28% chez les SANS (modification de stratégie) et 43% chez les AVEC (modification de stratégie), soit 15% de différence, et la différence va dans le même sens pour LES 4 COURS.
2. Dans 2 des 4 cours, la décision d'étudier en profondeur est caractérisée par une plus faible corrélation entre janvier et juin-septembre, c'est-à-dire le poids du passé, de la note du partiel (voir colonnes 11 et 12).

Que les étudiants reçoivent un feedback serait donc une condition nécessaire mais non suffisante pour que se produise la décision d'étudier en profondeur. Encore faut-il

- que ce feedback soit suffisamment « diagnostic » pour déclencher une prise de conscience (la distinction entre CONNAISSANCE – de mémoire – et UTILISATION ou COMPRÉHENSION semble un minimum suffisant pour produire un effet chez certains).
- que l'étudiant interprète ce feedback de façon à prendre une décision de modification de ses méthodes de travail vers l'étude en profondeur. Nous pourrions l'y aider (avec l'aide de Guidance-Etude par exemple).

5.5 Réserves

5.5.1 Réserve 1

L'analyse a porté sur 204 couples de notes (chaque couple ayant le «prétest» en janvier et le «post-test» en juin ou septembre), mais certains de ces examens concernent un même étudiant (ayant échoué, aux partiels de 2000 à plus d'un des 4 cours). Ces 204 examens ont été produits par 88 étudiants au total. On pourrait craindre que les observations sur les examens de chacun des 4 professeurs le soit sur un **ensemble d'étudiants identiques**, et qu'il s'agisse, en conséquence de «4 fois la même observation». Il n'en est rien comme le montre le tableau ci-dessous où chaque colonne correspond à l'un des cours (initiale du nom du titulaire) et où 0 signifie « avait échoué en janvier » et a donc fait partie des données pour ce cours-là ; 1 signifiant « avait réussi en janvier », et ne contribue pas aux données pour ce cours-là.

La colonne de gauche indique le nombre d'étudiants SANS qui sont dans la situation indiquée (par exemple 0100) et la colonne de droite le nombre d'étudiants dans la situation AVEC qui sont dans la même situation. Seuls 5 étudiants dans le groupe SANS et 4 dans le groupe AVEC sont dans la situation 0000 c'est-à-dire avaient échoué aux partiels de janvier chez les 4 professeurs **à la fois**.

SANS	L	M	C	B	AVEC	
11	0	1	1	1	7	Echec dans un seul cours
0	1	0	1	1	2	
0	1	1	0	1	0	
5	1	1	1	0	3	
4	0	1	1	0	7	Echec dans 2 cours
6	1	0	1	0	3	
0	1	1	0	0	3	
0	0	1	0	1	1	
0	0	0	1	1	3	
1	1	0	0	0	2	
0	1	0	0	0	0	Echec dans 3 cours
2	0	1	0	0	1	
8	0	0	1	0	9	
0	0	0	0	1	1	
5	0	0	0	0	4	Echec dans 4 cours
42					46	Total

5.5.2 Réserve 2

Il s'agit, en juin 2000, de **déclarations**. On pourrait craindre que le lien soit dû à une tendance à répondre (par désirabilité sociale ?) plus prononcée chez les « AVEC » que chez les « SANS ».

Il n'en est rien, comme le montre les taux de réponses positives aux questions sur l'informativité des données personnelles MOHICAN fournies en octobre 1999 :

Informativité des feedbacks des check-up en octobre ? % de OUI.	SANS N=46	AVEC N=51
Vocabulaire	47%	54%
Syntaxe	36%	33%
Compréhension	36%	31%
Lecture de graphiques	27%	37%
Histoire - économie - actualité	38%	37%
Connaissances artistiques	27%	25%
Avez-vous eu recours à une Guidance ?	15%	23%
FB de juin	62%	37%

5.5.3 Réserve 3

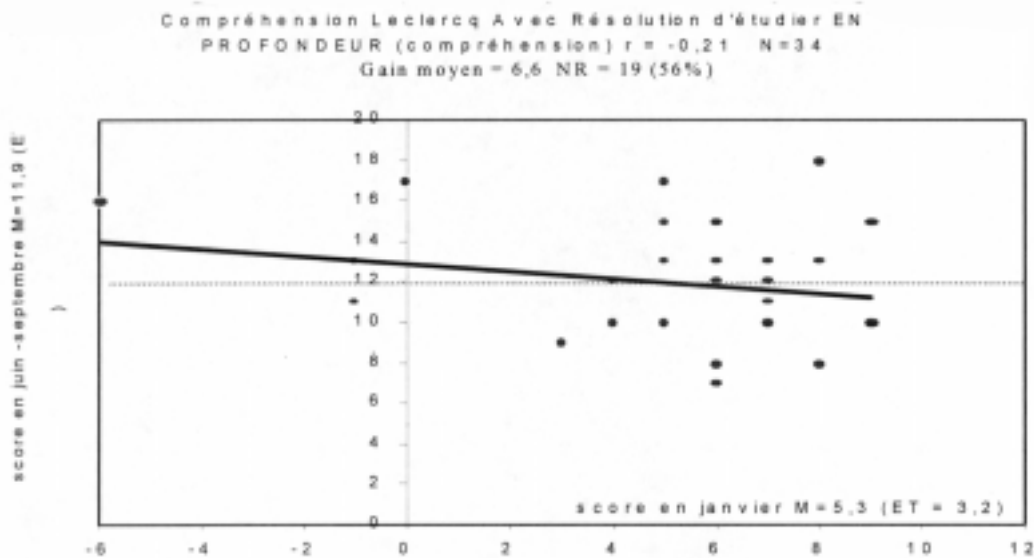
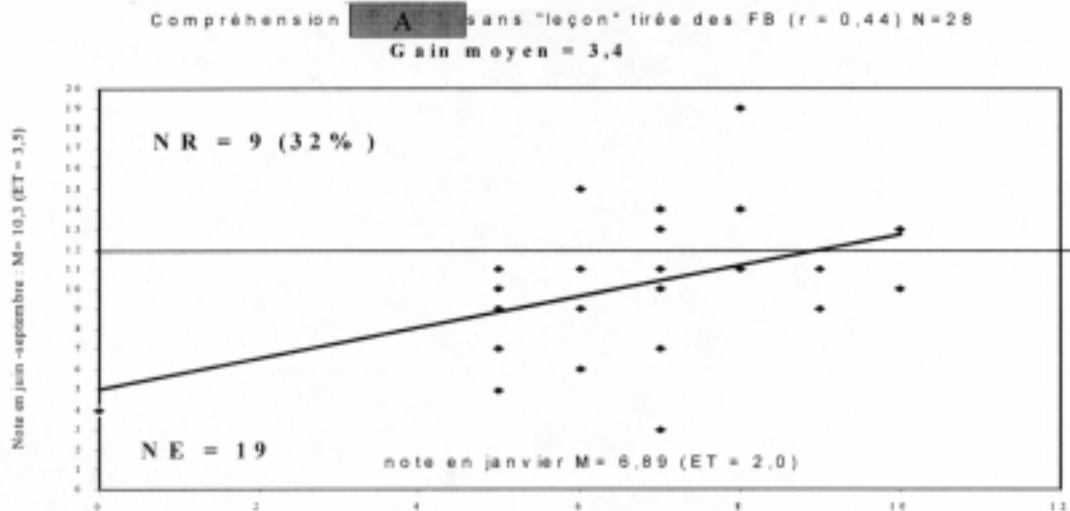
Le succès des deux groupes à **l'ensemble de l'année** est-il comparable aux observations faites sur les 4 cours seulement ? Oui globalement (les deux sessions confondues), ainsi que pour les abandons, mais pas dans la répartition des réussites entre les deux sessions :

	SANS N = 46	AVEC N = 51
Aux 2 sessions confondues	46,8%	50,9%
En 1 ^o session	34,04%	15,6%
En 2 ^o session	12,7%	35,3%
Abandons	27,6%	23,07%

Cette inégalité dans la répartition entre les sessions reste une question à approfondir. On peut émettre l'hypothèse que l'étude en profondeur a impliqué une plus fréquente décision de répartition de l'effort sur les deux sessions. Nous ne disposons pas de données permettant de tester cette hypothèse.

6. La représentation graphique détaillée par cours

Pour chacun de ces groupes, un nuage des points a été dressé (chaque point étant constitué par le couple de notes janvier d'une part et juin-septembre d'autre part). Chaque nuage a son centre (la rencontre des deux moyennes, non dessinée ici), sa droite de tendance et sa corrélation.



Dans le groupe SANS changement de méthode, la corrélation entre le prétest (janvier) et le post-test (juin-septembre) est positive (0,44) et pour le groupe AVEC changement de méthode elle est négative (-0,21).

Ce phénomène, bien connu en éducation (diminution de la corrélation entre prédicteurs et résultats suite à une intervention) s'observe dans un autre des 4 cours ; l'inverse (augmentation de la corrélation) dans un troisième et statut quo dans le quatrième (voir colonne 12 du tableau en E2).

7. Prospective

7.1 La portée des résultats

Ces résultats très encourageants, non pas par leur ampleur, limitée à quelques étudiants, mais par les hypothèses d'amélioration qu'ils permettent maintenant de formuler.

Ces résultats se reproduiraient-ils une autre année ? Rien ne permet de penser le contraire. Nous en faisons donc l'hypothèse. En gardant à l'esprit cependant que la relation causale «Je constate, donc je décide, donc je change, donc je réussis» soit fallacieuse. Il se pourrait que nous soyons devant un phénomène de **rationalisation a posteriori** pour bon nombre d'étudiants. Seules des analyses ultérieures plus approfondies dans des répliques de l'expérience pourront lever cette ambiguïté. Des facteurs non envisagés ici sont peut-être intervenus : les condisciples et leur effet d'entraînement (positif ou négatif), des circonstances externes (finale de la coupe de football ou leur équivalent), des constats faits par l'étudiant lui-même indépendamment des feedbacks MOHICAN et RESSAC.

Peut-on espérer améliorer ces résultats ? Nous pensons que oui, et de plusieurs façons. Une première consiste à être **plus explicite sur la signification des notes z**, en écrivant (par programme informatique) sur la feuille l'interprétation (ex «Mémorisation insuffisante»). Une deuxième amélioration consisterait à **pousser plus explicitement les étudiants à consulter un service de Guidance**, radiographie z en mains. C'est ce que nous testons en 2001-2002 (Leclercq, Lacaille et les 3 mêmes enseignants) dans un RESSAC 2002. Nous restons bien conscients cependant que tirer les conclusions («Vous devriez mémoriser plus» ou «Vous devriez étudier plus en profondeur») à la place des étudiants pourrait avoir un effet démobilisateur car un certain nombre d'étudiants pourraient n'«accepter» que les conclusions auxquelles ils sont arrivés seuls.

Une troisième amélioration consisterait à avancer les «partiels» ou des épreuves équivalentes (vers novembre par exemple), rendant le diagnostic plus précoce et le temps disponible pour remédier plus important.

Une quatrième amélioration consisterait à mener une telle opération **pour tous les cours** du premier semestre, et pour ceux du deuxième semestre (en mars par exemple), donc pour tous les cours de la première candi. On constate, en effet depuis plusieurs années que les cours ayant fait l'objet d'un partiel ont en moyenne un taux de réussite final supérieur aux cours sans partiel.

7.2 Le phénomène de résilience : on peut faire mentir les prédictions

Born et al. (1997, 684) définissent la **résilience** comme «une situation où un individu, bien qu'ayant été exposé à divers facteurs de risques, ne développe pas le comportement craint». Ici, nous traduirons «une situation où un étudiant, auquel les variables prédictives donnent peu de chances de succès y arrive quand même».

Cette capacité de surmonter le sort adverse (ou plutôt de faire mentir une prédiction défavorable), concept utilisé dans le domaine médical et celui de la délinquance, est l'œuvre conjuguée de l'étudiant, de ses enseignants et de l'institution qui parviennent, en conjuguant leurs efforts, malgré un mauvais **pronostic** de départ, à **faire mentir ce pronostic**. Dans un tel esprit, où l'on œuvre à la réussite de l'étudiant et non à son échec, c'est une force et non un handicap de connaître le diagnostic. Ceci va à l'encontre des conclusions que l'on tire généralement de la lecture de Rosenthal et Jacobson (1971) «il ne faut pas communiquer les prédictions aux enseignants, par crainte de déclencher le phénomène de réalisation automatique des prédictions ou Effet oedipien des prédictions». Au contraire, Boxus (1971) avait déjà démontré que faire mentir la prédiction de rende-

ment était possible, dans l'opération PREDIC entre l'école maternelle et l'école primaire. Pour cela, il faut alerter les acteurs de ce qui va arriver le plus probablement « si on ne change rien ». Les enseignants et l'apprenant se trouvent dans la situation d'un médecin qui, s'il n'intervenait pas, trouverait des corrélations quasi parfaites entre le fait d'être infecté et le fait de décéder. Mais voilà, les médecins soignent, faisant (heureusement) se dégonfler les mortelles corrélations : désormais on survit, même si on a été infecté... à condition d'intervenir à temps, en suffisance et de façon adéquate !!!

8. Bibliographie

- Born, M., Chevalier, V. et Humblet, I. (1997). Resilience, desistance and delinquent career of adolescent offenders. *Journal of Adolescence*, 20, 679-694.
- Boxus, E. (1971). Une méthode de prédiction du rendement en lecture en première année. *Education -Tribune Libre*. 127, 63-78.
- Rosenthal, R. & Jacobson, L. (1971). *Pygmalion à l'école*, Paris : Casterman.

Partie II : L'analyse check-up par check-up

Chapitre 9

Analyses psychométriques des questions des 10 check-up MOHICAN

Vue d'ensemble

*D. Leclercq, Président du Groupe de travail CIUF « Réussite en candidatures »
J.-L. Gilles, Directeur du SMART ULg*

1. Les indices de popularité des différentes solutions

Ces indices sont à la fois psychométriques et éducatifs. Ainsi, la popularité de la solution correcte à une QCM ou pourcentage de répondants ayant choisi cette solution, ou encore pourcentage de Réponses Correctes (%RC) est un indicateur psychométrique : plus ce taux de réponse correcte est proche de 50%, plus cette question contribue à discriminer les répondants entre eux. Il peut aussi s'interpréter éducativement en termes de distance à la perfection. Ainsi, un taux de réponses correctes de 80% indique une « non maîtrise » concernant 20% des étudiants et donne une idée de la taille du public concerné par l'utilité éventuelle d'une remédiation.

1.1 Les popularités des solutions correctes ou Facilité Objective de la question : Quel est le niveau objectif de réussite et d'erreur ?

Cette **popularité** de la solution correcte, ou taux de réussite, est l'indice de la **Facilité Objective de la question**, qui va de 0% à 100%. Dans la littérature, on trouvera souvent l'expression « p » ou « indice de difficulté », expression malheureuse puisque quand cet indice numérique est maximal (100%), la facilité est maximale mais la difficulté, elle, est minimale. C'est sur une échelle verticale de facilité qu'ont été positionnées les questions dans le chapitre 3. Ils sont donc ainsi situés (éducativement) en valeur absolue, c'est-à-dire par leur distance à la perfection (100%), et relativement les uns aux autres (le plus facile, le plus difficile, etc.).

1.2 Les popularités des distracteurs : quelles erreurs sont les plus fréquentes ?

Un distracteur est une solution incorrecte proposée parmi les choix de réponses, dont les solutions générales³² Aucune et Toutes. Nous appelons « **distracteur-vedette** » le plus populaire, le plus choisi, et son « challenger », celui qui est choisi en deuxième lieu, ou plutôt en deuxième proportion. Ces popularités renseignent sur l'ampleur des erreurs et surtout sur leur nature, voire leur cause. La **valeur diagnostique** d'un distracteur est due au talent du créateur de la question. Une méthode efficace de génération de distracteurs consiste à utiliser les réponses erronées données antérieurement par des étudiants lors de l'administration de la question de façon ouverte, c'est-à-dire demandant à l'étudiant de produire une réponse et non de la choisir. C'est de cette façon que les distracteurs de plusieurs des check-up de MOHICAN ont été générés (pour plus de détails, voir les auteurs de chaque check-up).

³² Ces solutions sont appelées « générales » parce qu'elles sont d'application à toutes les questions d'une épreuve. Elles seraient en outre appelées « implicites » si elles n'apparaissaient **pas** (dactylographiées) parmi les solutions de chaque question. Comme la grande majorité des étudiants n'avaient quasiment aucune expérience de cette technique, nous avons décidé de rendre explicites ces deux solutions, ce qui en diminue la valeur formative qui vise à entraîner la vigilance cognitive (Leclercq, 1986, 127-143).

2. Les indices (psychométriques) de discrimination de chaque solution de chaque QCM

2.1 Les erreurs et les réponses correctes viennent-elles d'étudiants de compétences différentes ?

Très souvent, dans une situation de testing, et c'est le cas pour MOHICAN, la meilleure information dont on dispose sur la compétence de chaque étudiant dans la matière est son score au total à l'épreuve. Chaque fois que l'on utilisera ci-après l'expression « étudiant fort » ou « étudiant faible », il faudra comprendre « dans la matière mesurée par le score total à l'épreuve dont la question fait partie ». On peut déterminer si le choix d'une solution (correcte ou distracteur) appartenant à une QCM est lié au score total à l'épreuve par le calcul d'une corrélation, la **corrélation point bisériale** (en raccourci **r. bis**). Cet indice prend des valeurs comprises entre -1 et +1.

Est-ce un groupe d'**étudiants « forts »** (au total de l'épreuve) qui ont choisi une solution ? Si oui, la corrélation est positive car le score moyen au total de l'épreuve de ceux qui ont choisi cette solution est supérieur au score moyen de ceux qui ne l'ont pas choisie. C'est la situation attendue pour la solution correcte d'une QCM.

Si ce sont des **étudiants « faibles »** (au total de l'épreuve) qui ont choisi cette solution, la corrélation (**r.bis**) est négative. C'est la situation à laquelle on s'attend pour les solutions incorrectes.

Quelle valeur du **rpbis** de la réponse correcte est satisfaisante ? Cela dépend d'une épreuve à l'autre, comme on va le voir.

2.2 La valeur-seuil ou valeur repère du **r.bis** : La question contribue-t-elle mieux que le hasard au score total à l'épreuve ?

On doit s'attendre à ce que la corrélation entre le choix d'une solution correcte à une question et le score total à toute l'épreuve soit positive. C'est une conséquence du fait que chacune des **NQ** questions de l'épreuve contribue, pour un poids de $1/NQ$ questions à ce score total. Autrement dit, chaque score (0/1) à une question est une partie du score total à l'épreuve. On parle aussi de « **recouvrement** » entre le score à une question et le score au total de l'épreuve, ou d'**inclusion** du premier dans le second. On peut calculer la corrélation « repère » qui découle de ce raisonnement par la formule $1/\sqrt{NQ}$ (Guilford & Fruchter, 1978). Cette corrélation « automatique » ou « repère » positive doit être atteinte et si possible dépassée par la valeur du **r.bis** de la réponse correcte pour considérer que la question « va dans le même sens », « mesure la même chose » que l'ensemble du test. Cet indice est d'ailleurs appelé **indice de cohérence interne**.

Dans une épreuve de 25 questions, la corrélation repère vaut $1/\sqrt{NQ}$ soit $1/\sqrt{25}$ soit $1/5$ soit 0,20. Avec 45 questions, la corrélation automatique ou repère pour la réponse correcte vaut $1/\sqrt{45}=1/6,7=0,15$. En deçà de cette valeur repère, le **r.bis** indique que la question pose un problème. Le **r.bis** n'indique pas la cause du problème. La raison peut en effet être une mauvaise formulation de la question, ou le fait qu'une question (de mathématique par exemple) n'est pas du même type que les autres (dans une épreuve de géographie par exemple) ou même que des erreurs ou des ambiguïtés ont été véhiculées par le cours oral ou par le support écrit.

2.3 Les valeurs négatives des r.bis des distracteurs : sont-ils choisis par des étudiants plus faibles que la moyenne générale ?

On s'attend à ce que le choix d'un **distracteur** ait, lui, une **corrélation négative** avec le score total parce que choisi par des étudiants dont le score total moyen à l'épreuve est inférieur au score total moyen des autres étudiants (dont les étudiants qui ont réussi).

Les valeurs des rpbis des distracteurs peuvent être comparées entre elles, ce qui permet de voir que certains distracteurs sont choisis par des élèves plus faibles (en moyenne) au total de l'épreuve que d'autres distracteurs. C'est souvent l'indication d'une différence de profondeur ou de **gravité des erreurs**.

L'**omission** est, elle, souvent le fait d'étudiants « intermédiaires » : moins forts que ceux qui ont choisi et moins faibles que ceux qui se sont trompés.

Les indices de discrimination r.bis, empruntés à la psychométrie, sont utilisés depuis des décennies par les gestionnaires d'épreuves pédagogiques, notamment ceux qui doivent garantir les propriétés métriques des tests et des examens. Quand ils ne prennent pas les valeurs attendues (voir ci-avant), ils servent de signal d'alarme qui invite le constructeur de l'épreuve à se demander pourquoi le signal s'est déclenché. Voici la vue d'ensemble des situations es 10 check-up à ce sujet.

	r.bis de la Réponse Correcte			r.bis des Rép. Incorrectes ou distracteurs		
	N	positif et non supérieur au seuil	positif ET supérieur au seuil	N	non négatifs	négatifs
Vocabulaire	45	2	43.	206	23	183
Syntaxe	12	0	12	72	0	72
Compr.Texte	6	1	5	36	1	35
Géo	10	1	9	56	3	53
Math	22	0	22	132	4	128
Physique	10	0	10	60	1	59
Chimie	8	1	7	48	2	46
Bio	10	3	7	60	3	57
Histoire	25	2	23	150	14	136
Art	25	2	23	150	25	125
	173			970		

Ce tableau récapitulatif montre qu'à quelques exceptions près, discutées en détails dans l'analyse de chaque épreuve, ces indices psychométriques (r.bis) sont satisfaisants.

3. Un indice psychométrique de la fidélité des tests : l’alpha de Cronbach

Cet indice mesure la cohérence interne des questions par rapport au score total au test, bref dans quelle mesure chaque question mesure la même chose que les autres.

La formule est la suivante : $\alpha = (NQ / (NQ - 1)) \cdot (1 - (\sum \sigma_q^2 / \sigma^2))$

Voici les valeurs des alpha de Cronbach pour chacun des check-up MOHICAN (extraite de Gilles, 2002, 308) :

	Vocabulaire	Syntaxe Compréh.	Compréh. Textes	Compréh. Gééo	Math	Physique	Chimie	Biologie	Hist-Act-Eco	Conn. Art
Alpha de Cronbach	0,833	0,572	0,393	0,530	0,769	0,473	0,414	0,410	0,668	0,707
Nbre de Questions	45	12	6	10	22	10	8	10	25	25
Allongement Nécessaire pour F = 0,80	-9	24	31	25	4	35	37	48	25	16

Pour le check-up **Vocabulaire**, on pourrait se permettre de supprimer 9 questions, et la fidélité de 0,8 ne serait pas remise en cause.

Pour les autres check-up, on voit que le nombre de questions devrait être largement augmenté, sauf pour le check-up de Math qui avec 26 questions (au lieu de 22) permettrait d’atteindre une fidélité de classement des étudiants entre eux (sur base de leur score total) avec une fidélité de 0,80.

Rappelons, cependant, que les check-up n’avaient pas pour ambition d’opérer un tel classement.

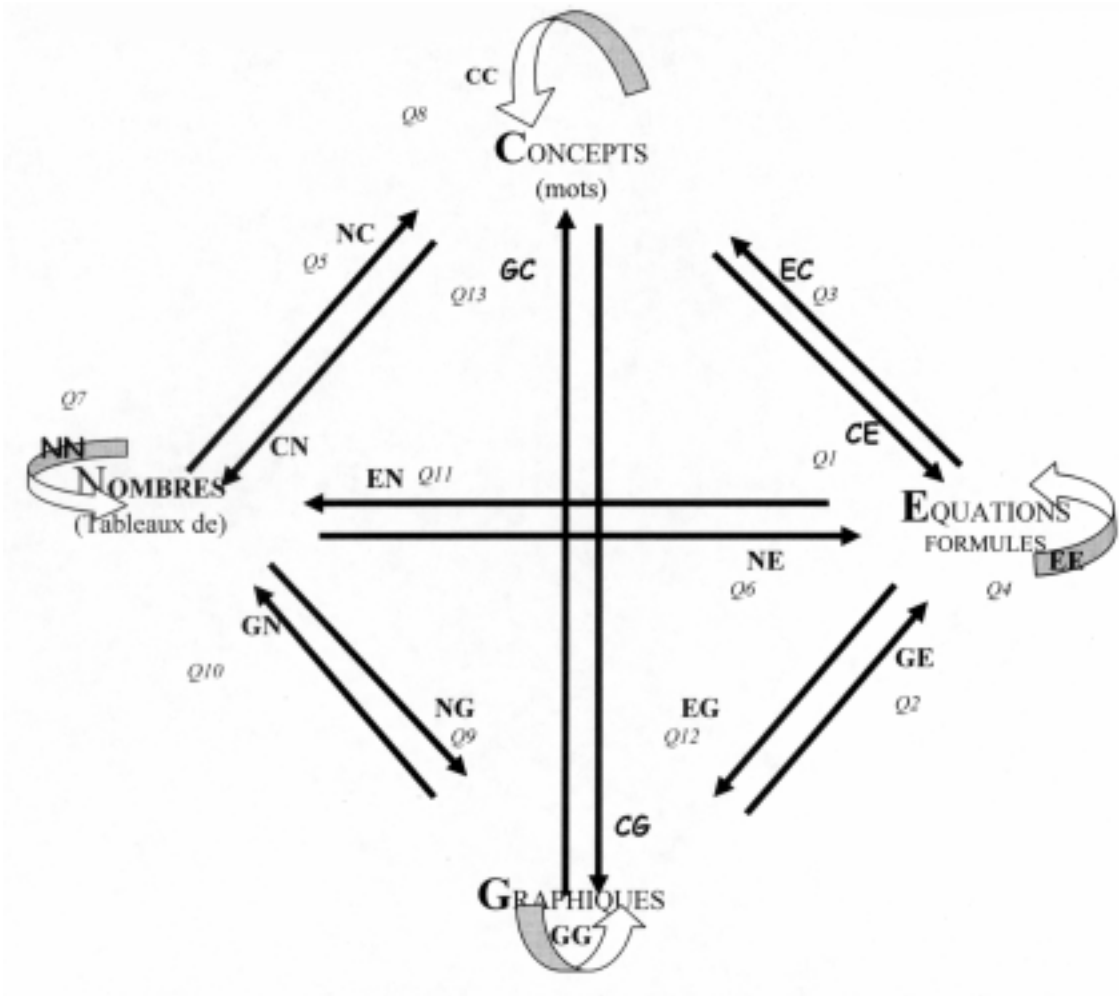
La préoccupation première était celle de la validité (ou pertinence de la mesure en fonction des objectifs de mesure) et non de la fidélité de classement.

Les critères de validité ont varié selon les check-up. Ainsi, pour le check-up **Vocabulaire**, la validité des questions présentes repose sur les observations systématiques faites par l’auteur (M. Monballin) des vocables non techniques de la langue française faisant difficulté pour les étudiants de 1^o année universitaire (sur base d’interviews de professeurs).

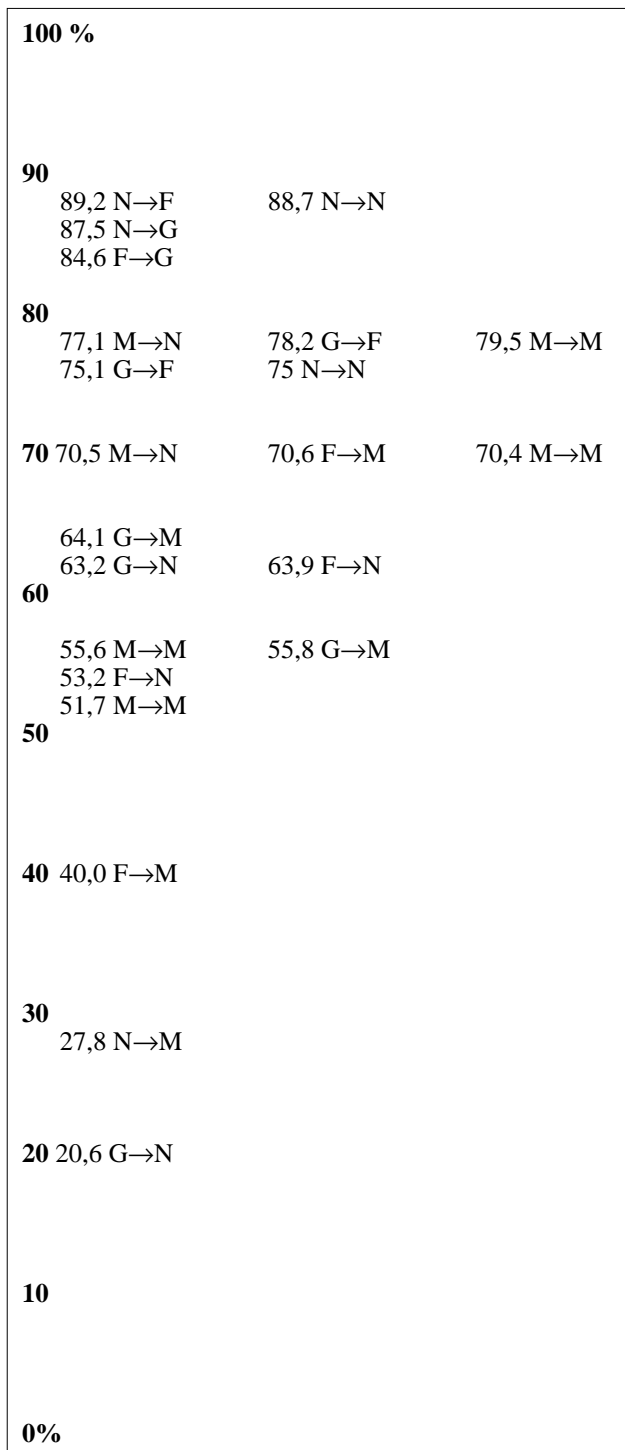
Pour les check-up de Syntaxe, Compréhension de Textes, Géographie et Lecture de cartes et graphiques, Physique, Chimie, biologie, Art et Histoire-Economie-Actualité, c’est la multiplicité (l’hétérogénéité) des domaines touchés qui a été le critère de choix, principe antagoniste avec le concept de fidélité.

Le check-up de **mathématique**, lui, porte sur l'application de principes, de connaissances, mais surtout sur la compréhension, dans le sens « traduction d'un langage dans l'autre » car ses questions peuvent être représentées comme dans un carrefour représentant les divers « tournants » possibles (du concept exprimé en mots au graphique, et vice versa, etc.).

Un tel « carrefour » comporte 16 « chemins », dont la plupart (voir les numéros correspondants des questions) ont été reflétés dans le « Check-up » Mathématique.



C'est cette signification, conçue dès la création des questions (ou plus exactement la sélection dans une banque de questions plus vaste des auteurs) **qui en fait la validité diagnostique** qui apparaît clairement au chapitre 3 page 14 (voir sa copie ci-contre) où il apparaît que ce sont les questions de traduction vers les MOTS qui sont les moins bien réussies !!!



Le lecteur jugera de l'intérêt diagnostique à la fois de la structuration en « carrefour des traductions » et de la présentation sur une échelle unique des questions par ordre de Facilité objective (FO) décroissante. On imagine que les 22 questions pourraient être situées dans un espace à deux dimensions, l'autre dimension étant celle des Facilités Subjectives, ou, mieux, des facilités subjectives des réponses Correctes, bref des Indices de Confiance, comme au chapitre 2, en section B.

4. Le Réalisme par Calibration par question ?

Il est tout à fait pertinent de calculer un indice de Réalisme par Calibration par question ; c'est ce qu'a fait Gilles (2002), auquel nous renvoyons.

5. Bibliographie

- Cronbah, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16
- Gilles, J.-L. (2002). Qualité spectrale des tests standardisés universitaires, Thèse de doctorat en Sciences de l'Education, Université de Liège.
- Guilford, J.-P. & Fruchter (1978). *Fundamental Statistics in Psychology and Education*. New York : Mac Graw Hill.

Chapitre 10

Analyses éduométriques et indices métacognitifs appliqués aux questions des 10 check-up MOHICAN

Vue d'ensemble

*D. Leclercq, Président du Groupe de Travail CIUF « Réussite en candidatures »
M. Poumay, Directrice du LabSET – ULg*

1. Distinctions entre psychométrie et éduométrie

Nous adoptons ci-après les définitions devenues célèbres de Ronald Carver (1974) sur les « dimensions psychométrique et éduométrique » des mesures fournies par des épreuves telles que des tests, des examens, des check-up, etc. Nous ne nous attacherons ici qu'à certains aspects de ces deux dimensions.

1.1 Les objets des mesures

La raison d'être des **indices psychométriques** est exposée par Cronbach (1971, 446) : « *le mouvement de testing en psychologie a démarré avec l'accent darwinien mis sur les **différences entre individus** et tout le travail théorique mené en arrière fond des scores aux tests a tenté de conceptualiser les différences en termes de **capacités** ou traits* »³³. Carver (1974, 512) ajoute, et cette précision est d'importance pour bien comprendre la différence avec l'approche éduométrique, qu'évaluer les propriétés psychométriques d'un test consiste à apprécier « *dans quelle mesure il reflète les différences inter-individuelles **stables*** »³⁴.

Les **indices éduométriques**, eux, sont conçus pour mesurer le mieux possible des **progrès intra-individuels**, comme le dit Mac Clelland (1973, 8) : « *l'école devrait tester des **compétences** et non des capacités... pour refléter des **changements** dans ce que les individus ont appris* »³⁵.

1.2 La validité des mesures

La validité **psychométrique** de mesures obtenues par un test peut être appréciée par la comparaison des différences interindividuelles mises en évidence par ce test avec des différences issues d'autres critères réputés mesurer valablement le « trait » considéré.

La validité **éduométrique** de mesures obtenues par un test peut être appréciée par leur sensibilité aux gains ou progrès intraindividu (ou inter-groupes) dans des situations où un gain évident est observé par des critères réputés le refléter valablement. Cela implique que l'épreuve (ou des épreuves parallèles) soi(en)t administrée(s) au moins deux fois aux mêmes sujets.

1.3 La fidélité des mesures

En **psychométrie**, la fidélité est synonyme de cohérence (*consistency*) ou de stabilité, c-à-d la capacité pour les mesures de discriminer entre les individus de la même façon d'une fois à l'autre. Dans une telle perspective, « *la variance des scores est très liée à cette cohérence, car plus la variance est petite plus les discriminations risquent de changer d'une occasion à l'autre* » (Carver, 1974, 514).

³³ (nous soulignons).

³⁴ (nous soulignons).

³⁵ (nous soulignons).

En **édumétrie**, la fidélité est aussi concernée par la cohérence ou stabilité, non pas de discrimination entre les personnes, mais de gain ou de croissance intraindividuelle. Une bonne façon de mesurer une telle cohérence serait d'appliquer (mais cela ne se fait quasiment jamais) tant au prétest qu'au post-test des formes parallèles du test et de comparer les gains chez une même personne. Cette concordance n'est pas dépendante des variances dans les scores des étudiants aux tests. Des effets planchers ou des effets plafonds, fréquents en éducation peuvent détériorer la fidélité des mesures édumétriques. Par contre, si tous les étudiants ont le même score à un test, sa fidélité psychométrique est nulle, alors que cela n'a pas d'impact sur sa fidélité édumétrique.

Ces deux types d'indices mesurent donc des choses différentes : « *Un test peut être très fidèle d'un point de vue édumétrique et être en même temps complètement non fidèle du point de vue psychométrique* ». (Carver, 1974, 513). Ceci mérite quelques explications.

1.4 Les questions les plus appropriées

D'un point de vue psychométrique, la question la plus efficace est celle dont le taux de réussite est de 50%, parce que c'est pour cette valeur que la variance est maximale dans les scores, et les questions auxquelles tous les étudiants réussissent et celles auxquelles tous les étudiants échouent sont éliminées ou revues » (Carver, 1974, 513). Depuis Davis (1946 et 1952), on va plus loin : dans une perspective psychométrique, on élimine ou révise les questions dont le taux de réussite est inférieur à 20% ou supérieur à 80%.

« ***D'un point de vue édumétrique***, la question la plus efficace est celle dont la sensibilité au gain est la plus grande... entre un pré et un post-test... avec un taux de réussite proche de 0% au prétest et de 100% au post-test... » (Carver, 1974, 513).

On reconnaît là la base conceptuelle des indices MUCER (*Maîtrise, Utilité, Conservation, Efficience, Rentabilité*) de De Ketele (1980) ou les indices édumétriques de Tourneur (1974)³⁶, ou encore l'indice de Gain Relatif de Mc Guigan (1967)³⁷.

Cependant, cette déclaration de Carver sur la question la plus efficace doit être nuancée sur plusieurs points.

1.4.1 Nuance 1

Des questions qui ont un taux de réussite de 100% au prétest et au post-test (effet plafond) méritent de figurer dans un test édumétrique. C'est le cas d'examens périodiques de vérification de compétences, pour les pilotes d'avion par exemple, qui doivent « passer sur simulateur » (Weiner et al., 1993) tous les deux ans, principe que l'on envisage de plus en plus pour les anesthésistes :

« *L'utilisation d'un entraînement par simulation au Management de Ressources en situation de Crise (MRC) est exigé par la loi pour les pilotes d'avion aux USA et dans beaucoup d'autres pays et est continué annuellement tout au long de la carrière sans considération de l'ancienneté (seniority)... il devrait en aller de même en anesthésiologie* » (Gaba et al., 2001, 188 & 189).

³⁶ décrits par Leclercq et al., 1977, 180-182

³⁷ décrit par Leclercq et al. 1977, 165-166

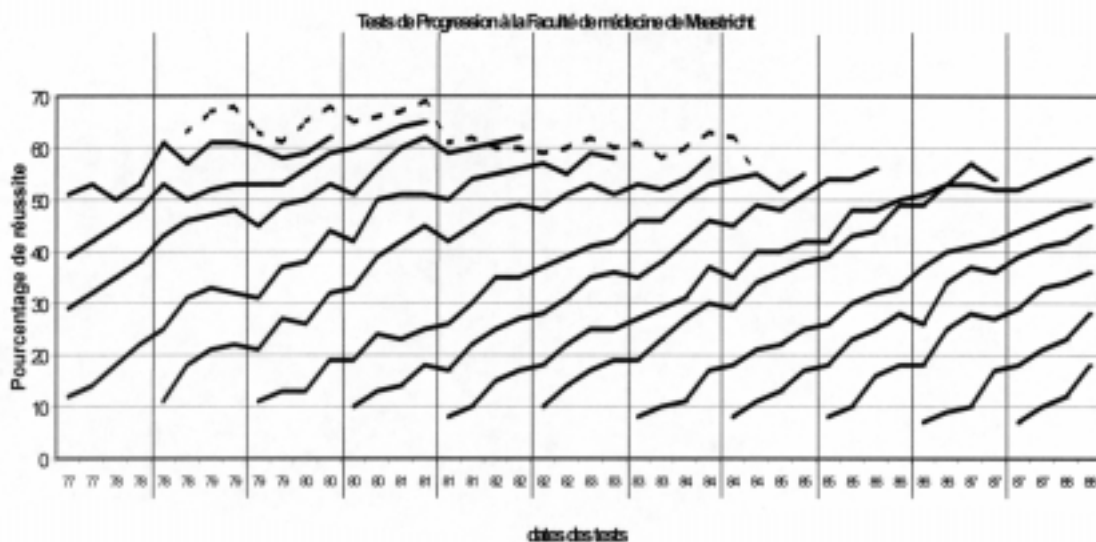
1.4.2 Nuance 2

Des questions qui auraient 0% au prétest et 0% au post-test (effet plancher) peuvent avoir leur place dans un test pédagogique. C'est le cas de plusieurs des 250 questions d'un des 4 tests de progression (*Progress Tests*) annuels obligatoires pour les étudiants en médecine de l'université de Maastricht (Leclercq et Vandervleuten, 1998, 199-201). Ces tests périodiques parallèles (les questions sont équivalentes mais non identiques d'un test à l'autre) portent sur TOUTE la médecine générale et sont passés le même jour par tous les étudiants, de la première à la dernière année. Certaines questions sont réussies à 0% lors du premier test par les étudiants de première année et le sont aussi à 0% lors du deuxième test par ces mêmes étudiants.

Leur pertinence éduométrique n'en est pas moins grande car

« On peut développer des tests éduométriques en présentant les questions à un groupe d'étudiants n'ayant jamais reçu d'enseignement sur un sujet et à un autre groupe d'étudiants ayant, eux, été formés sur le sujet. Les questions montrant une faible différence de taux de réussite devraient être éliminées » (Carver, 1974, 513).

C'est en effet une caractéristique des questions des *Progress Tests* d'être diversement réussies par des étudiants d'années différentes comme le montrent le graphique ci-dessous où chacune des lignes brisées représente les (24) moyennes successives (4 pendant 6 ans) d'une même cohorte d'étudiants. Selon les contenus, ces différences de réussite se produisent de façon continue ou, le plus souvent, par sauts.



Dans l'opération MOHICAN, nous n'avons adopté ni une étude longitudinale (qui aurait testé les mêmes étudiants à plusieurs reprises comme le font les *Progress Tests*) ni une étude transversale (qui aurait testé des groupes d'étudiants à divers niveaux scolaires, comme c'est le cas pour les *Progress tests*). Cela n'empêche pas d'avoir un « regard éduométrique » sur les données de chaque question et de chaque étudiant, mais à condition de recourir à de nouveaux indices éduométriques.

C'est ce que nous avons proposé via des concepts tirant parti des degrés de certitude accompagnant les réponses.

2. Indices métacognitifs de Confiance, Prudence et Discriminance

2.1 Les indices de confiance et de prudence : les étudiants sont-ils réalistes ?

Ces indices sont présentés pour chaque question dans les chapitres 11 à 20 et, au chapitre 3, pour l'ensemble de chaque épreuve. Ces indices constituent une innovation et c'est dans la présente recherche MOHICAN qu'ils sont utilisés à large échelle pour la première fois. Leur fondement théorique a été exposé au chapitre 2.

Nous parlons de **Prudence** lorsque une (ou des) réponse (s) incorrecte (s) est (sont) fournie (s) avec une certitude inférieure à 50%, ce que nous noterons $c < 50\%$ pour une réponse et plus souvent $C < 50\%$ quand il s'agit de la moyenne des certitudes de plusieurs réponses.

Nous parlons de **Confiance** lorsque la réponse correcte est fournie avec une certitude supérieure à 50%, que cela soit par une personne ou par la moyenne des personnes d'un groupe.

Nous parlons de **Discriminance** (mot nouveau) pour désigner la différence entre la Certitude Moyenne avec RC et la Certitude Moyenne avec RI.

Contrairement aux indices de discrimination, la première fonction des indices de Confiance et de Prudence n'est pas, même s'ils peuvent y contribuer, de renseigner sur les qualités d'une question et de ses distracteurs. Ils servent principalement à détecter, pour une population d'étudiants donnée, quelle est l'intensité subjective des réponses correctes et des erreurs et, dans les deux cas, d'attirer l'attention sur les mesures pédagogiques qu'il convient d'envisager (remédiation autonome ou assistée, approfondissement, etc.). D'où les deux **alarmes** de l'**imprudence** (signalée par une épine *) et du **manque de confiance** (signalée par un point noir ●).

2.2 Les alarmes d'imprudence et de manque de confiance : quelle suite leur donner ?

La définition des concepts de confiance et de prudence repose sur la distinction de 4 catégories de réponses, qui amène à inférer 4 types de connaissance (de gauche à droite) :

	Réponses Incorrectes		OM	Réponses Correctes	
	C > 50%	C < 50%		C < 50%	C > 50%
Statut cognitif	Méprise *	Mé-connaissance		Mi-connaissance ●	Maîtrise
Statut métacognitif	Imprudence	Prudence		Manque de confiance	Confiance
Synonyme (pour l'action) de connaissance...	...nuisible	...inutilisable ou confusion ou ignorance avouée.		... améliorable ou perfectible assurée
Définition	Erreur avec certitude supérieure à 50%. En anglais "fallacy" ou "misconception"	Erreur avec certitude inférieure à 50%. En anglais "dubiousness »		Réponse correcte avec certitude inférieure à 50%. En anglais "unassured knowledge"	Réponse correcte avec certitude supérieure à 50%. En anglais "Sure knowledge"
Implications sociales	Cette connaissance pourrait être utilisée sans vérification car "sûre" dans l'esprit de l'étudiant causant d'éventuels dangers sociaux	Cette connaissance ne sera probablement pas utilisée sans vérification ou précaution		Cette connaissance ne sera probablement pas utilisée sans vérification ou précaution	Cette connaissance sera très probablement utilisée sans vérification, ce qui constitue un gain de temps justifié.
Action à entreprendre par l'apprenant	Retravailler en profondeur et avec aide, voire désapprendre.	Retravailler (éventuellement en auto-apprentissage)		Creuser, faire des liens inter-matières, confrontation avec pairs	Rien.

2.3 Les quatre types de situations quant aux alarmes : leur répartition dans les différents check-up et leur interprétation à l'aide de la Taxonomie de Bloom des objectifs cognitifs

Dans la première colonne figurent les catégories de la taxonomie de Bloom prépondérantes dans l'épreuve : CN (Connaissance), CP (Compréhension) et AP (Application).

Catég. de Bloom	Matières	Total des questions	Manque de Confiance ● et Imprudence *	Confiance mais imprudence(s) *	Manque de confiance ● mais prudence	Confiance et Prudence
CN	Histoire-Economie-Actualité	25		1, 9, 11, 15 20, 23	6, 7, 16, 21 22, 25	2, 3, 4, 5, 8 10, 12, 13, 14, 17 18, 19, 24
CN	Connaissances artistiques	25		1, 3, 4, 9 11, 13	2, 7, 14, 17 18, 24	5, 6, 8, 10, 12 15, 16, 19, 20, 21 22, 23, 25
CN + CP	Vocabulaire	45	5	2,4,6,8,10,11,12, 14,15,16,19,30, 33, 35,37,39,41,42	3,7,9,17, 27,38,43, 44,45	1,13,18,20,21, 22,23,24,25,26, 28,29,31,32,34 36,40
CP	Syntaxe	12		1,2,3,4,5,6,7, 8,9,10,11,12		
CP	Compréhension de texte	6		1,2,4,5	3	6
CP	Géo et Lecture de cartes, graphiques	10		2, 3, 4, 5 9, 10	1, 7	6, 8
AP	Mathématique	22		1, 2, 3, 4, 5, 6 7, 8, 9, 10, 11, 12 13, 14, 15, 16, 17 18, 19, 20, 21, 22		
AP	Physique	10		2, 3, 4, 6, 7 9, 10	5	1, 8
AP	Chimie	8		1, 2, 3		4, 5, 6, 7, 8
AP	Biologie	10		2, 4, 6, 7, 8 9, 10	5	1, 3

Dans les deux épreuves de **connaissances (CN) de mémoire** de faits précis (*Connaissances en Arts, Histoire-Economie-Actualité*), la configuration « Confiance ET Prudence » s'observe pour la moitié des questions environ (13 sur 25) et les manques de confiance (6) et les imprudences sont minoritaires (6). Grosso modo, quand on sait, on sait qu'on le sait et quand on ignore, on en est conscient aussi.

L'épreuve de *Vocabulaire* combine la connaissance (CN) et la compréhension (CP) de concepts, acquis « **par immersion** », où le **sentiment de familiarité** de certains mots (déjà entendus ou déjà lus ou contenant un radical connu). Les indices d'auto-évaluation (confiance et prudence) correspondent à une configuration proche des deux épreuves de connaissances « pures », mais avec beaucoup plus d'imprudences (18) autant que de situations parfaites (17), alors que les manques de confiance sont deux fois moins nombreux (9) que les imprudences. Cela est probablement dû à un sentiment de connaissance partielle, l'impression (souvent trompeuse comme le montrent les résultats) d'être capable de fournir une interprétation avec une bonne approximation.

Les épreuves de **Compréhension-CP** (*Syntaxe, Compréhension de Texte et Géographie et compréhension de graphiques et de cartes*) totalisent 28 questions. Le sentiment (abusif) d'avoir compris semble encore plus présent puisque l'on observe massivement des situations d'imprudence (22 sur 28) et que les situations parfaites (3) et de manque de confiance (3) sont rares.

Les épreuves d'**application-AP** (processus qui, selon Bloom, requiert des connaissances et de la compréhension) que constituent les check-up de sciences (*Math, physique, chimie, biologie*), qui totalisent 50 questions sont caractérisées par l'absence quasi totale de manques de confiance (2), une minorité de situations parfaites combinant confiance et prudence (11, surtout dues à la chimie) et une écrasante majorité de situations d'imprudence (37), comme si les logiques (règles) donnaient des sentiments abusifs de confiance.

2.4 Nuances et réserves méthodologiques

2.4.1 Une première approximation

Les commentaires qui précèdent sont à considérer comme une première approximation, car à l'intérieur de chacune de ces épreuves, il serait possible de distinguer des sous-groupes de questions et de déboucher sur des interprétations plus nuancées. C'est pour faciliter cette tâche que nous avons mentionné les numéros des questions. Ces hypothèses pourraient en outre être confrontées à des données d'expériences qui seraient conçues pour les tester.

Ainsi, il nous paraît évident que certaines questions (spécialement en sciences et en compréhension de textes et de graphiques) relèveraient de ce que Bloom appelle l'**analyse**. Détecter ces questions amènerait sans doute un supplément de subtilité dans l'interprétation. Les solutions Aucune et Toutes pourraient jouer un rôle particulier dans ce type de questions.

Par contre, aucune question ne fait de place à ce que Bloom a appelé **synthèse** (ou formulation-expression-création), vu le format à choix multiple des questions.

Enfin, ce que Bloom appelle l'**évaluation** (le jugement sur base de critères internes) nous semble elle testée (et à 173 reprises car à chaque question) par une de ses formes : le réalisme de l'auto-évaluation de la performance via les degrés de certitude.

2.4.2 Des étudiants non entraînés à l'auto-évaluation via des degrés de certitude

La très grande majorité des étudiants qui ont répondu aux épreuves MOHICAN n'avaient jamais utilisé de degrés de certitude. Leur performance en réalisme serait probablement améliorée, en calibration, en confiance et en prudence si on les y entraîne comme certaines recherches permettent de l'espérer (Leclercq et al., 1993, 129).

Dans une telle perspective, les statistiques MOHICAN apparaissent comme un prétest qui pourra servir de repère pour apprécier les effets de pratiques d'auto-évaluation.

2.4.3 Des repères différents selon les épreuves et même selon les questions seraient justifiés

Tout d'abord, le degré de satisfaction quant au taux ou au degré de confiance accompagnant la réussite peut légitimement varier selon les disciplines. Ainsi dans les questions de vie ou de mort, on exigera des taux d'exactitude très élevés. C'est le cas à la Force Aérienne où, dans la préparation (en 1971) des pilotes et des mécaniciens au futur avion (Mirage à l'époque), la formation ne prenait fin que quand l'excellence (le maximum ou Maîtrise totale) était atteinte, certitude comprise, soit 100 %.

Dans l'évaluation des compétences de patients diabétiques, c'est un seuil de 90 % qui est pris en considération (Brutomesso et al., 2003). Les niveaux d'exigences, pour la prudence et la confiance, pourraient faire partie de plans éducatifs définis par des experts de chaque domaine, avec vérification des performances atteintes par des personnes jugées très compétentes dans leur domaine. En effet, il ne faudrait pas exiger des étudiants des performances que leurs enseignants dans la matière ne pourraient atteindre eux-mêmes.

3. Comment relever le défi de l'éducation à la métacognition ?

On se trouve ici devant un important défi de l'éducation : comment augmenter la confiance dans les réponses correctes et la prudence dans les réponses incorrectes ?

Pour certaines matières, on imagine facilement des pistes de solutions. Ainsi, pour la signification de mots, Leclercq et al. (2002) ont fait la démonstration expérimentale suivante à leurs étudiants (plus de 200) de première année d'université sur des mots d'un des chapitres du livre de base d'un de leurs cours.

Une expérience visant la prise de conscience d'une situation fréquente d'Imprudence

Dans un **premier temps**, Leclercq et al. ont présenté 10 mots par étudiants (2 formes parallèles totalisaient donc 20 mots) en demandant :

« Si vous aviez à fournir un synonyme de ce mot en seriez-vous capable ? »

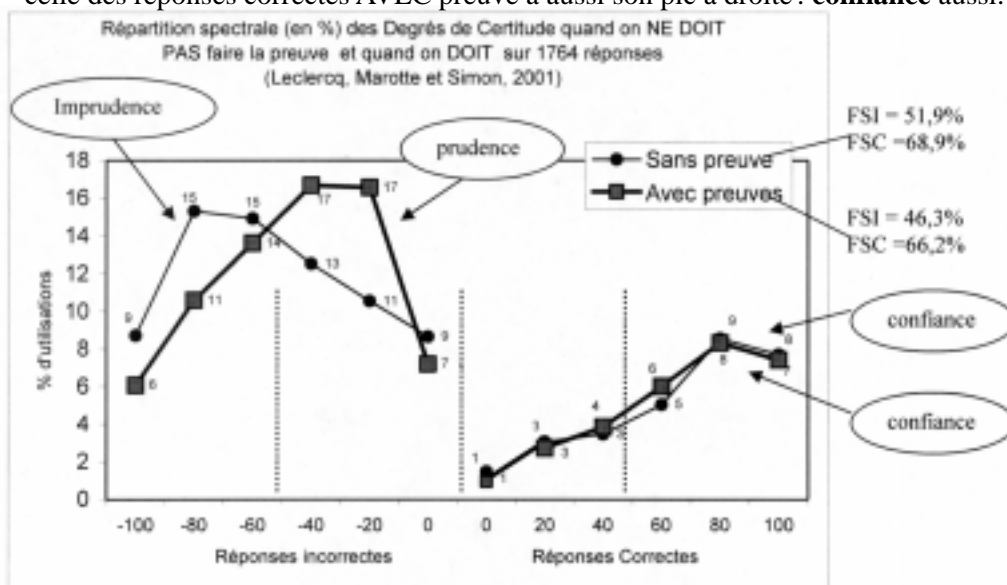
et ce **sans faire écrire ces synonymes**, mais en réclamant un degré de certitude (0%, 20%, 40%, 60%, 80% ou 100%) sur les chances de fournir une réponse correcte s'il fallait en fournir une.

Dans un **deuxième temps**, ils ont demandé **d'écrire les synonymes** (« de faire la preuve ») et de donner un degré de certitude. Le graphique qui suit (où ne figurent pas les omissions) montre que :

- quand, lors de la deuxième consigne, les réponses sont correctes, les premiers degrés de certitude pour ces mots étaient à peu près les mêmes (FSC = 68,9% et 66,2%).
- par contre, quand les réponses sont incorrectes, les degrés de certitude des premières réponses étaient largement plus élevés (51,9%) et imprudents en moyenne (supérieurs à 50%) que lors de la seconde fois (quand il a fallu faire la preuve : 46,3%), les degrés de certitude étaient fournis avec plus de prudence (inférieurs à 50%).³⁸

Les 4 courbes (ou plutôt lignes brisées) se lisent comme suit :

- la courbe des réponses incorrectes SANS devoir faire la preuve a un pic à gauche : **imprudence**.
- celle des réponses incorrectes AVEC preuve a son pic à droite : **prudence**.
- celle des réponses correctes SANS faire la preuve a son pic à droite : **confiance**
- celle des réponses correctes AVEC preuve a aussi son pic à droite : **confiance** aussi.



³⁸ En termes relatifs (voir chapitre 2 section C2), les pics des deux courbes de gauche montrent que l'imprudence quand on ne doit pas faire la preuve (SANS) s'est transformée en Prudence (AVEC) quand on doit faire la preuve.

Ces résultats ont été explicités en classe et débattus pour garantir leur compréhension en profondeur. Les auteurs espèrent que cette démonstration aux étudiants (sur LEUR comportement dans un cours qui pèsera sur LEUR réussite) aura un impact sur la façon de préparer les examens, bref d'étudier et, pourquoi pas, de lire en général, même quand il n'y a pas d'examen à la clé.

On pourrait attribuer ces résultats au fait que quand ils savent qu'ils ne doivent pas donner de preuve, les étudiants ont un comportement « stratégique » : donner une bonne image d'eux-mêmes (je pense que j'en serais capable) au professeur. Par contre, quand ils savent que la qualité de leurs réponses pourra être contrôlée par le professeur, la stratégie change (il vaut mieux paraître trop modeste que trop prétentieux).

Une troisième hypothèse (ci-après) est envisageable, et nous laissons à chaque étudiant le soin de peser entre les trois dans son cas particulier.

Cette troisième **hypothèse** est qu'en moyenne, quand ces étudiants lisent un texte sans devoir faire la preuve de sa compréhension, ils surestiment leur capacité d'en comprendre bon nombre de termes... et sont plutôt **imprudents** par non-recours au dictionnaire.

Pour d'autres matières que le vocabulaire, certaines techniques existent : la preuve par 9 dans les multiplications, par exemple en arithmétique, les « recoupements » historiques ou géographiques dans des questions culturelles.

Plus généralement, il importe de développer non seulement des techniques appropriées à cette augmentation de fiabilité, mais aussi, chez les étudiants, le souci de la confiance et de la prudence, d'où l'importance de disposer d'indices permettant de mesurer les progrès faits dans cet aspect de la métacognition.

4. Bibliographie

- Brutomesso, D., Gagnayre, R., Leclercq, D., Grazzolaro, D., Busata, E., d'Ivernois, J.-F., Casiglia, E., Tiengo, A. & Baritussio, A. (2003). Confidence degrees to evaluate knowledge in patients. Application to patients with type 1 diabetes, in Patient Education & Counseling, accepted nov. 2002.
- Carver, R.-P. (1974). Two dimensions of tests : Psychometric and Edumetric. *American Psychologist*, july 1974, 512-518.
- Cronbach, L.-J. (1971). Test validation, in R.L. Thorndike (Ed.). Educational measurement, Washington D.C. : American Council of Education.
- Davis F. B. (1966). *Analyse des items*,. Louvain : Ed. Nauwelaerts, traduction de la version anglaise de 1946 « Item Analysis ».
- Davis, F. B. (1952). Item analysis in relation to educational and psychological testing. *Psychological Bulletin*. vol. 49, 97-119.
- De Ketele, J.-M. (1980). *Observer pour éduquer*, Berne : Peter Lang.
- Fonzé, J. (2002). Evaluation de la nécessité d'une formation continue en réanimation du personnel de nursing d'une unité de soins intensifs néonataux, mémoire de licence en Sciences de la Santé Publique, Université de Liège.
- Gaba, D. M., Howard, S., Fish, K., Smith, D. & Sowb (2001). *Simulation-Based training in anesthesia crisis resource management (ACRM): A decade of experience, Simulation and Gaming*, vol. 32, n°2, June, 175-193.

- Guilford, J.-P. (1954). *Psychometric methods*. New York : Mac Graw Hill.
- Guilford J.-P. & Fruchter H. (1978). *Fundamental Statistics in Psychology and Education*. New York : Mac Graw Hill
- Leclercq, D. & Bruno, J. (1993), *Item Banking, Interactive Testing and Self assessment*, Heidelberg : Springer Verlag, NATO ASI series, F112.
- Leclercq, D., Simon F., Marotte, P., Verschueren, A. et Lacaille, C. (2002), *Former des étudiants de première candidature universitaire à des compétences transversales : lesquelles et comment ?*, Actes du 2^o Congrès des Chercheurs en Education, Louvain-La-Neuve.
- Mc Clelland, D.C. (1973). Testing for competence rather than for intelligence. *American Psychologist*, vol 28, 1-14.
- Weiner, E, Kanki, B. & Helmreich, R. (1993), *Cockpit Resource Management*, San Diego : Academic Press, cité par Gaba *et al.*, 2001.

Chapitre 11

Analyses des réponses au check-up de

Vocabulaire

M. Monballin, FUNDP

D. Leclercq, Ulg

N = 3905

rpbis repère = 0,15

Parmi les cinq mots proposés, choisissez celui qui reprend le plus exactement possible le sens que le mot en gras a dans la phrase donnée.

Question n° 1 : Ce nouvel élément **corrobore** les thèses

Surface de Compétence				
40,30%	%	r.bis	CM	
1. complète	16,1	-0,12	42,5	Comme on pouvait s'y attendre, l'erreur jugée la moins grave (1. Complète) est le distracteur-vedette, choisie par le plus grand pourcentage de sujets « en erreur » (16,1 %) et ces étudiants n'étaient pas les plus faibles au total du test (r.bis = -0,12). C'est le distracteur qui a la Certitude Moyenne la plus élevée (42,5 %) de tous ceux qui sont « dans l'erreur ». L'erreur la plus grave (4. Contredit) a attiré les moins bons étudiants (r.bis = -0,19) avec une certitude de 35,5 % (fracture subjective nette de 25 %). Confiance et prudence sont conformes aux attentes.
2. confirme	66,6	0,33	60,5	
3. nuance	2,2	-0,08	31,4	
4. contredit	9,7	-0,19	35,5	
5. détaille	2,1	-0,05	37,5	
6. aucune	0,4	-0,01	32	
7. toutes	0,2	-0,04	37,5	

Question n° 2 : Quand il fut nommé, il fit longuement valoir ses **prérogatives**.

Surface de Compétence				
16,30%	%	r.bis	CM	
1. droits	28,4	0,17	57,3	L'erreur la moins grave est (4. Revendications) parce qu'il contient la racine latine « rogare » (demander). Elle est d'ailleurs choisie par la majorité des étudiants (51,3 %) et les moins faibles (r.bis = 0,02) parmi les étudiants « en erreur », au score total, avec une certitude moyenne de 52,1 % (imprudence), presque égale à celle de la RC (57,3 %). La fracture subjective est très faible : 5 %. Cette réponse incorrecte constitue en moyenne une méprise (car sûre à plus de 50 %) pour ces étudiants. Confiance et prudence sont conformes aux attentes, sauf pour le distracteur-vedette.
2. relations	1,3	-0,06	40,8	
3. antécédants	8,4	-0,08	43,5	
4. revendications	51,3	0,02	52,1 *	
5. qualités	4,5	-0,13	40,8	
6. aucune	3,2	-0,01	34,2	
7. toutes	0,5	-0,03	36,8	

Question n° 3 : Le responsable des forces de l'ordre a **stigmatisé** l'attitude.

Surface de Compétence				
12,70%	%	r.bis	CM	
1. approuvé	13	-0,09	30,1	L'erreur jugée la moins grave (5. Commenté) a la même popularité (12,2 %) que deux autres (13 et 13,7 %). La forte proportion de (6. Aucune) (15,3 %) vient de synonymes imaginés par les étudiants (et absents de la QCM) tels que « éclairé », « mis en évidence », « désigné », « dénoncé ». La fracture subjective n'est que de 10 % (faible) avec le distracteur le plus grave (1. Approuvé). La réponse correcte n'est pas assurée, ce qui indique un manque de confiance . La prudence , elle, est conforme aux attentes.
2. exagéré	13,7	0,04	38,5	
3. condamné	31,8	0,17	40 *	
4. surveillé	7,4	-0,09	30,1	
5. commenté	12,2	-0,04	30,9	
6. aucune	15,3	0,04	30,6	
7. toutes	0,2	-0,02	17,1	

Question n° 4 : Il a présenté un projet sans en mentionner les difficultés **sous-jacentes**.

Surface de Compétence				
34,10%	%	r.bis	CM	
1. majeures	2,4	-0,16	38,5	L'erreur la plus populaire (27,2 % des choix) est (3. Annexes), ce qui est excusable sur le fond mais peut-être aussi explicable par la forme : dans un courrier, deux quasi synonymes de « Annexes » sont « ci-dessous » et « ci-jointes » qui toutes deux, sont proches, dans leur forme, de « sous-jacentes ». La fracture subjective n'est que de 8 % (très faible) avec ce distracteur-vedette (3. Annexes) qui est donné avec une certitude moyenne de 55,2 %, ce qui constitue une imprudence de la part de ces 27,2 % d'étudiants. La confiance est conforme aux attentes.
2. inévitables	5,7	-0,05	46,4	
3. annexes	27,2	-0,07	55,2 *	
4. cachées	54,1	0,26	63,1	
5. plausibles	4,8	-0,13	43,1	
6. aucune	3	-0,05	38,1	
7. toutes	0,8	-0,05	35,8	

³⁹ par les experts

Question n° 5 : Des haines **latentes** opposaient les deux hommes.

Surface de Compétence			
5,50%	%	r.bis	CM
1. profondes	19,7	-0,07	43,5
2. anciennes	52,2	0	51 *
3. farouches	2	-0,07	34,2
4. douloureuses	2,6	-0,1	37,2
5. patentes	6	0,12	51,8 *
6. aucune	11,9	0,15	46,1 *
7. toutes	2,3	0,02	50,2 *

* méprises (connaissances nuisibles) à désapprendre
 * mi-connaissance (connaissances douteuses) à améliorer

La réponse correcte (6. *Aucune*) a été très peu choisie (11,9 %), le r.bis est « limite » (0,15) et la facilité subjective, pour ces réponses correctes, faible (46,1 %), soit un **manque de confiance**. La **fracture subjective** est même inverse (-4,9 %), donc aberrante avec le distracteur-vedette (2. *Anciennes*) choisi par plus d'un étudiant sur deux (52,2 %) avec **imprudence** (certitude moyenne = 51%). La question est difficile, par son contenu et par sa forme (la solution correcte est *Aucune*).

Question n° 6 : Leurs interprétations **divergent** sur ce point.

Degré de Maîtrise			
58,50%	%	r.bis	CM
1. se renforcent	0,3	-0,06	50 *
2. s'embrouillent	0,8	-0,14	49,4
3. se contredisent	22	-0,18	74,6 *
4. s'écartent	74,1	0,27	78,9
5. sont fausses	0,2	-0,01	63,3 *
6. aucune	1,1	-0,04	62,7 *
7. toutes	0,1	-0,01	40

Trois étudiants sur quatre répondent correctement (74,1 %), avec **confiance** (la certitude moyenne est élevée : 78,9 %) et ce sont des étudiants plutôt forts au total du test (le r.bis est satisfaisant : 0,27). La **fracture subjective** est insignifiante (4,3 %) avec le distracteur-vedette (3. *Se contredisent*) choisi par un étudiant sur cinq (22 %), qui est donné avec une certitude très élevée (74,6 %), donc avec **imprudence**, comme trois autres distracteurs.

Question n° 7 : Personne n'a réagi à ses **dénégations**.

Surface de Compétence			
19,90%	%	r.bis	CM
1. mensonges	16,4	-0,06	39
2. interdictions	3,8	-0,13	30,3
3. démentis	39,9	0,28	49,9 *
4. aveux	15,4	-0,08	34
5. dénonciations	13,2	-0,06	35,4
6. aucune	5,7	-0,02	26,4
7. toutes	0,7	-0,06	26,9

Le faible taux de réussite (39,9 %) est accompagné d'une certitude moyenne faible elle aussi (49,9 %), à la limite du **manque de confiance**. Le distracteur (5. *Dénonciations*) a probablement attiré (13,2 %) par sa construction, alors que les distracteurs 1. *Mensonges* et 4. *Aveux* sont assez bien choisis alors qu'opposés sémantiquement. La **fracture subjective** est insignifiante (4,3 %) avec le distracteur-vedette (3. *Se contredisent*) choisi par un étudiant sur cinq (22 %). La **prudence** est conforme aux attentes.

Question n° 8 : Les événements du mois passé ont sans doute largement contribué au **discrédit** du gouvernement actuel.

Surface de Compétence			
35%	%	r.bis	CM
1. à la déconsidération	58,2	0,26	60,2
2. à la condamnation	4,9	-0,16	38
3. à l'impopularité	20,4	-0,02	57 *
4. à la dénonciation	2	-0,13	33
5. à la pénurie	1,6	-0,11	26,1
6. aucune	8,6	-0,05	45,9
7. toutes	1,6	-0,04	46,9

La première lettre de la solution correcte étant la même que celle du mot-question, le taux de réussite (58,2 %) est peut-être gonflé par la forme (QCM) de la question. La certitude moyenne qui l'accompagne (60,2 %), signe de **confiance**, n'est que de 3,2 % (**fracture subjective insignifiante**) supérieure au distracteur-vedette (3. *Impopularité*), qui est la conséquence du discrédit (et non son synonyme), et dont la certitude moyenne (57 %) constitue une **imprudence**.

La phrase incomplète qui suit chaque énoncé en reformule l'idée. Parmi les cinq fins de phrase proposées, choisissez celle qui conserve le plus exactement le sens que le mot souligné a dans l'énoncé de départ.

Question n° 9 : De la lecture du curriculum vitae de la candidate, le directeur **a inféré** qu'elle était apte à remplir les fonctions du poste à pourvoir.....le directeur ***** qu'elle était apte à remplir les fonctions du poste.

Surface de Compétence		%	r.bis	CM
16,40%				
1. a annoncé	6,6	-0,05	37,2	
2. a décidé	11,1	-0,04	41,3	
3. a admis	15,4	-0,06	42	
4. a conclu	34,6	0,18	47,3 *	
5. a fait savoir	20,6	-0,03	42,3	
6. aucune	2,3	0,03	29,2	
7. toutes	5,7	-0,04	48	

Un étudiant sur trois seulement (34,6 %) a réussi, avec une certitude faible (47,3 %) soit un **manque de confiance**. La **fracture subjective est très** faible (5 %) avec le distracteur-vedette (5.a. fait savoir) choisi par un étudiant sur cinq ou son challenger (3.a admis).
La **prudence** est partout conforme aux attentes.

Question n° 10 : Le conférencier a par ailleurs relevé certaines caractéristiques **inhérentes** à ce type de régime politique. Les caractéristiques relevées par le conférencier ***** ce type de régime.

Surface de Compétence		%	r.bis	CM
28,40%				
1. sont indispensables au bon fonctionnement de	4,7	-0,1	41,3	
2. sont à la source de	4,7	-0,06	42,3	
3. sont incompatibles avec	38	-0,38	50,4 *	
4. s'observent occasionnellement dans	2,7	-0,02	41,5	
5. sont inséparablement liées à	43,9	0,54	64,8	
6. aucune	2,8	-0,06	34,1	
7. toutes	0,2	-0,04	17,8	

Près d'un étudiant sur deux a bien répondu (43,9 %) avec une certitude satisfaisante (64,8 %), **confiante**, et, surtout, le r.bis (0,54) le plus élevé des 45 questions. La **fracture subjective** n'est hélas que modeste (14,4 %) avec le distracteur-vedette (3. Sont incompatibles avec), très populaire (38 %), dont la certitude moyenne est élevée (50,4 %), à la limite de l'**imprudence**, mais dont le r.bis est le plus négatif (-0,38) de tous les distracteurs de ces 45 questions. Heureusement, car ce distracteur est sémantiquement opposé à la réponse correcte.

Question n° 11 : Ces mesures accentueront les **clivages** sociaux. Suite à ces mesures, *** entre les classes sociales.

Surface de Compétence		%	r.bis	CM
26,90%				
1. les négociations vont s'intensifier	2,2	-0,15	28,7	
2. les conflits vont redoubler	15,7	-0,16	46	
3. les dérapages vont se multiplier	19,8	-0,11	50,9 *	
4. les échanges vont augmenter	4,6	-0,09	41,8	
5. les distances vont s'accroître	45,3	0,38	63,8	
6. aucune	8	0,01	41,4	
7. toutes	0,7	-0,05	42,8	

Près d'un étudiant sur deux a bien répondu (45,3 %), avec une certitude satisfaisante (63,8 %), donc avec **confiance**, et une **fracture subjective** faible de 12,9 % avec le distracteur-vedette (3. Les dérapages) dont la certitude moyenne (50,9 %) est juste au-dessus de la limite de l'**imprudence**. Ce distracteur, avec son challenger (2. Les conflits vont redoubler) sont des conséquences, des effets, de la réponse correcte.

Question n° 12 : Le valet de don Juan agit **à l'instar** de son maître, ce qui crée un décalage comique. Le valet de don Juan. *****

Surface de Compétence		%	r.bis	CM
11,30%				
1. désobéit effrontément à son maître	6,4	-0,14	45,9	
2. fait tout le contraire de ce que fait son maître	20,7	-0,08	55,1 *	
3. seconde et aide maladroitement son maître	9	-0,06	49,3	
4. cache à son maître tout ce qu'il fait	34,8	-0,01	66,9 *	
5. exagère ce que fait son maître	7,3	0,03	49,2	
6. aucune	18,6	0,3	60,9	
7. toutes	0,6	-0,05	48,8	

Suite à la confusion avec « à l'insu de » : le distracteur-vedette (4. Cache à son maître ...) est choisi deux fois plus (34,8 %) que la solution correcte (6. Aucune = 18,6 %) et avec une **fracture subjective inverse** = - 6 %, donc aberrante puisque la certitude moyenne est 66,9 %, ce qui en fait une **imprudence** importante. Le deuxième distracteur en popularité (2. Fait tout le contraire) a attiré un étudiant sur cinq avec une certitude moyenne assez élevée (55,1 %), **imprudente** elle aussi. La solution correcte, bien que choisie par peu d'étudiants (18,6%), l'a été avec **confiance** (60,9%)

Parmi les cinq mots proposés à la suite de chaque énoncé, choisissez celui qui exprime le plus exactement possible le **contraire** du mot qu'a le **mot** souligné dans la phrase donnée.

Question n° 13 : Dans l'interview à laquelle elle s'est prêtée, la chanteuse a **éludé** les Questions sur sa vie privée. LE CONTRAIRE de " éludé " =

Surface de Compétence				
19,60%		%	r.bis	CM
1. autorisé		13,9	-0,12	43,7
2. apprécié		2,3	-0,1	34,5
3. joué avec		2,4	-0,08	40,2
4. trié		2	-0,06	33,5
5. affronté		32,8	0,34	60,2
6. aucune		41,4	-0,08	49,1
7. toutes		0,4	-0,04	37,1

Le terme « éludé » est probablement peu connu car ceux qui fournissent la réponse correcte (32,8 %) le font avec une certitude moyenne modérée (60,2 %). « 6. Aucune » est le distracteur-vedette (41,4 %) le record du test pour ce distracteur général « Aucune ». La fracture subjective est faible (11,1 %) avec ce distracteur-vedette. **Confiance et prudence** sont conformes aux attentes.

Question n° 14 : La remarque qu'il a faite à son collègue était tout à fait **judicieuse**. LE CONTRAIRE de " judicieuse " =

Surface de Compétence				
49,50%		%	r.bis	CM
1. injuste		8,5	-0,15	57,4 *
2. absurde		71	0,26	69,7
3. blessante		1,3	-0,14	46,8
4. naïve		8,1	-0,03	58,9 *
5. pertinente		3,8	-0,1	57,8 *
6. aucune		4	0,02	53,6 *
7. toutes		0,3	-0,06	34

Un bon taux de réponses correctes (71 %) accompagné d'une bonne certitude moyenne (69,7 %), donc avec **confiance**, et une fracture subjective faible de 12,3 % avec le distracteur-vedette (1. *Injuste*), dont la structure formelle in-juste a probablement joué dans sa proximité à judicieuse, les deux mots étant porteurs de l'idée de justice. Quatre distracteurs sont choisis avec **imprudence**.

Question n° 15 : Dans son nouvel ouvrage, cet auteur **réfute** les thèses de ses prédécesseurs. LE CONTRAIRE de " réfute " =

Surface de Compétence				
62,50%		%	r.bis	CM
1. modifie		3	-0,15	45,3
2. reformule		1,8	-0,14	42,9
3. commente		0,8	-0,08	36,9
4. confirme		84,3	0,33	74,1
5. affine		1,4	-0,1	51,5 *
6. aucune		5,7	-0,1	48,7
7. toutes		0,2	-0,02	57,5 *

Très bon taux de réussite (84,3 %), avec une certitude moyenne plus décevante (74,1 %) mais une fracture subjective nette (25,4 %) avec le distracteur-vedette 6. *Aucune*.

Question n° 16 : Tous les chiffres de cette étude mettent en évidence la **précarité** des emplois de beaucoup de jeunes. LE CONTRAIRE de " précarité " =

Surface de Compétence				
33,90%		%	r.bis	CM
1. qualité		22,8	-0,03	61,7 *
2. diversité		4,5	-0,14	45,3
3. stabilité		46,6	0,39	72,7
4. rentabilité		5,3	-0,09	44,8
5. facilité		4,4	-0,12	43,1
6. aucune		11	-0,21	43,8
7. toutes		1,8	0	53,1

Près d'un étudiant sur deux a répondu correctement (46,6 %) avec un taux de certitude moyen assez élevé (72,7 %), signe de **confiance**, de 11 % supérieur seulement (fracture subjective faible) à celle du distracteur-vedette (1. *Qualité*) qui a été choisi par un étudiant sur cinq avec **imprudence** (certitude moyenne = 61,7%).

Question n° 17 : Dans sa plaidoirie, l'avocat a accumulé les **présomptions**. LE CONTRAIRE =

Surface de Compétence				
17,10%		%	r.bis	CM
1. témoignages	14,7	0,05	41,4	
2. contre-vérités	16,5	-0,14	33,8	
3. éloges	10,6	-0,1	36,1	
4. mensonges	5,2	-0,14	30,9	
5. détails	4	-0,04	42,2	
6. aucune	41,3	0,31	41,5 *	
7. toutes	0,4	-0,01	45,3	

Bien que la réponse correcte soit « Aucune », 41,3 % des étudiants l'ont choisie, mais avec une certitude moyenne très faible (41,5 %), donc un **manque de confiance**, et une **fracture subjective** très faible (8,7 %) avec le distracteur-vedette (2. Contre-vérités) et quasi nulle (0,1) avec le second distracteur-vedette (1. Témoignages). La **prudence** est partout conforme aux attentes.

Question n° 18 : L'histoire est la plus **partiale** des sciences. LE CONTRAIRE =

Surface de Compétence				
29,70%		%	r.bis	CM
1. objective	44,4	0,41	66,9	
2. complète	15,8	-0,2	47,2	
3. générale	11,1	-0,14	47,7	
4. ancienne	1,4	-0,1	30,7	
5. positive	0,8	-0,07	37,3	
6. aucune	21,7	-0,06	48,2	
7. toutes	0,6	-0,07	35,7	

Le distracteur-vedette « Aucune » a peut-être été choisi parce que certains attendaient « impartiale » ; **sa fracture subjective** avec la réponse correcte existe, tout en restant modeste, heureusement, élevée (18,7 %). Que le deuxième distracteur en popularité soit (2. Complète) vient peut-être de l'interprétation fautive (« parcellaire ») du mot « partiel ». La **fracture subjective** (19,7 %) est bonne elle aussi avec ce distracteur. **Confiance et prudence** sont conformes aux attentes.

Question n° 19 : Sa secrétaire était **habilitée** à délivrer certains documents officiels. **Au contraire**, sa secrétaire *** délivrer certains documents officiels.

Surface de Compétence				
52,90%		%	r.bis	CM
1. n'avait pas le droit de	70,6	0,42	74,9	
2. était lente à	1,2	-0,12	43,3	
3. n'avait pas coutume de	15,7	-0,24	57,9 *	
4. faisait des difficultés pour	4,6	-0,18	53 *	
5. s'était peu à peu habituée à	1,3	-0,1	50,6 *	
6. aucune	3,5	-0,04	46,3	
7. toutes	0,3	-0,04	42	

La majorité des étudiants ont réussi (70,6%), avec **confiance** (la certitude moyenne est 74,9%). La **fracture subjective** faible (14,1 %) avec le distracteur-vedette (2. N'avait pas coutume de), choisi avec **imprudence**, (57,9%) comme deux autres distracteurs.

Question n° 20 : A travers son oeuvre, l'auteur a dressé un tableau **exhaustif** de la France du XIXe siècle. **Au contraire**, dans le tableau qu'il a dressé, l'auteur *** la France du XIXe siècle.

Surface de Compétence				
19%		%	r.bis	CM
1. a établi un constat très critique de	19,6	-0,05	41,1	
2. a fait l'éloge de	8	-0,1	35,1	
3. a accentué les traits de	11,6	0,01	43,6	
4. a posé un regard pessimiste sur	14,3	-0,15	39,4	
5. a ironisé sur la situation de	5,7	-0,12	34,8	
6. aucune	34,5	0,36	55,2	
7. toutes	0,4	-0,04	26,3	

Un étudiant sur trois (34,5 %) seulement a réussi, mais avec une certitude moyenne assez faible (55,2 %). Sans doute l'effet du fait que la réponse correcte est « Aucune ». La **fracture subjective** est faible (14,1 %) avec le distracteur-vedette nettement faux qui a attiré un étudiant sur cinq (19,6 %). **Confiance et prudence** sont conformes aux attentes.

Question n° 21 : Il a formulé des griefs à l'égard de certains de ses collaborateurs.
 Au contraire, il *** ses collaborateurs.

Surface de Compétence		%	r.bis	CM
27,40%				
1. est resté très poli en s'adressant à		17,8	-0,21	40,7
2. fait des suggestions à		4,6	-0,09	36
3. a félicité certains de		48,1	0,38	56,9
4. a adressé des vœux de réussite à		8,1	-0,05	44
5. a laissé toute initiative à		5,8	-0,03	36,1
6. aucune		8,8	-0,05	37,9
7. toutes		1,5	-0,05	32,3

Près d'un étudiant sur deux a répondu correctement (48,1 %), mais avec une certitude moyenne assez faible (56,9 %), et une fracture subjective modeste (16,2 %) avec le distracteur-vedette « 1. Est resté poli », choisi par près d'un étudiant sur cinq, ce qui peut se comprendre, ce distracteur étant sémantiquement opposé à griefs, sans toutefois être son contraire.
Confiance et prudence sont conformes aux attentes.

Comparez un à un les mots des listes qui suivent avec le mot en gras auquel ils sont associés. Pour chacune des paires ainsi constituées, déterminez le rapport de sens en écrivant

- si, dans certains contextes, les deux mots peuvent avoir à peu près le MEME sens.
- si, dans certains contextes, les deux mots peuvent avoir des sens à peu près CONTRAIRES.
- si, dans aucun contexte, les deux mots n'entretiennent l'un des deux rapports précédents.

Question n° 22 : Assujettir- émanciper

Surface de Compétence		%	r.bis	CM
24,20%				
(1) Même		9,7	-0,2	32,9
(2) Contraires		41,3	0,41	58,5
(3) Aucun		44,4	-0,17	43,3

Moins de la moitié (41,3 %) des étudiants ont fourni la réponse correcte avec une certitude moyenne faible (58,5 %). Le distracteur-vedette (3. Ni contraire, ni même) est plus populaire (44,4 %) que la solution correcte mais heureusement, avec une certitude moyenne moins élevée (43,3 %) que la réponse correcte, d'où une fracture subjective modeste de 15,2 %.
Confiance et prudence sont conformes aux attentes.

Question n° 23 : Assujettir – subordonner

Surface de Compétence		%	r.bis	CM
27,40%				
(1) Même		50,1	0,39	54,7
(2) Contraires		15,9	-0,19	35,8
(3) Aucun		28,8	-0,13	36,9

La moitié des étudiants (50,1 %) ont bien répondu, avec une certitude moyenne faible (54,7 %). Le distracteur-vedette (3. Ni contraire, ni même) a, heureusement, une certitude moyenne plus faible (36,9 %), d'où une fracture subjective modeste de 17,8 %.
Confiance et prudence sont conformes aux attentes.

Question n° 24 : Assujettir – asservir

Surface de Compétence		%	r.bis	CM
41,30%				
(1) Même		67,7	0,42	61
(2) Contraires		12	-0,19	34,8
(3) Aucun		15,2	-0,19	27,5

Deux étudiants sur trois (67,7 %) ont répondu correctement, mais avec une certitude moyenne peu élevée (61 %) et, cependant, une fracture subjective élevée (33,5 %), la deuxième plus élevée (après la Q28) de tout le check-up Vocabulaire.
Confiance et prudence sont conformes aux attentes.

Question n° 25 : Assujettir - affranchir

Surface de Compétence		%	r.bis	CM
29,90%				
(1) Même		20,8	-0,24	40,3
(2) Contraires		47,9	0,48	62,4
(3) Aucun		26	-0,16	38,7

Près d'un étudiant sur deux (47,9 %) a réussi cette question avec une certitude moyenne passable (62,4 %) et une fracture subjective nette (23,7 %) avec le distracteur-vedette (3. Ni contraire, ni même).
Confiance et prudence sont conformes aux attentes.

Question n° 26 : Injonction - sommation			
Surface de Compétence			
35,90%	%	r.bis	CM
(1) Même	63,8	0,48	56,2
(2) Contraires	11,4	-0,23	31,1
(3) aucune	18,5	-0,22	30,2

Près d'un étudiant sur trois (63,8 %) a fourni la réponse correcte, mais avec une certitude moyenne peu élevée (56,2 %) et une fracture subjective nette (26 %) avec le distracteur-vedette (3. *Ni même, ni contraire*).
Confiance et prudence sont conformes aux attentes.

Question n° 27 : Injonction - invective			
Surface de Compétence			
15,70%	%	r.bis	CM
(1) Même	18,7	0,13	42,3
(2) Contraires	22,7	-0,09	27,8
(3) Aucun	49,2	0,13	31,9 *

Un étudiant sur deux (49,2 %) a fourni la réponse correcte, mais avec une certitude moyenne particulièrement faible (31,9 %), la plus faible pour une réponse correcte sur les 48 questions de ce check-up, donc avec **manque de confiance**. La fracture subjective, est insignifiante (4,1 %) avec le distracteur-vedette et négative avec la solution 1. La corrélation r.bis est inférieure à 0,15, ce qui devrait faire éliminer cette question (sans que l'on comprenne bien pourquoi) d'une épreuve sanctionnante.
 La **prudence** est conforme aux attentes.

Question n° 28 : Injonction - ordre			
Surface de Compétence			
51,10%	%	r.bis	CM
(1) Même	76,7	0,45	66,6
(2) Contraires	5,3	-0,16	29,1
(3) Aucun	12,4	-0,27	29,5

Trois étudiants sur quatre (76,7 %) ont fourni la réponse correcte, avec une certitude moyenne satisfaisante (66,6 %) et la fracture subjective élevée, est la plus importante (37,1 %), avec celle de Q30, des 45 questions de cette épreuve.
Confiance et prudence sont conformes aux attentes.

Question n° 29 : Injonction - addition			
Surface de Compétence			
42,40%	%	r.bis	CM
(1) Même	13,3	-0,26	39
(2) Contraires	7	-0,17	32,1
(3)Aucun	73,9	0,45	57,4

Trois étudiants sur quatre (73,9 %) ont fourni la réponse correcte, avec une certitude moyenne décevante (57,4 %), mais une fracture subjective modeste (18,4 %) avec le distracteur-vedette.
Confiance et prudence sont conformes aux attentes.

Question n° 30 : Prohibé - permis			
Surface de Compétence			
59,60%	%	r.bis	CM
(1) Même	11,8	-0,33	41,9
(2) Contraires	75,5	0,51	79
(3)Aucun	9,1	-0,21	55,8 *

Trois étudiants sur quatre (75,5 %) ont fourni la réponse correcte, avec une certitude moyenne élevée (79 %), donc avec **confiance**, et la fracture subjective (élevée) est la plus forte (37,1 %), avec celle de Q28, des 45 questions.
 Le distracteur "Aucune" est choisi avec **imprudence**.

Question n° 31 : Prohibé - inhibé			
Surface de Compétence			
35,10%	%	r.bis	CM
(1) Même	10,8	-0,05	49,3
(2) Contraires	23,8	-0,3	46
(3) Aucun	61,1	0,4	57,4

Près de deux étudiants sur trois (61,1 %) ont répondu correctement mais avec une certitude moyenne décevante (57,4 %) et une fracture subjective faible (11,4 %) avec le distracteur-vedette.
Confiance et prudence sont conformes aux attentes.

Question n° 32 : Prohibé - proscrit				
Surface de Compétence				
50,60 %				
(1) Même	71,1	0,48	71,2	
(2) Contraires	9	-0,15	45	
(3) Aucun	16,1	-0,32	35,7	

Bon taux (71,1 %) de réponses correctes et de certitude moyenne (71,2 %), avec une fracture subjective élevée (35,5 %) avec le distracteur-vedette.

Question n° 33 : Prohibé - légal				
Surface de Compétence				
52,60%				
	%	r.bis	CM	
(1) Même	11,4	-0,27	44	
(2) Contraires	69,8	0,47	75,3	
(3) Aucun	14,7	-0,18	51 *	

Deux étudiants sur trois (69,8 %) ont fourni la réponse correcte avec une bonne certitude moyenne (75,3 %), donc avec **confiance**. La fracture subjective est faible (14,3 %) avec le distracteur-vedette, choisi avec **imprudence**.

Question n° 34 : Différer - retarder				
Surface de Compétence				
49,70%				
	%	r.bis	CM	
(1) Même	69,4	0,52	71,6	
(2) Contraires	4,1	-0,15	47,9	
(3) Aucun	22,6	-0,36	47,4	

Deux étudiants sur trois (69,4 %) ont fourni la réponse correcte avec une bonne certitude moyenne (71,6 %) et une fracture subjective nette (24,2 %) avec le distracteur-vedette. C'est le r.bis le plus élevé (0,52) après celui de la Q10. **Confiance et prudence** sont conformes aux attentes.

Question n° 35 : Différer - contester				
Surface de Compétence				
38,20%				
	%	r.bis	CM	
(1) Même	30	-0,3	51,9 *	
(2) Contraires	5	-0,15	44,6	
(3) Aucun	60,6	0,46	63	

Le taux moyen (60,6 %) de réponses correctes est satisfaisant avec une certitude moyenne peu élevée (63 %), mais **confiante** tout de même. La fracture subjective est de 11,1 % (faible) avec le distracteur-vedette, choisi avec **imprudence** (51,9%).

Question n° 36 : Différer - ajourner				
Surface de Compétence				
36,90%				
	%	r.bis	CM	
(1) Même	54	0,51	68,3	
(2) Contraires	8,4	-0,16	41,7	
(3) Aucun	32,5	-0,3	44,1	

Un étudiant sur deux (54 %) a fourni la réponse correcte, avec une certitude moyenne acceptable (68,3 %) et une fracture subjective nette (24,2 %) avec le distracteur-vedette. **Confiance et prudence** sont conformes aux attentes.

Question n° 37 : Différer - anticiper				
Surface de Compétence				
30,60%				
	%	r.bis	CM	
(1) Même	5,8	-0,09	43,9	
(2) Contraires	46,7	0,39	65,5	
(3) Aucun	41,6	-0,22	51,1 *	

Moins d'un étudiant sur deux a fourni la réponse correcte, avec **confiance** : la certitude moyenne est passable (65,5 %). La fracture subjective est faible (14,4 %) avec le distracteur-vedette, choisi avec **imprudence**.

Choisissez, pour chacun des mots en majuscULES, LA DEFINITION QUI LUI CORRESPOND :

Question n° 38 : probité				
Surface de Compétence				
	9,80%	%	r. bis	CM
1. désir immodéré de posséder une chose		9,3	-0,08	17,4
2. observation rigoureuse des devoirs de la justice et de la morale		27,2	0,22	36,1 *
3. penchant effréné ou irrésistible pour la luxure		12,2	-0,03	20
4. sentiment violent qui pousse à vouloir du mal à quelqu'un		6,2	-0,08	17,2
5. période où l'on doit faire ses preuves		14,4	0,04	30,9
6. aucune		17,6	0	20,1
7. toutes		0,4	-0,02	15,7

Un étudiant sur quatre seulement (27,2 %) a fourni la réponse correcte, avec une certitude moyenne particulièrement faible (36,1 %) donc un **manque de confiance**, le plus prononcé après Q29 (Injonction – Addition). La **fracture subjective** est modeste (16 %) avec le distracteur-vedette (6. *Aucune*).

Question n° 39 : réactionnaire				
Surface de Compétence				
	13%	%	r. bis	CM
1. qui s'oppose au progrès et vise à rétablir un état antérieur		19,3	0,23	67,6
2. qui conteste l'ordre établi et est partisan de changements radicaux et rapides		56,2	0,05	58,7 *
3. qui possède des titres de propriété dans une entreprise		1,3	-0,11	35,4
4. qui exerce une force de propulsion		4,7	-0,14	39,1
5. qui est prompt à s'imiter, à s'emporter		10,7	-0,09	45,9
6. aucune		2,1	-0,04	28,6
7. toutes		1,9	-0,04	49,3

Cette question fait partie du trio (avec les questions 5 et 12) des questions les moins bien réussies (19,3 %) avec, pourtant, une certitude moyenne élevée (67,6 %), donc avec **confiance**, et une **fracture subjective** très faible de 8,9 % avec le distracteur-vedette (2. Qui conteste l'ordre établi) choisi par plus de la moitié des étudiants (56,2 %), ce qui en dit long sur la culture politique moyenne de l'échantillon... et avec **imprudence** (certitude moyenne = 58,7%)

Question n° 40 : exhorter				
Surface de Compétence				
	22,20%	%	r. bis	CM
1. empêcher quelqu'un de réaliser son projet		8,3	-0,1	33,5
2. arracher quelque chose à quelqu'un		16,8	-0,14	42
3. s'efforcer de persuader quelqu'un de faire quelque chose		35,7	0,37	62,1
4. élever très haut par ses discours, par ses enseignements		19,9	-0,05	45,6
5. émettre une odeur particulière		1,7	-0,08	27,2
6. aucune		9,8	-0,03	30,2
7. toutes		0,6	-0,01	28,2

Un étudiant sur trois (35,7 %) a fourni la réponse correcte, avec une certitude moyenne « passable » (62,1 %) et une bonne **fracture subjective** (16,5 %) modeste avec le distracteur-vedette (4. *Elever très haut par ses discours*) choisi par un étudiant sur cinq (19,9 %). **Confiance et prudence** sont conformes aux attentes.

Question n° 41 : arbitraire				
Surface de Compétence				
	42,70%	%	r. bis	CM
1. qui est prononcé par un ou plusieurs arbitres		3,5	-0,13	50,3 *
2. qui dépend du bon plaisir, du caprice de quelqu'un		57,1	0,41	74,8
3. qui est le résultat d'une concertation		10,8	-0,2	49,8
4. qui augmente progressivement		0,8	-0,1	44,5
5. qui ne peut être annulé		7,5	-0,13	50,1 *
6. aucune		15,8	-0,11	51,4 *
7. toutes		0,7	-0,04	50,3 *

Plus de la moitié des étudiants (57,10 %) réussit, avec une certitude moyenne assez élevée (74,8 %), donc avec **confiance**, et une **fracture subjective** nette de 23,4 % sur le distracteur-vedette (6. *Aucune*). Plusieurs distracteurs sont choisis avec **imprudence**.

Question n° 42 : empirique				
Surface de Compétence				
	39,20%	%	r.bis	CM
1. qui va de plus en plus mal	11	-0,14	52,2 *	
2. qui relève du style baroque	3,7	-0,12	36	
3. qui détient l'autorité absolue	17,4	-0,16	48,4	
4. qui s'appuie sur l'expérience	51,2	0,34	76,5	
5. qui est incertain	5,8	-0,01	47,3	
6. aucune	6,5	-0,03	43,3	
7. toutes	0,6	-0,03	49,6	

La moitié des étudiants (51,2 %) réussit, avec une certitude moyenne assez élevée (76,5 %), donc avec **confiance**, et une **fracture subjective nette** (27,9 %) avec le distracteur-vedette (3. Qui détient l'autorité absolue).
Un des distracteurs est choisi avec **imprudence**.

Question n° 43 : déperdition				
Surface de Compétence				
	14%	%	r.bis	CM
1. état de ce qui se détériore, se délabre, tombe en ruine	27	0,06	47,6	
2. état d'un navire en danger de naufrage	6,5	0,03	45,3	
3. perte qui se fait graduellement	31,1	0,09	45,1 *	
4. perte des principes moraux	14	-0,03	35,1	
5. acte de vandalisme	2,1	-0,06	28	
6. aucune	6,2	-0,05	21,7	
7. toutes	6,5	0,02	45,2	

Moins d'un étudiant sur trois (31,1 %) réussit, avec une certitude très faible (45,1 %), donc un **manque de confiance**, et une **fracture subjective négative aberrante** (-2,5 %) sur le distracteur-vedette (1. État de ce qui se détériore, se délabre) quasi aussi populaire (27 %) que la réponse correcte. La **prudence** est conforme aux attentes.

Question n° 44 : allégation				
Surface de Compétence				
	16,10%	%	r.bis	CM
1. affirmation quelconque	35,6	0,28	45,3 *	
2. obligation de fidélité et d'obéissance à une nation	15,6	0,04	41,8	
3. fait de rendre moins pesant, moins pénible	5,7	-0,14	30,7	
4. fait de donner à quelqu'un le droit d'agir au nom d'un autre	21,2	-0,08	34,3	
5. fait de minimiser un propos	7,4	-0,08	28,3	
6. aucune	6,4	-0,01	23,8	
7. toutes	0,6	-0,04	30	

Un étudiant sur trois (35,6 %) a réussi cette question, avec une certitude faible (45,3 %), donc un **manque de confiance** et une **fracture subjective** de 11 % (faible) avec le distracteur-vedette (4. Fait de donner à quelqu'un le droit d'agir au nom d'un autre) choisi par un étudiant sur cinq, probablement à cause de son caractère « juridique » fallacieusement repéré dans la séquence de lettres « léga » de al-lega-tion. La popularité (15,6 %) du deuxième distracteur-vedette (2. Obligation de fidélité et d'obéissance à une nation) est manifestement due à une confusion avec le terme « allégeance ». La **prudence** est conforme aux attentes.

Question n° 45 : subversion				
Surface de Compétence				
	13,20%	%	r.bis	CM
1. aide que l'État, qu'une association accorde à un groupement, à une personne	6,6	-0,17	36,7	
2. moyen habile et détourné pour échapper à une situation	34	0,08	43,5	
3. action de dérober avec adresse	6,3	-0,03	33	
4. bouleversement des idées et des valeurs reçues, renversement de l'ordre établi	31,6	0,18	41,8 *	
5. action de revenir à la surface	2,4	-0,11	27,8	
6. aucune	11	0	27,5	
7. toutes	0,6	-0,06	33,6	

Moins d'un tiers des étudiants (31,6 %) choisissent la solution correcte, avec une certitude faible (41,8 %), donc un **manque de confiance** et une **fracture subjective négative aberrante** ! (-1,7 %) avec le distracteur-vedette (2. Moyen habile et détourné pour échapper à une situation). La **prudence** est conforme aux attentes.

Bilan

Indicateurs psychométriques :

Les **indices de discrimination (r.bis)** de 43 des réponses correctes sur les 45 de l'épreuve sont positifs et supérieurs à 0,15. Les deux questions pour lesquelles ce n'est pas le cas (Q27 et Q43) ont des r.bis positifs mais inférieurs à 0,15.

Pour 31 des 45 questions, tous les r.bis des distracteurs sont nuls ou négatifs.

Peu de critiques peuvent donc être adressées à ces questions quant à leur cohérence interne avec l'ensemble de l'épreuve.

Indicateurs édumétriques :

La **confiance** est observée dans 34 des 45 réponses correctes aux 45 questions.

La **prudence totale** est observée dans 25 questions sur les 45. Pour toutes les autres, des distracteurs sont fournis avec **imprudence**, dans 11 questions, un seul distracteur est dans ce cas ; dans 9 autres, plusieurs distracteurs sont dans ce cas.

Chapitre 12

Analyses des réponses au check-up de Syntaxe

*J.-M. Defays, ULg
D. Leclercq, ULg*

Nombre d'étudiants = 3727

r.bis repère= 0,288

Question 1 : Cet étudiant fait des efforts considérables pour suivre la conférence du professeur ----- on voit qu'il ne comprend que quelques bribes.

Surface de Compétence				
18,60%		%	r.bis	CM
1. que l'		1,6	-0,15	52,3*
2. qu'		1,1	-0,11	43
3. duquel		12,1	-0,15	63,7*
4. laquelle		2,2	-0,06	57,8*
5. pour laquelle		49,5	-0,16	65,6*
6. aucune		31,8	0,38	58,4
7. toutes		0,3	-0,03	64*

La solution correcte (6. Aucune) n'est choisie que par un peu moins d'un étudiant sur trois (31,8 %), de bons étudiants (r.bis = 0,38), mais avec **confiance**, bien que la certitude moyenne soit faible (58,4 %). Le distracteur-vedette (5. Pour laquelle) attire un étudiant sur deux (49,5 %), ce qui s'explique par la confusion plausible entre les prépositions « à » et « pour ». Néanmoins, les étudiants qui optent pour cette réponse sont les plus faibles au total du test (r.bis = - 0,16) mais ils le font avec la plus haute certitude (65,6 %). La fracture subjective est négative (-7,2 %), donc aberrante. Le choix du distracteur 3 « duquel » peut s'expliquer par la proximité d'emploi des pronoms « dont » et « duquel » mais ce qui est plus grave c'est le peu d'attention portée au genre de l'antécédent (« professeur » à la place de « conférence »). De plus, cette réponse est choisie avec un taux de certitude très élevé (63,7 %). Quasiment tous les distracteurs (et surtout le distracteur-vedette) ont été choisis avec **imprudence** en moyenne (certitude moyenne supérieure à 50 %).

Question 2 : Une dernière chose ----- nécessaire que tu saches : les inscriptions doivent être renvoyées avant la fin de ce mois.

Surface de Compétence				
54,80%		%	r.bis	CM
1. qu'il est		69,9	0,44	78,4
2. qui est		24,9	-0,3	70,2*
3. qui y est		1,3	-0,28	60*
4. qui l'est		1,9	-0,11	60,3*
5. qui l'est		0,5	-0,14	63,2*
6. aucune		0,9	-0,06	44,7
7. toutes		0,1	-0,02	50*

Plus de deux étudiants sur trois (69,9 %) choisissent la réponse correcte avec **confiance** : la certitude moyenne est élevée (78,4 %) et ce sont de très bons étudiants (r.bis = 0,44). Le choix de la réponse 2 « qui est » (choisie par 24,9 % des étudiants - les plus faibles au total du test avec un r.bis de - 0,3 - avec une certitude élevée de 70,2 %) montre que les étudiants ne prêtent pas suffisamment attention à la deuxième partie de la proposition qui rend impossible ce type de construction. Une grande proximité euphonique peut également expliquer la confusion. La fracture subjective est très faible (8,2 %). Quasiment tous les distracteurs ont été donnés avec **imprudence** (certitude moyenne supérieure à 50 %) et tout spécialement le distracteur-vedette).

Question 3 : Il ne remettra jamais les pieds ici ----- tu ne lui présentes tes excuses.

Surface de Compétence				
66,30%		%	r.bis	CM
1. à moins que		85,4	0,4	77,6
2. d'autant plus que		2	-0,21	60*
3. jusqu'à ce que		8,1	-0,26	67*
4. alors que		0,3	-0,14	61,7*
5. pour que		0,3	-0,05	38,2
6. aucune		2,6	-0,12	46,7
7. toutes		0,6	-0,05	47,6

La solution correcte « à moins que » est choisie par un très grand nombre d'étudiants (85,4 %) par de bons étudiants (r.bis = 0,40) et avec **confiance** : la certitude moyenne est élevée (77,6 %). La réponse 3 « jusqu'à ce que », choisie par 8,1 %, aurait pu convenir sans la présence du « ne » explétif. Le choix de la solution générale implicite « 6. aucune » peut être expliqué par la simplicité plus grande de la construction avec « si » qui aurait néanmoins dû être suivie par une négation. La fracture subjective est de 10,6 % (faible). Les distracteurs principaux sont choisis avec **imprudence** (certitude moyenne supérieure à 50 %).

Question 4 : La conférence était ----- les participants à la formation n'ont pas vu le temps passer.

Surface de Compétence	%	r.bis	CM
36,10%			
1. passionnante à tel point que	8,4	-0,18	59,6*
2. passionnante de sorte que	0,7	-0,11	45,4
3. passionnante au point que	5,4	-0,14	61,1*
4. tellement passionnante que	20,6	-0,17	66,3*
5. si passionnante que	14,2	-0,11	65,8*
6. aucune	0,7	-0,16	57,9
7. toutes	48,9	0,44	73,9

La réponse correcte n'est donnée que par à peine un étudiant sur deux (48,9 %), mais avec **confiance** : la certitude moyenne est élevée (73,9 %) et par de bons étudiants (r.bis = 0,44). Les distracteurs 4 et 5 remportent un certain nombre de suffrages (20,6 % et 14,2 %) parce que ce sont des constructions plus simples qui contiennent une seule conjonction « que » plutôt que des locutions conjonctives. Les certitudes moyennes de ces réponses 4 et 5 sont élevées (66,3 % e 65,8 %), ce qui en fait des réponses nuisibles. La fracture subjective ne vaut que 7,6 % (très faible). Quasi tous les distracteurs sont choisis avec **imprudence** (certitude supérieure à 50 %).

Question 5 : Les solutions sont le lot des politiciens qui, prisonniers du court terme, en deviennent les otages. Leurs électeurs exigent au moins des promesses de solutions rapides. Ils ne s'en privent pas d'en distribuer. On se garderait bien de les en exempter. (V. Forrester, L'Horreur économique). Que représente le pronom *les* ?

Surface de Compétence	%	r.bis	CM
37,10%			
1. les otages	0,9	-0,1	38,2
2. les solutions	7,1	-0,18	45,9
3. les électeurs	13,7	-0,14	52,9*
4. les politiciens	58,5	0,41	63,4
5. les promesses	17,4	-0,18	56,6*
6. aucune	0,4	-0,12	32,9
7. toutes	0,1	-0,06	0

Plus d'un étudiant sur deux (58,5 %), de bons étudiants (r.bis = 0,41) ont fourni la réponse correcte avec **confiance**, bien que la certitude moyenne soit modeste (63,4 %). Le distracteur vedette (5. Les promesses), choisi par un étudiant sur six (17,4 %) est l'erreur jugée la plus grave, car elle reflète une confusion de la substitution pronominale entre l'article défini et indéfini. Elle est choisie par les étudiants les plus faibles à l'ensemble du test (r.bis = -0,18) et avec **imprudence** : la certitude est élevée (56,6 %). Quant à la réponse 3 « les électeurs », elle démontre sans doute une mauvaise compréhension du verbe « exempter ». La fracture subjective est très faible : 6,8 %. Les deux distracteurs principaux sont choisis de façon **imprudente** (certitude moyenne supérieure à 50 %).

Question 6 : On met de plus en plus ces résultats sur le compte de défauts inhérents aux tests ou à l'éducation reçue par ceux qui passent les tests. On n'y voit plus une preuve de l'existence de différences innées entre les races. Que représente le pronom *y* ?

Surface de Compétence	%	r.bis	CM
44,60%			
1. dans les défauts	3,5	-0,08	47,6
2. dans l'éducation	5	-0,28	48,9
3. dans les résultats	69,2	0,48	64,5
4. dans les tests	17,2	-0,27	52,7*
5. dans les races	0,7	-0,12	47,2
6. aucune	2,2	-0,1	42
7. toutes	0,2	-0,02	57,1*

Plus de deux étudiants sur trois (69,2 %) ont fourni la réponse correcte, de très bons étudiants (r.bis = 0,48), avec **confiance**, mais avec une certitude moyenne modeste (64,5 %). Le choix de la solution « 4. dans les tests » comme distracteur vedette (17,2 %) avec une certitude trop élevée (52,7 %) est compréhensible vu la grande proximité entre le pronom et le mot qu'il remplacerait. Le choix de la réponse 2 « dans l'éducation » est plus préoccupant car il est le fruit d'un contresens par rapport à la signification de l'adjectif « inné ». La fracture subjective 11,8 % est faible. Le distracteur principal est choisi avec **imprudence** (certitude moyenne supérieure à 50 %).

Question 7 : Déterminez quelle(s) conclusion(s) on peut tirer de la phrase suivante Personne n'est sans savoir que Nicolas n'est pas peu lésé dans cette affaire.

Surface de Compétence	%	r.bis	CM
69,80%			
1. Personne ne sait que Nicolas est largement lésé	1,5	-0,14	57,1 [*]
2. Personne ne sait que Nicolas n'est pas lésé du tout	0,7	-0,14	38,4
3. Tout le monde sait que Nicolas est largement lésé	87,6	0,4	79,7
4. Tout le monde sait que Nicolas n'est pas lésé du tout	6,6	-0,18	70 [*]
5. Certaines personnes bien informées savent que Nicolas est largement lésé	1,7	-0,25	56,3 [*]
6. aucune	0,7	-0,15	42,1
7. toutes	0,1	-0,02	60

Le pourcentage élevé de bonnes réponses (87,6 %) ainsi que la confiance : la certitude moyenne est élevée (79,7 %) obtenus à cette question est assez surprenant vu la complexité d'une phrase telle que celle-ci. L'erreur jugée la plus grave est la 2 « Personne ne sait que Nicolas n'est pas lésé du tout » car elle propose la signification opposée à celle attendue. Elle est heureusement choisie par un nombre infime d'étudiants (0,7 %). La fracture subjective avec le distracteur-vedette (4) est de 9,7 % (très faible). Les distracteurs principaux sont choisis avec imprudence (certitude moyenne à 50 %).

Question 8 : Quelle(s) fin(s) cohérente(s) pourrait-on donner à la phrase suivante La commission vient de décider de suspendre ses travaux, son président estimant ...

Surface de Compétence	%	r.bis	CM
57,60%			
1. que plusieurs éléments neufs au dossier justifiaient qu'elle les reprenne	4,7	-0,27	55,1 [*]
2. que les circonstances ne l'empêchaient désormais plus de les poursuivre	5,8	-0,21	54,4 [*]
3. que la situation politique les rendait actuellement impossibles	79,6	0,46	72,4
4. qu'il était urgent d'arriver à dégager des conclusions fiables	2,1	-0,15	50,4 [*]
5. qu'il lui fallait les mener le plus rapidement possible à leur terme	0,8	-0,1	37,2
6. aucune	2,7	-0,12	44,1
7. toutes	1	-0,05	52,4 [*]

Quatre étudiants sur cinq (79,6 %) ont bien répondu, de très bons étudiants (r.bis = 0,46) avec confiance : la certitude moyenne est modérée (72,4 %). Dans cette question, les réponses 1 « que plusieurs éléments neufs au dossier justifiaient qu'elle les reprenne », 2 « que les circonstances ne l'empêchaient désormais plus de les poursuivre » et 5 « qu'il lui fallait les mener le plus rapidement possible à leur terme » sont des erreurs graves car elles représentent des contresens manifestes. La fracture subjective avec le distracteur-vedette (2) vaut 18 % (modeste). La grande majorité des distracteurs sont choisis avec imprudence (certitude moyenne supérieure à 50 %).

Question 9 : Comment peut-on interpréter la phrase suivante ? Les questions du deuxième syllabus qui ont été traitées lors des travaux pratiques ne devront pas être révisées pour l'examen.

Surface de Compétence	%	r.bis	CM
42,90%			
1. Aucune question du deuxième syllabus ne doit être révisée pour l'examen	10,1	-0,27	71,1*
2. Il reste dans le deuxième syllabus des questions qui devront être révisées pour l'examen	58,6	0,39	73,2
3. Toutes les questions du deuxième syllabus ont été traitées lors des travaux pratiques	2,2	-0,14	58,1*
4. Aucune des questions du deuxième syllabus n'a été à la fois traitée lors des travaux pratiques et révisée pour les examens	13,2	-0,04	67,6*
5. Toutes les questions des travaux pratiques devront être révisées pour l'examen	0,9	-0,1	64,6*
6. aucune	12,5	-0,16	62*
7. toutes	0,3	-0,01	61,7*

La solution correcte est choisie par une majorité d'étudiants (58,6%), forts à l'ensemble du test (r.bis=0,39) et avec **confiance** (certitude moyenne = 73,2%). Cette question concerne la différence entre la proposition relative déterminative et la proposition relative explicative. Par rapport à cette différence, on peut donc considérer que les réponses 1 « Aucune question du deuxième syllabus ne doit être révisée pour l'examen » et 3 « Toutes les questions du deuxième syllabus ont été traitées lors des travaux pratiques » sont les plus graves ; la réponse 1 étant choisie par les étudiants les plus faibles au total du test (r.bis = -0,27) avec une certitude très élevée (71,1 %). La **fracture subjective** avec le distracteur-vedette (4) est très faible : 5,6 %. Tous les distracteurs sont choisis avec **imprudence** (certitude moyenne supérieure à 50 %).

Question 10 : Comment peut-on interpréter la phrase suivante ? Pierre aime Marie; Claude, Dominique.

Surface de Compétence	%	r.bis	CM
51,50%			
1. Pierre aime Claude	1,1	-0,15	63*
2. Pierre aime Marie, Claude et Dominique	9,2	-0,26	52,1*
3. Claude aime Dominique	63,3	0,46	81,4
4. Pierre aime Dominique	1,3	-0,09	62,4*
5. Marie aime Claude	0,4	-0,11	44
6. aucune	22,2	-0,21	63*
7. toutes	1,1	-0,16	66,3*

Près de deux étudiants sur trois (63,6 %) ont bien répondu et ce sont de bons étudiants (r.bis = 0,46), avec une **forte confiance** : la certitude moyenne est élevée (81,4 %). La solution générale implicite « aucune » est le distracteur-vedette choisi par 22,2 % des étudiants, qui ne sont pas les plus faibles au total du test (r.bis = 0,21), avec une certitude de 63 %. Il est évident que dans ce cas, l'absence de verbe a poussé les étudiants à considérer l'énoncé agrammatical. La **fracture subjective** vaut 18,4 % (modeste). L'erreur jugée la plus grave est la 2 « Pierre aime Marie, Claude et Dominique » qui témoigne d'une méconnaissance de l'emploi du point-virgule. Cette erreur grave a attiré près d'un étudiant sur dix (9,2 %), parmi les plus faibles (r.bis = -0,26). Quasiment tous les distracteurs ont été donnés avec **imprudence** (certitude moyenne supérieure à 50 %) et tout spécialement le distracteur-vedette.

Question 11 : Dans le texte suivant, certains mots ont été supprimés et remplacés par des lettres (a,b). Choisissez la combinaison qui convient pour obtenir un texte cohérent.
 Nous absorbons de l'eau sans penser, le plus souvent, que nous sommes en train de satisfaire un besoin physiologique; fréquemment, nous buvons (...a.) par habitude (...b.) par soif parce que nous sommes soumis à un véritable conditionnement familial et social.

Surface de Compétence	%	r.bis	CM
69,10%			
1. [. a.] toujours [. b.] jamais	3,2	-0,18	58,2*
2. [. a.] comme [. b.] si	0,8	-0,2	40
3. [. a.] rarement [. b.] et	1,3	-0,19	51,7*
4. [. a.] plutôt [. b.] que	90,4	0,42	76,4
5. [. a.] alors [. b.] que	1,6	-0,15	60,3*
6. aucune	2	-0,14	45,9
7. toutes	0,1	-0,01	0

Cette question est la mieux réussie (90,4 %) du check-up, avec cependant une **confiance** moindre : la certitude moyenne des réponses correctes est peu élevée (76,4 %). Le distracteur-vedette (1. toujours /jamais) choisi seulement par 3,2 % des étudiants est illogique et comporte une contradiction interne avec la présence de l'adverbe « fréquemment ». Sa certitude moyenne est cependant de 58,2 % et la **fracture subjective** est de 18,2 % (modeste). Les distracteurs principaux sont choisis avec **imprudence** (certitude moyenne à 50 %).

Question 12 : Dans le texte suivant, certains mots ont été supprimés et remplacés par des lettres (a,b). Choisissez la combinaison qui convient pour obtenir un texte cohérent. Lorsque nous transpirons, nous perdons de l'eau et sel. Or, nous ressentons le besoin d'eau et non le besoin de sel; logiquement, nous ne devrions boire (...a...) de l'eau pure (...b...) un liquide sucré, mais de l'eau légèrement salée.				
Surface de Compétence		%	r.bis	CM
64,80%				
1. [.a.] jamais	[.b.] plutôt	0,7	-0,12	41,6
2. [...a.] ni	[.b.] ni	86,7	0,43	74,7
3. [.a.] ni	[.b.] plutôt	2,2	-0,22	64*
4. [.a.] rarement	[.b.] ou	2,9	-0,18	58,1*
5. [.a.] que	[.b.] ou	3,6	-0,22	60,7*
6. aucune		3	-0,11	47,7
7. toutes		0,1	-0,04	0

Pour cette question, le taux de réussite (86,7 %) est très élevé, fourni par les étudiants les meilleurs (r.bis = 0,43), avec confiance, mais avec une certitude moyenne de 74,7 % seulement. La Surface de Compétence est également très élevée (64,8). Le choix de la réponse 5 (le distracteur-vedette) « que / ou » mène à penser que certains étudiants (3,6 %) n'ont pas lu la phrase jusqu'à la fin. La fracture subjective vaut 14 % (faible). Quant au choix de la solution générale implicite « aucune » (3 %), il témoigne sans doute du peu d'habitude qu'ont les étudiants des négations complexes. Les distracteurs principaux sont choisis avec imprudence (certitude moyenne à 50 %).

Bilan

Les indices de discrimination (r.bis) sont quasi parfaits :

- les r.bis des 12 réponses correctes sont TOUS positifs et TOUS largement supérieurs à la valeur repère 0,28 ;
- les r.bis de TOUS les distracteurs des 12 questions sont TOUS (les 72 !) négatifs !! Ce qui indique que cette épreuve est très bien construite.

Les indices de **confiance** (certitude moyenne de la réponse correcte supérieure à 50 %) sont TOUS bons. Par contre, les indices éduométriques sont *au rouge pour la prudence* pour TOUTES les questions,

- pour chacun des 6 distracteurs pour la question ;
- pour 5 des 6 distracteurs pour les questions 1, 2 et 10 ;
- pour 4 des 6 distracteurs pour les questions 4 et 8 ;
- pour 3 des 6 distracteurs pour les questions 3, 7, 11 et 12 ;
- pour 2 des 6 distracteurs pour la question 5 ;
- pour un des 6 distracteurs pour la question 6.

Quand ils se trompent, les étudiants le font avec une certitude excessive : pour 44 des 72 distracteurs en moyenne, la réponse erronée est **épineuse** et constitue une **méprise** (l'étudiant ne se rend pas compte de son erreur et est même convaincu qu'il a raison. On voit l'importance d'une remédiation pour ceux qui en ont besoin ; même si ceux-ci sont peu nombreux en valeur absolue, ils en ont **bien besoin** !

Chapitre 13

Analyse des réponses au check-up de Compréhension de textes

Ph. Hougardy, ULB

D. Leclercq, ULg

3469 étudiants

Seuil r.bis = 0,41

Question 1 : Choisissez le ou les titre(s) qui pourrai(en)t convenir pour le premier paragraphe (ligne 1 à 60) :			
Surface de Compétence			
	29%	%	r.bis CM
1. Quelle est la composition des cellules nerveuses ?		9,5	-0,19 49,6
2. Qui est l'inventeur de l'expression " cellule nerveuse " ?		6,3	-0,18 49,7
3. Quelle est l'origine du microscope électronique ?		2,1	-0,17 41,6
4. A quand remonte la découverte des premiers neurones ?		50,5	0,47 57,5
5. Le cerveau : une substance résistante		7,8	-0,11 53,3*
6. aucune		13,3	-0,05 49,4
7. toutes		2,6	-0,06 49

Un étudiant sur deux (50,5 %) ont fourni la réponse correcte, et ce sont de bons étudiants (r.bis = 0,47), avec **confiance**, bien que ce soit une certitude moyenne (57,5 %) faible pour une réponse correcte. Le distracteur-vedette (5. Aucune) a été choisi par des étudiants moyens (r.bis = -0,05) et avec une certitude faible (49,4 %) et une fracture subjective de 8,1 % (très faible). Cette question ne relève pas de la bonne connaissance ! Le distracteur 5 est choisi avec **imprudence**, et quatre autres le sont avec une certitude proche de l'imprudence.

Question 2 : Interprétation : Le cerveau a résisté longtemps à l'observation parce que, d'après cet article :			
Surface de Compétence			
	54,90%	%	r.bis CM
1. le microscope n'est pas un instrument permettant d'observer de fine partie du cerveau		8,9	-0,24 57,7*
2. les autorités religieuses de l'époque s'y opposaient fermement		2,7	-0,18 44
3. on craignait l'extension de maladies infectieuses particulièrement redoutées		1,9	-0,18 22,7
4. une technique de rigidification et de coloration des tissus était nécessaire pour l'observation		74,4	0,55 73,8
5. le cerveau forme un réseau continu		3	-0,16 44,5
6. aucune		4,8	-0,14 48,5
7. toutes		0,5	-0,07 50

Trois étudiants sur quatre (74,4 %) ont fourni la réponse correcte avec une bonne certitude moyenne (73,8 %), donc avec **confiance**, et ce sont les meilleurs étudiants (r.bis = 0,55). Le distracteur-vedette (1) est choisi par peu d'étudiants (8,9 %), faibles (r.bis = -0,24) et avec **imprudence** (la certitude moyenne est trop élevée pour une erreur : 57,7 %). La fracture subjective est de 16,1 % (modeste).

Question 3 : Interprétation : Que dit l'article sur la fonction du neurone ?			
Surface de Compétence			
	13,20%	%	r.bis CM
1. la composition asymétrique du neurone correspond rigoureusement à une séparation de ses fonctions		33,1	-0,09 61,8*
2. la fonction d'un neurone est d'interrompre un signal passant dans un réseau		5,4	-0,09 41,4
3. les neurones se présentent tous sous la même morphologie		5,3	-0,12 38,4
4. l'arbre dendritique des neurones est toujours extrêmement ramifié et étendu		12,9	-0,13 47,8
5. les cellules de la réine forment un petit réseau longiligne		7	0,02 57,4*
6. aucune		26,8	0,44 49,1*
7. toutes		2,3	-0,02 47,8

Seulement un étudiant sur quatre (26,8 %) a fourni la réponse correcte, avec une certitude moyenne faible (49,1 %), c-à-d un **manque de confiance**, et ce sont de très bons étudiants (r.bis = 0,44). Une partie de l'explication du faible taux de réussite tient sans doute au fait que la Réponse Correcte est « 6 . Aucune ». Le distracteur-vedette (1) est choisi par un étudiant sur trois (33,1 %), avec une certitude moyenne élevée (61,8 %) pour une erreur, ce qui donne une fracture subjective négative aberrante (-1,27 %). Deux distracteurs sont choisis avec **imprudence**.

Question 4 : Quelle est la nature du différend qui séparait Camillo Golgi de son confrère Santiago Ramon y Cajal ?			
Surface de Compétence			
	%	r.bis	CM
11,10%			
1. Les deux souhaitaient être reconnus comme " le précurseur	3,1	-0,11	42,3
2. une rivalité d'ordre personnel	13,2	-0,08	64,7*
3. une polémique d'ordre scientifique	12,3	-0,13	59*
4. l'opinion sur la forme (continue ou non) du réseau neuronal	40,2	-0,07	71*
5. la comparabilité du réseau neuronal à celui de la circulation sanguine	6,8	-0,07	60,9*
6. aucune	2,4	-0,1	44,3
7. toutes	16,3	0,48	68,2

Seulement un étudiant sur six (16,3 %) a fourni la réponse correcte, mais avec une certitude moyenne assez bonne (68,2 %), donc avec **confiance**, et c'est le fait de très bons étudiants (r.bis = 0,48). Le distracteur-vedette (4) est choisi par deux étudiants sur cinq (40,2 %), moyens au total du test (r.bis = -0,07), avec une certitude moyenne un peu plus élevée (71 %) que la solution correcte, ce qui donne une **fracture subjective** (légèrement) négative (-2,8 %).
Quatre des 6 distracteurs sont choisis avec **imprudence**.

Question 5 : La communication entre neurones a la caractéristique suivante :			
Surface de Compétence			
	%	r.bis	CM
12,60%			
1. Dans certains cas très limités, la transmission entre neurones est électrique	2,9	-0,07	44
2. Dans la plupart des cas, le signal est émis sous forme électrique et transformé pour sa réception sous forme chimique	36,7	-0,15	63,8*
3. Le signal chimique réceptionné est converti en signal électrique	16,9	-0,07	59,9*
4. Elle se fait à l'aide d'éléments chimiques " voyageurs " dans l'espace synaptique	10,1	-0,12	54,4*
5. De façon unidirectionnelle	5,9	-0,04	55,7*
6. aucune	4,1	-0,05	45,1
7. toutes	17,8	0,54	71

La réponse correcte est fournie par moins d'un étudiant sur cinq (17,8 %), mais ce sont de très bons étudiants (r.bis = 0,54), sûrs d'eux (certitude moyenne = 71 %), donc **confiants**. Le distracteur-vedette (2) est choisi par plus d'un étudiant sur trois (36,7 %) et avec une certitude moyenne élevée (63,8 %). Nous y voyons un signe d'impulsivité (dans le sens donné par Kagan) à savoir répondre avec précipitation, en choisissant la première solution acceptable sans lire les autres. Ce ne sont pas de bons étudiants (r.bis = -0,15) qui ont commis cette erreur. La **fracture subjective** est de 7,2 % seulement. Quatre des six distracteurs sont choisis avec **imprudence**.

Question 6 : Le mot synapse dans cet article désigne :			
Surface de Compétence			
	%	r.bis	CM
57,20%			
1. la région de quasi-contact entre deux neurones	75,6	0,53	75,6
2. les excroissances autour du soma	1,9	-0,16	31,9
3. l'arbre des neurones, ramifié et étendu	5	-0,23	45,3
4. une cellule bipolaire de la rétine	1,2	-0,14	30,5
5. le bout de l'axone qui reçoit le message de l'autre cellule	9	-0,24	48,5
6. aucune	3,4	-0,12	39,2
7. toutes	0,7	-0,05	43,3

Trois étudiants sur quatre (75,6 %) ont fourni la réponse correcte, de bons étudiants (r.bis = 0,53) avec une certitude moyenne élevée (75,6 %). Le distracteur-vedette (5) est fourni par des étudiants faibles (r.bis = -0,24) et avec une certitude faible (48,5 %). La **fracture subjective** est importante (27,1 %).
Confiance et prudence sont conformes aux attentes.

Bilan

Indicateurs psychométriques :

Les **indices de discrimination (r.bis)** sont tous sauf un conformes aux attentes, que ce soit pour les 6 solutions correctes (r.bis tous supérieurs à 0,41) ou pour les 36 distracteurs (un seul n'est pas négatif). Aucune critique ne peut donc être adressée à ces questions quant à leur cohérence interne avec l'ensemble de l'épreuve.

Indicateurs éduométriques :

La **confiance avec imprudence** est le schéma de 4 des 6 questions. Une seule allie confiance et prudence. Une autre manque de confiance et imprudence. Il s'agirait, dans tous ces cas, de « rectifier » ces idées fausses.

Chapitre 14

Analyses des réponses au check-up de Géographie et Lecture de graphiques, cartes et tableaux

F. Orban, FUNDP

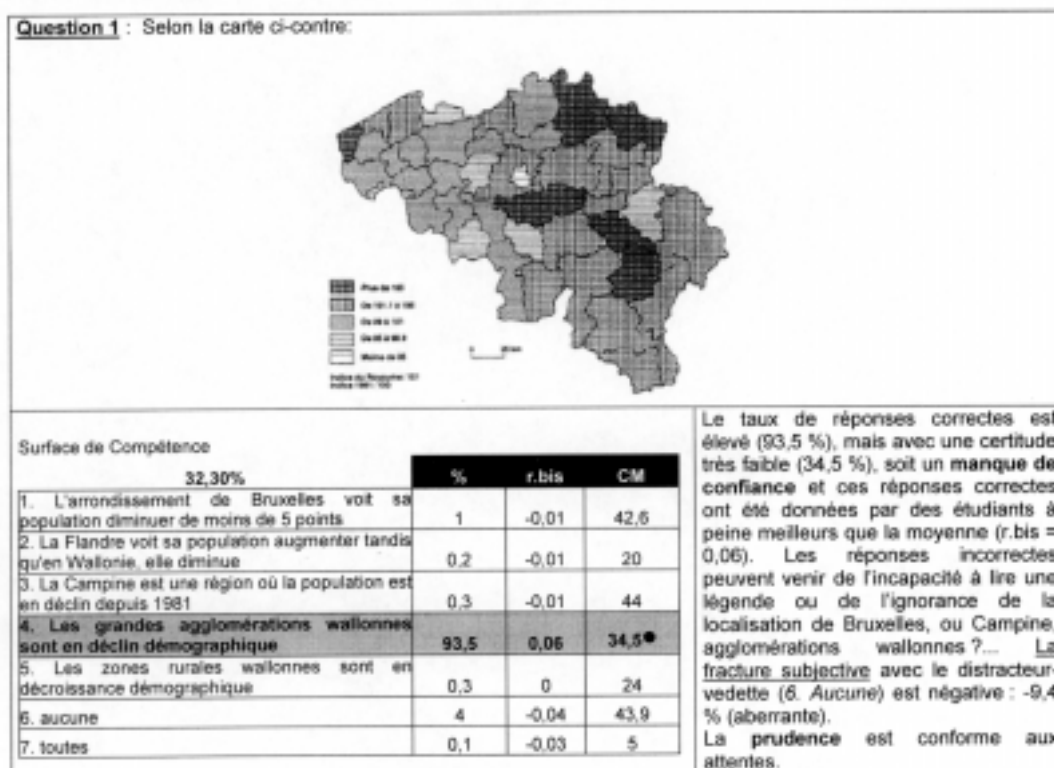
D. Leclercq, ULg

3738 étudiants

Seuil r.bis = 0,32

L'épreuve de « géographie » a été choisie (cela aurait pu être celle d'histoire ou d'économie) pour tester la capacité à lire des tableaux de nombres, des graphiques, des cartes (de géographie). Seules 4 questions (délocalisation, mégalopole, Andes, coordonnées géographiques) testent purement les connaissances géographiques.

On constate qu'une même tâche (lecture de carte) peut être très différemment réussie selon les questions (de 93% à 37%). Il en irait probablement de même si on avait diversifié plus les questions de « lecture de graphiques » ou si l'on avait posé des questions de « construction de graphiques ».



Question 2 : Quel titre donner à la carte ci-dessous ?

Cette question est la plus simple de toutes, car elle n'exige pas de prérequis : il suffit d'interpréter la carte.

Surface de Compétence 53%


	%	r. bis	CM
1. La revanche des villes du Nord	0,6	-0,09	35,2
2. Les grandes agglomérations les plus riches du monde	3,7	-0,12	60,4 *
3. Villes du monde	2,4	-0,11	52,6*
4. Les 20 plus grandes agglomérations du monde	75,7	0,43	70
5. La croissance des agglomérations urbaines de 1980 à 2000	10,7	-0,25	51,2*
6. aucune	5,1	-0,18	45,1
7. toutes	0,3	-0,04	25,5

Trois étudiants sur quatre (75,7 %) ont fourni la réponse correcte, de bons étudiants (r.bis = 0,43) avec une bonne certitude moyenne (70 %), donc avec **confiance**. La **fracture subjective** avec le distracteur vedette (5) est modeste : 18,8 %.

Trois des six distracteurs ont été choisis avec **imprudence**.

Les erreurs révèlent un manque de rigueur chez ces étudiants.

Question 3 : Selon la carte ci-dessous :



Surface de Compétence
35,99%

	%	r.bis	CM
1. La population d'Amérique du Sud observe un modèle d'opposition Centre / Périphérie	37,3	0,41	50,3
2. Tous les grands foyers de population sont dans la zone intertropicale	14,8	-0,05	48,33
3. Les zones très froides et très sèches attirent généralement les populations	2	-0,09	32,6
4. Les zones chaudes et humides sont toujours densément peuplées	12,3	-0,13	50,4
5. La population européenne est répartie de manière moins uniforme que la population d'Amérique du Nord	7,8	-0,15	42,9
6. aucune	19,3	-0,07	41,6
7. toutes	0,9	-0,04	52,73

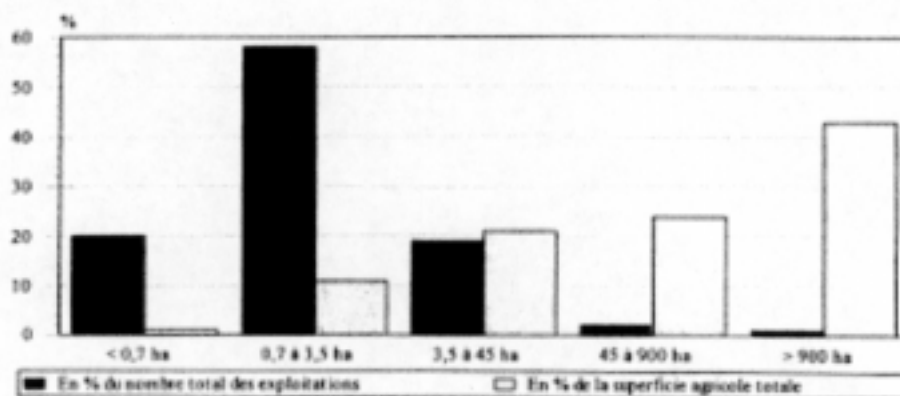
Cette question exige plus qu'une lecture de carte : la connaissance de notions telles que « zone froide » ou « zone sèche », de « centre » et de « périphérie ».

A peine plus d'un étudiant sur trois (37,3 %) a fourni la réponse correcte, avec une certitude TRES élevée (98,2 %), donc avec **confiance**, mais une **fracture subjective** insignifiante (1 %) avec le distracteur-vedette (6. Aucune) qui a attiré un étudiant sur cinq (19,3 %) dont la certitude moyenne vaut 95,2 %.

Tous les distracteurs ont été choisis avec une **imprudence** flagrante.

Il est supposé que les étudiants tentent de répondre sur base de leurs souvenirs quant aux relations sécheresse - population, etc, plutôt que sur base du document fourni.

Question 4 : Selon le graphique ci-dessous :



Surface de Compétence

	29,50%	%	r.bis	CM
1. Les petites exploitations (< à 3,5 ha) représentent plus de 70% de la superficie agricole totale		5,6	-0,12	43,5
2. Les grands domaines agricoles (> à 900 ha) occupent plus de la moitié de la surface agricole		7,9	-0,11	47,3
3. Les petites exploitations (< à 3,5 ha) représentent près de 80% des exploitations agricoles		42,7	0,52	69,2
4. Les exploitations agricoles supérieures à 45 ha représentent 25% de la surface agricole totale		20,6	-0,14	55,1*
5. Les grandes exploitations agricoles (> 45 ha) correspondent à plus de 10% du nombre total des exploitations		4,6	-0,14	39,1
6. aucune		9,8	-0,13	40,8
7. toutes		1,3	-0,07	39,2

Moins d'un étudiant sur deux (42,7 %) a réussi, mais ce sont les meilleurs (r.bis = 0,52) et avec confiance : la certitude moyenne = 69,2 %. La fracture subjective avec le distracteur-vedette (4) est faible : 14,1 %. Ce distracteur a attiré un étudiant sur cinq (20,6 %), avec imprudence.

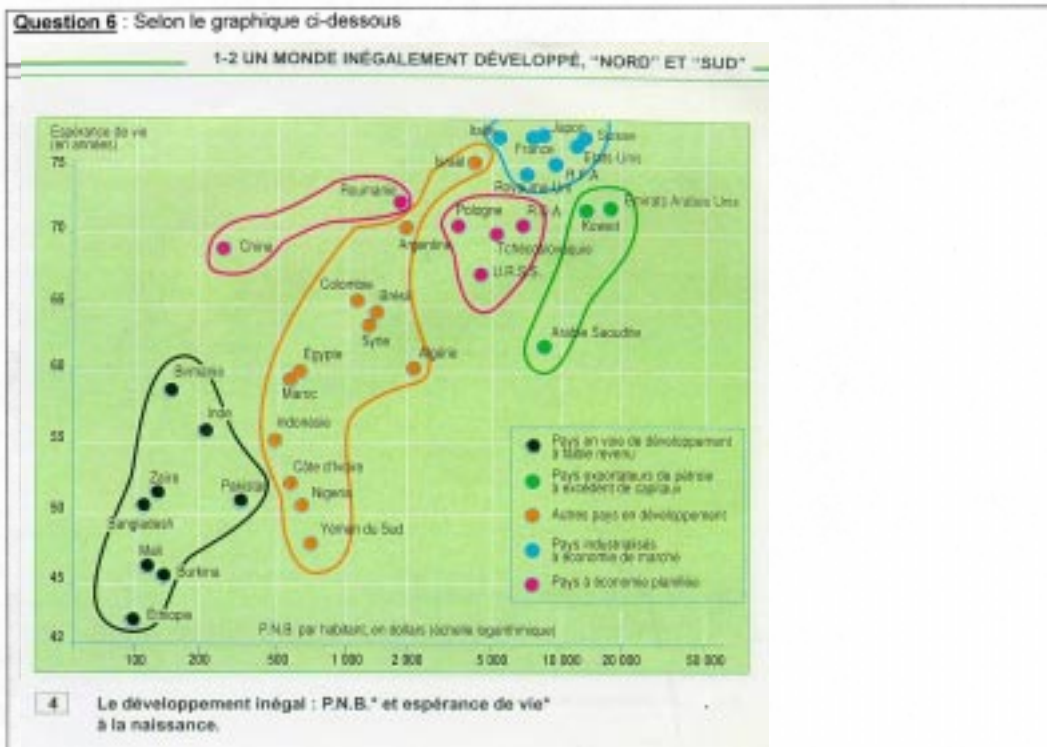
Cette question croise des acquis liés éventuellement aux cours de mathématiques ou de sciences expérimentales et la réussite devrait être corrélée avec celle à d'autres questions comportant des graphiques dans les autres matières.

Question 5 : Selon la carte ci-dessous

Surface de Compétence
35,60%

	%	r.bis	CM
1. Le plus grand importateur de pétrole est le Japon	2,2	-0,08	46,4
2. Les Etats-Unis sont exportateurs de pétrole	1,7	-0,1	44,2
3. Le Venezuela exporte principalement vers les Etats-Unis	53,5	0,57	66,6
4. L'OPEP regroupe exclusivement des pays Africains et Asiatiques	4,5	-0,13	49,9
5. L'Europe ne produit pas de pétrole	24,8	-0,31	59,4*
6. aucune	5,1	-0,14	46,2
7. toutes	2,1	-0,07	56,4*

A peine plus d'un étudiant sur deux (53,5 %) a choisi la solution correcte, mais ce sont les meilleurs à ce check-up ($r.bis = 0,57$) et ils l'ont choisie avec confiance : leur certitude moyenne est de 66,6 %. Ce niveau de certitude fait penser que ce type de représentation graphique peut être utilisée dans les médias. Le distracteur-vedette (5) attire, lui, près d'un étudiant sur quatre (24,8 %), avec une certitude élevée (59,4 %) pour une réponse incorrecte. La fracture subjective n'est que de 7,2 % (très faible). Deux distracteurs (dont le distracteur-vedette) ont été choisis avec imprudence.



Surface de Compétence
43,50%

	%	r.bis	CM
1. Lorsque le Produit National Brut (PNB) augmente, l'espérance de vie diminue	1,9	-0,11	33,1
2. La Pologne se situe en avance de développement socio-économique par rapport à l'Italie	1,7	-0,09	23,4
3. Le Pakistan possède une production économique par habitant moins élevée que celle de la Birmanie	5,8	-0,12	39,2
4. Si un pays a un PNB plus élevé qu'un autre, il a toujours une espérance de vie supérieure	8,3	-0,21	44,8
5. Les pays à économie planifiée ont une espérance de vie inférieure à celle des pays industrialisés à économie de marché	62,7	0,52	69,3
6. aucune	10,5	-0,13	43,1
7. toutes	0,8	-0,07	39,3

Cette question exige que l'on comprenne les deux axes, qu'on y interprète les positions des points (les pays). Malgré cela, les résultats constituent une heureuse surprise puisque presque deux étudiants sur trois (62,7 %), de très bons étudiants (r.bis = 0,52) ont fourni la réponse correcte avec une certitude moyenne assez bonne (69,3 %). La fracture subjective est nette (26,2 %) avec le distracteur-vedette (6).

Confiance et prudence sont conformes aux attentes.

A priori, cette question devrait être moins bien réussie dans la mesure où le graphique a deux dimensions. Or les étudiants la réussissent mieux que quand il faut combiner un double histogramme (p. 4).

Question 7 : Sachant que les coordonnées géographiques de villes sélectionnées ci-dessous sont : (1) Canberra : 36°S, 149°E (2) Madrid : 40°N, 40°O (3) Mexico : 19°N, 99°O (4) New York : 41°N, 74°O (5) Pretoria : 25°S, 28°E (6) Vorkouta : 67° N, 64°E - sachant que : (1) Tropique du Cancer : 23°27' N (2) Tropique du Capricorne : 23°27' S (3) Cercle Polaire Arctique : 66°33' N (4) Cercle Polaire Antarctique : 66°33' S. Cochez la(les) affirmation(s) exacte(s).
Note : Il peut être utile que vous visualisiez les situations en effectuant quelque croquis positionnant la terre par rapport au soleil au moment précisé.

Surface de Compétence		%	r.bis	CM
8,60%				
1. Le 1 décembre, le jour est moins long que la nuit à Pretoria.		5,1	-0,05	15,7
2. Le 1 juillet, au midi solaire local, le soleil culmine plus haut à Madrid qu'à New York		24,6	0,4	35,1
3. Le 1 février, le soleil se lève exactement à l'est de Mexico		10,4	-0,04	19,9
4. En avril, on peut voir le soleil de minuit à Vorkouta		13,3	-0,01	23,4
5. Pour bien fonctionner, des panneaux solaires doivent être orientés vers le Sud à Canberra		3	-0,04	19,8
6. aucune		12,7	-0,09	14,8
7. toutes		5,1	-0,03	14,8

Il s'agit d'une matière (de géographie physique) plus « anciennement » introduite dans les connaissances des étudiants. Cette question est la plus difficile du check-up, car elle nécessitait de faire un dessin sur le côté : (5 minutes sont nécessaires pour aboutir à la solution correcte). C'est sans doute pour cette raison que plus d'un étudiant sur quatre (25,8 %) a omis cette question et que le quart des étudiants seulement ont réussi (24,6 %), les meilleurs (r.bis = 0,4) avec une certitude faible (35,1 %), soit un manque de confiance. La fracture subjective est de 11,7 % (faible). La prudence est conforme aux attentes. On suppose que les étudiants ont acquis le fonctionnement du système solaire et la connaissance des coordonnées latitude des tropiques. C'est la plus complexe tant du point de vue prérequis que de la vision virtuelle dans l'espace à 3 dimensions.

Question 8 : La Cordillère des Andes s'est formée:

Surface de Compétence		%	r.bis	CM
30%				
1. Suite à la séparation de la plaque d'Amérique du Sud et de l'Australie		5,9	-0,13	25,9
2. Par la rencontre de la plaque Pacifique et de la plaque d'Amérique du Sud		52	0,53	57,6
3. Par la rencontre de la plaque Atlantique et de la plaque Amérique du Sud		12	-0,15	30
4. Par la rencontre de la plaque Amérique du Nord et de la plaque Amérique du Sud		10,9	-0,13	28,1
5. Par la rencontre de la plaque Antarctique et de la plaque d'Amérique du sud		3	-0,09	21,1
6. aucune		5,5	-0,13	22,7
7. toutes		0,3	-0,03	30,9

Cette question de géographie physique peut être résolue par réflexion. La solution correcte n'est fournie que par un étudiant sur deux (52 %), les meilleurs (r.bis = 0,53), avec une certitude moyenne assez faible (56,7 %). Le distracteur-vedette (3) peut être dû à la confusion entre les deux océans ... dont les étudiants semblent conscients car leur certitude moyenne est particulièrement faible (30 %) et la fracture subjective nette (27,6 %). Confiance et prudence sont conformes aux attentes.

Question 9 : Cochez la(les) définition(s) qui recouvre(nt) le concept de "Mégalopole"

Surface de Compétence		%	r.bis	CM
24,50%				
1. Plus grande ville d'un pays		16,9	-0,23	49,4
2. Réunion de plusieurs grandes agglomérations		42,2	0,36	58
3. Toujours un pôle de développement industriel		11,5	-0,11	42,7
		0,6	-0,09	20,9
		0,3	-0,04	20
6. aucune		16,9	0,04	51,6*
7. toutes		2,7	0	51,8*

Cette question est mal réussie (42,2 %) avec confiance, quoique la certitude moyenne soit faible (58 %), mais par d'assez bons étudiants à ce check-up (r.bis = 0,36). La fracture subjective est très faible (7,5 %) avec les distracteurs-vedettes 1 et 6. Deux des distracteurs sont choisis avec imprudence. La confusion vient du fait que l'étymologie peut faire penser à « la plus grande », bien que la géographie soit claire sur ce point et que l'actualité parle souvent de la mégalopole de l'Est ou de l'Ouest des USA en parlant de l'ensemble des grandes villes de ces côtes.

Question 10 : Cochez la(les) définition(s) qui recouvre(nt) le concept de " Délocalisation "			
Surface de Compétence	%	r.bis	CM
44,76%			
1. Transfert des activités de production ou de services d'une entreprise d'un pays industrialisé vers un autre pays	69,4	0,4	64,4
2. Transfert des bénéfices d'une entreprise nationale vers un autre pays	2	-0,09	32,8
3. Transfert du centre de décision d'une entreprise d'un pays vers un autre	8,4	-0,16	45,7
	0,4	-0,07	22,7
	0,4	-0,07	20
6. aucune	6,4	-0,13	39
7. toutes	5,1	-0,06	51,4*

Un peu plus de deux étudiants sur trois (69,4 %) ont répondu correctement avec confiance (la certitude moyenne est 64,4 %) et ce sont de bons étudiants (r.bis = 0,4). La fracture subjective avec le distracteur-vedette (3) est modeste (18,7 %) ; ce choix erroné est sans doute dû au manque de conscience du fait que dans les entreprises modernes, les centres de décision et de production peuvent être distants avant même toute délocalisation. Seul un distracteur est choisi avec imprudence.

Bilan

Indices de discrimination (r.pbis) :

Sauf pour la question 1, toutes les réponses correctes ont un indice de discrimination positif et supérieur à la valeur-repère (0,32).

Confiance et imprudence

sont la règle générale (4 sur les 6 questions de lecture de carte, de graphique, de tableaux et 2 sur les 4 questions de connaissances géographiques).

Chapitre 15

Analyses des réponses au check-up de Mathématiques

*J. Lega, UCL
M. Lebrun, UCL
D. Leclercq, ULg*

N = 2539

R.bis repère = 0,21

Question 1 : Pierre a actuellement 6 ans de plus que Paul. Il y a 10 ans, Pierre avait un âge double de celui de Paul. En supposant que Y représente l'âge de Pierre et X l'âge de Paul. Choisissez les équations représentatives du problème.

Surface de Compétence			
55,30%	%	r.bis	CM
1. $Y = X + 6$ et $2Y = X$	3,9	-0,19	56,9*
2. $Y = X - 6$ et $2Y - 10 = X - 10$	4,5	-0,21	59,6*
3. $Y = X + 6$ et $Y = 2X - 10$	8	-0,15	68,1*
4. $Y = X + 6$ et $(Y - 10) = 2(X - 10)$	70,5	0,43	78,4
5. $Y = X - 6$ et $(X - 10) = 2(Y - 10)$	5,9	-0,15	64,4*
6. aucune	4	-0,12	49,5
7. toutes	0	XXXXX	XXXXX

Les solutions 3 et 4 pourraient être considérées toutes deux comme correctes, la 3 étant la construction logique et la 4 une transformation. Les 8% d'étudiants qui ont choisi la 3 le font avec une certitude moyenne (68,1%) plus élevée que ceux qui ont choisi d'autres distracteurs. Les fractures subjectives avec la réponse correcte croissent à mesure que les popularités des distracteurs décroissent (à l'exception de "6. Aucune") : distracteur 3 : popularité 8,0% et Fract. Subj. 10,3% (faible).
distracteur 5 : popularité 5,9% et Fract. Subj. 14,0%
distracteur 2 : popularité 4,5% et Fract. Subj. 18,8%
distracteur 1 : popularité 3,9% et Fract. Subj. 21,5%
distracteur 6 : popularité 4,0% et Fract. Subj. 28,9%.
Confiance dans la solution correcte mais **imprudence** (certitudes supérieures à 50%) dans le choix des distracteurs.

Question 2 : Pierre a actuellement 6 ans de plus que Paul. Il y a 10 ans, Pierre avait un âge double de celui de Paul. En supposant que Y représente l'âge de Pierre et X l'âge de Paul. Donnez les solutions de ce problème.

Surface de Compétence			
65%	%	r.bis	CM
1. Y = 12 ans et X = 6 ans	4,5	-0,18	66,4*
2. Y = 22 ans et X = 16 ans	77,1	0,4	84,3
3. Y = 16 ans et X = 22 ans	7,5	-0,14	73,9*
4. Y = 18 ans et X = 12 ans	2,8	-0,16	54,9*
5. Y = 26 ans et X = 20 ans	1,2	-0,14	58,7*
6. aucune	3,9	-0,13	49,2
7. toutes	0,1	-0,06	46,7

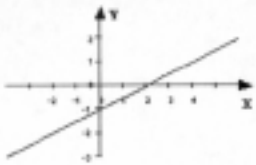
Trois étudiants sur quatre ont réussi (75%), avec **confiance** (la certitude moyenne est élevée : 84,1%) et ce sont de bons étudiants (r.bis = 0,36).
Le distracteur-vedette (2) est choisi avec une certitude moyenne plus basse (68,5%) d'où une fracture subjective de 15,6% (modeste). La solution 3, aberrante, est -heureusement- très peu choisie, et avec la certitude moyenne la plus faible : 37,8%.
Les distracteurs sont choisis avec **imprudence**.

Question 3 : La figure ci-contre représente un triangle dont l'angle a vaut 90°. Exprimez la valeur du côté c à l'aide des variables indiquées sur le dessin.

Surface de Compétence			
67,10%	%	r.bis	CM
1. $c = \sqrt{a^2 + b^2}$	8,2	-0,3	60,7*
2. $c = \sqrt{a^2 - b^2}$	78,2	0,49	85,8
3. $c = a^2 + b^2$	1,7	-0,15	56,7*
4. $c = \sqrt{b^2 - a^2}$	3,8	-0,12	70,5*
5. $c = a^2 - b^2$	3	-0,13	68,2*
6. aucune	3,9	-0,22	42,6
7. toutes	0	XXXXX	XXXXX

La « chanson » de Pithagore semble bien connue : le taux de réussite est élevé (78,2%) ainsi que la **confiance** (la certitude moyenne est de 85,8%).
Le distracteur-vedette (1), qui consiste à appliquer la chanson sans réfléchir, a attiré les étudiants les plus faibles au test (r.bis = -0,30), avec une fracture subjective nette (25,1%), plus importante que pour les deux distracteurs qui le suivent en popularité (4 et 5), le distracteur "6. Aucune" se caractérisant à nouveau par la certitude moyenne la plus faible (42,6%).
Cette question fait partie des quatre questions (avec la 16, la 21 et la 22) les plus discriminatives de toute cette épreuve (r.bis de la réponse correcte = 0,49).
Les distracteurs sont choisis avec **imprudence**.

Question 4 : Ecrivez l'équation de la droite représentée sur le graphique ci-dessous :



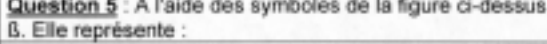
Surface de Compétence
64,70%

	%	r.bis	CM
1. $Y = 2X - 1$	7,6	-0,26	59,3*
1. $Y = 2X - 2$	1,2	-0,14	51,3*
3. $Y = (X / 2) - 2$	5,8	-0,19	77*
4. $Y = X - 1$	2,3	-0,12	67,1*
5. $Y = (X / 2) - 1$	75,1	0,45	86,1
6. aucune	7,1	-0,14	61,5*
7. toutes	0,1	-0,02	66,7*

La réponse correcte attire 3 étudiants sur quatre (75,1%) avec une **confiance élevée** (certitude moyenne = 86,1%). Le distracteur vedette (1) qui consiste en une erreur sur le coefficient de pente a attiré les étudiants les plus faibles ($r.bis = -0,26$) et avec une certitude moyenne faible (59,3%) donc une **fracture subjective** nette (26,8%). La solution "6. Aucune" est le deuxième distracteur vedette avec une popularité quasi égale (7,1% au lieu de 7,6%), mais avec une certitude moyenne légèrement supérieure (61,5% au lieu de 59,3%) ; ceci est probablement dû au fait que le dessin est imparfait car trop peu précis et ces étudiants ont pu considérer que la droite NE passe PAS par (0 ; -1).

Tous les distracteurs sont choisis avec **imprudence**.

Question 5 : A l'aide des symboles de la figure ci-dessus , précisez la signification concrète de la formule de tan B. Elle représente :



Surface de Compétence
23,60%

	%	r.bis	CM
1. La pente du segment de droite a	40	0,34	58,7
2. La longueur du côté b en face de l'angle B	13,7	-0,1	52,6*
3. La longueur du côté a qui touche l'angle B	3,5	-0,14	27,4
4. La valeur de l'angle B en degrés	6,2	-0,22	44,4
5. La grandeur de l'hypoténuse	5	-0,17	36,8
6. aucune	24,8	0,05	47,5
7. toutes	0,3	-0,05	34,3

Cette question est mal réussie (40% de réponses correctes seulement et avec une certitude moyenne faible elle aussi : 58,7%), ce qui est cependant de la **confiance**. Elle est aussi mal réussie quand elle est posée en question ouverte : "Dites le avec vos propres mots...". La solution "6. Aucune" a attiré un étudiant sur quatre (24,8%) , qui sont des étudiants "moyens" ($r.bis = 0,05$), mais avec une certitude moyenne faible (47,5%), donc avec **prudence**. Cette certitude est plus faible que celle de la réponse correcte (et une **fracture subjective** faible de 11,2%). Cela pourrait s'expliquer par le fait que dans l'énoncé, l'expression "la figure ci-dessus" pouvait faire référence aussi bien à la figure de la question 4 qu'à celle de la question 3 (qu'il fallait choisir). En outre, si l'on utilise le mot "pente" dans le sens de "coefficient angulaire", il faudrait mettre le dessin dans l'autre sens. Les distracteurs 2 à 4, qui sont aberrants, sont -heureusement- choisis avec une certitude faible, avec **prudence**. Seul le distracteur 2 a été choisi avec **imprudence**.

Question 6 : Soit la relation $Y = \log(X)$. De combien doit évoluer X pour que Y augmente d'une unité ?

Surface de Compétence
45,60%

	%	r.bis	CM
1. Il doit être multiplié par 2	2,5	-0,12	34
2. Il doit être multiplié par 10	70,6	0,32	64,6
3. Il doit être diminué de 10	0,9	-0,03	36,7
4. Il doit être divisé par 10	4	-0,07	46,1
5. Il doit être augmenté de 10	8,1	-0,17	47,5
6. aucune	8,5	-0,13	39,5
7. toutes	0,2	-0,01	16

Le taux de réussite (70,6%) est "moyen" pour cette question et la certitude moyenne l'est elle aussi : 64,6%. Pourtant ce sont plutôt de bons étudiants qui ont réussi ($r.bis = 0,32$). On ne comprend pas bien pourquoi "6.Aucune" est le distracteur-vedette (mais avec une certitude moyenne particulièrement basse : 39,5%), et une **fracture subjective** nette de 25,1 %.

Il est plus inquiétant que le deuxième distracteur-vedette (5) ait à peu près la même popularité (8,1%) mais une certitude moyenne plus élevée (47,5%), car il révèle une confusion entre les opérations de multiplication et d'addition. Heureusement, ce sont les étudiants faibles ($r.bis = -0,17$) qui l'ont choisi. **Confiance et prudence** sont conformes aux attentes.

Question 7 :

Un satellite de masse m tourne autour de la terre de masse M à une distance r du centre de celle-ci et avec une vitesse v . En admettant :

- l'hypothèse du mouvement circulaire, on trouve que la force centripète est : $F = m \frac{v^2}{r}$
- la loi de gravitation universelle, on trouve que cette force est : $F = \frac{GMm}{r^2}$

Etablissez en fonction de M , m , r et G :
la formule de la vitesse (v) du satellite

Surface de Compétence	%	r.bis	CM
1. $V = \sqrt{\frac{GM}{r^2}}$	4,1	-0,16	37,9
2. $V = \sqrt{\frac{GM}{r}}$	70,4	0,46	82,1
3. $V = \sqrt{\frac{F}{GM}}$	1,8	-0,13	25,3
4. $V = \sqrt{\frac{Fr}{m}}$	10,8	-0,22	61,8*
5. $V = \sqrt{\frac{GMm}{r}}$	5,6	-0,17	43,1
6. aucune	3	-0,14	27,9
7. toutes	0,3	-0,02	54,3*

Cette question part de la physique, mais constitue fondamentalement une manipulation mathématique. Le taux de réussite (70,4%) est "moyen" mais la certitude moyenne des réponses correctes (82,1%) est élevée, ce qui indique la confiance, ainsi que le r.bis de la réponse correcte (0,46), ce qui indique que c'est un groupe de très bons étudiants qui ont bien répondu. Le distracteur vedette (4), choisi par des étudiants faibles (r.bis = -0,22) l'est avec une certitude moyenne assez élevée (61,8%), plus élevée que celle qui accompagne les autres distracteurs, donc avec imprudence. La formule est correcte, mais on ne sait rien en faire : elle ne répond pas à la question. Ces étudiants n'ont pas bien lu la consigne et utilisé toutes les données. La fracture subjective nette vaut 20,3 %.

Question 8 : Calculez la fonction suivante :

$$\int x dx =$$

Surface de Compétence	%	r.bis	CM
1. $\ln(x) + C$	4,2	-0,23	43,8
2. $2x + \frac{2}{x}$	0,5	-0,06	31,7
3. $x^2 + C$	6	-0,18	58*
4. $\frac{x^2}{2} + C$	79,5	0,46	84,3
5. $2x + C$	2,9	-0,14	52,1*
6. aucune	5,3	-0,23	41,3
7. toutes	0,1	0	53,3*

Cette question est bien réussie (79,5%) avec une certitude moyenne élevée (84,3%), donc avec confiance, et ce par de bons étudiants (r.bis = 0,46). Il n'est pas étonnant que le distracteur vedette soit la solution 3 qui, elle aussi présente un carré, ce qui est logique dans ce calcul de surface. Sa certitude moyenne (58%), imprudente, est d'ailleurs plus élevée que la certitude moyenne des autres distracteurs. Sa fracture subjective est cependant nette (26,3%). Le choix du distracteur 5 s'est fait lui aussi de façon imprudente en moyenne.

Question 9 : Dans le tableau ci-dessous, on présente l'évolution d'un ensemble de données Y en fonction d'autres données X.
 X : 1 2 3 4 5 6 7
 Y : 2 4 8 16 32 64 128
 A quel type de série numérique, la progression des nombres de la série Y appartient-elle ?

Surface de Compétence				
	19,2	%	r.bis	CM
1. une série alternée	1,9	-0,08	20,9	
2. une série logarithmique	15,3	-0,03	54,2*	
3. une série géométrique	27,8	0,43	69,2	
4. une série arithmétique	31,4	-0,22	41,8	
5. une série cubique	3,7	-0,1	53,8*	
6. aucune	16,2	-0,07	46,7	
7. toutes	0,3	-0,05	40	

Cette question relève plus du vocabulaire mathématique (série et progression) que des mathématiques pures, mais ces notions sont utiles, notamment dans les cours de mathématique à l'université. Le taux de réponses correctes est très bas (27,8%), mais fournies par des étudiants forts (r.bis = 0,43) et avec une certitude moyenne (69,2%) passable pour ces réponses correctes, donc avec **confiance**.
 Le distracteur-vedette (4) est plus populaire (choisi par 31,4% des étudiants) que la réponse correcte, mais choisi par des étudiants nettement plus faibles (r.bis = -0,22) que ceux qui ont choisi d'autres distracteurs et avec une certitude moyenne (41,8%) nettement plus basse que celle de la réponse correcte, d'où une fracture subjective nette de 27,4%.
 Deux distracteurs sont choisis avec **imprudence**.

Question 10 : Dans le tableau ci-dessous, on présente l'évolution d'un ensemble de données Y en fonction d'autres données X. X : 1 2 3 4 5 6 7 Y : 2 4 8 16 32 64 128
 Laquelle des fonctions suivantes correspond le mieux à la relation entre les données X et Y :

Surface de Compétence				
	77,20%	%	r.bis	CM
1. $Y = X^2$	0,5	-0,11	63,3*	
2. $X = 2Y$	0,7	-0,09	60*	
3. $Y = \log(X)$	1,9	-0,16	37,1	
4. $Y = 2X$	1,7	-0,14	76,7*	
5. $Y = 2^X$	89,2	0,35	86,5	
6. aucune	4,8	-0,21	42,3	
7. toutes	0	XXXXX	XXXXX	

Cette question, proche de la précédente, ne fait pas intervenir le vocabulaire. Son taux de réussite (89,2%) est spectaculairement supérieur ainsi que la certitude moyenne (86,5%), donc une **confiance élevée**. Le distracteur-vedette (6. Aucune) est choisi par des étudiants faibles (r.bis = 0,36) et avec une certitude moyenne nettement plus basse (42,3%), donc avec **prudence**, d'où une fracture subjective très élevée : 44,2%.
 D'autres distracteurs sont par ailleurs choisis avec **imprudence**.

Question 11 : Réduisez ou développez les expressions suivantes : $(X^5)^3$

Surface de Compétence				
	77,10%	%	r.bis	CM
1. X^8	9,4	-0,2	77,3*	
2. X^{15}	88,7	0,26	86,9	
3. $3\log(X^5)$	0,4	-0,04	62*	
4. $\log(X^5) + \log(X^3)$	0,4	-0,07	70,9*	
5. $2X^4$	0,2	-0,12	66,7*	
6. aucune	0,2	-0,09	32	
7. toutes	0	XXXXX	XXXXX	

Le taux d'exactitude est élevé (88,7%) ainsi que la certitude moyenne (86,9%), donc avec **confiance**, d'où une Surface de Compétence importante (77,1%), et ce par de bons étudiants (r.bis = 0,26). Les erreurs se concentrent quasi sur un seul distracteur (1). Ce qui est préoccupant est à la fois la certitude élevée (77,3%), donc l'**imprudence**, de ceux qui ont choisi cette erreur grossière. Il s'agit d'étudiants faibles (R.bis = -0,20). La fracture subjective est très faible (9,6%).

Question 12 : Que vaut 1 m/s :				
Surface de Compétence				
	63,10%	%	r. bis	CM
1. 3,6 km/h		75	0,36	84,1
2. 1 (km/h) / 3,6		10,4	-0,11	68,5*
3. 10 km/h		0,4	-0,1	37,8
4. 0,36 km/h		4,4	-0,15	60,5*
5. 1 (km/h) / 0,36		1,6	-0,09	49,5
6. aucune		6,3	-0,24	42,9
7. toutes		0,2	-0,04	48,7

Trois étudiants sur quatre ont réussi (75%), avec une certitude élevée (84,1%), donc avec **confiance**, et ce sont de bons étudiants (r. bis = 0,36) au total du test.
Le distracteur-vedette (2) est choisi avec une certitude moyenne plus basse (68,5%) d'où une **fracture subjective** de 15,6% (modeste), mais cela n'en constitue pas moins une situation d'**imprudence**. La solution 3, aberrante, est -heureusement- très peu choisie, et avec la certitude moyenne la plus faible : 37,8%.

Question 13 : Parmi les affirmations ci-dessous, indique celle qui définit le mieux la notion de dérivée ?				
Surface de Compétence				
	33,30%	%	r. bis	CM
1. La limite de $y(x)$ quand x tend vers 0		12,4	-0,18	45,9
2. La dérivée représente la pente de la droite tangente à la fonction en l'origine		11,7	-0,16	48,2
3. La dérivée donne la pente de la fonction limite pour tout point de la fonction initiale		9,6	-0,09	42,2
4. La dérivée permet de calculer l'aire d'une surface délimitée par la fonction		4,5	-0,2	43,5
5. C'est la fonction qui donne la pente de la tangente à la fonction en tout point		51,7	0,43	64,4
6. aucune		5,3	-0,07	33,8
7. toutes		1,8	-0,03	43,9

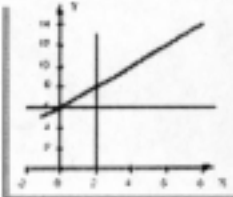
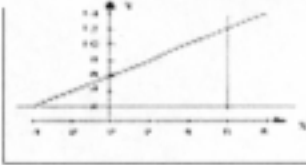
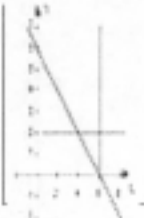
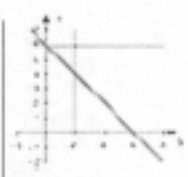
Un étudiant sur deux seulement (51,7%) a fourni la réponse correcte et avec une certitude moyenne modeste (64,4%), mais ce sont de bons étudiants (r. bis = 0,43). Trois distracteurs (1, 2 et 3) se disputent la vedette (choisis respectivement par 12,4%, par 11,7% et par 9,6% des étudiants), avec des taux de certitude bas et une **fracture subjective** modeste (18,5 %) avec la réponse correcte. Le choix du distracteur 2, qui a la certitude moyenne la plus élevée des distracteurs, est pardonnable mais reste fondamentalement faux.
Confiance et prudence sont conformes aux attentes.

Question 14 : Parmi les affirmations ci-dessous, indique celle qui définit le mieux la notion d'intégrale définie ?				
Surface de Compétence				
	37,30%	%	r. bis	CM
1. La surface totale comprise sous la courbe		5,6	-0,14	47,3
2. L'intégrale définie représente la pente de la droite tangente à la fonction en l'origine		0,9	-0,13	25,8
3. L'intégrale définie permet de calculer l'aire de la surface entre deux abscisses, l'axe des X et la fonction		55,6	0,39	67,1
4. L'intégrale est la fonction inverse de la dérivée		21,5	-0,26	58*
5. C'est la somme des surfaces élémentaires définies par la fonction d'origine		7,7	-0,01	54,2*
6. aucune		2,1	-0,06	40
7. toutes		3,7	-0,04	48,9

Un peu plus d'un étudiant sur deux (55,6%), ont fourni la réponse correcte, de bons étudiants (r. bis = 0,39), mais avec une certitude moyenne faible (67,1%), quoique relevant de la **confiance**. Un cinquième des étudiants (21,5%) ont choisi le distracteur 4, ce qui indique une confusion entre "inverse" et "réciproque". Ces étudiants, faibles (r. bis = -0,26), ont donné une certitude encore assez élevée (58%), **imprudente**, ce qui donne une **fracture subjective** très faible (9,1%). Dans la solution correcte (3), le terme "permet" devrait être remplacé par "représente".
Le distracteur 4 a lui aussi été choisi en moyenne avec **imprudence**.

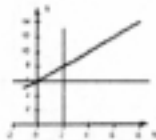
Question 15 : Identifiez le graphique qui représente les trois fonctions suivantes: $Y = X + 6$, $Y = 6$ et $X = 2$

Surface de Compétence 37,30%

	%	r.bis	CM
	87,5	0,38	85,2
	5	-0,2	74,1*
	0,9	-0,11	56,4
	2,5	-0,16	61,6*
	0,2	-0,03	72*
6. aucune	3,4	-0,21	53*
7. toutes	0,1	-0,02	0

La solution correcte a été choisie par une proportion très élevée (87,5%) des étudiants, forts (r.bis = 0,38) et avec une confiance élevée (certitude moyenne = 85,2%). Il n'est pas étonnant que le distracteur-vedette soit 2 (choisi avec une certitude moyenne élevée : 74,1% et une fracture subjective faible de 11,1 %) car les distracteurs 3 et 4 sont faux de façon bien plus flagrante (ceux qui les choisissent sont d'ailleurs moins sûrs (56,4 % et 61,6 %)). Dans l'énoncé de la question, l'expression "qui représente" devrait être remplacée par "où sont représentés". Des distracteurs sont choisis avec imprudence.

Question 16 : Quelle est la valeur de l'aire de la surface délimitée par les trois droites données ci-dessous (en unité de surface)

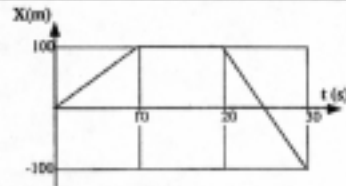


Surface de Compétence 49%

	%	r.bis	CM
1. 2	63,2	0,49	77,5
2. 4	11,1	-0,1	62,4*
3. 8	3,9	-0,19	33,6
4. 16	3,6	-0,17	33
5. 50	1,9	-0,09	64,3*
6. aucune	12,1	-0,24	44,2
7. toutes	0,2	-0,04	30

La solution correcte a été choisie par moins d'étudiants qu'à la question précédente, ce qui est normal, la question 16 étant "emboîtée" dans la question 15 qu'il fallait avoir réussie pour réussir la question 16. Elle a donc été réussie par 72,2% des 87,5% de ceux qui avaient réussi la question 5. Et ce avec une certitude élevée (77,5%), donc avec confiance, et par les étudiants les plus forts (r.bis = 0,49). Le distracteur-vedette (6. Aucune) a été choisi par les étudiants les plus faibles (r.bis = -0,24) et avec une certitude moyenne assez faible (44,2 %) d'où une fracture subjective élevée (de 33,3 %). Par contre, deux autres distracteurs ont été choisis avec imprudence.

Question 17 : La position X (exprimée en mètres) d'un mobile est donnée en fonction du temps t (exprimé en secondes) sur le graphique suivant,



que vaut la vitesse du mobile au temps. $t = 25$ s ?

Surface de Compétence

	14,30%	%	r.bis	CM
1. -20 m/s		20	0,36	71,8
2. 0 m/s		50,3	-0,35	66,5*
3. 5 m/s		1,5	-0,09	45,3
4. 10 m/s		2,9	-0,01	54,6*
5. 20 m/s		17,9	0,18	74*
6. aucune		5,7	-0,07	46,5
7. toutes		0,1	-0,02	50

Un cinquième des étudiants (20%) ont réussi, mais ce sont de bons étudiants ($r.bis = 0,36$), avec une certitude moyenne assez bonne (71,8%), donc avec **confiance**. La moitié des étudiants (50,3%) ont choisi le distracteur 2, avec une certitude élevée pour un distracteur (66,5%), donc avec **imprudence**, ce qui rend la **fracture subjective** tenue : 5,1 % pour des étudiants faibles ($r.bis = -0,35$). Le distracteur 5, par lequel on fait une erreur de signe, est choisie par près d'un cinquième des étudiants (17,9%), des étudiants plutôt bons ($r.bis = 0,18$) et une certitude moyenne (74%) tellement élevée pour un distracteur, donc une telle **imprudence**, que la **fracture subjective** aberrante entre ce distracteur et la réponse correcte est négative (-2,4%) !!!

Question 18 : Le tableau ci-dessous fournit quelques points des deux paraboles. Identifiez la ou les valeur(s) de X correspondant à d'éventuels points d'intersection des deux courbes.

$$Y_1 = -X^2 + 3X - 12 \quad \text{et} \quad Y_2 = X^2 + X - 36$$

X	-4	-3	-2	-1	0	1	2	3	4	5
Y1	-40	-30	-22	-16	-12	-10	-10	-12	-16	-22
Y2	-24	-30	-34	-36	-36	-34	-30	-24	-16	-6

Surface de Compétence

	51,40%	%	r.bis	CM
1. X = -3		11,6	-0,17	63,5*
2. X = -3 et X = 2		2,8	-0,17	44,8
3. X = 2 et X = 4		2,9	-0,15	20,5
4. X = 2		1,1	-0,12	29
5. X = -3 et X = 4		83,9	0,48	80,5
6. aucune		11,8	-0,19	49,6
7. toutes		0,6	-0,05	30,7

Près des deux tiers des étudiants (83,9%) ont choisi la solution correcte, de très bons étudiants ($r.bis = 0,48$) avec une certitude moyenne élevée (80,5%), donc avec une **confiance élevée**. Deux distracteurs (le 1 et le 6) se disputent la vedette (choisis par 11,6% et 11,8% des étudiants), par des étudiants du même niveau de faiblesse ($r.bis = -0,17$ et $-0,19$), mais avec des certitudes moyennes assez différentes : 63,5% (**imprudence**) et 49,6%. Cette dernière certitude moyenne est probablement plus faible que l'autre parce qu'elle porte sur le choix du distracteur "6. Aucune". La **fracture subjective** avec ce distracteur-vedette est de 30,9 % (élevée).

Question 19 : Trouvez directement une solution du système

$$\begin{cases} Y = X^2 - X - 6 \\ Y = X - 3 \end{cases}$$

Surface de Compétence

	45,20%	%	r.bis	CM
1. X = 0 et Y = 3		4	-0,14	56,8*
2. X = 2,2 et Y = -0,8		3,5	-0,12	47,3
3. X = 0 et Y = -3		8,4	-0,23	57*
4. X = 1,8 et Y = 1,2		1,7	-0,11	44,7
5. X = -1 et Y = -4		53,2	0,41	85
6. aucune		25,6	-0,11	60,1*
7. toutes		0,2	-0,04	40

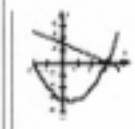
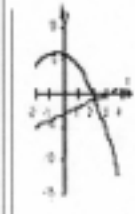
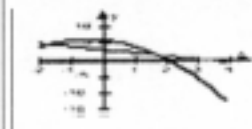
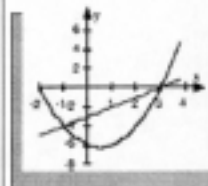
Un peu plus de la moitié des étudiants (53,2%) ont répondu correctement, de bons étudiants ($r.bis = 0,40$), avec une certitude moyenne élevée (85%), donc une **confiance forte**. Un quart des étudiants (25,6%) a choisi le distracteur "6. Aucune", des étudiants plutôt faibles ($r.bis = -0,11$), avec une certitude élevée pour la solution "Aucune" (60,1%), donc avec **imprudence**, comme d'ailleurs la solution 3. Il est assez étonnant que le taux de réponses correctes soit si bas pour un exercice que l'on peut faire "à l'envers", et tout spécialement la solution 3, choisie par 8,4% des étudiants (faibles il est vrai : le $r.bis = -0,23$) alors que l'on "voit" de suite que cela ne peut pas convenir. La **fracture subjective** nette est de 24,9 %.

Question 20 : Identifiez les graphes des deux fonctions suivantes

$$Y_1 = X^2 - X - 6$$

$$\text{et } Y_2 = X - 3$$

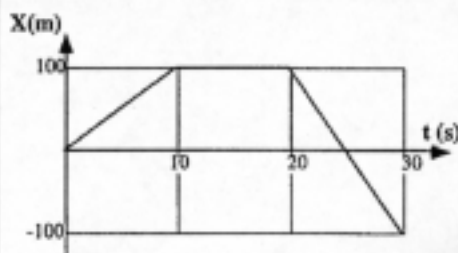
Surface de Compétence
68,40%



	%	r.bis	CM
	84,6	0,41	80,9
	1,1	-0,13	42,9
	4	-0,17	59,6*
	3,8	-0,19	59,8*
	0,2	-0,09	33,3
6. aucune	4,8	-0,2	48,1
7. toutes	0	XXXXX	XXXXX

Le taux d'exactitude est assez élevé (84,6%) ainsi que la certitude moyenne (80,9%), donc la confiance est élevée, et ce par des étudiants forts (r.bis = 0,41). Les erreurs sont dispersées sur trois distracteurs (3,4 et 6) avec une fracture subjective avoisinant toujours au moins les 20% (modeste à nette), bien que deux de ces distracteurs aient été choisis avec imprudence. La solution correcte peut être trouvée par élimination : on garde les solutions 1 et 4 car présentant des courbes ouvertes ; on garde ensuite celle dont la pente de la fonction linéaire est montante.

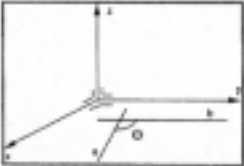
Question 21 : La position X (exprimée en mètres) d'un mobile est donnée en fonction du temps t (exprimé en secondes) sur le graphique suivant. Pour décrire le mieux possible le mouvement de ce mobile, on demande de compléter les phrases ci-dessous. Entre 10 et 20 secondes, le mobile ...



Surface de Compétence
44,3 %

	%	r.bis	CM
1. avance le long de l'axe des X positifs	2	-0,08	51,9*
2. avance le long de l'axe des X positifs à vitesse constante	35,5	-0,4	67,8*
3. avance le long de l'axe des X positifs en accélérant	1,6	-0,12	49,8
4. recule le long de l'axe des X positifs à vitesse constante	0,8	-0,06	47,6
5. est immobile	55,8	0,49	79,4
6. aucune	3	-0,08	46,8

Le taux de réussite est faible (55,8%), par d'excellents étudiants (r.bis = 0,49), avec une certitude moyenne élevée (79,4%), donc une confiance élevée. Le distracteur vedette (2) attire le tiers des étudiants (35,5%), très faibles (r.bis = -0,40) avec une certitude assez élevée : (67,8%), une imprudence élevée, nettement plus élevée que pour les autres distracteurs. La fracture subjective faible n'est que de 11,6%.

Question 22 : Comment peut-on qualifier l'angle θ formé par les droites a et b situées dans le plan (x, y) de la figure suivante ?			
			
Surface de Compétence	%	r.bis	CM
45,10%			
1. Aigu	64,1	0,49	70,4
2. Droit	8,4	-0,22	54,8 [*]
3. Obtus	19,5	-0,35	56,5 [*]
4. Alterne	0,7	-0,09	24,7
5. Interne	2,2	-0,1	38,9
6. aucune	3	-0,03	32,5
7. toutes	0,2	0	40

La solution correcte a été fournie par près de deux tiers (64,1%) des étudiants, d'excellents étudiants (r.bis = 0,49), avec une certitude moyenne passable (70,4%), signe de **confiance**. Le distracteur vedette (3. Obtus), conçu comme un piège (visuel) a en effet fonctionné comme un piège puisqu'il a attiré près d'un cinquième des étudiants (19,5%), mais des étudiants faibles (r.bis = -0,35), avec une certitude moyenne de 56,4%, donc avec **imprudence**, soit une fracture subjective faible de 13,9%.

Bilan

Indicateurs psychométriques :

Les **indices de discrimination (r.bis)** sont tous conformes aux attentes, que ce soit pour les 22 solutions correctes ou pour les 132 distracteurs. Aucune critique ne peut donc être adressée à ces questions quant à leur cohérence interne avec l'ensemble de l'épreuve.

Indicateurs édumétriques :

La **confiance** est observée dans TOUTES les 22 réponses correctes aux 22 questions.

La **prudence totale** est observée seulement pour deux questions : 6 et 13. Pour toutes les autres, des distracteurs sont fournis avec **imprudence**. Il s'agirait, dans tous ces cas, de « rectifier » ces idées fausses.

Chapitre 16

Analyses des réponses au check-up de Physique

P. Chapelle, FPMs

D. Leclercq, ULg

2520 étudiants

Seuil r.bis = 0,32

Question 1 : La Terre exerce une force d'attraction sur la Lune. Pourquoi, dès lors, la Lune ne tombe-t-elle pas sur la Terre ?

Surface de Compétence		%	r.bis	CM
28,80%				
1. parce qu'elle tourne autour de la terre		44,1	0,51	65,2
2. parce qu'elle est trop légère		1,2	-0,03	37,4
3. parce que l'attraction solaire compense l'attraction terrestre		24,3	-0,27	45,8
4. parce qu'elle n'est pas freinée par l'atmosphère terrestre		3	-0,1	46,1
5. parce qu'elle est repoussée par la masse de la Terre		13,4	-0,2	45,6
6. aucune		11,3	-0,1	45,1
7. toutes		0,4	-0,04	20

Moins d'un étudiant sur deux (44,1 %) ont fourni la réponse correcte, mais ce sont les meilleurs (r.bis = 0,51) avec une certitude moyenne de 65,2 %. Le distracteur-vedette (3) a attiré près d'un étudiant sur quatre (24,3 %), les plus faibles (r.bis = -0,27) et avec une certitude moyenne faible elle aussi (45,8 %), donc une fracture subjective de 19,4 % (modeste). Choisir cette solution, c'est oublier que la lune tourne ! Il est étonnant que la solution 5, l'erreur la plus grave, soit choisie par 13,4 % des étudiants ! **Confiance et prudence** sont conformes aux attentes.

Question 2 : Q2. Un mobile se déplace sur une trajectoire rectiligne. On relève la vitesse v à différents instants t :

t(s) : 0 2 4 6 8 10
V(m/s) 0 5 10 15 20 25
Calculez l'accélération en m/s²

Surface de Compétence		%	r.bis	CM
56,60%				
1. - 2,5 m/s ²		0,3	-0,09	30
2. 0 m/s ²		5	-0,09	64,9*
3. 2,5 m/s ²		79,1	0,4	71,5
4. 5 m/s ²		7,8	-0,22	46,2
5. 7,5 m/s ²		0,6	-0,04	43,8
6. aucune		3,5	-0,18	33,8
7. toutes		0,3	-0,05	20

Quatre étudiants sur cinq (79,1 %) ont choisi la réponse correcte, les meilleurs (r.bis = 0,4) avec une certitude moyenne assez bonne (71,5 %). L'erreur la plus grave (choisir la solution 2) n'est pas (5 %) la plus populaire (c'est la 4 choisie à 7,8%), mais celle qui est donnée avec la certitude la plus élevée (64,9 %). La fracture subjective avec le distracteur-vedette (4) est 25,3 % (nette). **Confiance et prudence** sont conformes aux attentes sauf pour le distracteur 2, choisi avec **imprudence**.

Question 3 : On considère un dl (décilitre) d'eau dans des conditions normales de température et de pression. Calculez la masse.

Surface de Compétence		%	r.bis	CM
54,10%				
1. 100,0 g		70,8	0,39	76,4
2. 100,0 N		2	-0,12	52,9*
3. 10,0 g		8,7	-0,2	57,2*
4. 1,0 Kg		11,4	-0,11	75,3*
5. 100,0 Pa		0,4	-0,09	28,9
6. aucune		4,4	-0,19	32,5
7. toutes		0	XXXXXX	XXXXXX

Le taux de réussite (70,8 %) est assez élevé, ainsi que la certitude moyenne (76,4 %), signe de **confiance** et est produit par de bons étudiants (r.bis = 0,39). Le distracteur-vedette (4) est choisi avec une certitude très élevée pour une erreur (75,3 %), sans doute par confusion entre dl et dm³. La fracture subjective est insignifiante (1,1 %). Trois distracteurs ont été choisis avec **imprudence**.

Question 4 : Quelle est la cause majeure d'éclatement des canalisations lorsque l'eau se congèle ?

Surface de Compétence		%	r.bis	CM
70,60%				
1. Le froid fragilise les canalisations		1,5	-0,12	47,6
2. En se congelant l'eau se dilate		86,5	0,37	81,6
3. Le frottement de la glace fragilise les canalisations		7,2	-0,25	55,4
4. En se congelant, l'eau se contracte		7,2	-0,25	56,4*
5. Le froid dilate les canalisations		1,7	-0,13	42,3
6. aucune		1,3	-0,09	41,3
7. toutes		0,1	-0,04	13,3

Le taux de réussite est très bon (86,5 %) ainsi que la certitude moyenne (81,6 %) des réponses correctes, signe de **confiance**. La fracture subjective (26,2 %) est nette avec le distracteur-vedette (4) qui constitue une exception, car en général, les corps se contractent sous le froid. Le distracteur-vedette est choisi avec **imprudence**.

Question 5 : Un objet est très éloigné d'une lentille mince convergente de 10 cm de distance focale. A quelle distance de la lentille se forme l'image de l'objet ?

Surface de Compétence				
19%		%	r.bis	CM
1. 0 cm		3,3	-0,06	23,6
2. 10 cm		41,3	0,45	46*
3. 100 cm		6,7	-0,08	22,8
4. à l'infini		10,1	-0,08	31,8
5. ça dépend de la taille de l'objet		21,2	-0,23	32,3
6. aucune		8,8	-0,09	22,4
7. toutes		0,2	-0,01	12

Le taux de réponses correctes est bas (41,3 %) ainsi que la certitude moyenne de ces réponses correctes (46 %), donc avec un **manque de confiance**, mais ce taux est produit par les meilleurs étudiants (r.bis = 0,45). Le distracteur vedette (5) qui a attiré un étudiant sur cinq (21,2 %) et les plus faibles (r.bis = -0,23) dénote une confusion entre taille et position de l'image, avec une certitude moyenne faible (32,3 %) et une **fracture subjective** de 13,7 % (faible). La **prudence** est conforme aux attentes.

Question 6 : Une personne se trouve dans un ascenseur qui est accéléré vers le haut. La force exercée par la personne sur le plancher est :

Surface de Compétence				
33,90%		%	r.bis	CM
1. nulle		3,4	-0,15	42,8
2. inférieure au poids de la personne		8,5	-0,11	44,9
3. égale au poids de la personne		26,2	-0,21	51,7*
4. supérieure au poids de la personne		51,9	0,46	65,4
5. sans rapport avec le poids de la personne		6,4	-0,17	37,8
6. aucune		1,7	-0,13	34,5
7. toutes		0	XXXXXX	XXXXXX

Un étudiant sur deux a réussi (51,9 %) avec **confiance**, bien que la certitude moyenne soit modeste (65,4 %), mais ce sont les meilleurs étudiants (r.bis = 0,46). Le choix du distracteur vedette (3) par un étudiant sur quatre (26,2 %) se fait par les moins bons étudiants (r.bis = -0,21) et avec une certitude trop élevée (51,7 %), soit avec **imprudence**, et donc une **fracture subjective** faible : 13,7 %.

Question 7 : On élève une masse m à une hauteur h du sol et on l'abandonne librement. Déterminer l'expression littérale de la vitesse v lorsque la masse retombe sur le sol.

Surface de Compétence				
17,10%		%	r.bis	CM
1. $v = \frac{1}{2}mv^2$		7,5	-0,08	54,3*
2. $v = \sqrt{2gh}$		24,1	0,43	71
3. $v = mgh$		48	-0,2	55,2*
4. $V = mc^2$		2,3	-0,09	39,6
5. $V = \frac{(mh)}{g}$		5,4	-0,11	29,6
6. aucune		9,1	0,01	42,8
7. toutes		0,2	-0,06	20

Un étudiant sur quatre (24,1 %) seulement a choisi la solution correcte, mais ce sont les meilleurs étudiants (r.bis = 0,43), avec **confiance** : la certitude moyenne est bonne (71 %). La majorité (48 %) a choisi le distracteur 3, avec une certitude moyenne inférieure (55,2 %) mais encore **imprudente** et une **fracture subjective** de 15,8 % (modeste). Ces étudiants sont les plus faibles (r.bis = -0,20). Ces étudiants se sont sans doute souvenus d'une formule proche. L'absence de m dans la solution correcte alors qu'il est présent dans l'amorce et dans les distracteurs constitue un piège. Le distracteur 1 est choisi avec **imprudence**.

Question 8 : Une résistance R de 10 ohms est reliée à une batterie de 12 volts. Calculer le courant dans la résistance.				
Surface de Compétence				
	49,1 %	%	r.bis	CM
1. 1,2 A		74,5	0,46	66,3
2. 1,2 w		2,9	-0,13	30,8
3. 12 A		4	-0,14	28,9
4. 1,2 W		6,8	-0,22	29,9
5. 1,2 C		2,6	-0,09	36,2
6. aucune		3,5	-0,14	24,7
7. toutes		0	XXXXX	XXXXX

Trois étudiants sur quatre (74,5 %) ont choisi la solution correcte, avec une certitude moyenne modeste (66,3 %), mais ce sont de très bons étudiants (r.bis = 0,46). La fracture subjective avec le distracteur-vedette 4 (qui résulte d'une confusion d'unité) est élevée : 36,4 %. **Confiance et prudence** sont conformes aux attentes.

Question 9 : Lors d'un orage, pourquoi percevons-nous le bruit du tonnerre après avoir vu l'éclair ?				
Surface de Compétence				
	67,90%	%	r.bis	CM
1. notre oreille est moins sensible que notre oeil		0,4	-0,09	22
2. la lumière de l'éclair se propage plus vite que le tonnerre		89,8	0,34	75,6
3. le bruit (tonnerre) se produit toujours après l'éclair		3,8	-0,19	57,5*
5. la pluie accélère la transmission lumineuse		1,2	-0,1	19,3
6. aucune		2,3	-0,12	45,4
7. toutes		0,2	-0,04	64*

Le taux de réussite est élevé (89,8 %), mais bizarrement, la certitude moyenne l'est moins (75,6 %), bien qu'elle relève de la confiance. La fracture subjective est modeste (18,1 %) avec le distracteur-vedette 3, erreur grave, peu choisie (3,8 %) et avec imprudence.

Question 10 : Pourquoi une bobine parcourue par un courant attire-t-elle une aiguille aimantée ?				
Surface de Compétence				
	67,40%	%	r.bis	CM
1. la bobine crée un champ magnétique		91,1	0,34	74
2. la bobine chauffe		0,6	-0,1	52*
3. la bobine crée un appel d'air		0,2	-0,1	24
4. parce que les corps chauds et les corps froids s'attirent		0,2	-0,06	26
5. par interaction électrostatique		6,4	-0,25	46
6. aucune		0,3	-0,07	17,5
7. toutes		0	XXXXX	XXXXX

Très bon taux de réussite (91,1 %), par de bons étudiants (r.bis = 0,34) mais avec confiance, bien que la certitude moyenne soit décevante (74 %). La fracture subjective avec le distracteur-vedette (5) est nette : 28 %. La prudence est conforme aux attentes sauf pour le distracteur 2, très peu choisi.

Bilan

Indicateurs psychométriques :

Les **indices de discrimination (r.bis)** sont tous conformes aux attentes, que ce soit pour les 22 solutions correctes ou pour les 132 distracteurs. Aucune critique ne peut donc être adressée à ces questions quant à leur cohérence interne avec l'ensemble de l'épreuve.

Indicateurs éducatifs :

La **confiance** est observée dans 7 des 10 réponses correctes.

La **prudence totale** est observée pour 3 questions. Pour toutes les autres, des distracteurs sont fournis avec **imprudence** (10 des 60 distracteurs). Il s'agirait, dans tous ces cas, de « rectifier » ces idées fausses.

Chapitre 17

Analyses des réponses au check-up de Chimie

A. Cornélis

D. Leclercq, Président du Groupe de Travail CIUF « Réussite en candidatures »

2522 étudiants

Seuil r.bis = 0,353

Question 1 : Un composé gazeux a pour formule moléculaire $(CH_3)_2CCH_2$. Vous en concluez que chaque molécule de ce composé comporte :

Surface de Compétence			
73,60%	%	r.bis	CM
1. cinq atomes d'hydrogène	1,3	-0,05	63,6 *
2. six atomes d'hydrogène	0,8	-0,11	48,4
3. sept atomes d'hydrogène	3,3	-0,16	70,2 *
4. huit atomes d'hydrogène	89,7	0,32	82
5. neuf atomes d'hydrogène	1	-0,08	64,6 *
6. aucune	3,3	-0,19	46,7
7. toutes	0,1	-0,02	100 *

La solution 5 rend la réponse « Toutes » absurde car comportant un atome de plus que le total. Le taux de réponses correctes est élevé (89,7 %) ainsi que la confiance (la certitude moyenne qui accompagne les réponses correctes est 82 %) et les étudiants sont les meilleurs (r.bis = 0,32). La fracture subjective avec le distracteur-vedette (3) n'est que de 11,8 % (modeste). La majorité des distracteurs sont choisis avec imprudence (certitude supérieure à 50 %).

Question 2 : Parmi les équations chimiques suivantes, quelle est celle où le (les) coefficient(s) stœchiométrique x a (ont) pour valeur 3

Surface de Compétence			
73,90%	%	r.bis	CM
1. $H_2 + Cl_2 \rightarrow x HCl$	0,4	-0,08	30,9
2. $N_2 + x H_2 \rightarrow 2 NH_3$	87,1	0,39	84,8
3. $x O_3 \rightarrow 3 O_2$	3,9	-0,2	46,7
4. $x NO + Cl_2 \rightarrow x NOCl$	0,6	-0,06	57,5 *
5. $x SO_2 + O_2 \rightarrow x SO_3$	1,3	-0,14	32,1
6. aucune	4,6	-0,2	51,5 *
7. toutes	0,2	-0,04	0

Le taux de réussite est élevé (87,1 %) ainsi que la confiance (certitude moyenne des réponses correctes = 84,8 %), données par les plus forts des étudiants (r.bis = 0,39). La fracture subjective avec le distracteur-vedette (6. Aucune) est élevée : 33,3 %. Les deux distracteurs les plus populaires sont choisis avec imprudence (certitude supérieure à 50 %).

Question 3 : La réaction de décomposition de N_2O_4 gazeux en NO_2 donne lieu à l'établissement d'un équilibre chimique. La réaction est endothermique dans le sens de la décomposition. Ce caractère endothermique est mis en évidence par le comportement suivant :

Surface de Compétence			
25,30%	%	r.bis	CM
1. la proportion de NO_2 dans le mélange à l'équilibre augmente quand la température augmente	44,8	0,48	56,5
2. deux moles de NO_2 sont produites par mole de N_2O_4 consommée	16,6	-0,18	41,8
3. la valeur de la constante d'équilibre de cette réaction est toujours positive à toute température	8,4	-0,12	37
4. NO_2 est plus intensément coloré en brun que N_2O_4	2,1	-0,08	24,8
5. la masse molaire de N_2O_4 est double de celle de NO_2	4,4	-0,14	34
6. aucune	11,6	-0,16	32,4
7. toutes	6,2	0,02	59,5 *

Toutes les solutions sont correctes dans l'absolu, mais seule la solution 1 est de l'ordre de la mise en évidence. Ce sont bien les meilleurs étudiants qui l'ont choisie (r.bis = 0,48), mais ils ne sont que 44,8 % et leur certitude moyenne n'est que de 56,5 % (confiance). La fracture subjective avec le distracteur-vedette (2) n'est que de 14,7 % (faible). La solution 4 est aussi une sorte de mise en évidence. Elle a été choisie par le groupe le moins faible (r.bis = -0,08) des étudiants qui se sont trompés. La solution parle de température ; on est donc dans le contexte de la question ; les étudiants qui l'ont choisie ne sont pas très faibles (r.bis = -0,12). **Confiance et prudence** sont présents partout où on les attend, à une exception près : la solution 7 !

Question 4 : Une solution aqueuse doit contenir les ions ci-dessous aux concentrations indiquées :

Ion	Concentration (mol/L)
Mg ²⁺	0,020
K ⁺	0,010
Na ⁺	0,030
Cl ⁻	0,040
(SO ₄) ²⁻	0,005
(NO ₃) ⁻	0,030

La façon la plus simple de réaliser cette solution est de dissoudre dans l'eau les quantités adéquates de :

Surface de Compétence	%	r.bis	CM
18,40%			
1. Mg(NO ₃) ₂ , K ₂ SO ₄ et NaCl	10,9	-0,12	35,3
2. MgSO ₄ , KCl et NaNO ₃	16,9	-0,1	46,7
3. Mg(NO ₃) ₂ , KCl et Na ₂ SO ₄	8,2	-0,12	27,3
4. MgCl ₂ , K ₂ SO ₄ et NaNO ₃	29,7	0,47	62
5. MgCl ₂ , KNO ₃ et Na ₂ SO ₄	6,6	-0,08	36,3
6. aucune	10,9	-0,05	33,1
7. toutes	8,2	-0,05	42,4

Cette question est la plus difficile des 8 questions du check-up de chimie : la réussite n'est que de 29,7 % mais par les meilleurs étudiants (r.bis = 0,47) et avec une certitude moyenne acceptable (62 %). Réussir cette question suppose que l'étudiant maîtrise la matière testée dans les questions 1 et 2 qui précèdent, et, en plus, la notion de concentration. D'où la grande sélectivité de cette question. La fracture subjective est de 15,3 % (modeste). **Confiance et prudence** sont présents partout où on les attend !

Question 5 : On dispose d'un ensemble de solutions aqueuses contenant chacune une et une seule des substances ci-après, toujours à la concentration de 0,010 mol/L :

- NH₃ (ammoniac)
- CH₃-CO₂H (acide acétique)
- HCl (chlorure d'hydrogène)
- KBr (bromure de potassium)
- NaOH (hydroxyde de sodium)

En ce qui concerne les valeurs de leur pH, ces solutions se classent dans l'ordre:

Surface de Compétence	%	r.bis	CM
38,30%			
1. a < c < e < b < d	3,2	-0,11	32,6
2. e < c < a < b < d	5,6	-0,13	34,3
3. d < c < a < e < b	8,9	-0,19	27,9
4. c < b < d < a < e	56,2	0,54	68,2
5. b < c < a < e < d	12,6	-0,22	44,5
6. aucune	8,6	-0,17	36,5
7. toutes	0,2	-0,02	8

On peut être déçu par le taux d'exactitude faible (56,20 %) ainsi que la certitude moyenne peu élevée (68,2 %) de cette question relevant de la matière de l'enseignement secondaire. Ce sont les meilleurs étudiants (r.bis = 0,54) qui ont fourni la réponse correcte. La fracture subjective est assez nette (23,7 %). **Confiance et prudence** sont présents partout où on les attend !

Question 6 : Une émulsion est une dispersion de particules très fines d'une substance dans une autre, liquide. Le " Petit Robert ", édition 1993, définit la margarine comme une " émulsion de corps gras alimentaires (surtout de graisses végétales) et d'eau ".

De la définition ci-avant, il ressort que la margarine est

Surface de Compétence	%	r.bis	CM
25,20%			
1. un corps pur	0,8	-0,06	29
2. une solution aqueuse	9,3	-0,1	48
3. un mélange homogène	34,3	-0,19	49,4
4. un corps simple	1,6	-0,08	25,5
5. un mélange hétérogène	48,1	0,37	52,3
6. aucune	2,8	-0,1	33,5
7. toutes	0,1	0,01	25,7

Cette question exige que les étudiants sachent faire la distinction entre corps pur et corps homogène, notions que le langage commun (ex : lait homogénéisé) tend à rendre confuses. Moins de la moitié des étudiants (48,1 %) ont fourni la réponse correcte, avec une certitude décevante (52,3 %), bien qu'il s'agisse de bons étudiants (r.bis = 0,37). La fracture subjective est insignifiante : 2,9 %. **Confiance et prudence** sont présents partout où on les attend !

Question 7 : Parmi les composés ci-dessous, quel est celui dont un des éléments constitutifs est dans un état caractérisé par un nombre d'oxydation (+III) ?

Surface de Compétence			
23%	%	r.bis	CM
1. MgSO ₄	3,3	-0,1	35,4
2. Na ₂ CO ₃	6,6	-0,1	33,1
3. (NH ₄) ₃ PO ₄	15,1	-0,13	43,7
4. Al ₂ (SO ₄) ₃	35,9	0,45	64
5. NH ₃	19,5	-0,12	48,8
6. aucune	11,1	-0,09	32,3
7. toutes	1,7	-0,05	36,8

Pour répondre correctement à cette question, il faut avoir une bonne vision de la structure atomique. Un peu plus d'un étudiant sur trois (35,9 %) ont réussi, et ce sont les meilleurs (r.bis = 0,45). La fracture subjective avec le distracteur-vedette (5) est de 15,2 % (modeste). **Confiance et prudence** sont présents partout où on les attend !

Question 8 : La matière grasse constituant la margarine est le plus souvent obtenue par une modification chimique des huiles végétales, auxquelles on fait subir une réaction appelée « hydrogénation catalytique ». Cette réaction est représentée par l'équation générale :
 $R-CH=CH-R' + H_2 \rightarrow R-CH_2-CH_2-R'$
 où R et R' représentent des chaînons et fonctions organiques divers.
 De l'équation chimique ci-dessus, on peut conclure que la réaction d'hydrogénation des huiles végétales est un exemple de réaction :

Surface de Compétence			
26,40%	%	r.bis	CM
1. d'hydrolyse	12,3	-0,2	31,1
2. d'addition	45,6	0,52	56,6
3. de polymérisation	18,7	-0,22	33,7
4. d'élimination	2,9	-0,1	27
5. de substitution	8,5	-0,07	34,9
6. aucune	5,7	-0,09	29,4
7. toutes	0,2	-0,03	25

Moins d'un étudiant sur deux (46,6 %) ont réussi cette question, avec une certitude faible (56,6 %), mais ce sont de très bons étudiants (r.bis = 0,52). La fracture subjective avec le distracteur-vedette (3) est nette : 22,9 % et toujours supérieure à 20 % avec les autres distracteurs. **Confiance et prudence** sont présents partout où on les attend !

Bilan

Indices de discrimination (r.pbis) :

Sauf pour la question 1 (dont le r.bis vaut 0,32), toutes les réponses correctes ont un indice de discrimination positif et supérieur à la valeur-repère (0,35).

Confiance et prudence :

Toutes ont un indice de confiance satisfaisant (degré de certitude moyen accompagnant la réponse correcte inférieur à 50 %). Les indices de prudence des distracteurs sont satisfaisants (quand la certitude moyenne accompagnant cette erreur est inférieure à 50 %) pour 41 des 48 distracteurs.

Chapitre 18


Analyses des réponses au check-up de Biologie

*J.-C. Verhaege, ULB
D. Leclercq, ULg*

N = 2 526 étudiants

r.bis repère = 0,32

Question n° 1 : Cytologie
 Sur cette photo et le schéma qui l'accompagne, donne le nom de l'organite indiqué par le numéro 2.



Surface de Compétence			
28,30%	%	r.bis	CM
1. Appareil de Golgi	9,7	-0,18	34,9
2. Centriole	8,6	-0,12	38,6
3. Vacuole	12,7	-0,13	40,2
4. Ribosome	12,8	-0,17	37,9
5. Mitochondrie	46	0,45	61,5
6. aucune	6,8	-0,04	47,1
7. toutes	0,2	-0,02	40

Le taux de RC est de 46 %, ce qui est faible, mais la réponse correcte est donnée par des étudiants forts (r.bis = 0,45), avec une certitude moyenne modeste (61,5 %). Cette matière est au programme de 5e. Les erreurs sont équitablement réparties. Quand l'étudiant se trompe, il choisit n'importe quel autre organite ; il n'y a pas de difficulté particulière. La fracture subjective avec les distracteurs est d'environ 20 % (23,6 %, soit nette, avec le distracteur-vedette), une ampleur assez caractéristique des questions de connaissance de mémoire. **Prudence et confiance** sont présents partout où on les attend.

Question n°2 : Génétique - biosynthèse des protéines
 La production d'une molécule d'ARN à partir d'un segment d'ADN est appelée :

Surface de Compétence			
37,60%	%	r.bis	CM
1. transcription	59,1	0,46	63,6
2. traduction	15,8	-0,19	51,3*
3. épissage de l'ARN	2	-0,13	33,2
4. réplication	14,6	-0,2	52,4*
5. recombinaison	2,7	-0,12	26,6
6. aucune	3	-0,12	32,5
7. toutes	0,4	-0,03	40

Le taux de RC est de 59,1 %, ce qui est relativement satisfaisant et le fait de bons étudiants. (r.bis = 0,46). Cette matière est au programme de 6°. Les distracteurs-vedettes « 2. Traduction » (15,8 %) et « 4. Réplication » (14,6 %), ont attiré 3 étudiants sur 10, ce qui est logique, tous ces processus étant liés à la transcription, avec une certitude moyenne inférieure de 10 % (fracture subjective très faible) à celle de la réponse correcte (12,3 % avec le distracteur-vedette), mais produit par de bons étudiants. **Prudence et confiance** sont présents partout où on les attend.

Question n° 3 : Embryologie - différenciation cellulaire.
Apparemment, nos cellules musculaires diffèrent de nos cellules nerveuses surtout parce qu'elles :

Surface de Compétence		%	r.bis	CM
17,20%				
1. expriment des gènes différents		28,5	0,5	60,4
2. contiennent des gènes différents		12,7	-0,14	34,7
3. utilisent des codes génétiques différents		19,4	-0,15	36,5
4. possèdent des ribosomes qui leur sont propres		13,7	-0,12	30,1
5. possèdent des chromosomes différents		5,5	-0,11	31,6
6. aucune		13,3	-0,03	38
7. toutes		1,9	-0,07	37,4

Le taux de RC est de 28,5%, ce qui est très faible (r.bis = 0,5). L'embryologie est au programme de 5e mais l'actualité (le "clonage" de Dolly) a récemment attiré l'attention sur le fait que toute cellule du corps renferme toute l'information génétique nécessaire à la formation d'un individu. Le distracteur-vedette est « 3. Utilisent des codes génétiques différents » est choisi par près d'un étudiant sur cinq (19,4 %), mais avec une certitude moyenne faible (36,5 %) et par des étudiants faibles à ce test (r.bis = -0,15). La fracture subjective est nette (23,9 %). Or le code génétique est au programme de 6e et sa principale caractéristique est d'être universel à tous les êtres vivants. C'est d'ailleurs ce qui permet de fabriquer des Organismes Génétiquement Modifiés (OGM). Ceci est donc a fortiori vrai pour toutes les cellules d'un être vivant. **Prudence et confiance** sont présents partout où on les attend.

Question n° 4 : Reproduction humaine
La fécondation de l'ovule chez l'Humain se produit le plus souvent dans :

Surface de Compétence		%	r.bis	CM
32,20%				
1. le vagin		1,5	-0,05	52,3*
2. l'ovaire		11,9	-0,21	54,6*
3. l'utérus		31,7	-0,29	59,4*
4. la trompe utérine		49,2	0,48	65,4
5. le conduit déférent		2,5	-0,03	51,6*
6. aucune		0,9	-0,08	35,5
7. toutes		0,2	-0,03	35

Le taux de RC est de 49,2 %, ce qui est faible mais avec une confiance : la certitude moyenne est assez bonne (65,4 %) et par de bons étudiants (r.bis = 0,48). La fracture subjective est très faible (6 % seulement). Le distracteur-vedette est « 3. dans l'utérus » ce qui est logique car après la bonne réponse (« 4. fécondation dans la trompe utérine »), c'est celle qui paraît la plus vraisemblable, quoiqu'elle indique une confusion entre fécondation et nidification. Par contre les 11,9 % de choix pour la fécondation dans l'ovaire et surtout les 2,5 % pour la fécondation dans le canal déférent interpellent d'autant plus que la certitude moyenne (supérieure à 50 %) est élevée pour des réponses incorrectes. La majorité des distracteurs sont choisis avec **imprudence** (certitudes moyennes supérieures à 50 %).

Question n° 5 : Écologie - niche écologique
La niche écologique est :

Surface de Compétence		%	r.bis	CM
8,20%				
1. la place de la population dans un ensemble de variables physiques, chimiques et climatiques		46,9	0,02	41,4
2. l'ensemble des adaptations morphologiques de la population à son environnement		11	-0,1	30,3
3. la place de la population dans une communauté, c'est-à-dire l'ensemble des relations biologiques		21,3	0,24	38,7*
4. la place de chaque individu au sein de sa population		8,4	-0,12	29,5
6. aucune		5,6	-0,05	26,5
7. toutes		0,8	0	32

Le taux de RC (21,3 %) est très faible et avec, en moyenne, un **manque de confiance** : la certitude moyenne = 38,7 % ; la réponse correcte est le fait d'étudiants plutôt meilleurs que la moyenne (r.bis = 0,24). C'est un concept central en écologie et l'étude des écosystèmes et des relations entre êtres vivants est explicitement prévue au programme de 6e Le concept de « niche écologique » est d'autant plus important qu'il fait le lien entre l'écologie et les théories de l'évolution. Il existe plusieurs définitions de la niche écologique, certaines incluant l'habitat dans la niche. Celle que nous avons testée nous semble la plus adéquate à l'enseignement dans le secondaire mais nous reconnaissons qu'il s'agit là d'un avis personnel. Le distracteur-vedette (1) (46,9%) est choisi. La **prudence** est présente là où on l'attend. La fracture subjective est négative : -2,7% (aberrante).

Question n° 6 : Évolution

Lorsque l'on verse un antibiotique dans un flacon contenant un très grand nombre de bactéries, cette population peut devenir résistante à cet antibiotique. Ceci est dû au fait :

Surface de Compétence		%	r.bis	CM
28,20%				
1. que l'antibiotique a provoqué des mutations donnant la résistance à cet antibiotique chez toutes les bactéries		9,4	-0,17	46,5
2. que l'antibiotique a provoqué des mutations donnant la résistance à cet antibiotique chez certaines bactéries; celles-ci se sont alors multipliées aux dépens des bactéries normales après l'ajout de l'antibiotique		51,8	0,29	54,4
3. que des bactéries résistantes aux antibiotiques existaient avant que l'on verse cet antibiotique et se sont multipliées aux dépens des bactéries normales après cette opération		28,2	-0,05	60,9*
4. que toutes les bactéries du flacon étaient résistantes aux antibiotiques		3,3	-0,11	32,3
		0,2	-0,05	48
6. aucune		1,5	-0,07	39,5
7. toutes		2,2	-0,07	39,3

Un étudiant sur deux a fourni la RC (51,8 %), ce qui est faible au regard de l'importance de l'enjeu ; la réponse correcte est fournie par de bons étudiants (r.bis = 0,29) avec **confiance** : la certitude moyenne est (légèrement) supérieure à 50%.
Le r.bis montre que la compréhension des lois de l'évolution est mal corrélée aux résultats généraux obtenus au test. L'évolution est prévue au programme de 6e et les arguments tirés de la résistance des bactéries sont explicitement cités pour le cours de biologie. Le résultat du test est faible alors que ces arguments sont incontournables aujourd'hui. Le distracteur-vedette (la théorie de Darwin) est fourni avec une certitude supérieure (60,9 %) à celle de la réponse correcte (54,4 %), ce qui donne une **fracture subjective négative** : -6,5 % (aberrante). Le distracteur-vedette a été choisi avec **imprudence** (certitude moyenne supérieure à 50 %).

Question n° 7 : Écologie – photosynthèse

Laquelle de ces formules représente le bilan chimique de la photosynthèse :

Surface de Compétence		%	r.bis	CM
48,50%				
1. $C_6H_{12}O_6 \rightarrow 2CH_3CH_2OH + 2CO_2$		3,5	-0,13	44,5
2. $C_6H_{12}O_6 \rightarrow 2CH_2CHOHCOOH$		1,8	-0,11	47,6
3. $6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2$		66,1	0,46	73,4
4. $C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O$		21,6	-0,31	63*
		0,2	-0,04	52*
6. aucune		3,9	-0,11	43
7. toutes		0,4	-0,05	60*

Le taux de RC est de 66,1 %, ce qui est bon et fourni par de bons étudiants (r.bis = 0,46) avec **confiance**. Le distracteur-vedette (4) fait référence à la respiration qui, au niveau du bilan, est exactement inverse de la photosynthèse; cela peut expliquer que cette solution est choisie par un étudiant sur cinq (21,6 %) ; avec une certitude moyenne élevée (63 %) pour une erreur et par des étudiants faibles (r.bis = -0,31). La **fracture subjective** est faible (10,4 %). Le distracteur-vedette a été choisi avec **imprudence** (certitude moyenne supérieure à 50 %).

Question 8 : Biochimie – Les protéines

Laquelle de ces substances est une protéine ?

Surface de Compétence		%	r.bis	CM
17,60%				
1. la cellulose		12,5	-0,13	42,7
2. l'hémoglobine		28,7	0,44	61,4
3. l'ADN		21	-0,12	54,4*
4. le cholestérol		8,2	-0,12	43,6
5. l'adrénaline		8,6	-0,01	41,9
6. aucune		12,8	-0,08	33,3
7. toutes		4,1	-0,03	54

Le taux de RC (28,7 %) est faible, mais avec **confiance** : la certitude moyenne est plus franche (61,4 %) et par de bons étudiants (r.bis = 0,44). La connaissance des macromolécules du vivant n'est pas une matière explicitement prévue au programme. Le r.bis indique cependant une bonne corrélation avec les résultats obtenus à l'ensemble du test. Certains professeurs en parlent probablement au cours de la partie du cours consacrée à la cytologie (programme de 5^e). Le distracteur-vedette (3. l'ADN) est tentant, car c'est à partir de l'ADN que sont fabriquées les protéines. Il est choisi par un étudiant sur cinq (21 %) moyennement faible (r.bis = -0,12) et avec une certitude moyenne trop élevée (54,4 %) pour une erreur, ce qui donne une **fracture subjective** de 7% seulement (très faible). Le distracteur-vedette a été choisi avec **imprudence** (certitude moyenne supérieure à 50 %).

Question n° 9 : Les hormones, l'insuline
Laquelle de ces propositions est exacte.
L'insuline :

Surface de Compétence			
24,80%	%	r.bis	CM
1. est une hormone stéroïde	10,7	-0,02	46,9
2. est sécrétée par le pancréas	41,2	0,42	69,3
3. fait augmenter la concentration de sucre dans le sang (glycémie)	31	-0,27	50,7*
4. est une maladie grave qui concerne la régulation de la concentration du glucose dans le sang	2	-0,08	53,3*
5. est un enzyme digestif	6,1	-0,08	41
6. aucune	4,2	-0,09	33,8
7. toutes	1,7	-0,02	53,6*

Le taux de RC est de 41,2 %, ce qui est un peu faible donné par de bons étudiants (r.bis = 0,42) et avec **confiance** (une certitude moyenne passable de 69,3 %). L'endocrinologie n'est pas explicitement au programme sauf en ce qui concerne le cas particulier du cycle hormonal féminin (programme de 5e). Le r.bis relativement élevé montre peut-être que certains professeurs étendent l'étude des hormones sexuelles aux autres hormones. Le distracteur-vedette «3. fait augmenter le taux de sucre dans le sang» est choisi par près d'un étudiant sur trois (31 %), des étudiants faibles (r.bis = -0,27) avec **imprudence** : la certitude trop élevée (50,7 %) et une **fracture subjective** très faible (9,6 %). Les élèves ont entendu parler de quelque chose à propos de l'insuline (peut-être pas au cours) mais confondent augmentation et diminution de glycémie.

Question n° 10 : La prévention du SIDA
Actuellement, le meilleur moyen de diminuer les risques d'être infecté par le SIDA par relation sexuelle est :

Surface de Compétence			
84,60%	%	r.bis	CM
1. de se faire vacciner contre le SIDA	0,3	-0,06	37,5
2. de faire préventivement une cure du médicament AZT	0,4	-0,05	72,7*
3. d'utiliser un préservatif (condom)	97	0,17	87,2
4. de se limiter aux relations hétérosexuelles (homme-femme)	0,4	-0,05	50
5. de se faire prescrire des antibiotiques la plus vite possible après toute relation suspecte	0,4	-0,05	65,5*
6. aucune	0,2	-0,04	32
7. toutes	0,3	-0,05	55*

Le taux de RC est de 97%, avec **confiance** (la certitude moyenne = 87,2 %), et ce par les meilleurs étudiants (r.bis = 0,17).

Les élèves montrent une maîtrise remarquable du problème de la prévention contre le SIDA. Le r.bis relativement faible (0,17) s'explique par l'important taux de RC obtenu à cette question. **Prudence et imprudence** s'observent selon les distracteurs, tous très peu choisis. La **fracture subjective** est d'au moins 15% (faible).

Bilan

Les Indices de discrimination (r.bis) des réponses correctes sont TOUS positifs et pour 7 sur 10 supérieurs à la valeur-repère (0,32). Pour 9 questions sur 10 TOUS les r.bis des distracteurs sont négatifs, conformément aux attentes.

La **confiance** dans la réponse correcte s'observe pour 9 questions sur 10 (sauf la question d'écologie)... et nous avons expliqué pourquoi.

On observe de nombreux cas d'**imprudence**. Seules deux questions y échappent.

Chapitre 19

Analyses des réponses au check-up de Connaissances en ART et culture

*D. Leclercq, ULg
A. Lamblin, Inspecteur Communauté Française*

$N = 1399$ étudiants

$Rpbis\ repère = 0,20$

On pourrait faire bien des reproches à l'épreuve qui suit. Pour chacun des thèmes abordés (peinture, sculpture, architecture, cinéma, musique, littérature, théâtre, poésie,...), le nombre de questions est réduit. Certains thèmes, propres à la culture moderne, font défaut : télévision, Bande Dessinée, Jeux Vidéos ?

Le check-up porte sur des connaissances verbales plus que sur l'analyse d'images (sauf les questions 2, 7 et 8) ou de textes (sauf les questions 16, 17, 18) ou de mouvements (pour la danse).

Enfin, aucune des questions ne teste fondamentalement le sens esthétique !

Pour les solutions choisies par moins de 1% d'étudiants, nous n'avons pas indiqué les « épines » (clignotant d'alarme d'imprudence) en gras, rappelant que nous ne leur donnerons pas le même poids qu'aux autres imprudences dans l'interprétation.

Question 1 : La cathédrale * Notre Dame de Paris * relève de l'architecture				
Surface de Compétence		%	r.bis	CM
29,20%				
1. baroque		21,1	-0,1	35,7
2. romane		14,3	-0,08	36,5
3. gothique		56,8	0,31	51,4
4. art nouveau		1,9	-0,11	34,8
5. jésuite		1,1	-0,16	54,7 [*]
6. aucune		1	-0,05	11,4
7. toutes		0,1	0,03	0

Question soit de connaissance, soit de confrontation de l'image que l'on en a avec des critères du gothique. La confiance dans la solution correcte est cependant faible (51,4 %). Le terme « baroque » est un terme fourre-tout. Bien que distracteur-vedette, il n'est pas étonnant qu'il ait été choisi avec une très faible certitude (35,7 %). La fracture subjective avec le distracteur-vedette est modeste (15,7 %). La prudence est générale à une exception près, pour le distracteur 5, peu choisi, avec une certitude supérieure à 50%.

Question 2 : Et celle d'Orléans que voici ?				
Surface de Compétence		%	r.bis	CM
25,50%				
1. baroque		24,9	-0,04	35,4
2. romane		10,9	-0,1	30,8
3. gothique		54,8	0,27	46,6 [*]
4. art nouveau		2	-0,08	28,6
5. jésuite		1,1	-0,1	37,3
6. aucune		1	-0,02	35,7
7. toutes		0,4	-0,11	55 [*]

L'intérêt de cette deuxième question est qu'elle permet de comparer à la première. Il est vraisemblable que les étudiants n'avaient pas d'image mentale préalable de la cathédrale d'Orléans, mais que l'image leur a fait assimiler celle-ci à la cathédrale de Paris. La solution correcte a été choisie autant (54,8%) que dans la question 1 (56,8%) , mais avec manque de confiance (certitude moyenne inférieure à 50%). La fracture subjective avec le distracteur-vedette (1) est faible (11,2 %). La prudence est conforme aux attentes.


Question 3 : Et le Parthénon, sur l'Acropole ?

Surface de Compétence
28,30%

	%	r.bis	CM
1. baroque	7,9	-0,11	27,4
2. romane	23,7	-0,17	34,2
3. gothique	7,2	-0,2	21,8
4. art nouveau	4,3	-0,14	25
5. jésuite	1,9	-0,06	15,4
6. aucune	47,8	0,51	59,2
7. toutes	0,6	-0,03	50*

Moins d'un étudiant sur deux ont choisi la solution correcte, mais ce sont des étudiants forts au total du test ($r.bis = 0,51$) et avec **confiance** (certitude moyenne = 59,2%). Pour un certain nombre d'étudiants (23,7 %), l'antiquité grecque et romaine sont confondues sur ce point (le Parthénon), avec une certitude faible (34,2 %). La **fracture subjective** est nette : 25 %. La **prudence** est conforme aux attentes.

Question 4 : Dans quelle ville est situé le Golden Gate bridge ?



Surface de Compétence
26,70%

	%	r.bis	CM
1. New York	16,4	-0,11	46,6
2. Los Angeles	35,8	-0,1	52,9*
3. Londres	4,1	-0,12	38,9
4. Pittsburgh	2,1	-0,09	28
5. Dallas	1,7	-0,11	32,5
6. aucune	35	0,4	76,3
7. toutes	0,2	-0,03	60*

Un étudiant sur 3 (35 %) associe le Golden Gate Bridge à San Francisco et avec **confiance** (certitude moyenne = 76,3 %). La confusion avec Los Angeles tant en popularité (35,8 %) qu'en **imprudence** (52,9 %) était prévisible. La **fracture subjective** est de 23,4 % (nette), avec Los Angeles et 29,7 % avec New York, ce qui est élevé : ceux qui savent se distinguent nettement de ceux qui ne savent pas.

Question 5 : Parmi les paires d'artistes constituées ci-après, laquelle (paire) peut-on associer à l'impressionnisme en peinture ?

Surface de Compétence
35,60%

	%	r.bis	CM
1. Rembrandt - David	5,4	-0,12	25,1
2. Monet - Sisley	58,7	0,46	60,7
3. Magritte - Alechinsky	13,4	-0,18	34,5
4. Bruegel - Bosch	5,4	-0,09	32,5
5. Picasso - Permeke	8,8	-0,15	35,3
6. aucune	2,9	-0,03	22
7. toutes	0,4	-0,06	23,3


Cette question mérite d'être comparée à la suivante parce que ceux qui ont fourni la réponse correcte l'ont fait avec la même **confiance** : la certitude moyenne vaut 60,7 % pour Q5 et 61,9 % pour Q6. Par contre, les taux de réussite sont assez différents : 58,7 % pour Q5 contre 41,8 % pour Q8. La **fracture subjective** de cette question 5 est nette (26,2 %). La **prudence** est conforme aux attentes.

Question 6 : Laquelle de ces 5 mêmes paires regroupe des peintres belges qui ont vécu au XX^e siècle ?

Surface de Compétence				
25,90%		%	r.bis	CM
1. Rembrandt - David		4,4	-0,15	29,5
2. Monet - Sisley		4,1	-0,17	33
3. Magritte - Alechinsky		41,6	0,33	61,9
4. Bruegel - Bosch		23,7	-0,12	45,6
5. Picasso - Permeke		4,9	-0,07	54,7
6. aucune		14,9	0,08	43,8
7. toutes		0,3	-0,04	10

La solution correcte a été choisie par des étudiants forts (r.bis = 0,33) et avec **confiance** (certitude moyenne = 61,4%).
Le distracteur-vedette (Bruegel-Bosch) récolte le quart des réponses (23,7 %), mais, heureusement, avec une certitude peu élevée (45,6 %), la fracture subjective étant de 16,3 % (modeste).
La **prudence** est conforme aux attentes.

Question 7 : La parabole des aveugles (ci-dessous) relève de la peinture




Cette question est difficile (% RC = 23,4 %) et les réponses correctes **manquent de confiance** : la certitude moyenne est de 38 %, en partie parce que la solution correcte est « aucune ». La fracture subjective est très faible (8,6 % avec les deux distracteurs-vedettes « roman » et « baroque ». Ce sont cependant plutôt les étudiants les meilleurs au total du check-up qui ont choisi la réponse correcte (r.bis = 0,3).

La **prudence** est partout où on l'attend.

Degré de maîtrise				
8,90%		%	r.bis	CM
1. cubiste		5,3	-0,11	27,6
2. roman		21,7	-0,04	30,1
3. gothique		12,9	-0,03	29,2
4. baroque		23,6	0,04	29,4
5. surréaliste		6,1	-0,12	32
6. aucune		23,4	0,3	38●
7. toutes		0,2	0,01	40

Question 8 : Et cette sculpture de Jacques Lipchitz ?

Surface de Compétence				
26,20%		%	r.bis	CM
1. cubiste		49,4	0,32	53,1
2. roman		1,6	-0,1	10,9
3. gothique		2,6	-0,14	26,3
4. baroque		2,5	-0,09	25,7
5. surréaliste		34,4	-0,1	46,8
6. aucune		3,2	0	33,8
7. toutes		0,1	0,03	0



Il y a peu de chances que les étudiants sachent que Jacques Lipchitz est considéré comme le « prototype » des sculpteurs cubistes comme Picasso l'est des peintres cubistes, et encore moins de chances qu'ils sachent que cette œuvre est de Lipchitz. C'est donc par raisonnement (sur les attributs d'une œuvre cubiste) que la moitié des étudiants (49,4 %) a pu choisir la solution correcte, avec **confiance** (certitude moyenne = 53,1 %) pour ce genre de question, mais avec une fracture subjective insignifiante (4,3 %) avec le distracteur-vedette (surréalisme) qui a attiré un étudiant sur 3 (34,4 %). Cette même question, posée à des enseignants (24) du supérieur a donné les mêmes résultats : 46 % de réponses correctes, le même distracteur-vedette choisi par 21 % des interrogés, mais 29 % se portant sur « Aucune ».
La **prudence** est conforme aux attentes.

Pour les 4 questions qui suivent, nous avons indiqué tous les r.bis et toutes les certitudes moyennes, même pour les solutions choisies par moins d'1% des étudiants, étant donné les taux d'exactitude élevés et, donc, le nombre élevé de distracteurs choisis par moins de 1% des étudiants.

Question 9 : Quelle est l'oeuvre cinématographique qui a été inspirée par le drame du SIDA ?				
Surface de Compétence				
	87,60%	%	r.bis	CM
1. Philadelphia	93,6	0,36	93,6	
2. La ligne rouge	1,9	-0,14	27,7	
3. Love Story	1,4	-0,17	24	
4. Basic instinct	0,6	-0,15	35,6	
5. Rain Man	0,4	-0,05	86,7*	
6. aucune	0,6	-0,11	40	
7. toutes	0,1	-0,02	50	

C'est la question la mieux réussie du check-up (93,6 %), avec la certitude moyenne la plus élevée (93,6 %), donc une **extrême confiance**. Il est vrai que d'autres questions (10, 11, 12) sur le cinéma actuel (de 1999) ont aussi été très bien réussies, à une exception (la question 13). Malgré ce taux de réussite élevé, la question garde un bon pouvoir discriminatif (r.bis = 0,36) et une **fracture subjective** très élevée de 65,9 % avec le distracteur-vedette (2). La **prudence** est générale, à une **seule exception** (Rain Man), peu fréquente, pour laquelle il y a méprise.

Question 10 : Et laquelle de ces 5 mêmes oeuvres a été inspirée par le problème de l'autisme ?				
Surface de Compétence				
	78,20%	%	r.bis	CM
1. Philadelphia	0,9	-0,1	25	
2. La ligne rouge	3,9	-0,19	38,9	
3. Love Story	1,1	-0,1	28	
4. Basic instinct	1,1	-0,1	32	
5. Rain Man	85,2	0,41	91,8	
6. aucune	4,6	-0,13	39,7	
7. toutes	0,4	-0,07	32	

Cette question est légèrement moins bien réussie que la précédente (le film Rain Man est plus ancien), mais avec une certitude semblable (91,8 % au lieu de 93,6 %) donc une **confiance extrême**. On aurait obtenu probablement un pourcentage aussi élevé de réponses correctes si la question était amorcée par une photo au lieu du thème (autisme) ou si l'on avait demandé le nom d'un acteur (Tom Cruise ou Dustin Hoffman). La **fracture subjective** avec le distracteur-vedette (6. Aucune) est très élevée (52,1 %). La **prudence** est conforme aux attentes.

Question 11 : Quel film a été réalisé par Steven Spielberg ?				
Surface de Compétence				
	64%	%	r.bis	CM
1. Il faut sauver le soldat Ryan	5,4	-0,19	79,2*	
2. Jurassic Park	13,9	-0,28	80,4*	
3. La couleur pourpre	0,4	-0,11	43,3	
4. Rencontre du troisième type	0,2	-0,07	46,7	
5. La liste de Schindler	2,8	-0,17	72,3*	
6. aucune	0,5	-0,05	51,4*	
7. toutes	73,1	0,49	87,6	

Dans le cas de questions où la réponse correcte est « Toutes », il n'est pas étonnant que la **fracture subjective** soit faible avec certains « distracteurs » (corrects en soi, mais pas dans la logique de la question) : la **fracture subjective** n'est que de 7,6 % avec « Jurassic Park » et 8,4 % avec « Il faut sauver le soldat Ryan ». Les certitudes moyennes sont très élevées non seulement pour la réponse correcte (87,6 %), signe de **confiance**, mais aussi pour les distracteurs-vedettes (80,4 %, 79,2 % et 72,3%), signes d'**imprudence**. La **fracture subjective** est 17,2 % (modeste). On pouvait s'attendre, parce que la réponse correcte est « Toutes » que les distracteurs (corrects en soi) soient choisis avec une certitude élevée, ce qui est le cas. Les étudiants devraient être plus habitués à ce genre de situation.

Question 12 : Et laquelle de ces 5 mêmes oeuvres a été inspirée par la Shoah (extermination du peuple juif) ?				
Surface de Compétence				
	84,20%	%	r.bis	CM
1. Il faut sauver le soldat Ryan	1,1	-0,11	48,8	
2. Jurassic Park	0,3	-0,11	40	
3. La couleur pourpre	2,3	-0,15	46,9	
4. Rencontre du troisième type	0,6	-0,09	27,5	
5. La liste de Schindler	91,6	0,39	91,9	
6. aucune	1,1	-0,08	41,3	
7. toutes	0,1	-0,02	30	

Ce sujet (film récent) est très bien connu par un grand nombre d'étudiants (91,6%), forts au total du test (r.bis = 0,39), et avec une **confiance** très élevée (certitude moyenne = 91,9%). La **prudence** est conforme aux attentes et la **fracture subjective** est très élevée (45 %).

Question 13 : Quel film a été réalisé par Stanley Kubrik ?			
Surface de Compétence			
	%	r.bis	CM
	31,80%		
1. Orange mécanique	39,9	-0,14	72,5*
2. 2001 Odyssée de l'espace	4,8	-0,15	34,3
3. Les sentiers de la gloire	2,3	-0,15	18,1
4. Full Metal Jacket	5,7	-0,13	43
5. Barry Lindon	1,1	-0,12	24
6. aucune	3,9	-0,06	17,4
7. toutes	36,7	0,48	86,7

La réponse correcte (Toutes) n'est connue que par un peu plus d'un étudiant sur trois (36,7 %), mais avec une **confiance** élevée : 86,7 %. Le distracteur-vedette (« Orange mécanique ») est plus populaire (39,9 %) que la réponse correcte (36,7 %), et donné avec **imprudence**, mais avec une certitude (72,5 %) quand même inférieure à celle de la réponse correcte. La fracture subjective est de 14,2 % (faible).

Question 14 : Lequel de ces mêmes 5 films traite de la " décimation " ?			
Surface de Compétence			
	%	r.bis	CM
	5,30%		
1. Orange mécanique	11,7	0,03	29
2. 2001 Odyssée de l'espace	8,1	-0,01	27,5
3. Les sentiers de la gloire	21,3	0,09	24,8*
4. Full Metal Jacket	22,7	0,08	38,9
5. Barry Lindon	9,3	0,01	25,2
6. aucune	10,9	-0,06	20,5
7. toutes	0,9	0,02	35,4

Au vu des résultats de l'épreuve de vocabulaire, le terme décimation, «châtiment qui consistait à faire périr un soldat sur dix; faire périr en grand nombre» est probablement lui-même mal connu. Il est peu surprenant que le film « Les sentiers de la gloire » qui lui est consacré, soit très peu connu (par 21,3% des étudiants seulement), et surtout avec un **fort manque de confiance** (un degré de certitude moyen de 24,8 % seulement !). Que la corrélation bisériale 0,09 soit largement inférieure à 0,20 indique que cette question porte probablement moins sur le cinéma que sur le vocabulaire. La fracture subjective aberrante avec le distracteur-vedette (4) est négative : -14,1 %. La **prudence** est conforme aux attentes.

Question 15 : Quel spectacle a été conçu par Bertold Brecht ?			
Surface de Compétence			
	%	r.bis	CM
	3,30%		
1. L'opéra de quatre sous	31,9	0,08	41,3
2. La résistible ascension d'Arturo Uy	12,6	0,01	47,5
3. Mère courage	11,4	-0,05	32,6
4. Le Brave soldat Schweick	11,7	-0,06	17
5. Le cercle de craie caucasien	3,6	0	29,4
6. aucune	10,2	0,01	17,9
7. toutes	6,2	0,2	54

A peu près un étudiant sur trois seulement (31,9 %) sait que Brecht est l'auteur de «L'opéra de quatre sous». Un sur huit (12,6 %) sait qu'il en va de même pour «La résistible ascension d'Arturo Uy», et un sur 10 (11,4 %) pour « Mère Courage ». Par contre, les 11,7 % qui ont choisi « Le brave soldat Schweick » ont répondu au hasard (certitude moyenne 17 %). La fracture subjective est de 12,7 % (faible). **Confiance et prudence** sont conformes aux attentes.

Question 16 : Qui a écrit la pièce de théâtre dont voici un extrait :

Le docteur : Mais vous n'avez jamais exercé.
Knock : Autre erreur.
Le docteur : Comment ? Ne m'avez-vous pas dit que vous veniez de passer votre thèse l'été dernier ?
Knock : Oui, trente-deux pages : *Sur les prétendus états de santé*, avec cette épigraphe, que j'ai attribuée à Claude Bernard « les gens bien portants sont des malades qui s'ignorent ».

Surface de Compétence

9%	%	r.bis	CM
1. Molière	34,7	-0,08	49,9
2. Racine	8,3	-0,1	29,7
3. Corneille	8,4	-0,12	20,5
4. Marivaux	13,9	0,02	29,5
5. Rostand	11,6	0,1	36,8
6. aucune	15,3	0,29	59
7. toutes	0,1	-0,02	0

Il est vrai que le style de Jules Renard dans *Knock* fait penser à Molière et ses satires de la médecine. La référence à Claude Bernard devrait pourtant faire éliminer cette solution. Les rares étudiants qui ont fourni la réponse correcte (15,3 %) avaient une certitude moyenne de 59 %, ce qui établit la fracture subjective de 9,1 % (très faible).
Confiance et prudence sont conformes aux attentes

Question 17 : Voici une strophe du poème " Napoléon II ". Quel est l'auteur de ce poème ?

Oui, l'aigle, un soir, planait aux voûtes éternelles.
 Lorsqu'un grand coup de vent lui cassa les deux ailes :
 Sa chute fit dans l'air un foudroyant sillon ;
 Tous alors sur son nid fondirent pleins de joie ;
 Chacun selon ses dents se partagea la proie ;
 L'Angleterre prit l'aigle, et l'Autriche l'aiglon.

Surface de Compétence

10,30%	%	r.bis	CM
1. Rimbaud	14,4	-0,09	22,3
2. Hugo	26,8	0,35	38,6
3. Baudelaire	18,4	-0,12	22,2
4. Hemingway	11,9	-0,03	22,6
5. Balzac	10,3	-0,03	23,1
6. aucune	6,7	0,01	20,4
7. toutes	0,2	-0,01	6,7

Un quart des étudiants (26,8 %) ont fourni la réponse correcte, avec un **manque de confiance** ; la certitude moyenne est faible (38,6 %), mais ce sont des étudiants forts au total du test : l'indice de discrimination (r.bis = 0,35) est assez bon.
 La **prudence** est conforme aux attentes.
 La fracture subjective est de 16,4 % (modeste) avec le distracteur-vedette (3. *Beaudelaire*).

Question 18 : Voici le début du poème " L'invitation au voyage ". Quel est l'auteur de ce poème ?

Mon enfant, ma sœur, Songe à la douceur
 D'aller là-bas vivre ensemble !
 Aimer à loisir, Aimer à mourir
 Au pays qui te ressemble !

Surface de Compétence

15,60%	%	r.bis	CM
1. Rimbaud	24,5	0,03	31,5
2. Hugo	12,2	-0,1	23,6
3. Baudelaire	35,3	0,26	44,2
4. Hemingway	6,6	0,01	27,5
5. Balzac	5,1	-0,1	26,8
6. aucune	3,6	0	15,2
7. toutes	0,1	-0,06	0

Une bonne année après la « poésie » (année du secondaire où l'on étudie les poètes), un étudiant sur trois (35,3 %) fournit la réponse correcte avec **manque de confiance** : la certitude moyenne est peu élevée (44,2 %).
 Que Rimbaud soit, pour un étudiant sur 4 (24,5 %) le distracteur-vedette n'est pas étonnant, avec une certitude plus faible (31,5 %), ce qui donne une fracture subjective de 12,7 % (faible).
 La **prudence** est conforme aux attentes.

Question 19 : Quel ouvrage Kant a-t-il écrit ?			
Surface de Compétence			
28,90%	%	r.bis	CM
1. Le discours de la méthode	18,7	-0,1	44
2. Mein Kampf	4,6	-0,16	48,6
3. La critique de la raison pure	51,8	0,34	55,8
4. Le capital	8,8	-0,02	32,2
5. Le meilleur des mondes	6,5	-0,17	26,6
6. aucune	5	-0,04	18,9
7. toutes	0,3	0	5

La moitié des étudiants (51,8 %) répond correctement avec une certitude moyenne de 55,8 %, donc **confiance**. Nous pensons que c'est dû au caractère « évidemment faux » des distracteurs. On est d'autant plus surpris que ceux-ci soient quand même choisis et surtout par 18,7 % pour le « Discours de la Méthode » (avec une certitude moyenne de 44 % !!). Heureusement, la **prudence** est conforme aux attentes.

La **fracture subjective** est 11,8 % (faible).

Question 20 : Quel ouvrage a été écrit par Aldous Huxley ?			
Surface de Compétence			
31,40%	%	r.bis	CM
1. Le discours de la méthode	6,9	-0,07	14,2
2. Mein Kampf	3,1	-0,12	15
3. La critique de la raison pure	7,5	-0,16	15,2
4. Le capital	14,7	-0,07	22,8
5. Le meilleur des mondes	48,2	0,39	65,2
6. aucune	6,9	-0,03	14,4
7. toutes	0,1	-0,01	100*

La moitié des étudiants (48,2 %) ont répondu correctement, avec une certitude assez élevée (65,2 %), donc **confiance**, et ce sont des étudiants forts au total de ce test : l'indice de discrimination est élevé lui aussi (r.bis = 0,39). La **fracture subjective** est très élevée (41,4 %) avec le distracteur-vedette choisi, étonnamment, par 14,7 %.

La **prudence** est forte et conforme aux attentes.

Question 21 : Quel ouvrage est antisémite ?			
Surface de Compétence			
64,40%	%	r.bis	CM
1. Le discours de la méthode	2,4	-0,14	22,4
2. Mein Kampf	77	0,46	83,7
3. La critique de la raison pure	3,2	-0,13	16
4. Le capital	3,1	-0,15	15,8
5. Le meilleur des mondes	5,3	-0,16	21,9
6. aucune	2,5	-0,1	9,1
7. toutes	0,2	0,02	53,3*

Cette question, assez bien réussie (77 %) avec une certitude élevée (83,7 %), donc une forte **confiance**, porte plus sur l'histoire que sur la littérature. Elle est très discriminative (r.bis = 0,46). La **fracture subjective** est la plus élevée de tout le check-up : 61,8 % avec le distracteur-vedette « Le meilleur des mondes ».

La **prudence** correspond aux attentes.

Question 22 : La chanson " We are the champions, my friend " (souvent entonnée sur les stades) a été lancée par quel interprète et quel groupe ?			
Surface de Compétence			
83,40%	%	r.bis	CM
1. Lou et les Hollywood Bananas	0,8	-0,14	47,3
2. Bono et U2	0,9	-0,08	51,7*
3. Mick Jagger et les Rolling Stones	1,9	-0,18	42,2
4. Paul Mc Cartney et les Beatles	1,7	-0,12	27,5
5. Freddy Mercury et le groupe Queen	91,2	0,38	91,5
6. aucune	0,6	-0,08	30
7. toutes	0,3	-0,12	25

Neuf étudiants sur dix (91,2 %) ont répondu correctement, avec un degré de certitude moyen élevé (91,5 %), donc une **extrême confiance** et avec une **fracture subjective** très élevée (49,7 %) avec le distracteur-vedette « Mick Jagger et les Rolling Stones ».

Cette question a été posée à des enseignants (24) du supérieur (universitaire et non universitaire) qui l'ont réussie à 83 %, le distracteur-vedette étant le même (Mick Jagger et les Rolling Stones), choisi par 8 % des répondants.

La **prudence** est générale, à une exception près (Bono et U2), qui concerne peu d'étudiants..


Question 23 : L'hymne " européen " aussi appelé l'Hymne à la joie, est l'œuvre de

Surface de Compétence
20,10%

	%	r.bis	CM
1. Mozart	21,3	-0,1	49
2. Berlioz	8,6	-0,05	29,4
3. Debussy	8,9	-0,05	26,3
4. Bach	22,9	-0,14	40,7
5. Dvorak	1,3	-0,1	31,1
6. aucune	30,2	0,44	66,7
7. toutes	0,1	0,03	10

En cette période de « tout à l'Europe », on est étonné de ce que moins d'un étudiant sur trois (30,2 %) sache qu'il s'agit de Beethoven, heureusement avec **confiance** (le degré de certitude moyen est décent : 66,7 %).
La **fracture subjective** est de 26 % (nette) avec le distracteur-vedette « Bach » et de 17,7 % avec le second distracteur en popularité, « Mozart ».
La **prudence** est conforme aux attentes.

Question 24 : Dans le film " Les uns et les autres ", le danseur, **Jorge Donn** (voir photo ci-dessous à côté de Tania Bari), qui y interprète le " Boléro " a fait partie



Surface de Compétence
3,60%

	%	r.bis	CM
1. Du ballet du XX ^e siècle	16,7	0,16	21,4●
2. De la troupe de danse de Serge Lifar	6,4	-0,03	11
3. De la troupe de danse de Rudolf Noureev	25,9	0,05	18
4. De la troupe de danse de Balanchine	12,4	0	14
5. De la troupe de danse de Patrick Dupont	9,6	-0,03	24,1
6. aucune	10,1	0	9,6
7. toutes	0,9	-0,04	25

L'œuvre de Maurice Béjart (nous avons évité de fournir son nom, ce qui aurait facilité la réponse) semble bien absente de l'esprit de cette population (16,7 % seulement de réponses correctes), et, pour ceux qui ont répondu correctement, avec une certitude particulièrement faible (21,4 %), donc un **fort manque de confiance**.
La **fracture subjective** est insignifiante (3,4 %) avec le distracteur-vedette Rudolf Noureev et -2,7 % avec le deuxième distracteur Patrick Dupont. Cet élément de la danse semble extérieur (r.bis = 0,16 seulement, inférieur au seuil de 0,20) au noyau de culture générale testé par ce check-up.
La **prudence** est conforme aux attentes.

Question 25 : De quel musicien est l'air "Le boléro", dansé dans le film " Les Uns et les Autres " par le danseur Jorge Donn ?

Surface de Compétence
33,70%

	%	r.bis	CM
1. Stravinsky	11	-0,12	17,9
2. Mozart	5,2	-0,12	27,4
3. Rodrigo	6,7	-0,12	16,8
4. Paco de Lucia	6,9	-0,09	14,4
5. Tchaikowski	14,3	-0,14	28
6. aucune	46,8	0,48	72,1
7. toutes	0,5	-0,06	68,6*

Cette question relève plus de la musique que de la danse. Le taux de réponses correctes (46,8 %) est élevé pour une question dont la réponse correcte est « Aucune », et la certitude moyenne qui l'accompagne est assez élevée (72,1 %), ce qui indique la **confiance**. C'est probablement en partie dû au fait que l'expression verbale « Le boléro de Ravel » est très connue, même par des personnes qui ne reconnaîtraient pas cet air à l'audition. La **fracture subjective** avec le distracteur-vedette (Tchaikowski) est très élevée : 44,1 %.
La **prudence** est conforme aux attentes.

Bilan

Indicateurs psychométriques

Les indices de **discrimination (r.bis)** sont positifs et pour TOUTES les **réponses correctes** sauf pour deux (Q14 et Q24), et supérieurs à la valeur-seuil (0,20) pour toutes les questions.

Sauf pour 7 questions sur les 25 (7, 14, 15, 16, 17, 18, 24), tous les **distracteurs** ont un indice de discrimination négatif. Ceci indique (sauf pour les exceptions) une bonne cohérence entre les questions et l'ensemble du check-up.

Indicateurs édumétriques

La **confiance** moyenne dans la réponse correcte est observée pour **20** questions sur les 25 (à l'exception des questions 2, 7, 14, 17, 18).

Une **prudence** moyenne générale (de tous les distracteurs) a été observée pour **14** questions sur les 25. Il resterait donc du travail de correction pour près d'une question sur 3 si elles étaient jugées importantes, au cas par cas, non seulement par question, mais par orientation d'étude.

Chapitre 20

Analyses des réponses au check-up de Connaissances générales en Histoire, Actualité, Economie

D. Leclercq, ULg

F. Georges, ULg

N = 1418

R.bis = 0,20

Les questions sur l'actualité telles que l'ONU et l'OTAN portent sur des contenus étudiés en dernière année du secondaire.

Question n° 1 : A l'heure actuelle, quels sont les 5 membres permanents de l'ONU ?				
Surface de Compétence				
18,30%		%	r.bis	CM
1. Japon - Allemagne - USA - Russie - France	21,4	-0,11	45	L'erreur jugée la plus grave (5. <u>Italie, France, USA, URSS, UK</u>) a été choisie par un quart (24,5 %) des étudiants, les plus faibles (r.bis = -0,19), avec prudence car leur certitude moyenne (43,7 %) est de 23,4 % inférieure à celle des Réponses Correctes, peu nombreuses (27,2%), mais fournies par de bons étudiants (r.bis = 0,43) et fournies avec confiance (67,1 %). La "fracture subjective" est donc de 23 %. L'erreur 4. <u>Chine, UK, France, USA, URSS</u> pourrait être considérée comme pardonnable (manque d'actualisation de "URSS" en "Russie"). Elle a d'ailleurs attiré des étudiants « moyens » ou les moins faibles (r.bis = 0,02) de ceux qui se sont trompés, et ils se sont trompés "de bonne foi" car leur certitude moyenne (52,5 %), imprudente , n'est que de 14,6 % inférieure (<u>fracture subjective faible</u>) à celle des RC (67,1 %).
2. République populaire de Chine - France - Grande-Bretagne - USA - Russie	27,2	0,43	67,1	
3. Brésil - Inde - République populaire de Chine - USA - URSS	0,8	-0,05	50 *	
4. République populaire de Chine - Grande-Bretagne - France - USA - URSS	13,8	-0,02	52,5 *	
5. Italie - France - USA - URSS - Royaume-Uni	24,5	-0,19	43,7	
6. aucune	8,9	-0,11	42,5	
7. toutes	0	XXXXX	XXXXX	

Question n° 2 : : Quel est le secrétaire général actuel de l'ONU ?				
Surface de Compétence				
18,30%		%	r.bis	CM
1. Javier Solana	10	-0,06	41,8	Le taux de RC est élevé (60,5 %) et la confiance aussi, la certitude moyenne valant 68,4 %, donnant une <u>fracture subjective</u> de 28,6 % (nette) avec le plus assuré des distracteurs (41,8 %). Confiance et prudence sont partout où on les attend.
2. Boutros Boutros Ghali	7,5	-0,11	39,8	
3. Jacques Santer	9,4	-0,2	26,3	
4. Antonio Samaranch	3,1	-0,08	31,4	
5. Kofi Anan	60,5	0,43	68,4	
6. aucune	5,2	-0,16	14,1	
7. toutes	0,2	-0,06	0	

Question 3 : Quel est le secrétaire actuel de l'OTAN ?				
Surface de Compétence				
2,20%		%	r.bis	CM
1. Javier Solana	35,6	0,36	60,4	Un tiers des étudiants (35,6 %), parmi les meilleurs (r.bis = 0,36) ont répondu correctement, avec une certitude moyenne modeste (60,4 %). La <u>fracture subjective</u> est modeste (18,6 %) avec le distracteur-vedette (5. <u>Kofi Anan</u>). En toute logique, cependant, les 9 % qui ont choisi « Aucune » pourraient être crédités de la réponse correcte, car Javier Solana a exercé cette activité pendant plusieurs années (actualité sédimentée par la répétition). En août 1999, Lord georges Robertson a été désigné comme successeur et il a pris ses fonctions la semaine qui a suivi le testing MOHICAN (le 14/10/2002). On en a donc parlé, mais sans "sédimentation par répétition", d'autant plus que c'était pendant les vacances. Confiance et prudence sont partout où on les attend.
2. Boutros Boutros Ghali	9,3	-0,02	29,8	
3. Jacques Santer	14	-0,1	31,6	
4. Antonio Samaranch	7,3	-0,06	30,2	
5. Kofi Anan	18,3	-0,17	41,8	
6. aucune	9	0,01	24,2	
7. toutes	0,3	-0,05	10	

Question 4 : Dans l'abréviation GATT (General Agreement ...), que signifie un des deux T ?				
Surface de Compétence				
29,70%		%	r. bis	CM
1. Tourisme	3,5	-0,13	29,8	
2. Total	4,7	-0,18	23,3	
3. Tolerance	8,7	-0,12	23,9	
4. Tax	13,3	-0,07	42,6	
5. Trade	59,1	0,38	50,3	
6. aucune	4	-0,1	29,1	
7. toutes	0,1	-0,02	0	

Les experts jugent que c'est une erreur dans la construction de la question de n'avoir pas précisé duquel des deux T on parle. Le taux de réponses correctes est néanmoins bon (59,1 %), mais la certitude moyenne qui y est associée est faible (50,3 %). Le distracteur le plus populaire est 4. Tax avec 13,3 % de choix avec le plus haut degré de certitude (42,6 %) des étudiants "en erreur". Ceci est un phénomène bien connu de "test wiseness" (LECLERCQ, 1986, 106) : les étudiants ignorants se réfugient dans les solutions "plausibles". La fracture subjective vaut 7,7 % (très faible).
Confiance et prudence sont partout où on les attend.

Question 5 : Dans l'abréviation OMS, que signifie S ?				
Surface de Compétence				
48,90%		%	r. bis	CM
1. Sport	3,2	-0,12	24,3	
2. Santé	72,8	0,4	67,2	
3. Sida	2,4	-0,13	35,9	
4. Sécurité	10,6	-0,17	30,7	
5. Société	5,5	-0,18	22,6	
6. aucune	1,6	-0,07	13,9	
7. toutes	0,1	0,02	60	

Le taux de réponses correctes est très élevé (72,8 %) ainsi que la certitude moyenne des RC (67,2 %). Le distracteur le plus populaire a été choisi par 10,6 % des étudiants. On peut y voir un phénomène de test wiseness comme dans la question 4. La fracture subjective est de 36,5 % (élevée).
Confiance et prudence sont partout où on les attend.

Question n° 6 : Dans l'abréviation U.N.E.S.C.O., que signifie le E ?				
Surface de Compétence				
18,50%		%	r. bis	CM
1. Enfants	26	-0,11	47,7	
2. Elèves	0,8	-0,12	41,7	
3. Education	37,8	0,31	49	
4. Environnement	16	-0,01	43,3	
5. Economie	12,4	-0,18	32,5	
6. aucune	1,8	-0,01	24	
7. toutes	0,1	-0,03	0	

Le distracteur le plus populaire (1. Enfants) était assez prévisible. La certitude moyenne (ou Facilité Subjective) avec laquelle les 37,8 % de Réponses Correctes (FSC) ont été données est basse (49 %), soit un manque de confiance et proche de celle du distracteur vedette (47 %). La fracture subjective n'est que de 2% (insignifiante) !!!
La prudence est partout où on l'attend.

Question n° 7 : Dans l'abréviation B.I.R.D., que signifie le B ?				
Degré de Maîtrise				
21,50%		%	r. bis	CM
1. Banque	50,6	0,24	42,4	
2. Bombe	3,9	-0,09	25,8	
3. Bataillon	3,7	-0,09	17,7	
4. Bilatéral	2,3	-0,03	20	
5. Bureau	21,7	-0,01	27,7	
6. aucune	9,2	-0,11	15,4	
7. toutes	0,3	0,02	15	

Le taux étonnamment élevé (par rapport à la question précédente) de Réponses Correctes (50,6 % sans doute lié à la facilité d'éliminer les distracteurs). La Facilité Subjective (ou Certitude Moyenne) de la Réponse Correcte (42,4 %) atteste du manque de confiance dans cette connaissance sur B (Qu'en aurait-il été sur le R de Reconstruction ?). Le distracteur-vedette (5) a été fourni avec une certitude très basse. La fracture subjective est de 14,7 % (faible).
La prudence est conforme aux attentes.

* Capacité de deviner les réponses correctes à des tests via des indices involontairement laissés dans les questions par leur auteur.

Question n° 8 : Dans l'abréviation U.N.I.C.E.F., que signifie le C ?			
Surface de Compétence			
42,90%	%	r.bis	CM
1. Children	76	0,34	56,5
2. Community	13,7	-0,22	32,9
3. Christian	0,9	-0,06	29,2
4. Country	3,2	-0,08	26,2
5. Communist	1,1	-0,1	47,5
6. aucune	1,3	-0,11	48,4
7. toutes	0,1	-0,05	0

Nous pensons que peu d'étudiants savent que le E signifie Emergency. Notre question a porté sur l'élément principal à nos yeux (children), qui semble bien connu (par 76% des répondants), mais avec une certitude encore assez faible (56,5%).
Le distracteur-vedette (2. Community) est fourni avec une certitude très faible (32,9 %). La fracture subjective vaut 23,6 % (nette).
Confiance et prudence sont partout où on les attend.

Question 9 : Voici 5 dates de la Construction Européenne. Laquelle correspond à l'entrée en vigueur de l'Euro ?			
Surface de Compétence			
47,30%	%	r.bis	CM
1. 1944	0,8	-0,14	51,7*
2. 1951	0,7	-0,1	28
3. 1972	0,6	-0,06	22,5
4. 1981	1,3	-0,11	31,1
5. 1999	65	0,33	72,8
6. aucune	29,6	-0,24	66*
7. toutes	0,3	-0,01	40

Le taux élevé de réponses correctes (65 %) et surtout de **confiance**, étant donné la certitude moyenne de celles-ci (72,8 %) ne surprendra pas, car cette question faisait, à l'époque, la une de l'actualité. La fracture subjective avec « Aucune » n'est cependant que de 6,8 % (très faible), probablement parce qu'on pourrait interpréter « date d'entrée en vigueur des billets et les pièces en euros », ce qui donne une date ultérieure.
Deux distracteurs ont été choisis avec **imprudence** (certitude moyenne supérieure à 50%).

Question 10 : Quel est le nom du nouveau Président de la Commission européenne (depuis trois mois) ?			
Surface de Compétence			
32,10%	%	r.bis	CM
1. SANTERRE	24	-0,19	40,7
2. PRODI	43,4	0,51	73,9
3. SOLANA	7,3	-0,11	26
4. FISCHLER	10,2	-0,11	29,3
5. LAFONTAINE	2,7	-0,11	37,9
6. aucune	7,5	-0,17	19,1
7. toutes	0,2	-0,05	20

On retrouve la distinction entre les connaissances « sédimentées par répétition et non remises à jour » avec la solution « Santerre » (qui avait été président durant plusieurs années) qui recueille 24 % des choix (mais avec une certitude faible : 40,7 %). Les étudiants dont la connaissance a été mise à jour sont sûrs d'eux-mêmes (73,9 % de certitude). La fracture subjective est élevée entre ces deux réponses (33,2 %).
Confiance et prudence sont partout où on les attend.

Question 11 : En quelle année Hiroshima a-t-elle été détruite par la première bombe atomique ?			
Surface de Compétence			
14%	%	r.bis	CM
1. 1940	4,3	-0,11	36,4
2. 1943	24,8	-0,14	44,8
3. 1946	36,6	0,05	56,8*
4. 1950	4,6	-0,11	31,7
5. 1955	6,3	-0,11	37,6
6. aucune	20,5	0,33	68,4
7. toutes	0,4	-0,01	76*

Il est regrettable que seulement 20,5 % des interrogés aient choisi la solution correcte, heureusement avec **confiance**, plus exactement avec une certitude moyenne de 68,4 % (donc une fracture subjective très faible de 7,6 % avec le distracteur-vedette « 1946 », choisi, lui, avec **imprudence**).
Le faible taux de réussite est probablement lié au recours par les auteurs du check-up à la solution « Aucune ». Le constat n'en est que plus grave : on peut soupçonner dès lors qu'un certain nombre de solutions correctes dans d'autres questions (de ce check-up et d'autres !) sont favorisées par la « reconnaissance » d'une solution que certains étudiants n'auraient pu « évoquer ». Posée à 24 enseignants du supérieur, cette question a été réussie à 48 %, le distracteur-vedette étant aussi 1946 (32 %) et le suivant aussi 1943 (16 %). Ces enseignants avaient majoritairement assez bien estimé le taux de réussite de ces étudiants, taux qu'ils ont jugé majoritairement

Question 12 : En quelle année le Congo a-t-il cessé d'être une colonie belge ?

Surface de Compétence : 31,80%

	%	r.bis	CM
1. 1940	3,7	-0,08	30,2
2. 1950	10,3	-0,13	32,2
3. 1960	51,9	0,4	61,3
4. 1970	12,1	-0,08	37
5. 1980	11	-0,17	38,3
6. aucune	7,5	-0,11	36,1
7. toutes	0,2	-0,01	33,3

Bien qu'il y ait des dates qui se retiennent « sans raisonnement », nous avons pris la précaution de séparer les solutions par une décennie. Heureusement, puisque la solution correcte n'a été choisie que par 51,9 % des étudiants et avec une certitude peu élevée (61,3 %). Comme on pouvait s'y attendre, les distracteurs sont d'autant plus populaires que la date est proche de la date correcte. Heureusement, la fracture subjective est nette : 24,3 % avec le distracteur « 1970 » qui a attiré 12,1 % des réponses.
Confiance et prudence sont partout où on les attend.

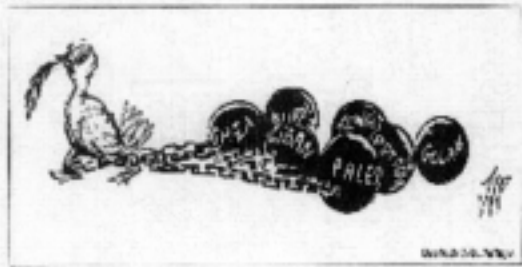
Question 13 : En quelle année a eu lieu la Révolution française ?

Surface de Compétence : 84%

	%	r.bis	CM
1. 1432	0,6	-0,07	42,5
2. 1524	0,6	-0,09	40
3. 1618	1	-0,11	32,9
4. 1789	91,5	0,35	91,8
5. 1815	3,8	-0,21	42,2
6. aucune	1,6	-0,16	38,3
7. toutes	0,1	-0,05	100

Il est rassurant que le taux de réponses correctes (91,5 %) et de certitude moyenne (91,8 %) et la fracture subjective (49,4 %) soient élevés.
 On peut cependant se demander comment Waterloo et la Révolution française peuvent être confondus par 3,8 % des étudiants, à vrai dire les plus faibles (le r.bis = -0,21). Sur 24 enseignants du supérieur interrogés, le plus grand nombre avaient prédit une réussite comprise entre 60 et 70 %. Ils ont donc été heureusement surpris, ce qui fait dire à 56 % d'entre eux que le résultat des étudiants (91,5 %) est très satisfaisant et à 33 % qu'il est satisfaisant, 7 % le jugent « limite » et 4 % (une personne) insuffisant. Pour notre part, nous le jugeons très insuffisant. La fracture subjective est de 49,6 % (très élevée).
Confiance et prudence sont partout où on les attend.

Question 14 : Quelle est la région à laquelle ce document fait allusion ?



Surface de Compétence : 62,80%

	%	r.bis	CM
1. L'Amérique du Sud	0,9	-0,08	36,9
2. Les Balkans	14	-0,25	46,8
3. La Méditerranée	0,9	-0,12	43,1
4. L'Asie du Sud-Est	4,6	-0,18	40,9
5. Le Proche et le Moyen-Orient	74,9	0,42	70,5
6. aucune	3,5	-0,13	32,2
7. toutes	0	XXXX	XXXXX

Nous faisons l'hypothèse que la signification même du distracteur vedette (Les Balkans) n'est pas connue de ces 14 % d'étudiants qui l'ont choisi, ce qu'ils reconnaissent eux-mêmes (certitude moyenne = 46,8 %). La fracture subjective avec ce distracteur est nette : 23,7 %.
Confiance et prudence sont partout où on les attend.

Question 15 : L'adjectif " Précolombien " est attribué :				
Surface de Compétence		%	r.bis	CM
26,10%				
1. à une civilisation, un objet, un art... de l'Amérique latine existant avant la découverte des Amériques par Christophe Colomb au 14ème Siècle		30,9	-0,12	50,8*
2. à un animal qui existait avant l'existence de la Colombe (Crétacé)		1,8	0	46,2
3. aux institutions, régime politique... qui régissaient l'Amérique latine avant l'indépendance de ces pays et la création de la Colombie au 19ème Siècle		11,7	-0,19	35,1
4. à une civilisation, un objet, ou un art de l'Amérique latine existant avant la découverte des Amériques au 15ème Siècle		42,8	0,35	60,9
5. aux civilisations dans l'état où elles existaient avant leur colonisation par l'Occident		8,1	-0,12	44,3
6. aucune		2,3	-0,02	33,1
7. toutes		0,1	-0,07	100

Que le distracteur-vedette (1) ait été choisi par 30,9 % des étudiants et avec une confiance (50,8 %), une **imprudence**, supérieure à celle des autres distracteurs est sans doute dû à deux effets combinés. D'abord, l'erreur, (prévisible) qui consiste à penser que 1492 est au 14^e siècle. Ensuite, une probable précipitation excessive dans le choix d'une solution qui a plusieurs points acceptables, qui « ressemble » à une solution correcte, sans examiner plus avant, puisqu'il apparaît en première position après l'énoncé de la question. La **fracture subjective** est de 10,1 % (faible).
A part ce cas (important), **confiance et prudence** sont conformes aux attentes.

Question 16 : Iconoclaste fait référence à :				
Surface de Compétence		%	r.bis	CM
8,10%				
1. quelqu'un qui détruit les images pieuses, les statues... à l'époque de la Réforme protestante au 16ème Siècle		28,7	0,06	48,2
2. un adorateur d'icônes dans le monde de la religion orthodoxe depuis le 11ème Siècle		15,7	-0,07	32
3. un Vandale, appartenant à ce peuple, lors des grandes invasions qui détruisaient l'Empire Romain d'Occident au 5ème Siècle et qui détruisait tout sur son passage		9,8	-0,07	28,9
4. un partisan d'une idéologie née dans l'empire byzantin au 9ème Siècle et qui refuse l'adoration des images pieuses, allant jusqu'à les détruire		23,3	0,17	34,7●
5. un destructeur des peintures religieuses dans les églises russes durant la période communiste (XXème Siècle)		5,6	0	29,1
6. aucune		9,3	-0,08	20,9
7. toutes		0	XXXXX	XXXXX

Tous les indicateurs statistiques condamnent cette question : le r.bis (0,17) de la solution correcte inférieur à la valeur-seuil (0,20) ; la **fracture subjective** négative aberrante (-13,5 %), puisque la réponse « 1 » est non seulement plus fréquente mais aussi plus assurée (48,2 %) que la réponse correcte (34,7 %), ce qui constitue un **manque de confiance**.
La **prudence** est conforme aux attentes.

Question 17 : Un Kolkhoze est :				
Surface de Compétence		%	r.bis	CM
35,60%				
1. une institution d'Etat en URSS au 19ème Siècle		10,1	-0,14	24,3
2. une grande ferme collective en URSS au 20ème Siècle		55,5	0,47	64,2
3. un membre des jeunesses communistes, en URSS, au 20ème Siècle		10,3	-0,19	24
4. une usine faisant partie d'un Combinat dans le système de l'économie planifiée mise en place en URSS au 20ème Siècle		9,2	-0,1	34,5
5. une communauté agricole en Israël au XXème Siècle		4	-0,1	36,1
6. aucune		4,9	-0,16	16,8
7. toutes		0,2	0,01	33,3

Nous nous attendions à une confusion entre kolkhoze et kibboutz, ce qui ne fut le cas que pour 4 % des étudiants. Le lien entre kolkhoze et URSS est assez fort mais flou (les réponses incorrectes se répartissent sur trois distracteurs comportant URSS). La **fracture subjective** est cependant élevée (30,2 %).
Confiance et prudence sont partout où on les attend.

Question 18 : L'Hégire est :			
Surface de Compétence			
25,90%	%	r.bis	CM
1. un phénomène météorologique qui permit au prophète Mohamed de conquérir l'Arabie	2	-0,08	19,3
2. en 622, la fuite de Mahomet de Médine, début du calendrier islamique	44,7	0,38	60,1
3. événement historique qui marque la fuite de Mahomet de la Mecque au 5ème Siècle et qui date la création de l'Islam	22,4	-0,13	31,4
4. un phénomène historique qui marque, en 742, la naissance de l'Empire musulman	11,1	-0,11	24,3
5. une cérémonie religieuse musulmane célébrée lors des mariages	7,6	-0,13	15,7
6. aucune	3,9	-0,07	12,9
7. toutes	0,1	0,03	0

L'Hégire semble robustement associé à l'Islam, mais de façon floue car les réponses erronées sont dispersées sur les solutions 1, 3, 4 et 5. La fracture subjective est 28,7 % (nette), ce qui est élevé : ceux qui ne savent pas le savent (qu'ils ne savent pas) : leurs certitudes moyennes sont très faibles. **Confiance et prudence** sont partout où on les attend.

Question 19 : Dans l'expression " Une vie spartiate ", le terme spartiate :			
Surface de Compétence			
43,20%	%	r.bis	CM
1. désigne la façon de vivre d'un habitant de la ville grecque de Sparte au 20ème Siècle	4,3	-0,12	34,4
2. fait référence aux habitudes des habitants de Sparte, dans la Grèce Antique, qui sacrifiaient confort et richesse... aux devoirs militaires et à la puissance de leur ville	72,8	0,35	59,4
3. désigne le sort d'un esclave romain qui avait suivi Spartacus lors de la révolte des esclaves qui embrasa la Péninsule Italienne au 1er siècle avant Jésus-Christ	4,9	-0,14	36,2
4. était le nom d'une exigeante combinaison (triple) d'épreuves des Jeux Olympiques de l'Antiquité, qui réunissait le marathon, le lancement du javelot et la lutte	1,6	-0,11	25,2
5. désigne une vie où tous les habitants d'une même ville sont tous traités également comme à Sparte	9	-0,14	30,6
6. aucune	3,3	-0,07	35,3
7. toutes	0,1	0	100*

Le lien entre « spartiate » et la Ville de Sparte est fort : les solutions 1, 2 et 5 attirent 86,1 % des réponses. La fracture subjective (28,8 %) est nette entre les réponses correctes (59,4 % de certitude moyenne) et le distracteur-vedette (30,6 % de certitude moyenne) bien que la certitude moyenne de la réponse correcte soit en lui-même peu élevée. **Confiance et prudence** sont partout où on les attend.

Question 20 : Quelle chaîne de TV crée et diffuse la séquence " No Comment " ?			
Surface de Compétence			
8,50%	%	r.bis	CM
1. La BBC	18,6	-0,18	30,5
2. FR3	2,8	-0,06	23,6
3. CNN	18,2	0,1	46,5
4. RTL	1,4	0	38
5. La Rai	0,5	-0,09	37,1
6. aucune	52,8	0,13	82,3
7. toutes	1,6	-0,06	76,5*

La réponse correcte est « Euronews », absent des solutions proposées ! Ce que semblent savoir 52,8 % des étudiants, très sûrs d'eux-mêmes (82,3 %). Le r.bis de la solution correcte (0,13) est cependant inférieur à la valeur repère (0,20). Sur cette base, cette question serait éliminée d'une épreuve sanctionnante. Or la certitude très élevée (82,3 %) des réponses correctes et la fracture subjective (51,8 % avec la solution 1 ; 35,8 % avec la solution 3 et 43,8 % avec la moyenne des deux) très élevée suggère de garder cette question. **Confiance et prudence** sont partout où on les attend, sauf pour le distracteur "Toutes", un peu comme si ces étudiants étaient persuadés que ce genre d'émission était générique, comme l'est le JT.

Question 21 : Quel journal appartient au groupe de presse Rossel ?

Surface de Compétence		%	r. bis	CM
9,90%				
1. Le Monde		22,1	0	18,5
2. Le New York Times		18,8	-0,07	22,4
3. Le matin		6,8	-0,02	20,2
4. Le Soir		22,7	0,29	43,5*
5. La Libre Belgique		9,9	-0,06	19,9
6. aucune		6,9	-0,04	12
7. toutes		0,5	-0,04	20

Rares sont ceux (22,7 %) qui connaissent la réponse correcte et ils ne sont pas sûrs (**manque de confiance** : la certitude moyenne vaut 43,5%).
Ceux qui lisent Le Soir à la maison sont sans doute minoritaires ... et avantagés dans la réponse à cette question, dont la **fracture subjective** avec le distracteur-vedette (1. Le Monde) est nette (25 %).
La **prudence** est conforme aux attentes.

Question 22 : Quel était, approximativement, le taux d'inflation annuel de la Belgique en 1998 ?

Surface de Compétence		%	r. bis	CM
11,50%				
1. 1%		25,1	0,43	45,7*
2. 5%		29,6	-0,05	25,1
3. 10%		16,6	-0,1	19,6
4. 15%		8,9	-0,06	28,7
5. 20%		3,6	-0,14	31
6. aucune		6,4	-0,11	25,9
7. toutes		0,1	-0,02	0

Le distracteur « 5 % » est plus populaire (29,6 %) mais beaucoup moins assuré (25,1 %) que le choix de la solution correcte (25,1 %) qui, lui, **manque de confiance** (45,7 %).
La **fracture subjective** vaut 20,6 % (nette).
La **prudence** est partout où on l'attend.

Question 23 : Combien a-t-on de FB (environ) pour un Euro ?

Surface de Compétence		%	r. bis	CM
84,20%				
1. Dix		1,1	-0,09	32
2. Vingt		1,1	-0,09	44
3. Trente		2,5	-0,07	51,1*
4. Quarante		92,8	0,24	90,7
5. Cinquante		0,8	-0,1	45,5
6. aucune		1,1	-0,14	44
7. toutes		0,1	-0,03	10

Cette question était au cœur de l'actualité de l'époque. Elle va devenir une donnée historique et devrait être marquée par l'oubli. C'est un des exemples où le but de l'éducation est qu'une population entière oublie une connaissance. La **fracture subjective** vaut 39,6 % (élevée).
Confiance et prudence sont partout où on les attend, sauf pour le distracteur-vedette (Trente), le plus proche sur le fond de la réponse correcte .

Question 24 : Qui délivre le minimex ?

Surface de Compétence		%	r. bis	CM
43,60%				
1. Le FOREM		6,4	-0,06	35,2
2. L'ONEM		13,7	-0,07	45,2
3. Le CPAS		66,3	0,28	66,7
4. La mutuelle		5,3	-0,13	30,9
5. Le syndicat		2,3	-0,11	29,4
6. aucune		3	-0,1	36,3
7. toutes		0,1	0	0

Deux étudiants sur trois (66,3 %) savent, et de façon moyennement assurée (66,7 %) comment fonctionne ce mécanisme social, avec une **fracture subjective** nette (21,5 %) par rapport à ceux qui pensent que c'est l'ONEM (distracteur-vedette).
Confiance et prudence sont partout où on les attend.

Question 25 : Le taux de chômage était, en Wallonie, en 1998, aux alentours de :			
Surface de Compétence	13%	%	r.bis
1. 5%	2,9	-0,05	43,9
2. 10%	17,3	0,07	38,8
3. 15%	34,1	0,28	38●
4. 20%	22,5	-0,08	33,4
5. 25%	11,8	-0,18	38,6
6. aucune	6,3	-0,14	36,4
7. toutes	0,1	0,02	0

Un tiers des étudiants (34,1 %) ont répondu correctement et ce sont bien des étudiants supérieurs à la moyenne (r.bis = 0,28). Cependant, leurs réponses étaient très peu assurées (leur certitude moyenne était 38 %), indiquant donc un manque de confiance. La fracture subjective est insignifiante (4,6 %) avec le distracteur-vedette. La prudence est conforme aux attentes.

Bilan

Indicateurs psychométriques :

Les indices de discrimination (r.bis) sont quasi tous conformes aux attentes, que ce soit pour les 25 solutions correctes ou pour les 150 distracteurs (seuls 9 ne sont pas négatifs). Aucune critique ne peut donc être adressée à ces questions quant à leur cohérence interne avec l'ensemble de l'épreuve.

Indicateurs édumétriques :

La confiance combinée à la prudence sont observées pour 13 des 25 questions.

La prudence totale est observée pour 6 des 12 questions qui restent et la confiance pour 5 de ces 12 questions qui restent.

Pour 6 questions, dans lesquels 8 distracteurs sont fournis avec imprudence, il s'agirait, de « rectifier » les idées fausses.

Partie III : Synthèse et Conclusions

Chapitre 21

Synthèse des résultats

D. Leclercq, Président du Groupe de Travail CIUF « Réussite en candidatures »

1. **MOHICAN : Quelles pertinences et quelles faisabilités ?**
2. **MOHICAN : Quels résultats ?**
3. **La synthèse des résultats check-up par check-up**
4. **L'impact de variables biographiques (du passé individuel) sur la Réussite en 1^{re} candi**
5. **Utiliser les données MOHICAN pour améliorer la réussite via des variables changeables**
6. **Prédiction et sélection : mises en garde**

Résumé :

Ce chapitre traite des apports et des limites du projet MOHICAN, laissant entrevoir des perspectives de prolongements qui tireraient parti de ces points positifs. Au terme de nos analyses, nous jugeons cette expérience à la fois pertinente pédagogiquement, réalisable politiquement et institutionnellement, informative tant pour les professeurs que pour les étudiants, de nature à contribuer à renforcer les compétences métacognitives des étudiants et surtout à favoriser leur réussite en candidature.

Pour inspirer de futures expériences aux niveaux communautaire ou local, cette synthèse aborde les limites du projet et les améliorations conceptuelles et pratiques à lui apporter.

Enfin, après avoir spécifié le champ d'usage des données récoltées, champ axé sur la guidance (et l'auto-guidance) plutôt que sur une prédiction visant la sélection (et l'auto-sélection), ce chapitre rejoint les réflexions en cours sur le rôle des universités et les droits de leurs étudiants.

1. MOHICAN : quelles pertinences et quelles faisabilités ?

1.1 La pertinence pédagogique d'un projet comme MOHICAN

En 1998, les résultats décevants en Sciences⁴⁰ (Monseur, 1998) des étudiants de 14 ans de la Communauté française ont attiré l'attention du public et des médias. Plus récemment, ce sont les résultats relatifs des étudiants de 15 ans en Compréhension de lecture (Lafontaine et al, 2001) qui constituent une nouvelle onde de choc remontant jusqu'au Parlement de la Communauté française qui, en conséquence, inscrit à son programme l'examen des causes de ces observations (Dupont, 2002). Par ailleurs, les taux de réussite en première année universitaire restent toujours aussi décevants (en moyenne 40 % pour les primants) avec une grande constance sur ces vingt dernières années (voir Introduction, section A).

De plus, très nombreuses sont les voix s'élevant pour « constater » une grande différence de préparation selon l'école secondaire d'origine et surtout pour proclamer l'existence d'une « baisse du niveau » des étudiants lors de l'entrée à l'université. Cette dernière conviction ne peut être confrontée à des données objectives dans la mesure où aucune épreuve externe (comme le Baccalauréat en France) n'existe à l'échelle de la Communauté Française Wallonie Bruxelles.

Enfin, les centres de guidance universitaire ne voient affluer les étudiants préoccupés par leur méthode d'étude qu'après les interrogations partielles de janvier, ce qui est déjà fort tard et laisse peu de temps pour des actions de remédiation qui prétendraient porter leurs fruits en juin ou septembre de la même année.

Dans un tel contexte, est-il pertinent de concevoir des épreuves qui, de façon précoce (dès l'entrée à l'université) donneraient des informations diagnostiques aux étudiants afin de leur permettre, avec l'aide de leurs enseignants, de prendre des mesures, précoces elles aussi, quant aux méthodes et contenus d'étude ? C'est le pari qu'a voulu relever le projet MOHICAN.

Le groupe de pilotage du projet MOHICAN ne considère pas que l'enseignement secondaire ait pour objectif principal la préparation à l'enseignement supérieur. Il trouve cependant légitime que les enseignants du supérieur désirent savoir en quoi ceux qui ont choisi d'entrer dans ce niveau d'étude et dans une filière particulière y sont préparés. Par ailleurs, que les étudiants disposent en premier lieu et de façon confidentielle des informations qui les concernent personnellement nous paraît un principe évident pour des raisons déontologiques et pédagogiques.

Le projet MOHICAN a conçu des épreuves intitulées « check-up » mais en en connaissant certaines faiblesses dès le départ, liées, pour une bonne part à la contrainte de temps. En effet, il a été clair dès le début que 3 heures seraient le maximum octroyé par les universités à cette opération de réponses aux check-up. D'où, pour la majorité des check-up le petit nombre de questions par matière⁴¹, entraînant des choix difficiles dans leur sélection et une restriction de leur pouvoir diagnostique.

Le check-up de **vocabulaire** portait sur le langage abstrait commun non technique que l'on devrait connaître à l'entrée d'une faculté de Lettres. Il n'était ciblé sur aucun domaine particulier et devait donner une idée générale des compétences vis-à-vis du lexique de la langue française, **le versant métacognitif** (*Est-ce que j'ai conscience de ce que je ne comprends pas ?*) étant aussi important.

⁴⁰ Third International Mathematics and Science Study (TIMSS)

⁴¹ à deux exceptions près : le vocabulaire (45 questions) et les mathématiques (22 questions) ; les 25 questions sur l'Art portent sur des matières très différentes allant de la peinture à la musique en passant par la danse et l'architecture. Les 25 questions sur l'Histoire, l'actualité et l'économie elles aussi portent sur des matières très hétérogènes.

Le check-up de **mathématique**, grâce à un nombre suffisant de questions (22) a pu « couvrir » des exemples de toutes les modalités de traduction d’un langage (tableau, graphique) à l’autre, ce qui permet des diagnostics quant aux sens des traductions les moins maîtrisées (voir chapitre 3). Plutôt que de couvrir le programme de mathématique du niveau CESS (Certificat de l’Enseignement Secondaire Supérieur), cette épreuve visait à tester des compétences transversales telles que traduire des formules en phrases, en graphiques, en tableaux et vice versa, bref la capacité de « parler mathématique » dans tous les sens.

Les capacités de **lecture de graphiques, de cartes, de tableaux** ont été mesurées chez tous sur un contenu **géographique**. Cela aurait pu être sur un autre contenu (histoire, économie, sciences, etc.) et même idéalement sur plusieurs.

Le check-up de **syntaxe**, très structuré, permet lui aussi des diagnostics, mais limités vu le faible nombre de questions (12), limite plus marquée encore dans le check-up « **Compréhension de Textes** » (6 questions). Y multiplier le nombre de questions eût entraîné un allongement du texte à lire, du temps nécessaire à cette lecture... et ce check-up à lui seul aurait consommé l’entièreté du temps disponible.

Les épreuves de **physique, chimie et biologie** (respectivement 10, 8 et 10 questions) sont, on s’en doute, des « coups de sonde » sur des matières très variées.

Il en va de même pour le check-up des **Connaissances artistiques** qui, malgré ses 25 questions, reste une collection de coups de sonde en littérature (5 questions), théâtre (2 questions), cinéma (6 questions), musique (2 questions), architecture (4 questions), peinture-sculpture (4 questions), danse (2 questions).

De même, le check-up **Histoire-Economie-Actualité** se décompose en Histoire et vocabulaire à référence historique (9 questions), Economie (4 questions), Organismes internationaux (6 questions), Actualités (6 questions).

Conclusions : Dès l’origine du projet, le comité de pilotage du projet MOHICAN savait que, vu la constitution même de ces check-up, ils représentaient plutôt une série de **coups de sonde** dans lesquels **le volet métacognitif était souvent prépondérant** par rapport au contenu-matière. Il savait aussi que si, à l’avenir, les universités acceptaient le principe de consacrer une journée entière à ce genre d’opération (idéalement le même jour dans toutes les sections), on pourrait se focaliser plus sur l’évaluation de **compétences transversales**, y compris la **capacité d’apprendre**. Enfin, le comité était conscient dès le départ que **des check-up ciblés par section** (médecine, pharmacie, etc.) plutôt que communs à plusieurs facultés apparaîtraient comme **plus pertinents pour les acteurs de terrain** que des check-up communs à plusieurs facultés. Malgré ces restrictions, le comité a choisi de mener cette action MOHICAN potentiellement porteuse d’autres informations précieuses. On en jugera dans les sections qui suivent.

1.2 La faisabilité politique d’épreuves externes en Communauté française de Belgique

La Communauté française de Belgique, par son absence quasi totale de sélection à l’entrée de l’enseignement supérieur, contraste avec la plupart des pays qui l’entourent, qui, eux, pratiquent une sélection le plus souvent via les principes du *numerus clausus*. Le site Eurybase d’Eurydice nous permet de présenter le panorama qui suit ⁴².

⁴² Nous avons mis en italique nos traductions des textes d’Eurybase

En **Allemagne**, l'accès est, en principe, garanti à tous, mais « *Par suite du grand nombre de candidats et du nombre insuffisant de places, il y a actuellement (semestre d'été 1999) des restrictions aux cours universitaires dans les domaines suivants : architecture, administration des affaires, biologie, économie alimentaire, chimie alimentaire, médecine, pharmacie, psychologie, droit, médecine vétérinaire et dentisterie. Les places disponibles sont attribuées via une procédure centralisée de sélection. Les cours qui en font partie peuvent varier de semestre en semestre. Plusieurs institutions d'enseignement supérieur restreignent aussi l'admission de façon locale... Le critère principal est la moyenne des notes obtenues à l'Abitur, examen de sortie du secondaire et de capacité à l'enseignement supérieur* »⁴³.

En **Angleterre**, au Pays de Galles et en Irlande du Nord, le principe de libre accès est le même, à condition d'avoir obtenu un certain résultat minimum au General Certificate of Education Advanced Level ; mais « *en pratique, suite à la compétition pour les places, de nombreuses institutions exigent des niveaux de qualification nettement au-dessus du minimum... Les candidatures sont habituellement soumises en automne, un an avant le début des cours. En 1998-99, le coût d'inscription en candidature était de 1000 £.* ».

En **Hollande**, « *il existe un système centralisé d'admission pour distribuer les places aux diverses universités. Ce sont les institutions ou le Ministère de l'Education, de la culture et de la science, qui décident annuellement quels cours sont l'objet de quotas d'admission (numerus clausus).* »

En **Espagne**, pour obtenir le droit d'accès au premier cycle (nos candidatures) il n'est pas requis de passer les examens externes du *Bacchilerato*, mais « *si l'on garde à l'esprit que les étudiants qui ont passé ces examens ont la priorité sur ceux qui ne l'ont pas passé et dans la mesure où à peu près toutes les filières d'étude ont un nombre de places limité, en pratique, passer ces examens constitue un passage obligé* ».

En **Italie**, les étudiants sont soumis à la « Maturita », épreuve nationale externe. Certaines études (médecine, pédagogie, etc.) sont contingentées au niveau national : chaque institution universitaire est informée du nombre de places qui lui est alloué, avec mission de sélectionner sur base d'un concours d'entrée spécifique. Dans les disciplines non contingentées par le pouvoir national, la majorité des universités ont instauré leur propre système de sélection.

Enfin en **France**, en plus de l'examen externe national que constitue le Baccalauréat, des concours d'entrée sont systématiques pour les Grandes Ecoles (Polytechnique, Ecoles de mines pour ingénieurs, ENA, etc.) et pour certaines sections universitaires comme par exemple la médecine vétérinaire, la kinésithérapie, la pharmacie et la logopédie. Ceci explique un flux d'étudiants français dans nos universités.

Ce phénomène de « dernier bastion de l'accès libre à l'enseignement supérieur »⁴⁴, combiné au constat répété depuis deux décennies du faible taux de réussite en première année (40% pour les primants) dans les universités de la **Communauté Française Wallonie Bruxelles** amène certains à réclamer l'instauration chez nous aussi d'une sélection à l'entrée. Les fédérations d'étudiants redoublent donc de méfiance vis-à-vis de toute épreuve standardisée passée avant ou juste au début du supérieur qui pourrait être utilisée à des fins de recherche d'abord pour établir la prédictivité des résultats, puis à des fins de sélection sur base de cette prédictivité si elle était démontrée grâce à cette opération. L'année 1999 avait été l'objet d'affrontements à ce sujet entre certaines autorités académiques et les fédérations d'étudiants.

⁴³ Texte extrait du site www.eurydice.org/Eurybase/Description des systèmes éducatifs nationaux, 1999/6B. Les citations (traduites) pour la Hollande, l'Espagne et l'Angleterre ont la même source

⁴⁴ Le fait de constituer une exception n'est pas à nos yeux un défaut, au contraire même quand il s'agit de législation sociale marquée fortement par la solidarité collective, ou d'universalité de la compétence juridique pour les crimes contre l'humanité.

Dans un tel contexte, est-il possible de proposer des épreuves standardisées communes à tous les étudiants entrant dans les universités de la Communauté française de Belgique ?

Le projet MOHICAN a montré que cela était possible, même s'il ne s'est déroulé que dans 8 des 9 universités et s'est limité lui-même à un tiers des étudiants environ (mais de toutes facultés et sections). Il faut cependant insister sur le fait que ce projet s'est déroulé dans des circonstances politiques favorables, puisque le gouvernement venait de s'engager, par l'article 17 de sa déclaration gouvernementale, à ne pas instaurer de nouvelles épreuves de sélection durant la législature. En outre, les responsables du pilotage du projet, professeurs issus des 9 universités, se sont engagés (voir Introduction) à ne pas utiliser les résultats de la recherche MOHICAN dans une perspective sélective. Les fédérations d'étudiants ne se sont pas opposées à MOHICAN, ce qui n'a pas réduit les craintes à néant, comme en témoignent les 2% d'étudiant(e)s qui, lors d'un sondage (cf. chapitre 5, section c) ont choisi, parmi les raisons (proposées) pour lesquelles ils n'ont pas consulté le feedback qui leur avait été distribué, l'explication « Je ne souhaite pas cautionner un système de tests, quel qu'il soit, en début de cursus universitaire ». De telles circonstances politiques favorables à une semblable opération perdureront-elles ? L'avenir le dira.

Conclusion : Une évaluation externe interuniversitaire non sanctionnante n'est pas à exclure par principe en Communauté française de Belgique puisque MOHICAN a démontré que des accords nécessaires entre diverses parties étaient possibles et que ces accords ont été respectés. Néanmoins, le souci démocratique exprimé par cette Communauté via la liberté d'accès et par le projet MOHICAN via la volonté d'information précoce ne constitueront des mesures efficaces que si elles s'assortissent de mesures pédagogiques telles que l'accompagnement personnalisé des étudiants ou telles que des réformes méthodologiques donnant de l'importance aux façons d'apprendre, aux compétences démultiplicatrices (méthodes d'apprentissage) tout autant sinon plus qu'aux compétences spécifiques (propres à un domaine disciplinaire).

1.3 La faisabilité institutionnelle d'épreuves communes interuniversitaires

La conception et la mise au point d'épreuves standardisées communes est une tâche de longue haleine. La récolte des données sur le terrain dans de nombreux sites et le retour rapide d'informations personnalisées et confidentielles est un autre défi. Enfin le recueil de résultats complémentaires un an après les premiers, suivi du traitement statistique multivarié des résultats constitue une troisième entreprise fastidieuse. La lourdeur de ces opérations n'a rien à envier à celles des enquêtes internationales telles que TIMSS⁴⁵ et PISA⁴⁶.

Une telle entreprise était-elle possible ?

Dans le projet MOHICAN, grâce au CIUF et aux subsides ministériels de l'enseignement supérieur, les 9 universités ont pu unir leurs forces. Tout d'abord dans la conception des épreuves, ce qui a été facilité par l'implication d'experts, qui heureusement ne manquent pas dans les universités, selon un principe de partage des tâches et d'addition des talents. Ensuite dans la concentration des traitements informatiques et de la production des feedbacks personnalisés en un lieu unique (le SMART – ULg), puis dans la distribution des feedbacks dans les nombreuses sections par le membre du groupe de pilotage représentant son université. L'exploitation des résultats était laissée à la liberté de chaque institution.

⁴⁵ Third International Mathematics and Science Study, menée par l'International Association for the Evaluation of Educational Achievement (IEA).

⁴⁶ PISA = Programme International pour le Suivi des Acquis des élèves de 15 ans.

Conclusion : Les check-up les plus intéressants n'étant pas les mêmes dans les différentes sections, il serait pertinent, ultérieurement, de concevoir des opérations MOHICAN plus « ciblées » sur des facultés ou sections particulières, chacune ayant son groupe de conception et de pilotage plus restreint que celui de MOHICAN, tout en gardant éventuellement des épreuves « transversales » communes à tous. Ce que l'ensemble des 9 universités a pu faire, un groupe plus restreint de 3 ou 4 le peut. En outre, chaque partenaire serait ainsi plus impliqué dans l'exploitation des données sur le terrain, avec les étudiants, ce qui augmenterait l'efficacité de l'opération.

1.4 La valeur informative de feedbacks pour les professeurs

Les avis sont souvent partagés entre enseignants sur le fait de savoir s'ils retireraient une quelconque information de feedbacks comme ceux qui ont été fournis par MOHICAN. Si nous nous focalisons sur le seul exemple de la compréhension par les étudiants des mots du vocabulaire général de la langue française, certains enseignants ont la conviction qu'ils savent, par expérience, quels mots posent problème à leurs étudiants. C'est certainement vrai pour les termes techniques propres à leur discipline et sur lesquels ils interrogent par l'écrit et par l'oral. Mais pour les termes généraux, en va-t-il de même ? En partie sans doute puisque dans diverses universités des linguistes ou des spécialistes de la pédagogie universitaire font des enquêtes auprès de leurs collègues en leur demandant de fournir ces mots « à problème ».

Qu'apportent de plus en précision, en information, aux professeurs, les taux de réussite des étudiants de leur section pour chacune des questions (chacun des mots du check-up de vocabulaire par exemple) ? Avec quelle précision ceux-ci sont-ils capables de prédire les difficultés de leurs étudiants objectivées par une épreuve standardisée ?

Un sondage, dans une faculté (chapitre 6) montre que ces professeurs estiment assez bien l'ordre ($r=0,56$) des difficultés de compréhension de divers termes pour leurs étudiants. Cependant, en valeur absolue, ils sous-estiment la difficulté des mots concernés pour leurs étudiants puisqu'ils surestiment leurs taux de réussite en moyenne de 20%. Enfin, de nombreux mots constituent des « surprises » heureuses (les étudiants réussissent mieux que ce que leurs enseignants pensaient) ou, beaucoup plus souvent, des surprises malheureuses (les étudiants réussissent moins bien qu'espéré). Donc les prédictions subjectives des professeurs ont, dans ce domaine et dans notre expérience limitée à une section, une certaine validité statistique globale ; cependant une prédiction valide est impossible pour un mot en particulier. Les professeurs sont donc dans la même situation que les médecins qui savent quelles maladies ont les patients qu'ils n'ont pas encore vus et qui sont dans leur salle d'attente. Cette boutade de Lindley (1971) veut simplement dire que les médecins savent beaucoup de choses sur les statistiques de maladies en général et que, dans leur région et à cette période de l'année, c'est surtout tel type de maladie qui amène les patients à leur consultation. Cela ne les libère nullement de l'obligation d'ausculter les patients un par un car la prédiction individuelle, elle, n'est pas possible sans l'examen de chaque cas.

Conclusion : Si l'on extrapole à l'ensemble des check-up ce qui a été observé pour l'épreuve Vocabulaire à une section, des sondages avec feedback aux enseignants paraissent informatifs. Dans le cas précis du Vocabulaire, c'est évidemment sur des mots utilisés dans le cours de chaque enseignant que ce genre de feedback devrait porter. C'est pourquoi chaque enseignant pourrait (devrait ?) développer son propre check-up de vocabulaire, dans le cadre d'une stratégie complexe (voir chapitre 10). Cette stratégie pourrait porter dans un premier temps, sur l'instauration du réflexe de recourir au dictionnaire quand il n'est pas encore suffisamment présent, en se donnant les moyens d'observer l'évolution de cette pratique pour son propre cours.⁴⁷ Dans un deuxième temps, le professeur pourrait assurer la maîtrise de certains de ces termes de la langue française, en relevant devant les étudiants l'intérêt de telle ou telle nuance, la richesse de telle précision ou les problèmes posés par telle erreur dans le choix d'un terme, d'une conjonction, etc. L'enjeu est ici de sensibiliser les étudiants à l'importance de la maîtrise de la langue, à la rigueur dans l'expression et dans son interprétation.

1.5 Feedbacks détaillés, personnalisés, rapides et confidentiels

On sait⁴⁸ que si les réponses correctes sont communiquées aux étudiants, on ne doit pas craindre un effet pervers⁴⁹ des QCM qu'est la mémorisation possible de solutions fausses. Pour cette raison, les réponses correctes ont été rendues disponibles par affichage et par distribution de photocopies dès la semaine qui a suivi les check-up (pas avant, pour éviter leur éventuelle circulation auprès d'étudiants non encore évalués); ces réponses correctes ont aussi été systématiquement fournies avec les feedbacks individuels. On sait par ailleurs que si le feedback est important pour la consolidation des savoirs⁵⁰, un délai de plusieurs jours, voir de plusieurs semaines n'est pas un handicap⁵¹. L'intention des check-up MOHICAN n'était pas « *d'augmenter la fréquence des réponses correctes* » (car les questions des check-up ne sont pas les mêmes que celles auxquelles les étudiants seront soumis dans leurs examens universitaires) mais visait à assurer « *le rôle informatif joué par une procédure de feedback capable d'aider l'élève à localiser puis à dépasser ses difficultés* » (Depover, 1987, 32). Pour ce faire, les feedbacks doivent être très diagnostiques, dépasser la seule « note », par exemple en indiquant pour chaque question avec quelle **confiance** l'étudiant a réussi et avec quelle **prudence** il s'est trompé. Cette mesure du réalisme des étudiants est tout spécialement importante pour les connaissances fallacieuses (méprises) où les étudiants sont sûrs à 100% de réponses qui pourtant sont erronées.

Dans le contexte des universités de la Communauté française de Belgique, est-il possible de fournir des détails diagnostiques dans des délais raisonnables tout en respectant la confidentialité promise ?

Dans les trois semaines qui ont suivi le check-up (donc **avant fin octobre**), tous les étudiants ont reçu leur enveloppe scellée (confidentielle) contenant une page par check-up. Plus précisément, tous ont reçu les 4 pages des épreuves communes, comportant des feedbacks sur 73 questions au total. Les étudiants en sciences humaines ont reçu 2 pages supplémentaires (Art et Histoire-Economie-Actualité) comportant des feedbacks sur 50 questions et les étudiants en sciences de la nature ont reçu 4 pages supplémentaires (math, physique, chimie, biologie) comportant des

⁴⁷ comme l'ont fait Leclercq et al. (2002).

⁴⁸ Cette démonstration expérimentale est due à Karraker (1967) cité par Leclercq (1986, 38-40).

⁴⁹ Cet effet a été démontré par Preston (1965) cité par Leclercq (1986, 36-37)

⁵⁰ Selon les travaux de Kulhavy et Anderson (1972) et de Surber et Anderson (1975), cités par Depover (1987, 30)

⁵¹ Selon les travaux de Sassenrath et Yonge (1969) et de Sturges (1969) cités par Depover (1987, 31)

feedbacks sur 50 questions au total. Dans chaque feuille de résultat figurait la position spectrale de chaque réponse, c'est-à-dire son exactitude (erronée à gauche et exacte à droite), sa certitude : (de -100% à +100%, le signe « moins » indiquant que la réponse était erronée), et son réalisme. Il s'agit d'un **degré de précision** jamais fourni jusqu'ici à un tel nombre d'étudiants (qui n'étaient, malheureusement pas préparés à les interpréter). Nous espérons que la présente étude contribuera à familiariser avec ce genre de données.

Par ailleurs, si les épreuves MOHICAN ne s'appellent pas « tests » mais « check-up », c'est pour affirmer leur intention formative et « de routine », de même que l'on se soumet à un check-up médical « en routine », sans qu'aucune alerte de santé ne se soit manifestée, sans être motivé par une pathologie existante. On sait que les check-up médicaux sont souvent pratiqués de manière volontaire, dans l'intérêt total de la personne examinée, et qu'ils débouchent le cas échéant sur des conseils de consultations approfondies, jamais sur des conduites imposées à la personne. Dans ces check-up médicaux, la **confidentialité** est une caractéristique fondamentale à laquelle les personnes qui se soumettent à l'examen sont très attentives car elles savent que certaines institutions (les assurances par exemple) sont tentées de se servir de données de cette nature pour sélectionner ou rejeter leurs clients ou leurs employés.

MOHICAN permet de répondre par l'affirmative quant à la capacité de respecter la confidentialité, après avoir distribué environ 4000 dossiers individuels dans 8 institutions et plus de 50 sections différentes. Aucune plainte pour rupture de confidentialité ni utilisation abusive de données n'a été rapportée aux membres du groupe de pilotage. Ni de la part d'étudiants, ni de la part de sections ou d'universités auxquels la confidentialité avait été promise dès le départ. Les institutions identifiables dans les rapports (chapitres 4, 5 et 6) ont marqué leur accord pour qu'il en soit ainsi. Dans ces chapitres, il s'agit de contenus sur lesquels une comparaison inter-institutions n'est pas possible, les données n'étant pas disponibles.

Conclusion : Il est possible de fournir des feedbacks détaillés et personnalisés à un grand nombre d'étudiants et de sections dans des délais raisonnables, même dans une procédure inter-institutions. La formule « un responsable informé par université, tenu lui-même par la confidentialité » a permis de conjuguer efficacité sur le terrain et respect de la confidentialité promise à tous les niveaux.

1.6 La consultation des feedbacks MOHICAN par les étudiants

Noël et Romainville (1998, 143) notent, en parlant des programmes de remédiation, que « *les étudiants hésitent à s'y engager (Denef et al., 1990). Plusieurs raisons peuvent être avancées pour expliquer que des étudiants qui en auraient pourtant bien besoin n'ont pas recours aux accompagnements pédagogiques proposés (Romainville, 1992). D'abord, ils craignent d'être catalogués comme « étudiants à problème » et donc marginalisés. [...] Ensuite, la **logique étudiante de première année**⁵² n'est pas celle de l'apprentissage mais celle de la réussite. Comme la plupart d'entre eux perçoivent cette dernière comme étant en partie une sorte de loterie, ils ont tendance à ne pas s'inscrire dans des programmes remédiatifs en cours d'année : « j'essaye quand même, on ne sait jamais » [...] Enfin, l'étudiant en difficulté n'ose pas toujours se l'avouer : admettre, dès le milieu de l'année, que l'échec est l'issue la plus probable demande un certain courage. Certains préfèrent ignorer cette menace pourtant quasi certaine et n'avouer leur échec à leur parents, à leur entourage voire à eux-mêmes que lorsque celui-ci ne peut plus être camouflé ».* Ce qui vient d'être dit sur les contenus des cours et au milieu de l'année nous paraît s'appliquer, *mutatis mutandis*, aux mécanismes présents au début de l'année pour les feedbacks MOHICAN.

⁵² C'est nous qui soulignons

Ce contexte étant rappelé, les étudiants consultent-ils l'information précoce mise à leur disposition et qu'en font-ils ?

Trois stratégies institutionnelles différentes ont été appliquées suite à la communication des résultats individuels aux check-up MOHICAN. A la FUCAM, tous les étudiants ont été reçus un par un par B. Noel, dossier MOHICAN en mains, pour un entretien relatif aux méthodes d'étude et à l'adaptation universitaire, entretien qui a largement débordé les données MOHICAN. A l'ULg a eu lieu l'opération RESSAC (voir chapitre 8). Dans les autres universités, l'exploitation des données a été laissée au soin des étudiants. Qu'ont-ils fait de ces informations ?

Un sondage sur ce point (voir chapitre 5) a eu lieu, par questions fermées⁵³. Il s'est déroulé dans deux institutions différentes auprès de deux sections contrastées : une en Psycho-pédagogie (les réponses de 76 étudiants ont été obtenues) et l'autre en Polytechnique-Ingénieurs (les réponses de 154 étudiants ont été obtenues). Selon les réponses (non exemptes d'un possible phénomène de « désirabilité sociale »⁵⁴), 80% des étudiants ont consulté les feedbacks (respectivement 72% et 84% dans ces deux sections), soit 209 sur les 260 étudiants qui ont répondu.

Les motivations à consulter les plus choisies sont, d'abord « Par intérêt pour les résultats personnels » (63 et 86%), puis, très loin derrière, « Par curiosité pour les réponses correctes » (27% et 6%).

Ce qui les a intéressés dans ces informations est

1. « Mon taux de réponses correctes pour les différentes parties » (33% et 42%),
2. « Ma position à l'intérieur du groupe » (20% et 23%),
3. « Mon taux global de réponses correctes » (25% et 18%),
4. « Les informations relatives à mes degrés de certitude » (5% et 11%)
5. et enfin « Les solutions correctes aux questions » (16% et 5%).

Les raisons de la non-consultation choisies (parmi les raisons proposées) par ceux qui annoncent ne pas avoir consulté (20% soit 51 sur 260 répondants) sont, dans l'ordre,

1. « Les questions ne sont pas représentatives, dans leur contenu, de ce qui est attendu de moi à l'université » (28% et 78%),
2. « Idem dans leur niveau de difficulté » (10% et 77%),
3. « Quels que soient mes résultats, ils ne changeront pas ma manière de travailler » (56% et 27%),
4. « J'ai oublié » (42% et 27%),
5. « Je n'ai pas répondu avec tout le sérieux nécessaire et les résultats ne sont pas le reflet de mes capacités » (21% et 20%),
6. « Le simple fait de répondre aux questions me donne une indication sur mon niveau » (32% et 10%),
7. « J'ai perdu mon numéro de check-up » (10% et 20%),
8. et enfin « Je ne souhaite pas cautionner un système de tests, quel qu'il soit, en début de cursus universitaire » (21% et 0%).

Les arguments (3) et (5) sont à rapprocher de ce que Noël et Romainville (1998, 143) ont appelé « la logique estudiantine de première année ».

⁵³ les étudiants doivent choisir parmi des solutions proposées.

⁵⁴ mécanisme par lequel des répondants à un questionnaire ont tendance à fournir comme réponse non pas ce qu'ils pensent, mais ce qu'ils pensent que l'interrogateur attend que soit ces réponses, et ce dans le but de donner une certaine image d'eux-mêmes.

Une large majorité des étudiants (63% et 72%) jugent les résultats **conformes à leurs prévisions**. Pour un tiers (15% et 43%), ils sont rassurants. Aucun n'a été découragé par les résultats. Dans leur très grande majorité, les étudiants (90%) n'ont pas l'intention de consulter une guidance ni (75%) d'augmenter leur quantité de travail.

Quant à la difficulté d'interpréter les feedbacks, elle est signalée par la majorité des étudiants (80% et 49%). Il faut reconnaître qu'il y a là une « avalanche de chiffres », pour des étudiants qui n'ont pas été préparés à les interpréter. Etant donné la forme numérique et graphique des feedbacks, il n'est pas étonnant que les Ingénieurs déclarent moins avoir des difficultés à interpréter.

Conclusion : Les étudiants disent être intéressés par les feedback et les consulter, mais des améliorations devraient être apportées quant à la pertinence de leur contenu, à leur lisibilité et au suivi de leur utilisation réelle à des fins formatives.

2. MOHICAN : quels résultats ?

2.1 *Le niveau d'entrée en première candidature baisse-t-il ?*

Cette question « du niveau » fait l'objet, dans tous les pays -et le nôtre n'y échappe pas- de débats périodiques entre formateurs et dans le grand public. Pour les tenants d'une position alarmiste⁵⁵, « le niveau baisse », et même depuis très longtemps. A l'inverse, Beudelot et Establet (1989), représentant le courant « relativiste », ont rédigé un livre entier pour prouver le contraire : « Le niveau monte » (c'est le titre de leur livre). Or ce débat est difficile pour trois raisons.

Première source de difficulté : **le niveau requis change**. Il n'y a pas de consensus sur la pérennité de la valeur de certains savoirs ou de certaines compétences et donc pas de consensus sur la pertinence de toute comparaison avec le passé sur ces matières⁵⁶. Le monde change rapidement ; les objectifs de l'éducation aussi.

Deuxième source de difficulté : **les données comparatives sont rares**. Même quand on est d'accord sur des points « communs », on dispose rarement des données de base (les résultats de deux cohortes comparables⁵⁷, éloignées de deux décennies par exemple) qui permettraient les comparaisons. De telles données existent cependant, mais elles sont contradictoires. Le collectif d'enseignants français « *Sauver les lettres* » fait état d'une recherche menée en 1996 qui a consisté à représenter à 6000 élèves le certificat de fin d'études des années 1920, et qui a montré que « les élèves d'aujourd'hui font deux fois et demie plus de fautes et qu'ils maîtrisent beaucoup moins bien la grammaire qu'il y a 75 ans ». ⁵⁸ Rappelons que ces deux cohortes ne sont pas du tout comparables. Séry (2002, 76) évoque l'expérience de Manesse et Chervel qui ont soumis à 3000

⁵⁵ «A sonder quelques quotidiens des dernières années, on récolte une brassée de titres apocalyptiques tels que «Le français s'appauvrit», «Comment que ça s'écrit ? », «Faut-il encore enseigner la langue maternelle ? », «Parlez-vous bédéen ? », etc. Quant aux grandes lignes du diagnostic, elles sont les mêmes partout : « pauvreté et imprécision générales du vocabulaire,... confusion dans l'expression, tant à l'oral qu'à l'écrit » (Klinkenberg, 2001, 98). (NB : L'auteur, J-M. Klinkenberg, tenait une opinion « relativiste », a une position plus nuancée que ce qu'il relève dans les journaux).

⁵⁶ «Les performances exigées des élèves à un âge donné ne cessent de se modifier». (Beudelot et Establet, 1989, 22-23).

⁵⁷ «il existe peu de séries homogènes sur longue ou même courte période» (Beudelot et Establet, 1989, 19). La comparabilité même des cohortes est un problème redoutable et il bien plus «sûr» de comparer des personnes à elles-mêmes, par des épreuves de «progrès» ou de «progression». «... le niveau mythique de jadis ne concernait en fait qu'une moindre partie de la population» (Romainville, 2000, 82-83).

⁵⁸ *Sauver les lettres*, Des professeurs accusent, Paris : Editions Textuel, 2001, 30.

écoliers la dictée que l'inspecteur général Beuvrain avait administrée à 6000 écoles primaires entre 1873 et 1877, recherche d'où il ressort que le niveau a monté. Plane (1994, 21), commentant les mêmes données, constatait que « *les enfants de notre époque font en moyenne moins de fautes que leurs ancêtres, et surtout que les fautes qui prouvent une mauvaise compréhension de la langue sont moins nombreuses.* » Pourquoi dès lors ce sentiment persistant chez une partie du corps professoral que le niveau baisse ? Cet auteur l'explique par plusieurs facteurs : « *le décalage culturel, le fait d'avoir chaque année à faire face à une nouvelle génération d'élèves avec laquelle « tout est à refaire », le fait que les exigences culturelles augmentent et que les attentes vis-à-vis de l'enseignement sont de plus en plus fortes, la diversification des objets culturels, l'écart important qu'ils estiment déceler entre les compétences qu'ils maîtrisaient dans leur jeunesse et celles de leurs élèves (alors que le niveau que leurs condisciples atteignaient à l'époque n'était pas nécessairement identique au leur)* » (Id. p. 19).

Enfin, troisième source de difficulté : **les données sont trop imprécises**. Même quand on dispose de données, il arrive assez souvent qu'elles soient relativement grossières : des pourcentages de réussite par exemple, incapables de distinguer les réponses correctes obtenues par chance et celles qui sont l'indice d'une connaissance assurée. Or les spécialistes des recherches internationales savent que les attitudes quant au « guessing » (répondre à une QCM en état de connaissance partielle plutôt que d'omettre) varient fort d'une culture à une autre. Elles varient aussi probablement dans le temps.

Malgré les difficultés énoncées ci-devant, les résultats sur les degrés de maîtrise des divers contenus des check-up peuvent-ils être intéressants à titre de « référence pour le futur » ?

Diversement ! En effet, certains contenus se modifieront probablement peu d'ici une vingtaine d'années (ex : les fonctions linéaires en mathématique) alors que d'autres perdront leur pertinence en tout ou en partie (ex : le nom de l'actuel président de la Commission de l'Union européenne ou encore la conversion FB-Euros qui, pour les étudiants qui auront 20 ans en 2020 sera de l'histoire... ancienne).

La **Surface de Compétence** (SurfC), obtenu (e) par multiplication de la **Facilité Objective** (FO) et de la **Facilité Subjective** des Réponses **Correctes** (FSC ou Indice de Confiance moyenne), se justifie pour ceux qui considèrent qu'il ne suffit pas d'obtenir un taux élevé de réponses correctes (%RC). Encore faut-il que cela soit avec un Degré de Certitude (DC) moyen élevé. Pour les 10 check-up MOHICAN, ces Surfaces de compétence moyennes vont de 29,4% (la plus faible) à 50,4% (la plus élevée). Les valeurs moyennes prises par ces indices donnent une idée du « chemin qu'il reste à parcourir » ou plutôt « de la surface qu'il reste encore à couvrir ».

La **Surface d'Incompétence** (SurfI) est calculée selon le même principe : la multiplication de la Difficulté Objective (DO ou 1-FO) par la Facilité Subjective des Réponses Incorrectes (FSI ou Indice de Prudence moyenne), à nouveau sur les proportions et probabilités multipliées par 100. Cet indice se justifie pour ceux qui pensent que l'ignorance reconnue est moins grave qu'une erreur tenue pour vraie avec une certitude élevée, de telles connaissances constituant des dangers. Comme disait Mark Twain, « *Ce n'est pas ce que nous ignorons qui nous cause des problèmes, c'est ce que nous savons... mais qui est faux.* »

Pour les 10 check-up MOHICAN, ces Surfaces d'Incompétence vont de 13,1% (la plus faible, donc le meilleur cas puisqu'il s'agit d'incompétence !) à 24,5% (la plus élevée, donc la pire). A nouveau, ces indices permettent d'imaginer les progrès à faire, soit au niveau objectif (diminuer le nombre de réponses incorrectes) soit au niveau subjectif (diminuer l'ampleur des degrés de certitude accompagnant les réponses incorrectes).

La **Surface de Maîtrise** (SurfM) est obtenue par SurfC-SurfI. On pourrait donc trouver des SurfM négatives. Cet indice se justifie pour ceux qui pensent qu'une mesure du niveau de qualité d'une performance à un check-up doit tenir compte de SurfC et de SurfI, et que cette combinaison (SurfM) est la mesure de qualité la plus valide sur le plan conceptuel (*construct validity*). On a vu (au chapitre 7) que cet indice a aussi une validité prédictive. Aucun des check-up pris dans leur ensemble n'a de SurfM négative : elles vont de 5,4% à 35,2%. Mais des SurfM négatives s'observent, hélas, pour des questions particulières et des étudiants particuliers à certains check-up.

Ces indices sont nouveaux, et ne permettent donc pas de répondre rétrospectivement à la question « le niveau d'entrée a-t-il baissé ? », mais ils pourraient le permettre ultérieurement. Ils ont surtout été développés pour faire prendre conscience du caractère peu assuré des réponses correctes, donc de la fragilité des connaissances, et du caractère souvent imprudent des réponses correctes, donc de la vulnérabilité des connaissances. Les valeurs faibles des Surfaces de maîtrise devraient fonctionner comme des signaux d'alarme !

Or SurfM et **MS (Moyenne Spectrale)**, la simple moyenne des scores spectraux) ont la même valeur. On constate au chapitre 7, section F, que ce score (MS) est **plus prédictif de la réussite** que le simple Taux d'Exactitude pour les facultés de Philo et Lettres, de Droit, d'Economie, de Médecine Vétérinaire, et interviennent comme **deuxième variable prédictive** en Philo et lettres, en Psycho, en Economie (où elle est aussi la troisième variable la plus prédictive), en Sciences, en Sciences Appliquées et en Médecine Vétérinaire.

Plus que les valeurs absolues des moyennes aux check-up (moyennes qui dépendent des questions), ce sont d'une part le caractère de signal d'alarme des valeurs de MS, de SurfC, etc. et d'autre part le caractère prédictif des indices mêlant la performance objective et le regard métacognitif sur la gestion de cette performance, qui méritent que l'on accorde plus d'attention au traitement subjectif qu'opèrent les étudiants sur leur propre connaissance.

2.2 Attentes légitimes de compétences

Les **résultats métacognitifs** du sondage MOHICAN nous paraissent offrir un « ancrage », une référence stable pour le futur.

Certaines matières font légitimement l'objet d'une **attente sociale de compétence** chez une personne (dans sa profession, dans la conduite d'une voiture, dans la connaissance des événements importants pour ses proches, etc.). Les autres (patrons, gendarmes, parents, etc.) sont en droit de s'attendre à une connaissance assurée (indice de **confiance** élevé) de sa part sur certains contenus. Le degré de certitude avec lequel on avance des réponses fausses, lui, par contre, devrait TOUJOURS être le plus faible possible : quand on se trompe, quel que soit le contenu, on devrait toujours le faire, en moyenne, avec **prudence**.

D'autres matières ne font pas l'objet d'attentes légitimes de compétence ; dès lors, le degré avec lequel on est sûr de ses réponses correctes (indice de confiance) dépend du degré de familiarité avec la matière.

Par contre, la plupart des matières enseignées font légitimement l'objet chez les enseignants d'**attentes scolaires de compétence** chez les étudiants ayant reçu une formation. On y vise alors un maximum de confiance dans les réponses correctes aux questions d'évaluation. Cet indice (de même que d'autres indices métacognitifs liés aux réponses correctes) est une manière de rendre plus subtile la mesure des connaissances. Les check-up MOHICAN donnent un éclairage à ce sujet (voir chapitres 9 à 20).

Etant donné ces divers types d'attentes vis-à-vis des connaissances, MOHICAN apporte-t-il des éléments d'information ou de réflexion ?

Nous avons pu constater, grâce à MOHICAN, des statistiques de réussite interpellantes telles que « 91% des étudiants interrogés connaissent la date (1789) de la révolution française ». Cette statistique est interpellante parce que la moitié des enseignants que nous avons interrogés informellement se réjouissent de ce score qu'ils jugent « élevé » et l'autre moitié sont déçus de ce que 9% des étudiants de 18 ans de notre pays l'ignorent.

De même que le parlement de la Communauté française a été amené à préciser les objectifs de l'éducation (dans le Décret-Missions de 1997), peut-être sera-t-on amené à préciser que certaines matières devraient être connues avec un certain **degré de certitude minimal**. Puisqu'il y a **divergences de vues sur les exigences** parmi les enseignants eux-mêmes, il ne faut pas s'étonner de divergences de résultats chez les étudiants.

Ces exigences reflèteront souvent des options culturelles. Pour rester dans le même exemple (la date de la révolution française), nous avons été frappé de ce que bon nombre de Nord Américains, eux, l'ignorent carrément, sans en faire le moindre complexe. La tentation est grande de taxer ces Américains d'incultes comme nous, les Européens, le faisons si souvent quand nous les « mesurons » à l'aune de NOTRE culture. Or il faut avouer que bien peu d'Européens seraient capables de dire ce qui marqua l'histoire des Etats-Unis en cette même date de 1789. Eux le savent : l'élection de leur premier président, Georges Washington. Ce n'est qu'un exemple. On pourrait multiplier les réflexions à partir des résultats de MOHICAN à bien d'autres questions.

Certains résultats des check-up MOHICAN peuvent nourrir la réflexion d'une part sur les minima de compétences à atteindre et d'autre part sur un des objectifs de l'enseignement fondamental : la capacité de la personne à contribuer au développement d'une société démocratique... et l'ouverture aux autres cultures.

2.3 La vigilance cognitive

La façon de poser les questions est elle aussi susceptible d'apporter un éclairage intéressant sur les habiletés intellectuelles. Les étudiants de 18 ans sont-ils suffisamment habitués à fournir des réponses là où on ne leur en suggère pas ? Ou à considérer que plusieurs réponses peuvent être correctes à la fois, résistant ainsi au « curriculum caché » de l'école qui habitue trop souvent à penser qu'il y a toujours une (et d'ailleurs une seule) réponse à toute question ? Ou à détecter les questions contenant des absurdités ou manquant de données suffisantes pour pouvoir répondre ? Pour tester ces deux dernières capacités, il eut suffi que nous ajoutions à chaque QCM des check-up MOHICAN les solutions générales « **Manque de données** pour pouvoir trancher entre plusieurs réponses qui s'excluent » et « **Absurdité** dans l'énoncé ». Nous nous en sommes bien gardés, estimant que les étudiants n'y étaient (hélas) pas entraînés. Nous nous sommes contentés de deux solutions générales moins novatrices : « Aucune » et « Toutes ».

Bien nous en prit, car, on a pu le constater (chapitre 3), surtout dans les check-up de Vocabulaire, de Syntaxe, de Compréhension de texte et de Connaissances artistiques, les questions dont la réponse correcte est le choix de la solution « **Aucune** » ou « **Toutes** » sont en moyenne nettement moins bien réussies que les autres questions.

Ce constat peut avoir deux explications qui se cumulent. Une première explication est que les questions qui ont pour solution correcte «Aucune» ne permettent pas de fournir la réponse correcte par «reconnaissance» de la bonne solution, et les questions dont la réponse correcte est «Toutes» exigent que l'étudiant inspecte méticuleusement toutes les solutions. Une deuxième explication est que les étudiants ont été peu ou pas entraînés à ce type de question, pourtant justifié (voir chapitre 1).

Devant ces attentes légitimes d'une capacité des étudiants à faire preuve de vigilance cognitive, qu'apportent les résultats MOHICAN dans la connaissance du niveau d'entrée des étudiants (du moins ceux de l'échantillon de cette cohorte) ?

Les étudiants sont très peu performants dans les questions qui demandent de déjouer des pièges élémentaires (Aucune, Toutes). Nous faisons l'hypothèse qu'ils ont été peu entraînés systématiquement à exercer leur vigilance cognitive.

2.4 L'auto-évaluation des compétences est-elle stable d'un check-up à l'autre ?

La passation des check-up MOHICAN avec Degrés de Certitude (DC) ne semble avoir posé de problème à aucun endroit. Ce mode de réponse ne semble donc pas incongru aux étudiants qui le rencontrent pour la première fois. Pourquoi ? probablement parce qu'ils sont bien conscients de **l'inévitable présence du doute**, que la consigne leur demande ou non de l'exprimer. Ce ne sont pas les Degrés de Certitude qui créent le doute. Ils permettent, au contraire, de nuancer des réponses que d'habitude on les force à présenter comme si elles ne différaient en rien entre elles quant à la certitude, bref **d'échapper à un manichéisme imposé**.

Il existe plusieurs façons de mesurer la capacité d'un étudiant à estimer sa compétence avec réalisme lors d'une épreuve.

Les indices de **prudence**, de **confiance** et de **discriminance** (voir chapitres 3 et 10) en sont une. Une autre consiste à comparer deux nombres : d'une part, la Certitude Moyenne ou Facilité Subjective (FS) moyenne (pour toutes les réponses données au test) et d'autre part l'Exactitude Moyenne ou Facilité Objective (FO) ou encore le Pourcentage de Réponses Correctes (%RC). Cette différence est appelée «Erreur Moyenne de Centration» (EMC). Si elle est positive, il y a surestimation de soi et si elle est négative, il y a sous-estimation de soi. Attention ! Une EMC très proche de 0 (la perfection) peut être le résultat de compensation des sous-estimations par les surestimations !! Pour les dix check-up, **l'Erreur Moyenne de Centration (EMC) globale pour un test va de 1% à 17%, avec la moyenne des moyennes égale à 7%, toujours dans le sens de la surestimation de sa capacité (ou sous-estimation de la difficulté)**. On pourrait s'estimer heureux d'une si faible erreur moyenne, mais pour nombre d'étudiants existent de fortes surestimations et de fortes sous-estimations (dépassant les 20% en plus ou en moins).

L'indice de **réalisme par Calibration**, lui, ne permet pas de compensation de surestimations par des sous-estimations. Il se calcule par la formule **Réalisme = 100-EMAC⁵⁹**. Dans MOHICAN, **l'erreur moyenne absolue (EMAC) des étudiants vaut 24%** en valeur absolue (donc le **Réalisme moyen est 76**), mais c'est tantôt par surestimation, tantôt par sous-estimation, les deux se compensant en partie pour donner une EMC moins ample. Ces valeurs sont des repères quant aux progrès que l'on pourrait faire puisque l'on a montré ⁶⁰ que **cette capacité d'auto-estimation peut être améliorée avec l'entraînement**.

⁵⁹ EMAC est l'Erreur Moyenne Absolue de Certitude, plus ample que EMC.

⁶⁰ Leclercq, 1993, 129

Les observations faites dans le cadre de MOHICAN permettent-elles d'avancer dans la question de savoir si la Calibration, ou l'Erreur Moyenne de Centration, ou la Confiance, ou la Prudence ou la Discriminance constituent des caractéristiques individuelles stables à travers les divers tests ou si, au contraire, cette capacité métacognitive varie fort d'un test à l'autre ?

Schraw (1997) avait observé (sur 95 étudiants de première année de psychologie de l'université du Nébraska-Lincoln testés dans 4 tests différents) que les corrélations moyennes intratests entre les taux de réussite et les certitudes moyennes étaient de 0,14 en moyenne alors que les corrélations entre certitudes moyennes intertests étaient de 0,43, ce qui plaide en faveur de l'hypothèse d'une tendance générale chez la personne. Cette étude ne permet cependant pas de dire si ces étudiants se sont surestimés ou sous-estimés. Or la corrélation entre la Facilité Objective (FO) et la Facilité Subjective est forte. Ainsi, pour le check-up de mathématique (2539 étudiants), calculée sur les 22 questions, la corrélation est de 0,76.

Les données MOHICAN montrent (chapitre 7 section G) que les EMC (Erreurs Moyennes de Certitude), complètement indépendantes des Facilités Objectives, par définition, sont **corrélées entre elles d'un check-up à l'autre en moyenne à 0,43⁶¹**.

Une telle observation, plus encore que celle de Schraw, indique une certaine « **stabilité** » de **chaque personne dans sa façon de se surestimer ou sous-estimer à travers les épreuves**, et confirme l'intérêt de travailler sur ce facteur métacognitif.

Conclusion : Nous disposons maintenant de repères métacognitifs sur des matières diverses pouvant servir de référence à des activités visant l'amélioration non seulement de la métacognition, mais aussi de la connaissance car le doute est, en principe, le moteur de la vérification, de l'approfondissement. Comme la majorité des recherches rapportées dans la littérature, ont été observés un manque de prudence⁶² et une prépondérance de la surestimation sur la sous-estimation. Par contre, il semble exister une **tendance personnelle stable** dans l'auto-évaluation à travers les épreuves, ce qui donne des **espoirs d'efficacité de remédiations et d'entraînement**, en faisant l'hypothèse qu'il est plus facile d'avoir un effet sur un mécanisme systématique que sur un phénomène aléatoire.

2.5 La Confiance et la Prudence

A l'heure où nous devons tout autant « gérer notre ignorance » que « mobiliser nos connaissances » (Albertini, 1997), il importe que les étudiants soient le plus conscients possible de ce qu'ils savent et de ce qu'ils ignorent, bref des limites et des forces de leurs connaissances. Nous avons encapsulé ces idées dans les concepts de **Confiance** (réalisée si la facilité Subjective des réponses Correctes est supérieure à 50%) et de **Prudence** (réalisée si la facilité Subjective des réponses Incorrectes est inférieure à 50%). Dans les échanges professionnels ou intimes, quand le sort des uns dépend des autres, la fiabilité des réponses peut être cruciale. Souvent il n'est pas grave d'ignorer à condition d'en être conscient : les ressources documentaires permettront de lever les incertitudes. Par contre, ce que De Finetti a appelé les « prétentions non fondées de compétence » coupent court aux vérifications et approfondissements et peuvent mener à la catastrophe.

⁶¹ ce qui explique 18,5% de la variance inter-individus

⁶² variable selon les check-ups

Devant cette problématique du réalisme, critère manifeste d'employabilité, qu'apportent les données MOHICAN ?

On peut être déçu du (grand) nombre d'**alarmes** déclenchées par les **manques de confiance** (absolue) et par les **imprudences** (absolues) tout au long des chapitres 11 à 20.

L'indice de **Fracture Subjective (FractS)** ou de Discriminance est une façon de rendre compte de l'information supplémentaire sur la maîtrise *d'une question* par une population d'étudiants, apportée par l'expression de leur certitude dans leur réponse. L'indice FractSQ fait intervenir la Confiance ou FSC (Facilité Subjective des Réponses Correctes) et la Prudence ou FSI (Facilité Subjective des Réponses Incorrectes), c'est-à-dire, pour un test (ou pour un étudiant), **l'écart** (la fracture) entre la moyenne des certitudes des réponses correctes et celles des réponses incorrectes. Bien que cet indice soit surtout précieux pour chaque question, nous l'avons calculé pour l'ensemble de chaque check-up, ce qui permet de mettre en évidence certaines différences qui pourraient bien être liées aux matières⁶³ et d'avancer des hypothèses explicatives quant à ces différences. Pour les 10 check-up MOHICAN, cet indice moyen de discriminance va de 11,6% à 34,7% de différence entre la FSC moyenne et la FSI moyenne du check-up. Ces différences sont faibles, et révélatrices à nos yeux d'une insuffisance tantôt de prudence tantôt de confiance tantôt des deux chez les étudiants. Or il s'agit d'un **indice à la fois de réalisme et d'apprentissage**, ce dernier devant consacrer une fracture entre « ceux qui ont appris » et « ceux qui n'ont pas appris » ou « ce qui a été appris » et « ce qui ne l'a pas été ».

Conclusions : Les séries de résultats obtenus donnent une idée des améliorations qui restent à apporter dans la **métacognition** car si le taux de réussite importe peu dans certaines matières. Par contre, quelle que soit la matière, prudence et confiance devraient caractériser les configurations de réponse, car elles constituent des preuves de la capacité d'auto-évaluation de l'étudiant, de sa **maîtrise réaliste** de la matière. Ces indices constituent des signes d'utilisabilité des réponses, dans un monde professionnel exigeant de plus en plus la « qualité totale ». MOHICAN fournit des indices simples et des repères pour permettre aux enseignants d'interpréter les valeurs de ces indices de compétences cognitives et métacognitives. **Trop d'imprudences et de manques de confiance** (en moyenne) sont à déplorer, avec, pour conséquence, des **fractures subjectives ou discriminances faibles** (en moyenne).

2.6 Le manque de confiance et l'imprudence : des habitudes d'apprentissage ?

Les étudiants de la Faculté de Psychologie de l'ULg (qui ont répondu aux check-up d'octobre 1999), confrontés en 2000 à des épreuves certificatives (sanctionnantes) commettaient encore beaucoup d'imprudences lors de la première d'entre elles (janvier). Cependant, ces imprudences « chutent » lors des deux épreuves suivantes (juin et août). Voici les nombres de questions dont les distracteurs-vedettes ont été choisis en moyenne (collective) avec imprudence à ces 3 occasions, dans l'épreuve à Livres Ouverts (LO) et dans l'épreuve à Livres Fermés (LF) :

	LO janvier	LO juin	LO août	LF janvier	LF juin	LF août
Nombre d'imprudences dans le choix du distracteur vedette	19	10	7	13	1	3
Nombre de questions de l'épreuve	24	20	20	30	30	18 ⁶⁴

⁶³ Ce qu'il faudrait vérifier par d'autres constats

⁶⁴ Plusieurs questions ayant eu un taux de réussite de 100%, le calcul de la certitude moyenne (FSI) d'un distracteur est impossible. Les calculs n'ont été fait que sur 18 questions (au lieu de 24).

Nous faisons les **hypothèses**⁶⁵ suivantes :

La tendance à l'imprudence, donc fournir des réponses incorrectes avec une certitude élevée chez certains étudiants **est une habitude...**

- **ancienne** (qui remonte à de nombreuses années)
- **peu consciente**: quels sont les étudiants qui ont reçu un feedback à ce sujet ?
- **jugée peu grave** par les étudiants: les résultats (ci-dessus) au premier test sanctionnant semblent montrer que jusque là cette tendance n'avait pas été corrigée («on va toujours essayer, on verra bien»⁶⁶), modifiée uniquement sous la sanction des résultats «qui comptent» (principe du renforcement opérant)
- **en œuvre lors de l'apprentissage** et pas seulement quand ces étudiants répondent à des épreuves, mais aussi et surtout quand ils s'y préparent, bref quand ils étudient. Nous nous appuyons dans ce raisonnement sur les résultats mentionnés au chapitre 10 sur l'estimation de la capacité d'étudiants à comprendre des mots (fournir des synonymes) quand on ne doit pas faire la preuve. Il faut donc craindre une surestimation de cette compétence.
- **plus manifeste dans les épreuves de compréhension** que de mémoire (les étudiants réussissant mieux dans ces dernières et s'y surestimant moins)

Pour toutes ces raisons, nous recommandons

- de dissocier les épreuves de mémoire des épreuves de compréhension afin de garantir le caractère diagnostic des feedbacks (comme ce fut le cas pour RESSAC),
- de confronter de façon précoce les étudiants à des épreuves sanctionnantes,
- de rédiger les feedbacks en termes métacognitifs d'imprudence, de manque de confiance, de discriminance, etc.

3. La synthèse des résultats check-up par check-up⁶⁷

Vocabulaire. On peut être étonné que des termes tels que *éludé, inféré, stigmatisé, allégation, inhérent, présomption, anticiper* n'atteignent pas 50% de réussite dans l'échantillon testé. Le véritable enjeu n'est pas d'enseigner ces mots un à un afin qu'ils soient tous compris. Ce qui importe, c'est que les étudiants soient **conscients de leur degré de compréhension (ou d'incompréhension) de ces mots**, comme ils devront l'être pour bien d'autres mots qu'ils rencontreront ultérieurement, et en tirent les conséquences, notamment en **consultant le dictionnaire**. Or certaines recherches montrent que c'est bien à propos de ce dont on doute que l'on consulte⁶⁸, et d'autres permettent de penser que quand on lit sans devoir faire la preuve de sa compréhension, on a tendance à surestimer sa capacité de compréhension⁶⁹. Heureusement, un entraînement à cet aspect de la métacognition que constitue l'auto-évaluation de la qualité des réponses est nécessaire, possible et efficace⁷⁰.

⁶⁵ Le fait que ce sont des hypothèses devrait nous faire écrire «probablement» à chacune des phrases qui suit.

⁶⁶ Nous avons déjà eu l'occasion (en section A6 ci-avant) d'évoquer ce genre de raisonnement signalé par Noel et Romainville.

⁶⁷ On se référera au chapitre 4 pour une analyse détaillée des indices par check-up et au site du CIUF pour une analyse des indices par question.

⁶⁸ Leclercq et Boskin, 1990

⁶⁹ Leclercq et al., 2002

⁷⁰ Leclercq, 1993, 129.

Les très nombreuses situations d'erreurs avec imprudence justifieraient un entraînement rendant cet enseignement particulièrement justifié, même si c'est pour une minorité (importante) d'étudiants.

Syntaxe. Il est rassurant de constater que le tiers des questions est réussi à plus de 80%, qu'une moitié des questions l'est à plus de 50% et que deux seulement ont un taux de réussite inférieur à 50%. Contrairement à chacun des mots de vocabulaire testés dans le check-up, le diagnostic de ces deux incompétences en syntaxe devrait donner lieu à une remédiation systématique, car **leur portée est générale**. Les erreurs sont commises avec imprudence pour TOUTES les questions, ce qui rend le problème grave, même si c'est pour une minorité (importante) d'étudiants.

Compréhension de textes. Bien qu'il y ait trop peu de questions pour conclure fermement, ces données (taux de réussite faible) vont dans le même sens que les résultats du survey PISA (Lafontaine, 2001). Entraîner à la **compréhension en profondeur** n'est pas une mission réservée aux niveaux d'enseignement primaire et secondaire. Elle incombe aussi au supérieur qui doit en faire une de ses priorités et devrait s'imposer de mesurer périodiquement les gains qu'il a entraînés chez ses étudiants.

Que les erreurs aient été en plus généralement marquées d'imprudence montre l'importance des enjeux de formation à la « compréhension en profondeur ».

Compréhension de graphiques, tableaux et cartes en Géographie. Les deux questions relatives à la **lecture de graphiques** sont mal réussies (respectivement 42% et 62% de réussite) et avec imprudence pour les erreurs commises dans la première. C'est une des tâches évidentes du supérieur d'améliorer cette capacité et d'en suivre systématiquement l'évolution.

Mathématique. Les questions de « **traduction vers les mots** » (que l'on parte de graphiques, de nombres, de formules ou de mots eux-mêmes) ont des taux de réussite inférieurs à la moyenne des questions. Il y aurait là à approfondir non seulement les causes de cette difficulté à dire dans une langue naturelle ce que l'on manipule avec les symboles mathématiques, mais aussi les conséquences de cette difficulté : incompréhension et démotivation grandissantes à mesure que l'on avance « dans la matière » (effet boule de neige). Que pour TOUTES les questions, les erreurs soient généralement marquées par l'imprudence nous paraît alarmant.

Physique et chimie. Ces deux check-up se distinguent par la très grande dispersion des résultats entre les questions : (de 24% à 91% de réussite pour la physique ; de 30% à 90% pour la chimie). Nous y voyons l'effet des contenus, certains étant plus facilement maîtrisables ou plus enseignés que d'autres. Que les erreurs (surtout en physique) soient fournies avec imprudence est inquiétant.

Biologie. On observe ici encore le même phénomène de grande dispersion des résultats (de 21% à 97%), avec la même explication. Notons qu'une matière comme le SIDA, qui est « vitale » et qui a été traitée pédagogiquement de diverses façons a un taux de réussite très élevé (97%). Que les erreurs soient en grande majorité fournies de façon imprudente est inquiétant.

Histoire. La question sur l'année de la révolution française (1789) présente un taux de réussite élevé (91,5%). Doit-on s'en réjouir ou regretter que 8,5% des étudiants de 18 ans entrant dans les universités ignorent cette date ? En **Economie**, la valeur d'un euro en francs belges a fait l'objet de cette période (fin 1999) d'une telle insistance tant à l'école que dans les médias qu'il n'est pas étonnant que la question ait un taux de réussite (93%) très supérieur à tous les autres. En **Actualité**, certains éléments sont mal connus, mais l'importance (le caractère utile) de cette connaissance est relative aux événements mondiaux de l'époque et doit donc être relativisée... subjectivement !

Connaissances artistiques. A nouveau, on constate une très grande dispersion (de 6% à 94%). Le cinéma d'aujourd'hui (les films Philadelphia et Rain Man, le réalisateur Spielberg) et la musique moderne (une chanson du groupe Queen) sont les mieux connus (entre 85% et 93%). Le théâtre du XX^e siècle (Brecht, Renard) est le moins bien connu (15% et 6%). A propos de Brecht, on peut y voir la nécessité de continuer à informer les jeunes générations sur le nazisme et ses opposants. Beaucoup moins d'imprudences sont ici à déplorer que dans d'autres check-up sans doute parce que « quand on ne sait pas, on le sait (qu'on ne sait pas) ».

Conclusion : Certains contenus des check-up MOHICAN 1999 sont déjà obsolètes en 2002. D'autres auront une plus grande pérennité. L'intérêt des résultats MOHICAN à des fins de comparaison diffère donc d'une question à l'autre. Leur intérêt à des fins de guidance des étudiants dépend lui aussi fortement d'un check-up à l'autre, les épreuves ayant trait aux disciplines scientifiques et celles faisant moins intervenir de connaissances factuelles permettant sans doute de dégager les meilleurs indicateurs pour des remédiations nécessaires à la réussite en première candidature.

4. Liens entre variables biographiques (du passé individuel) et Réussite

De très nombreuses études ont montré combien la réussite est liée à des caractéristiques relatives au passé de l'étudiant. Ainsi, Beguin (1991)⁷¹ montre, dans une étude faite à l'UCL, qu'un taux de réussite peut varier en moyenne de 20% selon que l'on est à l'un ou l'autre des deux extrêmes des catégories d'origine socio-économique⁷², en moyenne de 30% selon que l'on entre « à l'heure » ou avec un an ou plus de retard, de 20% selon les études secondaires préalables, etc. Depuis, les données concernant l'origine socio-économique (profession des parents) ou socio-culturelle (diplôme le plus élevé des parents) n'avaient plus été introduites systématiquement dans la banque de données du Conseil des Recteurs Francophone (CreF). Elles l'ont été dans MOHICAN (les analyses détaillées figurent au chapitre 7).

Qu'en est-il dans MOHICAN ?

Il ressort que

1. Le diplôme du père est lié au choix des études d'abord, puis aux taux de réussite ensuite avec une ampleur de variation de 20% entre la catégorie des diplômés (du père) les plus élevés et la catégorie des diplômés (du père) les plus faibles.
2. Le niveau de diplôme des parents (voir chapitre 7, section A) est lié aux résultats à l'entrée. En effet, les performances moyennes (tant en Taux d'Exactitude qu'en Moyenne Spectrale) sont inférieures, pour le groupe 1 à celles des groupes 2 et 3, et ce pour les 10 check-up.
3. Les étudiants qui n'ont passé que 5 ans dans le primaire (ils constituent 6% de la cohorte testée)
 - réussissent à 58% (au lieu de 45% pour 6 ans et 23% pour 7 ans).
 - se répartissent à 13% en sciences de la nature et 87% en sciences humaines.
 - se caractérisent par un taux de réussite en 1^{re} session (R1) et particulièrement élevé (38% au lieu de 27% pour les « 6 ans de primaire ») et leur taux d'abandon particulièrement bas (8% au lieu de 19%).

⁷¹ cité dans Leclercq et al. 1998, 38-40

⁷² Par rapport aux étudiants ayant une origine socio-économique modeste, les étudiants dont l'origine socio-économique est « moyenne » ont un taux de réussite de 10% plus élevé, et les étudiants d'origine socio-économique élevée de 20% plus élevé.

4. Les étudiants qui ont passé 6 ans (et pas plus) dans le secondaire ont un taux de réussite (50%) nettement supérieur à ceux qui en ont passé 7 (29%), 8 (22%) et 9 (24%), et légèrement plus bas que ceux qui en ont passé 5 (54%). Ils ont surtout un taux d'abandons plus faible (17%) que pour 7 ans (25%), 8 ans (33%), 9 ans (23%), alors qu'avec 5 ans il n'est que de 10%.
5. Les 3650 étudiants issus du Général de Transition réussissent beaucoup mieux (48%) et abandonnent beaucoup moins (17%) que les 230 issus du Technique de Transition (16% de réussite et 41% d'abandons).
6. Les 73 étudiants réorientés après une année passée dans le Supérieur Court réussissent légèrement moins bien (41%) que les 3048 qui n'y sont pas passés (46%). Par contre, les (40) « passerelles », qui ont passé 2 ou 3 ans dans le Supérieur Court réussissent mieux (75%).
7. Les 132 étudiants qui ont passé un an à l'étranger réussissent mieux (62%) que les autres qui ne l'ont pas fait (45%), mais il ne faudrait pas y voir une causalité simple (due à ce seul prédicteur).

Conclusion : Nous disposons de repères (dont plusieurs peuvent être comparés avec ceux produits par Droesbeke et al., 2001) qui permettent de tenir compte des liens entre certaines variables déterminées avant l'entrée à l'université et la réussite en 1^{re} année. Ces repères montrent clairement la légitimité et la nécessité de mesures d'aide aux étudiants en difficulté dans les institutions qui ne se contentent pas de la « reproduction sociale » mise en évidence par Bourdieu (1970).

5. Utiliser les données MOHICAN pour améliorer la réussite via des variables changeables

Il est légitime de viser à réduire le « *gaspillage, tant social qu'individuel que représentent les échecs importants en première année* » (Romainville, 1998, 32). Or la difficulté de prédiction INDIVIDUELLE fait que la sélection à l'entrée « *n'est guère docimologiquement praticable* » (idem) car « *parler de prédiction dans le domaine de l'éducation apparaît sans doute comme largement optimiste... l'objet de recherche -le sujet humain- est aussi acteur, agent de son propre devenir. C'est un sujet intentionnel qui peut, par exemple, décider d'investir la réussite scolaire alors qu'il l'avait peu fait auparavant* » (idem, p. 33).

La recherche belge PREDIC (Boxus, 1971, citée dans Debry et al., 1998, 60-61) a montré que ces propos concernant les étudiants pouvaient s'appliquer aussi aux professeurs. On sait en effet que Rosenthal et Jacobson (1971) ont fort marqué les esprits par leur livre « *Pygmalion à l'école* », si bien que s'est instauré une sorte de tabou de prédiction. C'est en opposition totale à ce tabou et avant son émergence que Boxus a communiqué aux enseignants, en 1970, les prédictions⁷³ quant aux performances en lecture d'enfants entrant en première primaire, avec la volonté partagée par les enseignants de **faire mentir la prédiction**, ce qu'ils ont réussi à faire d'une façon très importante !

Dans un contexte d'opinions partagées quant aux fonctions et efficacités des prédictions, pourrait-on utiliser les données MOHICAN pour prédire la réussite... et faire mentir la prédiction ? Sous quelle forme, dans quelles conditions et avec quelle ampleur ?

⁷³ établies par des études longitudinales dans le prolongement de celles d'Inizan, Decroly et VanWayenbergh

L'opération RESSAC, menée en première candidature de la faculté de Psychologie et des Sciences de l'Éducation de l'Université de Liège, montre de la même façon que Boxus (1971) que faire mentir les prédictions est **faisable et justifié**, mais à *condition d'intervenir, comme le font des médecins* qui, eux aussi, par leurs soins, font mentir la prédictivité -pourtant fort élevée si on n'intervient pas- entre la maladie et la mort. La différence entre les opérations PREDIC et RESSAC est que dans la dernière, étant donné l'âge des sujets (18 ans au lieu de 6 ans), ce sont les étudiants eux-mêmes et non les enseignants qui ont été les principaux artisans de ces démentis de prédiction.

L'expérience RESSAC⁷⁴ a consisté à fournir juste avant la période de « bloqué »⁷⁵ (début mai 2000), des informations présentées dans un tableau appelé « radiographie Z » combinant des données MOHICAN d'octobre 1999 et celles des interrogations partielles de janvier 2000 dans ces 4 cours. Par les valeurs absolues (% de réussite) et relatives (position Z par rapport à la moyenne des résultats du groupe⁷⁶) et par la représentation graphique des notes Z, on a essayé de mettre en évidence les forces et les faiblesses dans deux performances contrastées : celles faisant appel à la connaissance de **mémoire** et celle faisant appel à **l'utilisation de connaissances ou compréhension** en profondeur. Ces deux rubriques correspondent à des « méthodes d'étude » sur lesquelles l'étudiant a les pleins pouvoirs, à propos desquelles il devrait s'attribuer la responsabilité personnelle⁷⁷ et qu'il devrait considérer comme **modifiables**. Le suivi de ces étudiants (réussites / échecs en juin et septembre) et un questionnaire d'avis fin juin 2000 (après tous les examens de première session mais avant la délibération) a permis de mesurer, sur 4 cours, l'impact de cette information.

Pourquoi le choix de **notes Z, un indice normatif** faisant référence à la moyenne du groupe ? Parce que les étudiants disposent AUSSI des notes brutes (% de réussite) et que nous n'avons pas demandé à chaque enseignant quel niveau ils jugeaient « suffisant » (il est probable qu'ils aient répondu 60% ou « 12/20 », le seuil habituel). La position dans le groupe est une information supplémentaire qui, en outre, nous a paru susceptible d'interpeller les étudiants aux résultats faibles. Enfin, cette unité de mesure était commune pour les résultats aux interrogations partielles de janvier 2000 et aux check-up MOHICAN d'octobre 1999. Bien entendu, ce type d'indice comporte aussi ses effets pervers, comme ceux d'induire des sous-estimations⁷⁸ ou surestimations⁷⁹ non fondées. La présence des données « brutes » (% de réussite) est de nature à minimiser des effets pervers.

Un deuxième aspect important de RESSAC est que l'information est fournie pour les quatre cours séparément, de telle sorte que l'étudiant puisse constater s'il existe **un phénomène systématique**, qui se produit dans plusieurs cours⁸⁰.

Parmi les étudiants ayant échoué aux partiels de janvier (et qui ont reçu toute l'information), à peu près la moitié ont déclaré⁸¹ avoir **modifié leurs stratégies d'étude en conséquence des feedbacks** ainsi fournis (selon les cas, soit plus d'étude en profondeur donc centrée sur la compréhension, soit plus de mémorisation, soit les deux).

⁷⁴ Leclercq et al., 2002.

⁷⁵ Période durant laquelle les cours sont suspendus pour que les étudiants puissent étudier.

⁷⁶ on se rappellera qu'une note z indique à quelle distance (négative ou positive, donc en moins ou en plus) de la **moyenne** se situe le score d'un étudiant.

⁷⁷ **attribution interne** en termes de Rotter (1966).

⁷⁸ Chez des étudiants forts inférieurs à la moyenne du groupe (plus fort encore).

⁷⁹ Chez des étudiants faibles, supérieurs à la moyenne du groupe (plus faible encore).

⁸⁰ Souvent l'un ou l'autre résultat dans un cours fait exception dans une tendance générale, mais l'étudiant a souvent une explication à cette exception.

⁸¹ Nous n'avons pas vérifié la correspondance entre la réalité (ont-ils VRAIMENT modifié leur méthode d'étude comme ils l'annoncent) et leurs déclarations.

Pour cette moitié des étudiants (ceux qui ont modifié leur stratégie d'étude suite aux feedbacks), le **taux de réussite est nettement plus élevé** (en moyenne 43% de réussite au lieu de 28%, soit une différence de 15%) que pour ceux qui n'ont tiré des feedbacks « radiographie Z » aucune conséquence en termes de stratégie d'étude. Et **ce phénomène a été observé dans chacun des 4 cours considérés**, avec plus ou moins d'ampleur selon les cours.

Conclusion : Dans les feedbacks aux étudiants, il importe d'atteindre un niveau diagnostique et de crédibilité qui provoque une prise de conscience et une motivation suffisante pour que, quand c'est justifié par les résultats, les étudiants modifient leur METHODE d'étude (notamment étudier en profondeur alors qu'ils ne le faisaient pas ou mémoriser plus que ce qu'ils faisaient) ? Cependant cette condition nécessaire n'est pas suffisante. Encore faut-il que l'étudiant (e) interprète ces informations, les fasse siennes (les « admette ») et les mette en pratique. Les services de Guidance pour l'Etude pourraient contribuer à favoriser ce processus.

Etant donné la méthode de recueil des avis (par questionnaire), nous ne pouvons pas exclure les biais que constituent la désirabilité sociale (des changements sont déclarés et non réels) ou réduction de la dissonance cognitive⁸² (moins consciente) et la rationalisation a posteriori (les changements sont réels mais ne sont pas dûs aux feedbacks MOHICAN). La cause de ces différences (positives) de réussite serait donc à vérifier par des expériences ultérieures.

6. Prédiction et sélection : mises en garde

Périodiquement ressurgit le souhait de procéder à des prédictions en vue de sélectionner. Or une telle prédiction est intéressante pour intervenir en vue de la faire mentir si elle est défavorable. Par contre, elle ne peut être utilisée dans une opération de sélection car l'erreur de prédiction est trop grande et entraînerait d'importantes erreurs de type 1 (refuser à l'entrée des étudiants qui auraient réussi).

Quelle est la prédictivité des résultats aux check-up sur la réussite ?

Au chapitre 8, les corrélations simples (entre chaque check-up pris comme prédicteur et la réussite/échec considérée comme une dichotomie 0/1) et les corrélations multiples (combinant plusieurs résultats de check-up différents pour prédire la réussite) mettent en évidence les faits suivants :

1. **La variance (de réussite) « expliquée »** en moyenne par UNE variable isolée est très faible (4%⁸³), de même que par la variable la plus prédictive et même par des combinaisons de plusieurs variables les plus prédictives est **faible : 32%** au maximum.
2. C'est **en Sciences Appliquées que la prédictivité⁸⁴ des scores aux check-up est la plus faible** (0,28). Cela n'a rien d'étonnant : par l'examen d'entrée, il y a garantie que les compétences de base sont présentes. Les différences inter-individus de réussite ou d'échec dépendent d'autres facteurs (tels que la persévérance par exemple que de ces compétences « relativement uniformes chez tous les admis »).
3. C'est **en médecine vétérinaire que la prédictivité est la plus élevée**. Elle atteint au maximum 0,55 pour l'épreuve de mathématique (soit 32% de variance expliquée). Cependant la présence massive d'étudiants Français (40% dans cette section, 3% dans les autres) qui entrent avec un an d'études et de maturité supplémentaire biaise le calcul de la prédictivité (leur taux de réussite moyen en mathématiques par exemple est de 73% alors qu'il est de 59% pour les Belges).

⁸² Phénomène par lequel la personne met en concordance la réalité et ses croyances ou explications (à elle-même ou à autrui), en modifiant ces dernières pour éviter la dissonance.

⁸³ Chap 7 p. 14 point 1.

⁸⁴ Sur les différences inter-individus quant à la réussite.

4. **Les variables les plus prédictives** sont **aussi souvent** des variables prenant en compte l'expression de **degrés de certitude** que n'en tenant pas compte⁸⁵. Il y a donc un intérêt à poursuivre l'exploration de la contribution que la subjectivité des étudiants appliquée à leurs performances (leur auto-évaluation subjective) peut apporter à la compréhension des concepts de compétence, d'apprentissage et de réussite.
5. Pour obtenir des corrélations multiples significatives (donc des **combinaisons de variables indépendantes ou prédictives**), les variables retenues dans les combinaisons sont **différentes selon les facultés**, comme anticipé, avec, cependant, des variables communes à plusieurs combinaisons facultaires, surtout Syntaxe et Mathématique.
6. **La longueur des check-up** (en nombres de questions) **ne semble pas peser dans le caractère prédictif**, sauf quand ce nombre est vraiment trop petit (par exemple le check-up de Compréhension de texte qui ne compte que 6 questions).

Conclusion : Dans les prédictions de la réussite en juin – septembre 2000 par les performances aux épreuves MOHICAN en octobre 1999, les erreurs de type 1⁸⁶ et de type 2⁸⁷ sont importantes car de très nombreux autres facteurs interviennent dans la réussite en plus du niveau de compétence à l'entrée ; en outre, les épreuves n'étaient pas conçues dans ce but prédictif. Enfin, et cet argument crucial ne dépend en rien de la qualité des épreuves, se servir des prédictions pour sélectionner serait accepter l'idée que les dés sont jetés avant même d'avoir commencé. Or les choses peuvent changer, comme le montre l'opération RESSAC.

7. Bibliographie

- Albertini, J.M. (1997). *Innovations pédagogiques et nouvelles technologies*, in Boxus et al, Actes du 15^e colloque AIPU, 27-41
- Beudelot, C. et Establet, R. (1989). *Le niveau monte*. Paris : Seuil.
- Beguïn, A., 1968 : une révolution inutile, Quelques chiffres et quelques réflexions à propos de la démocratisation du recrutement des étudiants à l'université, *Socio*, Périodique d'information et d'échange des professeurs de sciences sociales, Louvain-la-Neuve, 1991, 20-26.
- Bourdieu, P. (1970), *La reproduction*, Paris : Editions de Minuit.
- Boxus, E. (1971). Une méthode de prédiction du rendement en lecture en première année. *Education - Tribune Libre*. 127, 63-78.
- Debry, M., Leclercq, D. & Boxus, E. (1998). De nouveaux défis pour la pédagogie universitaire, in D. Leclercq (Ed.), *Pour une pédagogie universitaire de qualité*, Sprimont : Mardaga, pp. 55-80.
- Defays et al. *La maîtrise du français. Du niveau secondaire au niveau supérieur*. Bruxelles : De Boeck, 82-83.
- Depover, C. (1987). *L'ordinateur media d'enseignement. Un cadre conceptuel*. Bruxelles : De Boeck Université. Collection Pédagogies en développement, problématiques et recherches, 30.

⁸⁵ Si l'on examine, pour chacune des 9 sections la variable la plus prédictive, c'est dans 4 cas sur 9 une variable incorporant les degrés de certitude. Quand on examine les équations de régression multiple «maximales» (incorporant le maximum de variables ayant des caractéristiques prédictives suffisantes en combinaison entre elles), 21 sur 31 sont de ce type.

⁸⁶ elles consistent à refuser l'entrée à un étudiant qui aurait réussi (Leclercq, 1998, 58-59)

⁸⁷ elles consistent à admettre des étudiants qui échoueront

- Droesbeke, J.-J. (Ed.) (2001). *La population étudiante. Descriptions, évolution, perspectives*. Bruxelles : Editions de l'Université de Bruxelles. Collection Statistique et Mathématiques Appliquées.
- Dupont, C. Maîtrise du français en Communauté française : état de la question et pistes de remédiation éventuelles, Rapport introductif, Parlement de la Communauté française, session 2001-2002. Karraker (1967).
- Karraker, R.-J. (1967), Knowledge of results and incorrect recall of plausible of multiple choice alternatives, *Journal of educational Psychology*, 58, 11=14.
- Klinkenberg, J.-M. (1998). *La langue et le citoyen*, PUF.
- Kulhavy, R.W., Anderson, R.C. (1972). Delay-retention effect with multiple-choice-tests. *Journal of Educational Psychology*, vol. 63, n° 5, 505-512.
- Lafontaine, D. et al. (2001). PISA (Programme International pour le Suivi des Acquis des élèves de 15 ans), rapport final sur la campagne 2000 (1^{er} cycle) en Communauté française de Belgique, Université de Liège : Service de Pédagogie expérimentale.
- Leclercq, D. (1998). *Pour une pédagogie universitaire de qualité*. Sprimont : Mardaga.
- Leclercq, D. et Boskin, A. (1990). Note Taking behaviors studied with the help of hypermedia, in Estes, Heene et Leclercq (Eds). *Proceedings of the 7th International Conference on Technology and Education*, Brussels.
- Leclercq, D., Simon, F., Marotte, P., Lacaille, C. (2002). Former des étudiants de première candidature universitaire à des compétences transversales : lesquelles et comment ?, 2^e Congrès des chercheurs en éducation, Louvain-la-Neuve, Mars.
- Leclercq, D., Boxus, E., de Brogniez, Ph., Wuidar, H., Lambert, F., The TASTE Approach : General Implicit Solutions in MCQs, *Confidence Marking, Open Books Exams and Interactive Testing*, in *Leclercq and Bruno*, 1993.
- Leclercq, D. et Van der Vleuten, C. (1998). PBL – Problem Based Learning ou APP – Apprentissage Par Problèmes, in D. Leclercq (Ed.), *Pour une pédagogie universitaire de qualité*. Liège : Margada.
- Lindley, D. (1971). *Making decisions*. Londres : Wiley.
- Monseur, C. (1998). L'enseignement des sciences en Communauté française de Belgique est-il dans le 36^e dessous ? Résultats de la 3^e étude internationale en mathématique et en sciences de l'IEA. Liège : ULg – Service de Pédagogie expérimentale.
- Noël, B. & Romainville, M. (1998), Accompagner les étudiants, in M. Frenay *et al.*, *L'étudiant-apprenant. Grilles de lecture pour l'enseignant universitaire*, Bruxelles : De Boeck & Larcier, chap. 8, pp. 129-147.
- Plan, S. (1994). *Ecrire au collège*. Paris : Nathan.
- Preston, R.C. (1965), Multiple choice test as an instrument of perpetuating false concepts, *Educational and Psychological measurement*, 25, 111=116.
- Romainville, M. (1992). Savoir parler de ses manières d'apprendre. Contribution à l'étude de la relation entre la métacognition et la performance chez l'étudiant universitaire de première année, thèse de doctorat en Psychopédagogie. Louvain-la-Neuve, Université Catholique de Louvain.
- Romainville, M. (1998). L'enseignement universitaire a les étudiants qu'il mérite, in M. Frenay *et al.*, *L'étudiant-apprenant. Grilles de lecture pour l'enseignant universitaire*, Bruxelles : De Boeck & Larcier, chap. 6, 95-107.
- Romainville, M. (2000). Et si on arrêta de tirer sur le pianiste ? in

- Rotter, J. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80 (1, whole, 609).
- Sassenrath, J.-M., Yonge, G.D. (1969). Effects of delayed information feedback cues in learning on delayed retention. *Journal of Educational Psychology*, 60 (3), 174-177.
- Schraw, G. (1997). The effects of generalized metacognitive knowledge on test performance and confidence judgments, in *Journal of Experimental Education*, 65 (2), 135-146.
- Sturges, P.T. (1969). Verbal retention as a function of the informativeness and delay of informative feedback. *Journal of Educational Psychology*, 60, 174-177.

Chapitre 22

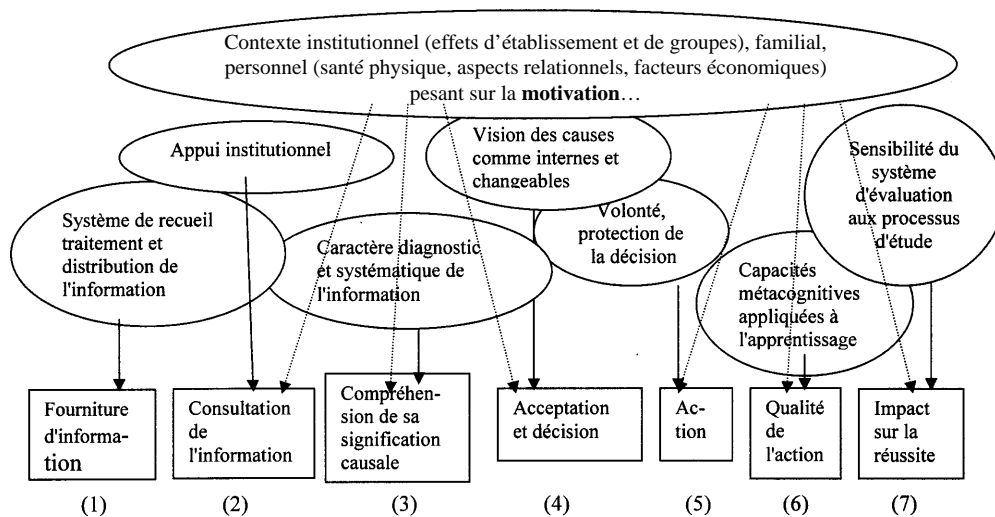
Conclusions générales et perspectives sur les rapports entre MOHICAN et la réussite en première année

D. Leclercq, Président du Groupe de Travail CIUF « Réussite en candidatures »

Le caractère individuel et **métacognitif** de l'approche MOHICAN vise la modification des comportements des étudiants⁸⁸ sans changer les méthodes d'enseignement. Ce n'est qu'une des stratégies possibles pour optimiser l'apprentissage dans ses composantes cognitives et affectives. **D'autres stratégies** consistent soit à modifier le curriculum⁸⁹, soit à modifier la méthodologie,⁹⁰ soit à combiner ces approches. Notons que ces autres stratégies **peuvent se combiner avec l'approche informative de MOHICAN**. C'est d'ailleurs ce qui se fait en faculté de médecine à l'université de Maastricht où les Tests de Progression qui ont une forte composante « RESSAC » par le feedback très détaillé et personnalisé qui est fourni chaque trimestre sur tous les contenus⁹¹ s'intègre dans l'Apprentissage Par Problèmes qui donne à l'étudiant, intégré dans des groupes tutoriels d'une durée de 6 semaines, la responsabilité de son apprentissage... sans enseignement. La phrase suivante de Konrad Lorenz⁹² résume *grosso modo* les étapes du processus visant à l'impact d'informations sur les résultats :

- Said but not heard (1)
- Heard but not understood (2)
- Understood but not accepted (3)
- Accepted but not put into practice (4)
- Put into practice but for how long ? (5)

Cette phrase est traduite dans le schéma chronologique suivant où les « bulles » représentent certains des facteurs affectant chacune des étapes. Par rapport à la phrase de Lorenz, une sixième et une septième étape sont introduites : (6) **la qualité de l'action** (de la nouvelle stratégie d'étude : est-elle pratiquée efficacement ?) et (7) **l'impact sur la réussite** (est-ce suffisant, étant donné l'existence d'autres facteurs ?).



En termes de Lorenz, ces deux ajouts pourraient être traduits par (6) « *Put into practice, but how effectively ?* » et (7) « *Put effectively into practice, but with which effect ?* ». Un des facteurs qui pèsent sur cette dernière étape est la sensibilité de l'évaluation aux processus d'étude. A quoi sert par exemple d'étudier « intelligemment » si l'examen ne favorise que l'étude par cœur ?

⁸⁸ d'où les expressions « Régulation de l'Etude Personnelle par l'Information » (REPI) ou « Résultats d'Epreuves Standardisées au Service des Apprentissages en Candidature » (RESSAC)

⁸⁹ par exemple en instaurant la semestrialisation avec examens en proximité plus grande avec les cours

⁹⁰ par exemple en instaurant, la Pédagogie du projet (PDP) ou l'Apprentissage Par Problèmes (APP) ou l'Enseignement à Distance (EAD), etc.

⁹¹ Leclercq et Van der Vleuten, 1998,200-201

⁹² Les numéros entre parenthèses sont de nous.

Ce schéma énonce les conditions nécessaires (mais pas toujours suffisantes) à un impact positif sur les chances de réussite. L'absence de tout feedback est le niveau 0 de la phrase de Lorenz ; c'est en quelque sorte du « *not said* ». Un feedback non consulté est du « *said but not heard* ». Dans le sondage du chapitre 5, ce fut le cas pour 20% des étudiants interrogés sur les feedbacks MOHICAN ; pour les 80% qui disent avoir consulté, une partie seulement semble n'avoir pas eu de difficulté d'interprétation (*understood*). Dans l'opération RESSAC décrite dans le chapitre 8, seulement la moitié de ceux qui auraient dû le faire ont dit avoir été convaincus (*accepted*) et avoir changé leur comportement (*put into practice*). **Pour ceux-là, le taux de réussite a été nettement supérieur** au taux de réussite des autres. Il ne s'agit là que d'une observation, susceptible de nombreux biais et qui mériterait d'être recoupée par des expériences vérificatrices. Moyennant ces réserves, la portée de ce dernier constat reste considérable. Il rappelle qu'en éducation, comme en médecine, l'absence de mesure « corrective » adaptée à chaque individu laisse s'exercer sans contre force avec une prédictivité accablante le poids de variables prédictives connues, dans un cas la maladie et dans l'autre cas des déterminants tels que le parcours antérieur, les prérequis, le statut socio-culturel de la famille.

Par contre, les interventions, ici informatives, à condition d'être suffisamment diagnostiques pour être déclenchantes d'une auto-médication, donnent de sérieux espoirs d'une amélioration additionnelle⁹³ des chances de « franchir le cap des candis »⁹⁴.

C'est sur ce message d'espoir et de mobilisation (car beaucoup reste à faire) que se termine le présent ouvrage qui espère avoir montré, comme tout bon Mohican, sur quelles pistes et avec quels pères il est raisonnable et responsable de s'engager.

Sans une information diagnostique, personnalisée, précoce sur sa maîtrise de la matière, l'étudiant n'a pas les éléments nécessaires pour procéder à une régulation efficace quand elle est indispensable. Or nous savons depuis des décennies que **ce besoin concerne la majorité des étudiants de première année** (une part importante des 60% qui échouent). On peut faire l'hypothèse que si on leur fournit cette information d'une certaine façon, **une part significative en tirera profit**. Notre conviction est que ce bénéfice pourrait concerner **plus d'étudiants encore via des mesures** complémentaires, telles qu'un **accompagnement de guidance plus serré** sur base d'un dossier sur les performances issu **d'opérations MOHICAN plus ciblées, plus diagnostiques et plus centrées sur les processus d'apprentissage et d'études**.

Il est logique et sain que l'université mette dans les mains des étudiants la responsabilité de l'apprentissage autonome, à charge pour elle (l'université) de préciser les objectifs à atteindre, les contenus à maîtriser, les modalités d'évaluation à satisfaire, le niveau à atteindre. La façon la plus claire de préciser le tout est de permettre à l'étudiant de se confronter aux épreuves d'évaluation elles-mêmes.

Au moment où ils acceptent leur part du contrat qui implique leur prise en charge autonome de leur apprentissage, **les étudiants devraient considérer comme un droit la possibilité de disposer des informations précoces, personnalisées et diagnostiques sur leur maîtrise de la matière, condition déterminante dans la possibilité pour le grand nombre de réaliser leur partie du contrat**.

⁹³ Car ce type d'intervention est susceptible de combiner ses effets à ceux d'autres interventions, notamment celles où les encadrants ont l'initiative.

⁹⁴ Leclercq et al., 1997