

# Utilisation des degrés de certitude et normes de réalisme en situation d'examen et d'auto-estimation à la Faculté de Psychologie et des Sciences de l'Education de l'Université de Liège

Jean-Luc GILLES

Centre d'AutoFormation et d'Evaluation Interactives Multimédias  
Faculté de Psychologie et des Sciences de l'Education de l'Université de Liège - Belgique  
Tél. : +32 - 4 - 366.20.78 Fax : +32 - 4 - 366.29.53 Internet : jl.gilles@ulg.ac.be

## 1. Contexte de la Faculté de Psychologie et des Sciences de l'Education (FA.P.S.E.)

Depuis l'année académique 1994-1995, le Centre d'AutoFormation et d'Evaluation Interactives Multimédias de la Faculté de Psychologie et des Sciences de l'Education (C.A.F.E.I.M.-FA.P.S.E.<sup>1</sup>), sous l'impulsion de son Doyen, le Professeur DE KEYSER, aide les enseignants, en priorité du 1er cycle (là où les étudiants sont les plus nombreux), dans la réalisation de leurs examens. Il met entre autres à leur disposition un dispositif soutenu par des techniques d'évaluation informatisée. Deux types d'outils d'évaluation sont proposés. Le premier est interactif et repose sur l'utilisation d'un logiciel intitulé WINCHECK développé au Service de Technologie de l'Education (S.T.E.). Ce logiciel permet de poser des Questions à Choix Multiple avec Solutions Générales Implicites<sup>2</sup> (QCM-SGI) via l'écran d'un ordinateur tout en gérant l'utilisation du "Double Check"<sup>3</sup> ainsi que l'utilisation des degrés de certitude selon la consigne et les principes mis au point par LECLERCQ (1975, 1983, 1986). Le second dispositif recourt à la lecture optique de marques : les étudiants répondent à des QCM-SGI en utilisant des feuilles spéciales sur lesquelles ils cochent leurs réponses en les accompagnant systématiquement d'un degré de certitude. Ces feuilles sont lues par un lecteur optique de marques qui génère un fichier. Après transcodage, le traitement des réponses est effectué à l'aide du logiciel CERT réalisé en 1991 par le S.T.E. avec l'aide de la Commission des Communautés européennes dans le cadre du programme EUROTECNET (BOXUS & al., 1991).

## 2. Les enjeux liés à l'auto-estimation de ses compétences

### a) Nuancer les remédiations en fonction des états de connaissance partielle

On considère habituellement une réponse fournie à une question de façon binaire : soit elle est correcte, soit elle est incorrecte, sans se préoccuper de nuances liées à la conviction avec laquelle l'étudiant a répondu. Des pionniers tels que DE FINETTI (1965), VAN NAERSEN (1965), SHUFORD & al. (1966), LECLERCQ, (1975) suivis par de nombreux chercheurs (BRUNO, 1993; DIRKZWAGER, 1993; FABRE, 1993; GILLES, 1995, HUNT, 1993; JANS, 1995; PLUNUS, 1996; VAN LENTHE, 1993) recommandent de traiter différemment l'ignorance reconnue (réponse incorrecte et peu sûre), la connaissance incomplète (réponse correcte mais peu sûre), la connaissance assurée (réponse correcte et très sûre), l'ignorance ignorée (réponse incorrecte et très sûre). En effet, les différents états de connaissance partielle qui découlent de l'association d'une réponse et d'un degré de certitude autorisent un diagnostic plus subtil et par là différents niveaux de remédiation. Par exemple, pour un même objectif, l'énergie et la forme de la

<sup>1</sup> Le C.A.F.E.I.M.-FA.P.S.E. (Président : Prof. D. LECLERCQ, directeur : J.-L. GILLES) s'est fixé pour objectifs :

- de promouvoir l'apprentissage autonome et coopératif recourant aux multimédias;
- de faciliter les évaluations de grands groupes d'étudiants en recourant aux technologies appropriées (QCM, lecture optique, etc);
- de permettre les animations et/ou les évaluations interactives individuelles ou collectives (théâtre électronique);
- de permettre et de développer une activité de recherche scientifique en matière d'application des technologies de l'information à l'apprentissage, l'enseignement et l'évaluation;
- de permettre des activités de développement (conception) et de production (réalisation de produits de formation).

<sup>2</sup> Les Solutions Générales Implicites (LECLERCQ, 1986) autorisent, en plus des solutions habituellement proposées, les quatre possibilités suivantes : Rejet (aucune solution proposée n'est correcte), Toutes (toutes sont correctes), Manque (il manque des données dans l'énoncé pour que l'on puisse choisir UNE solution comme correcte), Absurdité (il y a une contre-vérité dans l'énoncé à dénoncer en priorité!).

<sup>3</sup> Cette procédure d'évaluation interactive consiste à poser une question en deux volets *prim* et *bis*. L'étudiant reçoit une première (*prim*) question (QCM-SGI) où la réponse correcte attendue peut, par exemple, être « 8. Manque de données dans l'énoncé ». Après avoir répondu, l'étudiant reçoit la réponse puis la deuxième partie de la question (*bis*), par exemple : « quelle donnée manque ? ». Suivent à nouveau une série de propositions. Les performances des étudiants se présentent alors selon différents cas de figure qui peuvent ensuite donner lieu à des procédures de remédiation adaptées selon le diagnostic.

remédiation consacrée à la correction d'une ignorance ignorée devrait être différente de celle qui permettra d'améliorer un état de connaissance incomplète.

#### b) Contre efficacement la divination au pur hasard pour les QCM

Il est inadmissible que la formule d'attribution des points à une QCM contribue à encourager la réponse au hasard. C'est malheureusement le cas lorsque sont utilisés des barèmes de correction qui n'intègrent pas les degrés de certitude. Même la correction pour divination<sup>4</sup> classique (*correction for guessing*) opérée pour atténuer l'effet des réponses correctes apportées par hasard aux QCM est sujette à caution. Citons sur ce point DELANDSHEERE (1979) : "*La correction pour divination fait l'objet de nombreuses critiques. Elle repose notamment sur l'hypothèse gratuite que tous les sujets ont également deviné. De plus, on établit pas la distinction entre l'élimination de certains choix sur base de connaissances réelles et la divination au pur hasard. Une correction beaucoup plus adéquate est assurée quand le sujet indique dans quelle mesure il est certain de sa réponse*". Pour garantir une utilisation sans biais des degrés de certitude, une série de règles méthodologiques doivent être respectées. Ces règles sont exposées succinctement ci-après (3. Conditions d'utilisation des degrés de certitude).

#### c) Mesurer sa certitude pour réduire l'incertitude

Il n'y a pas que lors des épreuves scolaires que nous sommes amené à nous poser des questions. Constamment des choix se présentent à nous, nous amenant à peser le pour et le contre des décisions qui font notre quotidien. Les enjeux sociaux liés à l'auto-évaluation du degré de certitude sont réels, la formation d'individus avisés est souhaitable quelque soit la profession et passe par l'entraînement systématique à l'auto-estimation de ses compétences, à la mesure de sa certitude. Cette capacité constitue une compétence métacognitive d'ordre stratégique qui soulève le problème de "l'après diagnostic", de la prise de décision en situation d'incertitude. Comment réagir face au doute ? D'une part, en réduisant l'incertitude par la prise d'informations (voir à ce sujet les travaux de JANS et LECLERCQ, 1996), et, d'autre part, en reconnaissant son incompétence lorsque la situation l'impose (on ne s'improvise pas médecin, pilote de ligne, ...).

*Il ne sait pas et ignore qu'il ne sait pas, c'est un niais, évite-le.*

*Il ne sait pas et sait qu'il ne sait pas, c'est un enfant, enseigne-le.*

*Il sait mais ignore qu'il sait, c'est un endormi, éveille-le.*

*Il sait et sait qu'il sait, c'est un sage, suis-le.*

*proverbe arabe<sup>5</sup>*

### **3. Conditions d'utilisation des degrés de certitude pour l'auto-estimation des compétences**

Une série de conditions méthodologiques doivent être rencontrées pour garantir un recueil sans biais des données liées à l'auto-estimation de ses compétences. Le non respect des règles ci-après explique l'échec de nombreuses études menées aux Etats-Unis entre 1960 et 1975. Le lecteur trouvera des informations plus complètes à ce propos dans LECLERCQ (1993, pp. 141-143) et BOXUS & al. (1991, pp. 56-70).

#### a) Le barème de tarifs doit être conforme à la théorie des décisions

Il s'agit de gratifier une réponse correcte accompagnée d'un degré de certitude élevé d'un meilleur score que si elle était accompagnée d'une certitude faible et inversement pour les réponses incorrectes. Les tarifs du barème des points doivent être calculés de manière à favoriser une seule stratégie : celle qui consiste à dire la vérité. Le barème des points ci-après garantit que l'expression de son intime conviction rapporte plus de points que tout autre stratégie.

<sup>4</sup> Il existe plusieurs formules de correction pour divination. Celle qui est habituellement utilisée (appelée « classique ») consiste à retirer  $1/(k-1)$  points en cas de réponse incorrecte,  $k$  étant le nombre de solutions proposées dans la QCM.

<sup>5</sup> Cité par E. H. SHUFORD dans D. LECLERCQ & J. BRUNO, *Item Banking : Interactive Testing and Self-Assessment*, p. 85, NATO ASI Serie F : Computer and Systems Sciences, Vol. 112, Springer Verlag, Heidelberg, 1993.

b) La consigne doit être "probabiliste".

Demander à l'étudiant d'indiquer sa certitude par des termes vagues du type "peu sûr", "moyennement sûr", "très sûr", etc. est à proscrire car ces expressions recouvrent des réalités différentes en fonction des sujets. De plus, avec des termes aussi flous la variabilité est telle chez un même étudiant qu'on ne peut même pas recourir à des traitements ordinaux intra-sujets. Voici la consigne mise au point par LECLERCQ (1983, 1993, 1995) et qui est utilisée à la FA.P.S.E.

| Si vous considérez que votre réponse a une probabilité d'être correcte comprise entre | Ecrivez | Vous obtiendrez les points suivants en cas de |                         |
|---|---------|---|-------------------------|
|   |         | réponse correcte (RC)                         | réponse incorrecte (RI) |
| 0 % et 25 %   | 0       | + 13  | + 4                     |
| 25 % et 50 %  | 1       | + 16  | + 3                     |
| 50 % et 70 %  | 2       | + 17  | + 2                     |
| 70 % et 85 %  | 3       | + 18  | + 0                     |
| 85 % et 95 %  | 4       | + 19  | - 6                     |
| 95 % et 100 %   | 5       | + 20  | - 20                    |

c) Des zones correspondant à la précision humaine

Les coupures sur l'axe ne sont pas équidistantes ce qui permet une expression du degré de certitude plus nuancée à l'extrémité supérieure de l'échelle. Ainsi, l'étudiant peut faire la distinction entre 90 % (valeur centrale de la certitude 4) et 97,5 % (valeur centrale de la certitude 5) bien que la différence soit de 7,5 % seulement. Dans le premier cas (90 %) il n'a qu'1 chance sur 10 (1/10) de se tromper tandis que dans le second (97,5 %) il n'a qu'1 chance sur 40 (1/40), soit 4 fois moins. Etablir la même différence au milieu de l'échelle, par exemple entre 40 % (1/1,7) et 47,5 % (1/1,9), n'est pas pertinent car nous sommes pas capable de distinguer ces deux derniers « rapports »...

d) Le calcul d'indices métacognitifs doit être possible

La consigne utilisée autorise le calcul d'un indice de réalisme basé sur les différences entre les taux d'exactitude et les valeurs centrales des intervalles de probabilité ainsi que le calcul d'un indice de centration basé sur la différence entre la certitude moyenne et le taux d'exactitude moyen et dont le signe détermine la surestimation (+) ou la sous-estimation (-).

Il existe différentes variantes pour la formule de réalisme (LICHTENSTEIN & al., 1975; LECLERCQ, 1975, 1983; LECLERCQ & al., 1993). Dans le cadre de cette étude nous avons adapté la formule de LECLERCQ afin que le minimum de réalisme soit égal à 0 et le maximum égal à 1:

$$\text{REALISME} = \left( \left( 1 - \sqrt{\text{MEM}} \right) - \beta \right) \alpha$$

avec : 
$$\text{MEM} = \frac{\sum_i (\text{TE}_i - \text{VC}_i)^2 \text{NU}_i}{\text{NR}}$$

$\text{TE}_i$  = Taux d'exactitude de la certitude  $i$ ,  $\text{VC}_i$  = Valeur centrale,  $\text{NU}_i$  = Nombre d'utilisations,  $\text{NR}$  = Nombre total de réponses et  $i$  = indice (ici de 0 à 5) des degrés de certitude.

$\beta$  = correction pour l'erreur minimale, dans notre cas  $\beta = 0,025$ .

$\alpha$  = correction d'étendue, dans notre cas, pour étendre la plage<sup>6</sup> de 0,95 à 1,  $\alpha = 1,0526$

L'expression du doute à l'aide de degrés de certitude permet aussi de mettre à jour différents états de connaissance partielle (voir le point 2.a) ci-avant). LECLERCQ & PLUNUS (1996) ont présenté sur un

<sup>6</sup> Dans la procédure d'utilisation des degrés de certitude définie par LECLERCQ,  $\text{ERR} = 0,975$  pour un sujet qui se trompe systématiquement avec le degré de certitude 5, et,  $\text{ERR} = 0,025$  pour un sujet qui répond correctement à toutes les questions avec un degré de certitude 5, ceci détermine une plage de 0,95 ( $0,975 - 0,025$ ).

vecteur ces différents états répartis en 6 catégories : "compétence parfaite (RC<sup>7</sup> et DC<sup>8</sup> élevé : 4 ou 5) - compétence insuffisante (RC et DC moyen : 2 ou 3) - compétence ignorée (RC et DC faible : 0 ou 1) - ignorance reconnue (RI<sup>9</sup> et DC faible) - ignorance insuffisante (RI et DC moyen) - ignorance ignorée (RI et DC élevé)". La ventilation des réponses des étudiants dans ces différentes catégories est intéressante à observer. Les pourcentages de réponses fournies qui relèvent de ces catégories varient-ils d'un examen à l'autre ou restent-ils stables ? La difficulté des épreuves influence-t-elle la répartition des réponses au sein de ces catégories ? Nous tenterons de répondre à ces questions par une analyse de régressions multiples.

e) Les sujets doivent être préalablement entraînés à l'utilisation des degrés de certitude

Avant d'employer les degrés de certitude tels que décrits ci-avant lors des examens, les étudiants se familiarisent avec leur utilisation, d'une part, à l'aide du logiciel GUESS, et, d'autre part, lors de QUIZZ en fin de cours.

Le programme GUESS<sup>10</sup> a été décrit en détail par LECLERCQ & GILLES (1994). Chaque étudiant s'entraîne individuellement à l'utilisation des degrés de certitude à l'aide d'un jeu où il doit deviner les lettres successives d'un texte d'au moins cent lettres (jeu inspiré des travaux de SHANNON, 1951 & d'ATTNEAVE, 1959). Le joueur effectue une prédiction en tapant une lettre qu'il accompagne de la probabilité subjective de réussite exprimée à l'aide d'un degré de certitude. Il est ensuite informé de la réponse correcte qui s'affiche dans la zone réservée au texte. Lettre par lettre, le texte s'affiche ainsi à l'écran. Evidemment, le début des mots est plus difficile à deviner que leur fin (c'est ce que SHANNON voulait démontrer). Après un nombre donné de réponses un graphique de réalisme se construit dans le coin supérieur droit de l'écran, l'étudiant peut y observer la justesse des probabilités subjectives de réussites qu'il a attribuées à ses prédictions (en abscisse) en les confrontant aux taux de réussite (en ordonnée) pour chaque degré de certitude. Les étudiants du premier cycle sont invités dès le 1er semestre à s'entraîner avec GUESS (c'est une activité obligatoire dans le cadre du partim I *Approche technologique de l'éducation et de la formation* du cours *Introduction aux sciences de l'éducation*).

Les étudiants de la faculté de Psychologie et des Sciences de l'Éducation ont aussi la possibilité de s'entraîner à l'utilisation des degrés de certitude lors de QUIZZ organisés en fin de cours où ils sont invités à répondre à quelques questions (entre 5 et 8). Ils répondent sur des feuilles spéciales fournies par le C.A.F.E.I.M.-FA.P.S.E. sur lesquelles ils cochent les réponses qu'ils accompagnent d'un degré de certitude (comme à l'examen). Après le ramassage des feuilles de réponses, les réponses correctes sont communiquées. Les feuilles sont traitées dans la semaine à l'aide du dispositif de lecture optique de marques. Au cours suivant, le professeur commente les statistiques de résultats de l'ensemble du groupe, commente les erreurs les plus fréquentes et réexplique éventuellement la matière mal comprise. Enfin, chaque étudiant reçoit une feuille de feedback personnalisée avec entre autres informations : les réponses et les certitudes qu'il a fournies, les réponses correctes attendues, son graphique de réalisme, les indices de son réalisme et de sa centration. Cette procédure et les résultats ont été décrits par GILLES & LECLERCQ (1995) et GILLES (1995).

Ces deux types d'entraînement, avec chacun leurs avantages et leurs inconvénients, se complètent bien. Grâce au logiciel GUESS, beaucoup de réponses accompagnées de certitudes sont fournies en peu de temps et la visualisation d'un feedback immédiat est possible. Cependant, l'exercice proposé est assez éloigné de la situation d'examen qui, elle, porte sur des contenus liés à la psychologie, aux sciences de l'éducation et comporte des QCM à 9 solutions maximum. À l'inverse, les QUIZZ ne permettent pas de poser beaucoup de questions et il faut attendre le cours suivant avant d'obtenir un feedback. Par contre, les questions y sont fort proches de celles qui sont posées à l'examen, tant du point de vue de la forme que du contenu.

---

<sup>7</sup> RC = réponse correcte

<sup>8</sup> DC = degré de certitude

<sup>9</sup> RI = réponse incorrecte

<sup>10</sup> Le logiciel GUESS a été programmé dans le langage OpenScript de TOOLBOOK par M. HURARD, ingénieur informaticien.

#### 4. Questions et hypothèses liées aux normes de réalisme en situation d'examen et d'auto-estimation à la FA.P.S.E.

Le dispositif qui vient d'être décrit soulève toute une série de questions auxquelles les données de nombreuses interrogations permettent d'apporter des débuts de réponses. Les données que nous avons recueillies concernent les étudiants de diverses facultés de l'Université de Liège : Droit, Médecine Vétérinaire, Philosophie et Lettres, Psychologie et Sciences de l'Education, Sciences Appliquées. Pour l'ensemble de ces facultés, lors des années académiques 1994-1995 et 1995-1996, nous disposons des données de 69 examens et quizz où les réponses ont été accompagnées de degrés de certitude, selon les modalités décrites ci-avant. Etant donné l'entraînement poussé dont ont bénéficié les étudiants de la faculté de Psychologie et des Sciences de l'Education ainsi que la proportion des évaluations réalisées avec ces sujets, 41 épreuves sur un total de 69, nous avons préféré, dans le cadre de cette étude, nous limiter à l'analyse des épreuves liées à ce groupe d'étudiants.

a) Comparaison examens-quizz : les performances métacognitives lors des examens sont-elles équivalentes avec les performances métacognitives des étudiants lors des quizz ?

Les 41 épreuves récoltées depuis 1994-1995 se répartissent en 28 examens et 13 quizz. La réalité que recouvrent les quizz diffère de celle des examens dans la mesure où les quizz constituent un entraînement sans impact en ce qui concerne les scores récoltés et où peu de questions sont posées. Le réalisme, la centration et la répartition des réponses dans le vecteur des degrés de compétence sont-ils équivalents lors des examens et lors des quizz ?

b) Comparaison du niveau de performance d'un individu ou d'un groupe : comment les individus se répartissent-ils dans les différents niveaux de scores aux indices métacognitifs ?

Les normes qui suivront seront locales à la FA.P.S.E. où elles permettront d'informer les étudiants sur leur situation dans la population de cette faculté ainsi que les enseignants sur la position de leurs évaluations en ce qui concerne les performances métacognitives des étudiants. Quelle est la proportion d'étudiants qui obtiennent les meilleurs scores à l'indice de réalisme et de centration ? Comment se répartissent les sous-estimations et les surestimations ? Nous faisons ici l'hypothèse qu'il y a proportionnellement moins de sujets qui se sous-estiment étant donné la tendance humaine à la surestimation souvent décrite dans la littérature.

c) Comparaison des niveaux de performance en fonction de la difficulté des épreuves : la difficulté des épreuves influence-t-elle les performances métacognitives des étudiants ?

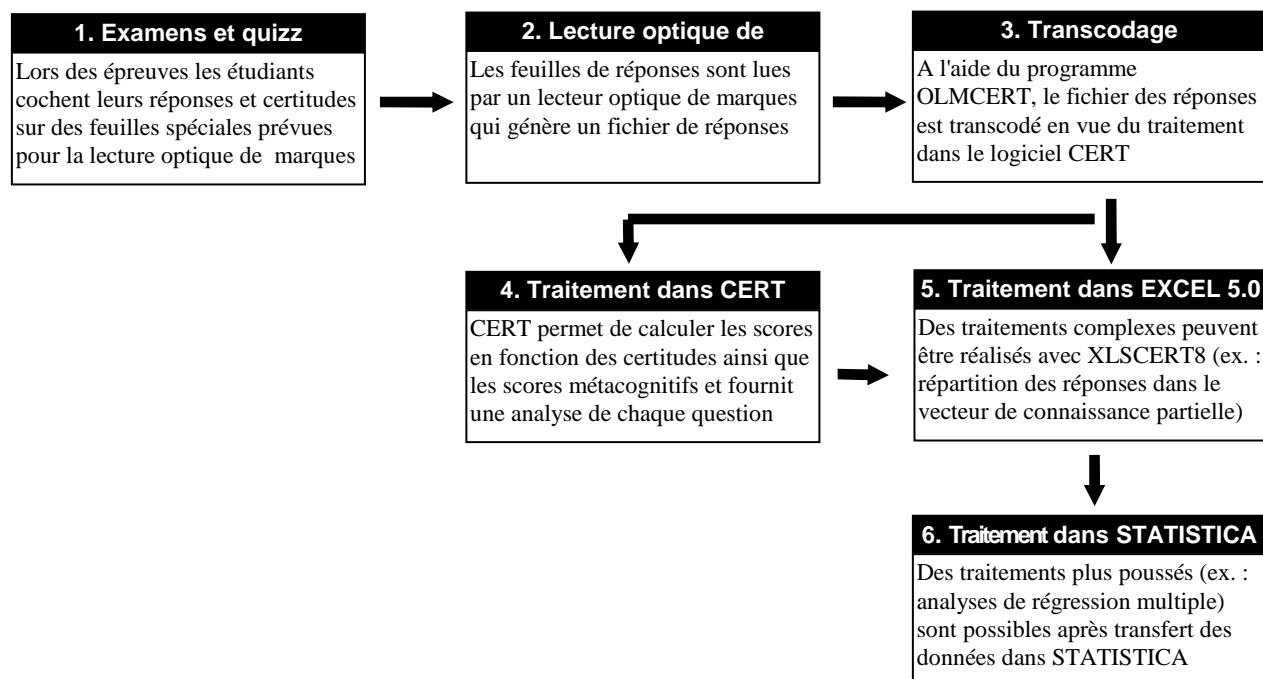
Les données recueillies depuis deux ans devraient nous permettre, à l'aide des taux de réponses correctes et des certitudes moyennes, d'observer l'impact de la difficulté sur les performances métacognitives liées à l'auto-estimation de ses compétences.

d) Comparaison de la progression des performances en fonction des sous-groupes "classes" : lorsqu'on compare les performances métacognitives des 1ère et 2ème candidature, les scores des 2ème candidature sont-ils meilleurs ?

Il s'agit ici d'évaluer la progression des performances des étudiants en fonction de l'année d'étude dans laquelle ils se situent. On peut émettre l'hypothèse que les étudiants de la FA.P.S.E. inscrits en 2ème année obtiendront de meilleurs scores métacognitifs que les étudiants de 1ère année qui, globalement, ont été moins confrontés aux situations d'évaluation avec auto-estimation des compétences, et ce, d'autant plus qu'une sélection s'opère à la fin de la 1ère année.

## 5. Dispositif de recueil et de traitement des données

Dans le cadre de cette étude, ce sont les données des examens et des QUIZZ, récoltées à l'aide du dispositif de lecture optique de marques depuis l'année académique 1994-1995 qui ont été utilisées. Nous avons complété la chaîne de traitement par un ensemble de macros de calcul adaptées au logiciel Microsoft EXCEL 5.0<sup>11</sup>. Ce module récupère les fichiers destinés au logiciel CERT qu'il complète pour la réalisation de traitements complexes ainsi que pour la mise en forme des données. Il autorise également la récupération des données dans le logiciel STATISTICA. Le schéma ci-dessous reprend les différentes étapes du recueil et du traitement des données.



## 6. Résultats des traitements pour les données recueillies au sein de la FA.P.S.E.

### a) Comparaison des niveaux moyens de performance métacognitive aux examens et aux quizz

#### 1. Scores à l'indice de réalisme

Rappelons que le score maximum à l'indice de réalisme vaut 1 et que le score minimal vaut 0.

|                                      | Examens<br>(N = 28 examens) | Quizz<br>(N = 13 quizz) | différences |
|--------------------------------------|-----------------------------|-------------------------|-------------|
| Moyennes pour l'indice de réalisme : |                             |                         |             |
| Moyenne                              | 0.763                       | 0.706                   | 0.057       |
| Ecart-type                           | 0.046                       | 0.028                   | 0.018       |
| Minimum                              | 0.411                       | 0.305                   | 0.106       |
| Maximum                              | 0.989                       | 0.966                   | 0.023       |
| Plage                                | 0.578                       | 0.661                   | 0.083       |
| Coefficient d'asymétrie              | -0.477                      | -0.446                  | 0.031       |

<sup>11</sup>Il s'agit d'une série de procédures de calcul réalisées dans le langage Visual Basic pour EXCEL 5.0 et regroupées dans une macro principale intitulée XLSCERT8 (analyse et programmation : J.-L. GILLES)

Jean-Luc GILLES, Utilisation des degrés de certitude et normes de réalisme en situation d'examen et d'auto-estimation à FA.P.S.E. - ULG, Colloque de l'ADMEE-EUROPE "Dix années de travaux de recherche en évaluation", 18, 19 et 20 septembre 1996.

On constate à la lecture du tableau ci-dessus que le réalisme aux examens est en moyenne meilleur que le réalisme aux quizz. Ceci pourrait s'expliquer par un effort d'auto-estimation accru lorsqu'il y a un enjeu à la clé et par l'effet d'entraînement qui amènerait une meilleure performance à l'examen (postérieur aux quizz). Ces interprétations sont à prendre avec prudence car le nombre de quizz est moins élevé et le nombre de questions posées par quizz l'est aussi. Les étudiants reçoivent de 30 à 50 questions lors des examens, alors que ce nombre varie de 5 à 10 lors des quizz. De plus, tous les étudiants inscrits ne sont pas forcément présents lors des séances de cours où ont lieu les quizz. En ce qui concerne la facilité des épreuves, on observe des taux de réponses correctes très proches : en moyenne 55,2 % pour les 28 examens et 56,6 % pour les 13 quizz.

Le coefficient d'asymétrie moyen indique une courbe de réalisme très décentrée vers la droite (c'est à dire en forme de J) aux quizz et aux examens, la courbe de ces derniers est encore plus décentrée.

## 2. Scores à l'indice de centration

Rappelons que le score idéal à l'indice de centration vaut 0 et que le score de sous-estimation maximale vaut -100, tandis que le score de surestimation maximale vaut 100.

| Moyennes pour l'indice de centration : | Examens<br>(N = 28 examens) | Quizz<br>(N = 13 quizz) | différences |
|--|-----------------------------|-------------------------|-------------|
| Moyenne                                | 2                           | 3                       | 1           |
| Ecart-type                             | 16                          | 20                      | 4           |
| Minimum                                | -43                         | -53                     | 10          |
| Maximum                                | 49                          | 56                      | 7           |
| Plage                                  | 92                          | 109                     | 17          |
| Coefficient d'asymétrie                | 0.005                       | -0.138                  | 0.133       |

La centration est en moyenne légèrement meilleure aux examens car plus proche de 0 et moins prononcée en ce qui concerne les sous-estimations (min. = -43 contre -53) et les surestimations (max. = 49 contre 56).

Les normes présentées ci-après sont basées uniquement sur les données des 28 examens.

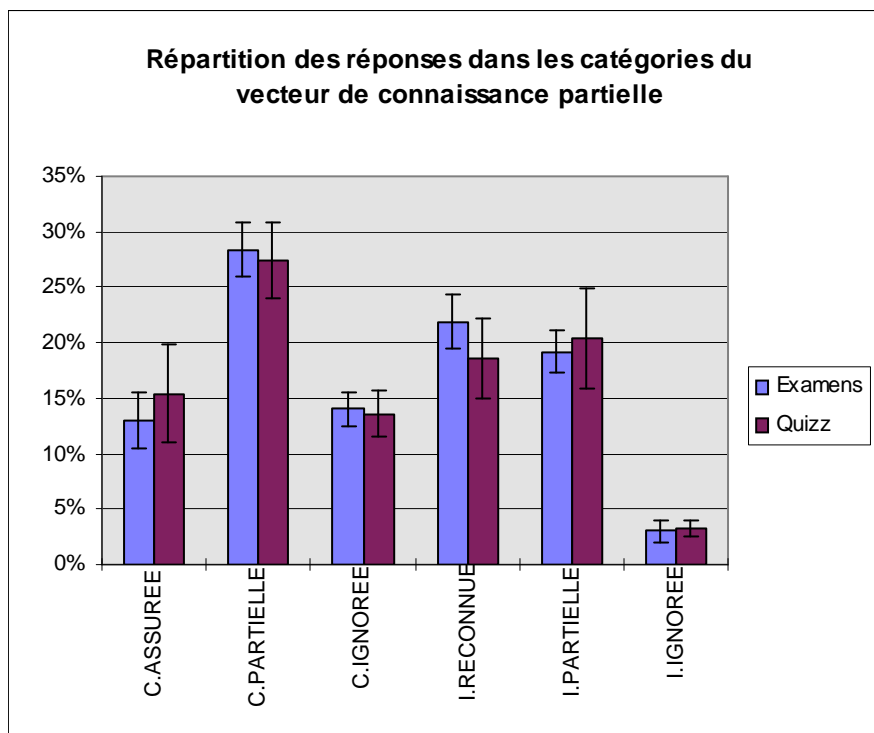
## 3. Etats du vecteur des degrés de compétence

Les réponses fournies par les étudiants peuvent se répartir dans un vecteur des degrés de compétence (LECLERCQ & PLUNUS, op. cit). Nous avons calculé les pourcentages de réponses relevant des catégories du vecteur des degrés de compétence, voici le tableau récapitulatif :

|                                | Réponse CORRECTE avec CERTITUDE |                      |                    | Réponse INCORRECTE avec CERTITUDE |                     |                   |
|--------------------------------|---------------------------------|----------------------|--------------------|-----------------------------------|---------------------|-------------------|
|                                | 4 ou 5                          | 2 ou 3               | 0 ou 1             | 0 ou 1                            | 2 ou 3              | 4 ou 5            |
|                                | Compétence assurée              | Compétence partielle | Compétence ignorée | Ignorance reconnue                | Ignorance partielle | Ignorance ignorée |
| Moy. Examens :<br>(28 examens) | <b>13%</b>                      | <b>28.4%</b>         | <b>14%</b>         | <b>21.9%</b>                      | <b>19.2%</b>        | <b>3%</b>         |
| Moy. Quizz :<br>(13 quizz)     | <b>15.4%</b>                    | <b>27.4%</b>         | <b>13.6%</b>       | <b>18.6%</b>                      | <b>20.4%</b>        | <b>3.2%</b>       |

Ces pourcentages sont forts proches, cependant, on observe une inversion des tendances en ce qui concerne les catégories "ignorance reconnue" et "ignorance partielle". La majorité des réponses incorrectes est en effet accompagnée d'un degré de certitude faible lors des examens et d'un degré intermédiaire lors des quizz. Les étudiants semblent plus prompts à reconnaître leur ignorance en situation d'examen qu'en situation de quizz.

Le graphique ci-après traduit ces données en histogrammes accompagnés des écarts type aux moyennes.



b) Normes pour la comparaison du niveau de performance métacognitive d'un individu ou d'un groupe lors des examens

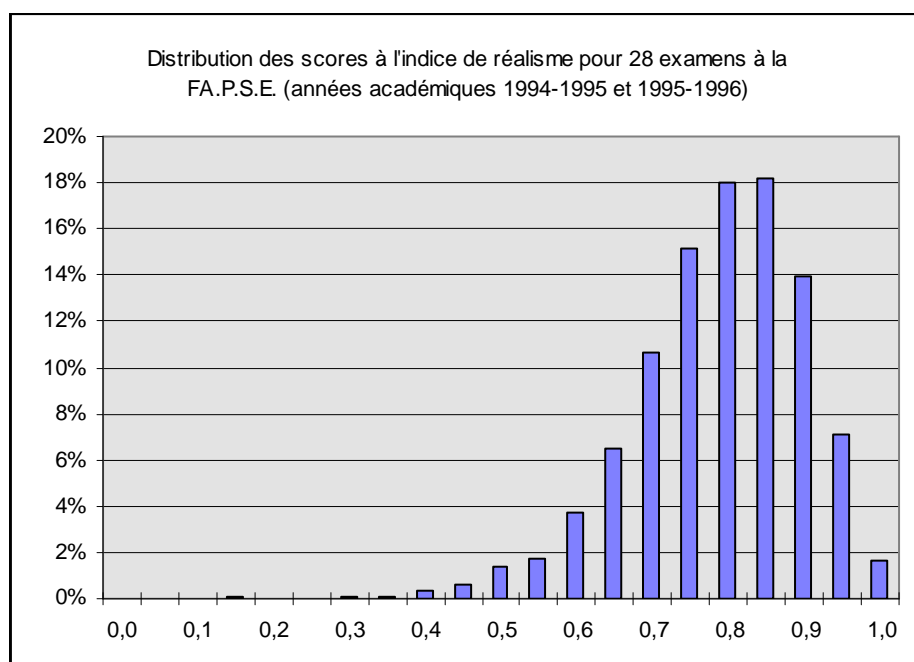
1. Normes pour l'indice de réalisme

Le tableau ci-dessous reprend les fréquences en pourcentages des scores de réalisme pour les 28 examens ayant eu recours aux degrés de certitude à la FA.P.S.E. entre 1994 et 1996..

| R    | %1   | %2   | %3   | %4   | %5   | %6   | %7   | %8   | %9   | %10  | %11  | %12  | %13  | %14  | %15  | %16  | %17  | %18  | %19  | %20  | %21  | %22  | %23  | %24  | %25  | %26  | %27  | %28  | %M   |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 0.00 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 0.05 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 0.10 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| 0.15 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.3  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 2.0  | 0.1  |
| 0.20 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.3  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.5  | 0.0  | 0.0  | 0.0  |
| 0.25 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.3  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.5  | 0.0  | 0.0  | 0.0  |
| 0.30 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.7  | 0.0  | 0.0  | 0.0  | 0.6  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.5  | 0.0  | 0.0  | 0.0  | 0.1  |
| 0.35 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 1.0  | 0.2  | 0.0  | 0.0  | 0.6  | 0.9  | 0.0  | 0.0  | 0.3  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.1  |
| 0.40 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.4  | 2.0  | 0.0  | 0.2  | 0.0  | 1.2  | 2.1  | 0.7  | 0.6  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 3.4  | 0.0  | 0.0  | 0.4  |
| 0.45 | 0.0  | 0.3  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.5  | 0.0  | 0.0  | 3.3  | 0.2  | 0.3  | 0.6  | 3.7  | 5.4  | 1.4  | 0.3  | 0.3  | 0.0  | 0.0  | 0.3  | 0.0  | 0.0  | 0.0  | 0.0  | 0.6  |      |
| 0.50 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.9  | 0.9  | 0.0  | 4.6  | 1.0  | 0.2  | 1.5  | 4.9  | 10.2 | 3.4  | 1.1  | 0.3  | 0.0  | 0.3  | 2.0  | 0.0  | 1.3  | 4.3  | 0.0  | 1.5  | 1.4  |
| 0.55 | 1.1  | 1.0  | 0.0  | 0.0  | 1.1  | 0.0  | 0.0  | 0.5  | 0.9  | 0.3  | 0.0  | 9.3  | 1.5  | 0.7  | 2.3  | 5.7  | 4.8  | 6.6  | 4.5  | 1.0  | 0.4  | 0.6  | 1.3  | 0.0  | 0.0  | 5.3  | 0.8  | 1.5  | 1.8  |
| 0.60 | 2.2  | 3.8  | 1.0  | 0.0  | 2.2  | 3.0  | 4.0  | 2.4  | 1.9  | 2.4  | 1.1  | 10.3 | 2.5  | 3.5  | 1.5  | 7.9  | 14.4 | 9.3  | 7.3  | 3.0  | 2.2  | 2.6  | 2.0  | 3.0  | 0.0  | 7.2  | 0.4  | 2.9  | 3.7  |
| 0.65 | 5.1  | 7.3  | 4.0  | 0.6  | 5.0  | 5.0  | 2.0  | 6.6  | 3.9  | 2.1  | 1.9  | 11.9 | 6.1  | 5.1  | 7.0  | 13.4 | 12.9 | 13.8 | 11.5 | 8.6  | 6.9  | 4.6  | 8.1  | 4.5  | 10.0 | 8.2  | 3.8  | 2.9  | 6.5  |
| 0.70 | 9.7  | 8.4  | 8.0  | 1.9  | 12.8 | 12.0 | 6.0  | 10.9 | 7.8  | 8.6  | 8.4  | 13.9 | 9.1  | 10.8 | 12.3 | 13.4 | 13.8 | 19.3 | 13.7 | 10.9 | 7.8  | 7.2  | 13.1 | 10.0 | 13.8 | 11.1 | 11.1 | 12.7 | 10.7 |
| 0.75 | 18.8 | 12.9 | 14.0 | 10.6 | 17.9 | 17.0 | 17.0 | 19.0 | 11.3 | 13.8 | 11.4 | 14.6 | 16.9 | 17.2 | 14.9 | 14.6 | 13.2 | 13.4 | 17.1 | 18.2 | 16.9 | 13.2 | 16.4 | 12.4 | 17.5 | 15.9 | 15.3 | 13.2 | 15.2 |
| 0.80 | 22.0 | 19.6 | 19.0 | 22.4 | 19.6 | 15.0 | 14.0 | 20.9 | 20.0 | 20.5 | 14.8 | 11.9 | 17.7 | 20.5 | 20.5 | 16.5 | 9.9  | 13.8 | 16.5 | 17.8 | 22.9 | 21.6 | 17.1 | 19.4 | 20.0 | 12.5 | 16.9 | 21.0 | 18.0 |
| 0.85 | 21.7 | 22.0 | 24.0 | 23.6 | 17.9 | 21.0 | 26.0 | 14.7 | 22.4 | 26.9 | 26.2 | 6.0  | 18.0 | 19.7 | 19.9 | 9.1  | 5.4  | 9.7  | 14.6 | 19.8 | 15.6 | 19.0 | 16.1 | 23.9 | 15.0 | 13.5 | 20.3 | 17.1 | 18.2 |
| 0.90 | 14.4 | 15.0 | 19.0 | 24.8 | 11.7 | 17.0 | 18.0 | 13.3 | 17.0 | 14.1 | 23.6 | 5.6  | 16.1 | 13.1 | 11.7 | 5.5  | 3.9  | 4.1  | 8.1  | 11.9 | 16.5 | 17.5 | 12.1 | 16.4 | 16.3 | 9.1  | 19.5 | 14.1 | 13.9 |
| 0.95 | 3.6  | 7.3  | 8.0  | 13.7 | 9.5  | 8.0  | 10.0 | 8.1  | 11.0 | 8.0  | 8.7  | 3.3  | 8.6  | 6.6  | 5.6  | 4.9  | 1.2  | 3.1  | 3.6  | 6.3  | 7.4  | 10.1 | 9.7  | 8.0  | 3.8  | 5.3  | 6.9  | 8.3  | 7.1  |
| 1.00 | 1.1  | 1.7  | 3.0  | 1.9  | 1.7  | 1.0  | 2.0  | 3.3  | 1.9  | 1.8  | 3.0  | 1.0  | 1.3  | 1.5  | 1.5  | 0.0  | 0.3  | 1.0  | 0.8  | 1.3  | 1.3  | 2.3  | 1.0  | 2.0  | 0.0  | 1.9  | 4.2  | 1.0  | 1.6  |



Voici le graphique de la distribution des données de la dernière colonne (%M) reprenant les fréquences moyennes pour l'indice de réalisme lors de ces 28 examens :



On observe une distribution très décentrée (par son mode) vers la droite (coefficient d'asymétrie = -0,477), donc plutôt en J. En nous basant sur les données des 28 examens, nous recommandons pour les épreuves qui ont lieu à la FAPSE, d'interpréter l'indice de réalisme calculé selon la formule décrite ci-avant (voir 3. c) de la façon suivante.

#### L'INDICE DE REALISME EST

|                         |                         |                     |                           |                            |
|-------------------------|-------------------------|---------------------|---------------------------|----------------------------|
| <b>IDEAL</b>            | si il est compris entre | <b>0,96 et 1</b>    | Performance accomplie par | <b>1,6 %</b> des étudiants |
| <b>EXCELLENT</b>        |                         | <b>0,91 et 0,95</b> |                           | <b>7,1 %</b>               |
| <b>TRES BON</b>         |                         | <b>0,86 et 0,90</b> |                           | <b>13,9 %</b>              |
| <b>BON</b>              |                         | <b>0,81 et 0,85</b> |                           | <b>18,2 %</b>              |
| <b>MOYEN</b>            |                         | <b>0,76 et 0,80</b> |                           | <b>18 %</b>                |
| <b>FAIBLE</b>           |                         | <b>0,71 et 0,75</b> |                           | <b>15,2 %</b>              |
| <b>INSUFFISANT</b>      |                         | <b>0,66 et 0,70</b> |                           | <b>10,7 %</b>              |
| <b>TRES INSUFFISANT</b> |                         | <b>0 et 0,65</b>    |                           | <b>14,7 %</b>              |

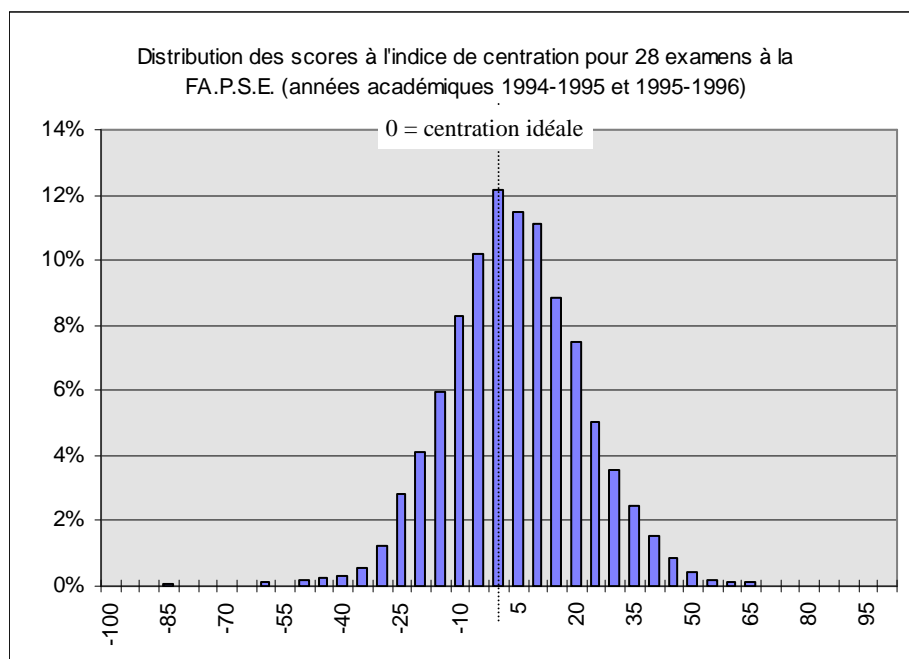
#### 2. Normes pour l'indice de centration

Le tableau ci-après reprend les fréquences en pourcentages des scores de centration pour les 28 examens ayant eu recours aux degrés de certitude à la FA.P.S.E. entre 1994 et 1996.

| C    | %1  | %2  | %3  | %4  | %5  | %6  | %7  | %8  | %9  | %10 | %11 | %12 | %13 | %14 | %15 | %16 | %17 | %18 | %19 | %20 | %21 | %22 | %23 | %24 | %25 | %26 | %27 | %28 | %M  |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| -100 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| -95  | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| -90  | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| -85  | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,1 |
| -80  | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| -75  | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,5 | 0,0 | 0,0 | 0,0 |
| -70  | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,5 | 0,0 | 0,0 | 0,0 |
| -65  | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| -60  | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 5,4 | 0,0 | 0,0 | 0,1 |
| -55  | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,5 | 0,0 | 0,0 | 0,0 |
| -50  | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,2 | 0,3 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 3,4 | 0,0 | 0,5 | 0,2 |
| -45  | 0,4 | 0,0 | 0,2 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,5 | 0,6 | 0,6 | 0,0 | 0,0 | 0,0 | 0,3 | 0,0 | 0,0 | 0,0 | 0,0 | 1,3 | 2,9 | 0,0 | 0,5 | 0,3 |
| -40  | 0,0 | 0,0 | 0,5 | 0,0 | 0,6 | 0,0 | 1,6 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,9 | 2,4 | 0,6 | 0,0 | 0,0 | 0,3 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 1,0 | 0,0 | 1,5 | 0,0 | 0,3 |
| -35  | 0,0 | 0,0 | 0,9 | 0,0 | 0,0 | 0,0 | 0,0 | 0,5 | 0,4 | 0,6 | 0,4 | 0,0 | 0,5 | 1,3 | 1,2 | 1,2 | 0,0 | 0,3 | 0,3 | 0,0 | 0,0 | 1,4 | 0,3 | 0,0 | 0,0 | 3,4 | 0,0 | 2,0 | 0,5 |
| -30  | 1,1 | 0,7 | 2,1 | 0,6 | 1,7 | 1,7 | 2,4 | 0,9 | 1,1 | 0,9 | 0,8 | 0,0 | 1,2 | 1,3 | 1,8 | 2,4 | 1,2 | 1,0 | 0,3 | 0,7 | 0,4 | 0,9 | 0,0 | 0,5 | 2,5 | 1,4 | 1,5 | 2,9 | 1,2 |

|     |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |             |
|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-------------|
| -25 | 2,9  | 0,7  | 4,8  | 0,6  | 2,2  | 1,3  | 6,4  | 1,4  | 1,8  | 3,1  | 4,2  | 0,0  | 2,3  | 3,5  | 3,2  | 4,9  | 1,8  | 0,3  | 0,6  | 1,3  | 1,3  | 2,9  | 2,3  | 3,5  | 5,0  | 3,8  | 4,6  | 8,3  | <b>2,8</b>  |
| -20 | 5,1  | 1,4  | 9,2  | 2,5  | 6,7  | 3,1  | 5,6  | 3,3  | 4,1  | 2,8  | 6,1  | 1,3  | 5,1  | 4,8  | 5,0  | 4,9  | 0,6  | 7,7  | 1,1  | 1,7  | 2,2  | 4,9  | 2,7  | 5,0  | 3,8  | 7,2  | 3,8  | 9,8  | <b>4,1</b>  |
| -15 | 7,6  | 3,5  | 9,5  | 5,6  | 6,1  | 6,6  | 8,8  | 6,6  | 7,2  | 6,1  | 10,6 | 3,0  | 6,0  | 5,8  | 7,3  | 7,9  | 1,2  | 1,4  | 2,5  | 2,3  | 3,9  | 8,3  | 4,0  | 3,5  | 10,0 | 6,3  | 3,8  | 11,7 | <b>6,0</b>  |
| -10 | 9,0  | 7,7  | 12,2 | 9,3  | 9,5  | 4,4  | 9,6  | 4,3  | 9,9  | 12,8 | 8,7  | 2,3  | 8,4  | 12,3 | 11,4 | 4,9  | 2,4  | 2,4  | 2,5  | 4,6  | 5,6  | 14,4 | 8,1  | 9,5  | 12,5 | 7,2  | 11,9 | 15,1 | <b>8,3</b>  |
| -5  | 9,4  | 10,1 | 14,3 | 14,3 | 11,2 | 6,6  | 12,8 | 12,3 | 10,8 | 11,6 | 14,1 | 9,3  | 10,1 | 14,9 | 11,7 | 10,4 | 4,2  | 3,8  | 3,9  | 12,9 | 4,3  | 11,5 | 8,1  | 3,0  | 13,8 | 6,3  | 16,5 | 8,8  | <b>10,2</b> |
| 0   | 15,5 | 11,9 | 12,9 | 14,9 | 9,5  | 12,2 | 12,8 | 14,2 | 11,5 | 16,8 | 15,2 | 8,3  | 13,2 | 11,3 | 11,7 | 9,1  | 4,5  | 7,2  | 9,2  | 10,2 | 13,9 | 13,8 | 14,1 | 14,9 | 12,5 | 11,1 | 13,4 | 14,1 | <b>12,1</b> |
| 5   | 12,3 | 11,9 | 9,5  | 14,3 | 11,2 | 10,0 | 12,0 | 12,8 | 11,0 | 9,8  | 9,9  | 7,3  | 11,4 | 12,4 | 11,7 | 10,4 | 7,8  | 7,9  | 10,6 | 13,5 | 16,0 | 13,5 | 17,1 | 13,4 | 12,5 | 10,6 | 11,5 | 9,3  | <b>11,5</b> |
| 10  | 9,7  | 11,5 | 11,1 | 16,1 | 8,9  | 14,8 | 10,4 | 11,4 | 11,8 | 15,3 | 14,4 | 8,9  | 12,6 | 7,3  | 7,6  | 11,6 | 8,7  | 13,4 | 10,4 | 11,6 | 13,9 | 8,6  | 11,1 | 13,9 | 6,3  | 7,7  | 14,2 | 7,8  | <b>11,1</b> |
| 15  | 11,6 | 10,5 | 6,2  | 14,3 | 10,1 | 10,5 | 8,0  | 10,9 | 10,1 | 5,5  | 7,6  | 11,3 | 8,3  | 6,5  | 7,3  | 6,1  | 7,8  | 13,1 | 13,7 | 13,5 | 13,0 | 4,0  | 7,4  | 10,0 | 5,0  | 6,3  | 6,9  | 2,9  | <b>8,9</b>  |
| 20  | 6,9  | 8,7  | 3,9  | 5,6  | 7,8  | 9,6  | 4,0  | 7,6  | 8,1  | 4,6  | 4,2  | 9,3  | 6,3  | 8,3  | 7,0  | 7,3  | 12,3 | 11,7 | 15,7 | 8,6  | 10,8 | 5,2  | 8,1  | 9,5  | 7,5  | 3,8  | 5,4  | 1,5  | <b>7,5</b>  |
| 25  | 4,0  | 7,7  | 0,7  | 1,2  | 5,0  | 8,3  | 2,4  | 6,6  | 4,6  | 4,3  | 1,1  | 10,3 | 5,6  | 5,3  | 3,2  | 4,3  | 10,5 | 11,0 | 9,8  | 5,3  | 5,2  | 4,0  | 4,7  | 4,0  | 3,8  | 4,8  | 3,1  | 0,5  | <b>5,0</b>  |
| 30  | 1,8  | 4,9  | 0,2  | 0,0  | 5,6  | 4,8  | 0,8  | 3,3  | 1,9  | 1,8  | 0,8  | 6,6  | 3,3  | 1,5  | 2,6  | 3,7  | 9,9  | 7,9  | 7,3  | 6,6  | 5,6  | 3,4  | 4,7  | 3,0  | 1,3  | 4,8  | 1,1  | 0,5  | <b>3,6</b>  |
| 35  | 2,5  | 3,1  | 0,9  | 0,0  | 1,1  | 3,5  | 0,8  | 1,9  | 2,1  | 1,5  | 0,8  | 7,3  | 2,8  | 1,2  | 2,0  | 1,8  | 7,5  | 7,6  | 5,0  | 3,6  | 2,6  | 0,9  | 3,7  | 1,0  | 0,0  | 1,0  | 1,9  | 0,0  | <b>2,4</b>  |
| 40  | 0,4  | 3,5  | 0,5  | 0,0  | 1,1  | 1,3  | 0,0  | 0,5  | 1,1  | 0,9  | 0,4  | 5,6  | 0,8  | 1,0  | 1,8  | 3,0  | 6,3  | 3,4  | 3,6  | 1,7  | 0,9  | 1,4  | 1,7  | 0,0  | 1,3  | 1,0  | 0,0  | 0,0  | <b>1,5</b>  |
| 45  | 0,0  | 1,4  | 0,2  | 0,0  | 0,6  | 0,4  | 0,8  | 0,9  | 1,4  | 0,3  | 0,0  | 3,0  | 0,8  | 0,2  | 0,3  | 1,8  | 4,5  | 2,4  | 2,0  | 0,7  | 0,0  | 0,3  | 1,0  | 0,0  | 0,0  | 0,5  | 0,0  | 0,0  | <b>0,8</b>  |
| 50  | 0,0  | 0,3  | 0,0  | 0,0  | 0,6  | 0,4  | 0,0  | 0,0  | 0,7  | 0,6  | 0,0  | 2,0  | 0,3  | 0,2  | 1,2  | 0,6  | 3,0  | 1,4  | 0,3  | 0,0  | 0,0  | 0,0  | 0,0  | 0,7  | 0,0  | 0,0  | 0,0  | 0,0  | <b>0,5</b>  |
| 55  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,2  | 0,3  | 0,0  | 1,3  | 0,3  | 0,0  | 0,0  | 0,0  | 2,1  | 0,7  | 0,6  | 0,0  | 0,0  | 0,3  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | <b>0,2</b>  |
| 60  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,2  | 0,0  | 0,4  | 1,3  | 0,2  | 0,2  | 0,0  | 0,0  | 1,5  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | <b>0,1</b>  |
| 65  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,7  | 0,2  | 0,2  | 0,0  | 0,0  | 0,6  | 0,7  | 0,3  | 0,3  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | <b>0,1</b>  |
| 70  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,3  | 0,0  | 0,0  | 0,0  | 0,0  | 0,3  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | <b>0,0</b>  |
| 75  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | <b>0,0</b>  |
| 80  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | <b>0,0</b>  |
| 85  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | <b>0,0</b>  |
| 90  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | <b>0,0</b>  |
| 95  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | <b>0,0</b>  |
| 100 | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | 0,0  | <b>0,0</b>  |

Voici le graphique de la distribution des données de la dernière colonne (%M) reprennant les fréquences moyennes pour les 28 examens :



On observe plus d'étudiants qui se sur-estiment (52,2%) que d'étudiants qui se sous-estiment (34,1 %) et une minorité d'étudiants dont l'auto-estimation est excellente (12,1%). La tendance à la surestimation a souvent été décrite dans la littérature. En nous basant sur les données des 28 examens, nous recommandons pour les épreuves qui ont lieu à la FAPSE, d'interpréter l'indice de centration (calculé selon la formule décrite ci-avant) de la façon suivante.

## L'INDICE DE CENTRATION REFLETE :

### UNE AUTO-ESTIMATION

**EXCELLENTE** si elle est comprise entre **-4 et +4** Performance accomplie par **12,1 %** des étud<sup>ts</sup>  
(avec surestimation très modérée si signe positif, ou sous-estimation très modérée si signe négatif)

### UNE SOUS-ESTIMATION

|                    |  |  |
|--------------------|--|--|
| <b>TRES LEGERE</b> | si elle est comprise entre <b>-5 et -9</b> | Performance accomplie par <b>10,2 %</b> des étud <sup>ts</sup> |
| <b>LEGERE</b>      | <b>-10 et -14</b>                          | <b>8,3 %</b>   |
| <b>ASSEZ FORTE</b> | <b>-15 et -19</b>                          | <b>6 %</b>   |
| <b>FORTE</b>       | <b>-20 et -24</b>                          | <b>4,1 %</b>   |
| <b>TRES FORTE</b>  | <b>-25 et -29</b>                          | <b>2,8 %</b>   |
| <b>EXTREME</b>     | <b>-30 et -100</b>                         | <b>2,7 %</b>   |

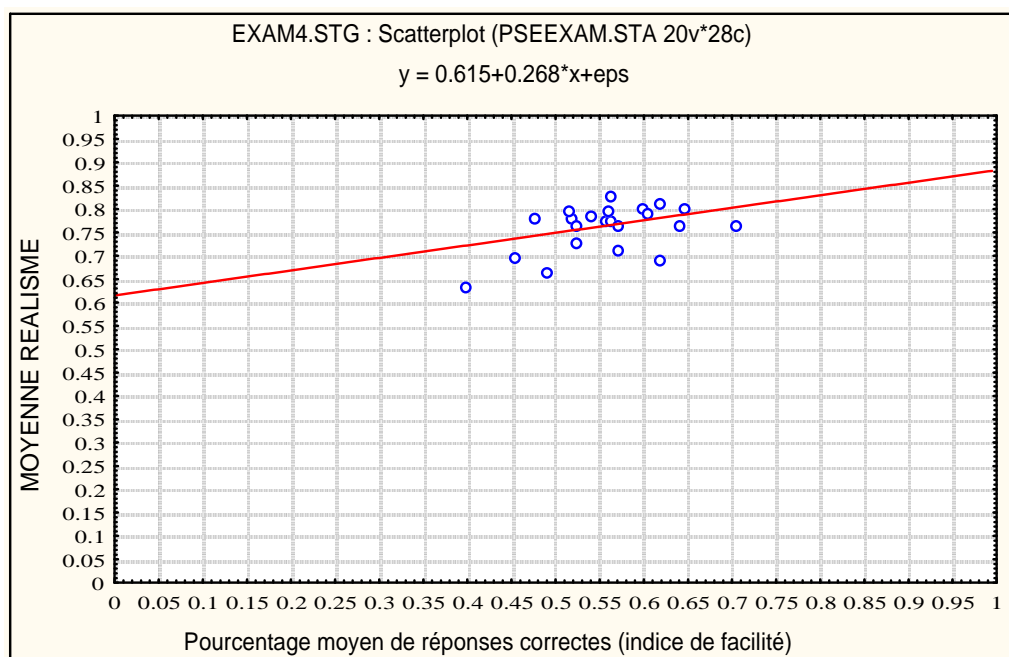
### UNE SURESTIMATION

|                    |  |  |
|--------------------|--|--|
| <b>TRES LEGERE</b> | si elle est comprise entre <b>+5 et +9</b> | Performance accomplie par <b>11,5 %</b> des étud <sup>ts</sup> |
| <b>LEGERE</b>      | <b>+10 et +14</b>                          | <b>11,1 %</b>  |
| <b>ASSEZ FORTE</b> | <b>+15 et +19</b>                          | <b>8,9 %</b>   |
| <b>FORTE</b>       | <b>+20 et +24</b>                          | <b>7,5 %</b>   |
| <b>TRES FORTE</b>  | <b>+25 et +29</b>                          | <b>5 %</b>   |
| <b>EXTREME</b>     | <b>+30 et +100</b>                         | <b>9,2 %</b>   |

- c) Comparaison des niveaux de performance métacognitive en fonction de la difficulté des examens (données des 28 examens ayant eu recours aux degrés de certitude entre 1994 et 1996 à la FA.P.S.E.)

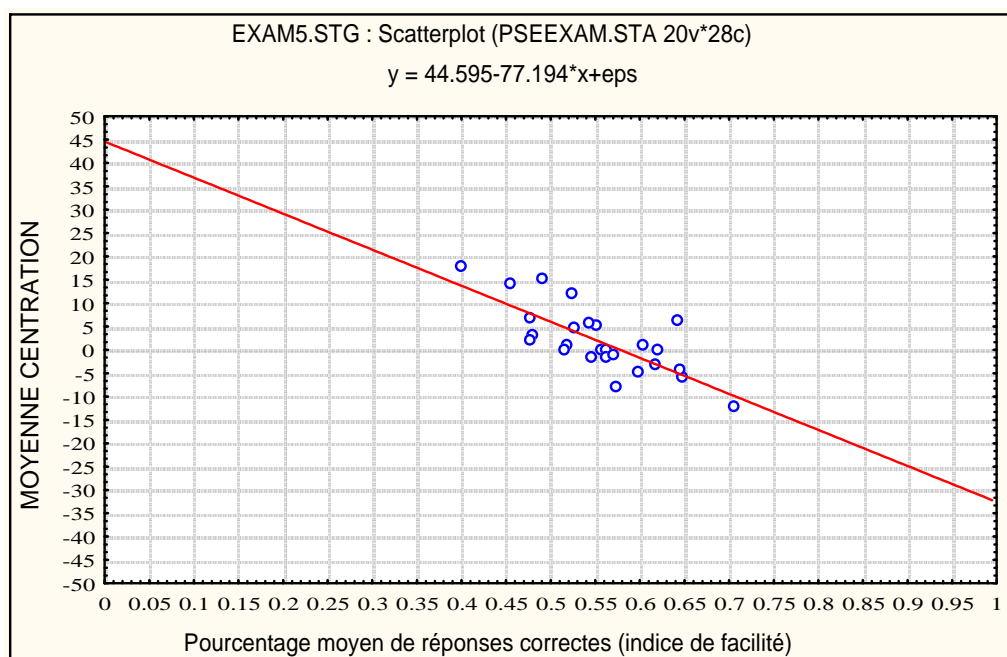
#### 1. Indice de réalisme et facilité de l'épreuve

On observe sur le graphique ci-après une relation entre la facilité des examens et les scores moyens à l'indice de réalisme. Un examen facile s'accompagne en général d'un meilleur score moyen de réalisme qu'un examen difficile. La difficulté de la tâche semble donc être liée à ce type de performance métacognitive.



## 2. Indice de centration et facilité de l'épreuve

On observe sur le graphique ci-après une relation entre la facilité des examens et les scores moyens à l'indice de centration. Plus une évaluation est facile, plus les étudiants se sous-estiment et inversement.

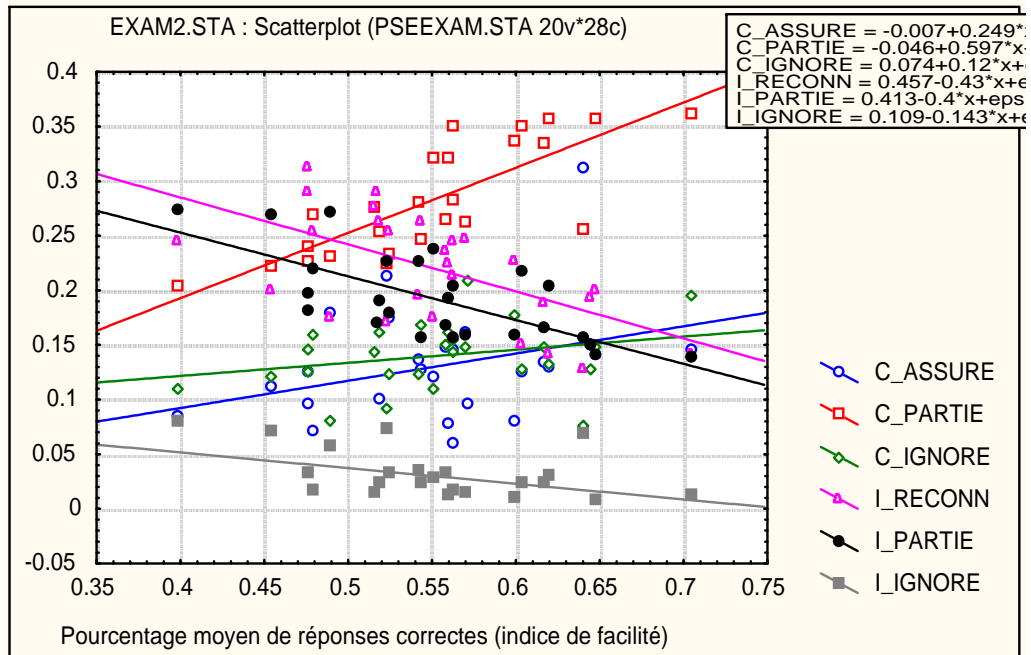


Cette relation est liée au mode de calcul de l'indice de centration qui consiste à soustraire le pourcentage de réponses correctes (taux de réussite qui est aussi notre indice de facilité d'une épreuve) à la certitude moyenne. Dès lors, comme le montre le tableau ci-dessous, il est logique d'observer une tendance à la sous-estimation lors d'examens avec un taux de réponses correctes élevé et une tendance à la surestimation lorsque le taux d'exactitude est faible.

|                                | Certitude moyenne <u>élevée</u>            | Certitude moyenne <u>faible</u>              |
|--------------------------------|--|--|
| % Rép. correctes <u>élevé</u>  | Centration proche de 0                     | Centration négative : <b>Sous-estimation</b> |
| % Rép. correctes <u>faible</u> | Centration positive : <b>Surestimation</b> | Centration proche de 0                       |

### 3. Etats du vecteur des degrés de compétence et facilité de l'épreuve

Les 6 catégories de compétence sont liées entre elles de façon complexe, parfois semblables du point de vue de l'exactitude et non du point de vue de la certitude et inversement.



Les catégories compétence assurée, partielle et ignorée étant liées aux réponses correctes, il est logique d'observer des pourcentages de fréquence systématiquement plus élevés à mesure que l'épreuve est plus facile et, inversement, des pourcentages plus élevés pour les catégories d'ignorance reconnue, partielle et ignorée à mesure que l'épreuve est plus difficile (droite descendante).

#### a. Degrés de certitude élevés

En terme de fréquence, c'est l'ignorance ignorée et la compétence assurée qui sont les catégories les moins fréquentes. Probablement les étudiants osent-ils moins recourir aux degrés de certitude élevés étant donné les scores récoltés en cas de réponse incorrecte (-20 avec certitude 5 en cas d'erreur). Ces deux catégories « démarrent » pratiquement ensemble sur le graphique, mais l'une, la compétence assurée, croît avec la facilité des épreuves, tandis que la fréquence de la catégorie ignorance ignorée décroît de façon presque symétrique.

#### b. Degrés de certitude intermédiaires

Les catégories ignorance partielle et compétence partielle ont des « pattern » pratiquement inversés l'une compense l'autre. Remarquons l'augmentation de la fréquence de la catégorie compétence partielle, proportionnellement plus forte que l'augmentation de la catégorie compétence assurée (également moins fréquente).

#### c. Degrés de certitude faibles

La compétence ignorée, inverse de l'ignorance reconnue est moins fréquente lorsque les épreuves sont difficiles, la catégorie compétence ignorée devient moins fréquente lorsque les examens sont plus faciles.

d) Comparaison des performances métacognitives en fonction des sous-groupes « classes » (1<sup>ère</sup> et 2<sup>ème</sup> candidature à la FA.P.S.E.)

1. Evolution de l'indice de réalisme et de centration : comparaison des performances en 1<sup>ère</sup> candidature et en 2<sup>ème</sup> candidature

| Examens de 1994 à 1996 à la FA.P.S.E. | Réalisme moyen | Centration moyenne |
|---------------------------------------|----------------|--------------------|
| 1ère candidature (15 examens) :       | 0.747          | 4.348              |
| 2ème candidature (7 examens) :        | 0.770          | 1.451              |
| Différences :                         | + 0.023        | - 2.897            |

On observe une légère amélioration à l'indice de réalisme ainsi qu'une baisse de presque 3 points de la tendance à la surestimation.

2. Evolution des états du vecteur des degrés de compétence : comparaison des performances en 1<sup>ère</sup> candidature et en 2<sup>ème</sup> candidature

On observe 53 % de réponses correctes en moyenne pour les 15 examens de 1ère candidature contre 57,3 % de réponses correctes pour les 7 examens de 2ème candidature. Le tableau ci-dessous reprend la répartition des réponses correctes et incorrectes dans le vecteur des degrés de compétence.

|  | Réponse CORRECTE avec CERTITUDE  |                                  |                                  | Réponse INCORRECTE avec CERTITUDE |                                  |                                |
|--|----------------------------------|----------------------------------|----------------------------------|-----------------------------------|----------------------------------|--------------------------------|
|  | 5 ou 4                           | 3 ou 2                           | 1 ou 0                           | 0 ou 1                            | 2 ou 3                           | 4 ou 5                         |
|  | Compétence assurée               | Compétence partielle             | Compétence ignorée               | Ignorance reconnue                | Ignorance partielle              | Ignorance ignorée              |
| Moy. 1 <sup>ère</sup> candi (15 examens) | <b>14.7%</b><br>(/53 = 27,7 %)   | <b>25%</b><br>(/53 = 47,1 %)     | <b>13.3%</b><br>(/53 = 25,1 %)   | <b>23.1%</b><br>(/47 = 49,1 %)    | <b>19.4%</b><br>(/47 = 41,3 %)   | <b>3.9%</b><br>(/47 = 8,3 %)   |
| Moy. 2 <sup>ème</sup> candi (7 examens)  | <b>11.6%</b><br>(/57,3 = 20,2 %) | <b>32.2%</b><br>(/57,3 = 56,2 %) | <b>13.5%</b><br>(/57,3 = 23,5 %) | <b>19%</b><br>(/42,7 = 44,5 %)    | <b>20.7%</b><br>(/42,7 = 48,5 %) | <b>2.3%</b><br>(/42,7 = 5,4 %) |

On constate en 2<sup>ème</sup> candidature une baisse de fréquence des catégories de compétence assurée (-3,1%), d'ignorance reconnue (-4,1%) et d'ignorance ignorée (-1,6%) au profit des catégories de compétence partielle (+7,2%) et d'ignorance partielle (+1,3%), la fréquence de la catégorie de compétence ignorée reste pratiquement stable (+0.2%).

3. Comparaison des indices de réalisme et de centration des étudiants de 1<sup>ère</sup> candidature en 1994-1995 en comparaison avec les étudiants de 2<sup>ème</sup> candidature en 1995-1996

Le groupe des étudiants inscrits en 2<sup>ème</sup> année en 1995-1996 obtient-il globalement de meilleurs scores métacognitifs que le groupe inscrit en 1<sup>ère</sup> année ?

|  | Réalisme moyen | Centration moyenne |
|--|----------------|--------------------|
| Examens de 1ère candidature en 1994-1995 (7 examens) : | 0.765          | 3.001              |
| Examens de 2ème candidature en 1995-1996 (5 examens) : | 0.766          | 0.342              |
| Différences :  | + 0.001        | - 2.659            |

On observe des scores de réalisme très proches. Par contre, en ce qui concerne la centration, le groupe des 2<sup>ème</sup> candidature en 1995-1996 confirme la baisse de la tendance à la surestimation (- 2.659).

4. Evolution des états du vecteur des degrés de compétence des étudiants de 1<sup>ère</sup> candidature en 1994-1995 en comparaison avec les étudiants de 2<sup>ème</sup> candidature en 1995-1996

Le pourcentage moyen de réponses correctes pour les 7 examens de 1<sup>ère</sup> candidature en 1995-1996 vaut 53,1% contre 59,2 % de réponses correctes pour les 5 examens de 2<sup>ème</sup> candidature lors de l'année académique suivante.

|  | Réponse CORRECTE avec CERTITUDE  |                                  |                                  | Réponse INCORRECTE avec CERTITUDE |                                  |                                |
|--|----------------------------------|----------------------------------|----------------------------------|-----------------------------------|----------------------------------|--------------------------------|
|  | 5 ou 4                           | 3 ou 2                           | 1 ou 0                           | 0 ou 1                            | 2 ou 3                           | 4 ou 5                         |
|  | Compétence assurée               | Compétence partielle             | Compétence ignorée               | Ignorance reconnue                | Ignorance partielle              | Ignorance ignorée              |
| Moy. 1 <sup>ère</sup> candi<br>(7 examens) | <b>13.3%</b><br>(/53,1 = 25 %)   | <b>26.2%</b><br>(/53,1 = 49,3 %) | <b>13.6%</b><br>(/53,1 = 25,6 %) | <b>24.4%</b><br>(/46,9 = 52 %)    | <b>19%</b><br>(/46,9 = 40,5 %)   | <b>3%</b><br>(/46,9 = 6,3 %)   |
| Moy. 2 <sup>ème</sup> candi<br>(5 examens) | <b>12.4%</b><br>(/59,2 = 20,9 %) | <b>33.3%</b><br>(/59,2 = 56,25%) | <b>13.5%</b><br>(/59,2 = 22,8 %) | <b>18,1%</b><br>(/40,8 = 44,3 %)  | <b>19.8%</b><br>(/40,8 = 48,5 %) | <b>2.2%</b><br>(/40,8 = 5,4 %) |

Les tendances décelée lors de la comparaison des performances en 1<sup>ère</sup> candidature et en 2<sup>ème</sup> candidature de 1994 à 1996 se confirment lorsqu'on compare uniquement les étudiants de 1<sup>ère</sup> candidature en 1994-1995 avec les étudiants de 2<sup>ème</sup> candidature en 1995-1996. On observe aussi en 2<sup>ème</sup> candidature une baisse de fréquence des catégories de compétence assurée (-0,9%), de l'ignorance reconnue (-6,3%) et d'ignorance ignorée (-0,8%) au profit des catégories de compétence partielle (+7,1%) et d'ignorance partielle (+0,8%), la fréquence de la catégorie de compétence ignorée reste pratiquement stable (+0.1%).

## 7. Conclusions

a) Globalement, performances métacognitives légèrement meilleures aux examens qu'aux quizz

L'indice de réalisme (tableaux) passe de 0.706 lors des quizz (N = 13 quizz) à 0.763 pour les examens (N = 28 examens). Les scores de centration montrent une plage plus réduite et plus centrée sur la valeur idéale 0 aux examens. Les étudiants semblent également plus prompts à reconnaître leur ignorance lors des examens (21.9% pour la catégorie ignorance reconnue du vecteur des degrés de compétence contre 18.6%). Ceci peut s'expliquer par un effet d'entraînement et peut-être aussi par un effort d'auto-estimation accru lorsqu'il y a un enjeu en termes de points à comptabiliser pour la réussite de l'année académique.

b) Des normes métacognitives pour les examens qui auront lieu à la FA.P.S.E.

Les normes établies devraient permettre aux étudiants de mieux se situer dans la population de la FA.P.S.E., d'une part, en ce qui concerne les scores obtenus à l'indice de réalisme et de centration, et, d'autre part, du point de vue des états du vecteur des degrés de compétence. Remarquons que des progrès restent encore à faire, seuls 1.6% des étudiants obtiennent un score idéal à l'indice de réalisme et, en ce qui concerne la centration, à peine 12.1 % font partie de la catégorie « auto-estimation excellente ».

Ces informations intéressent aussi les enseignants soucieux de former des individus avisés capables de s'auto-évaluer. De plus, le vecteur des degrés de compétence est informatif quant à la profondeur avec laquelle la matière a été assimilée. La lecture des données et du graphique au point 6.a)3. ci-avant, montre une situation qui est loin d'être idéale. En effet, en ce qui concerne les réponses correctes lors des examens, nous souhaiterions voir une inversion des tendances des fréquences des catégories compétence partielle (28.4%) et compétence assurée (13%) ainsi qu'une diminution de la compétence ignorée (14%) compensée par une augmentation de la compétence assurée.

Ceci pose le problème de l'amélioration des performances liées à l'auto-estimation de ses compétences. Comment améliorer le réalisme chez les étudiants ? Comment les encourager à moins se surestimer et à moins se sous-estimer ? Comment augmenter la fréquence de la catégorie compétence assurée dans le vecteur des degrés de compétence ? Des stratégies de changements peuvent être mises en place. Une première consiste à fournir systématiquement un feed-back sur les performances métacognitives. Une seconde, complémentaire et plus incitative, pourrait consister à associer au niveau de

compétence une cote prise en compte dans le résultat final de l'étudiant. Nous formulons l'hypothèse que cela contribuerait à favoriser la lecture et la prise en considération des feed-back métacognitifs et contribuerait à une amélioration des performances. L'étude comparative des effets de procédures de cotation avec ou sans « bonus » pour la performance métacognitive devrait permettre de mesurer l'impact de cette seconde stratégie.

c) L'impact de la facilité des examens sur les scores métacognitifs

Une évaluation facile s'accompagne en général d'un score moyen à l'indice de réalisme plus élevé qu'un examen difficile. Quant à l'indice de centration, il est logique d'observer une tendance à la sous-estimation lors d'examens faciles et une tendance à la surestimation lors d'examens difficiles.

Du point de vue des fréquences des catégories des états du vecteur des degrés de compétence, on observe, avec la facilité des examens :

- une croissance forte pour la catégorie compétence partielle ;
- une croissance intermédiaire pour la catégorie compétence assurée ;
- une croissance très faible pour la catégorie compétence ignorée ;
- la baisse très forte pour les catégories ignorance reconnue et ignorance partielle qui ont des pentes pratiquement parallèles ;
- une baisse, avec pente plus faible que pour les catégories ignorance reconnue et ignorance partielle, et, avec une décroissance presque symétrique à la catégorie compétence assurée, pour la catégorie ignorance ignorée.

Lorsqu'on observe les points de départ et d'arrivée des droites de régression du graphique 6.c)3. ci-avant, on constate un relatif étalement des fréquences des 6 catégories lorsque l'épreuve est difficile. Par contre, lorsque l'examen est facile, c'est une séparation assez nette aux extrêmes des catégories compétence partielle et ignorance ignorée qui est observée, les autres catégories se rassemblant entre les deux. On constate également un changement dans l'ordre d'importance des fréquences pour les réponses correctes : lorsqu'un examen est difficile, la compétence partielle est suivie de la compétence ignorée puis de la compétence assurée, tandis que lorsque l'examen est facile, c'est la catégorie compétence assurée qui précède la compétence ignorée.

d) Les scores aux indices de réalisme et de centration ainsi que les états du vecteur des degrés de compétence évoluent-ils avec le niveau d'étude des étudiants de la FA.P.S.E. ?

On observe une légère amélioration de l'indice de réalisme lorsqu'on compare tous les examens de 1<sup>ère</sup> année avec tous ceux de 2<sup>ème</sup> année (+0.023) et un *status quo* lorsqu'on compare uniquement les scores aux examens des étudiants qui en 1994-1995 étaient en 1<sup>ère</sup> année avec ceux de 2<sup>ème</sup> année en 1995-1996. En ce qui concerne l'indice de centration, on assiste à une progression moyenne de presque 3 points vers la valeur idéale 0 dans les deux cas.

Lorsqu'on compare 1<sup>ère</sup> et 2<sup>ème</sup> année, l'évolution des fréquences des catégories du vecteur des degrés de compétence montre, pour les réponses correctes, une augmentation de l'ordre de 7 % pour la catégorie compétence partielle. Pour les réponses incorrectes, il y a inversion des tendances : en 2<sup>ème</sup> année c'est l'ignorance partielle qui est plus fréquente que l'ignorance reconnue, à l'inverse de ce qui se passe en 1<sup>ère</sup> année.

Il semble donc que des changements s'opèrent avec le niveau de familiarisation des étudiants à la procédure d'auto-évaluation des compétences utilisée à la FA.P.S.E. Dans une recherche en cours, nous envisageons l'étude des données récoltées depuis deux ans dans une perspective longitudinale. Nous espérons mettre en évidence des profils de progression chez nos étudiants universitaires, tant du point de vue des indices de réalisme et de centration, que des états du vecteur des degrés de compétence.



## 8. Bibliographie

- ATTNEAVE, F. (1959). Application of information theory to psychology. New York: Holt, Rinehart and Winston.
- BOXUS & AL. (1991). Principes communs pour évaluer les résultats cognitifs de la formation. Commissions des Communautés européennes, programme Eurotecnet.
- BRUNO, J. (1993). Using testing to provide feedback to support instruction: a reexamination of the role of assessment in educational organizations. NATO ASI Series, Item Banking: Interactive Testing and Self Assessment, Berlin: Springer Verlag, 1993, Vol. 112, pp. 190-209.
- DE LANDSHEERE, G., (1979). Dictionnaire de l'évaluation et de la recherche en éducation, Presses universitaire de France, Paris.
- DE FINETTI, B. (1965). Methods for discriminating levels of partial knowledge concerning a test item, British Journal of Mathematical and Statistical Psychology 18, pp. 87-123.
- DIRKZWAGER, A. (1993). A computer environment to develop valid and realistic predictions and self-assessment of knowledge with personal probabilities. NATO ASI Series, Item Banking: Interactive Testing and Self Assessment, Berlin: Springer Verlag, 1993, Vol. 112, pp. 146-166.
- FABRE, J.-M. (1993). Subjective uncertainty and the structure of the set of all possible events. NATO ASI Series, Item Banking: Interactive Testing and Self Assessment, Berlin: Springer Verlag, 1993, Vol. 112, pp. 99-113.
- GILLES, J.-L. (1995). Entraînement à l'autoévaluation : une comparaison filles/garçons à l'université. Actes Colloque de l'AIPU « Enseignement supérieur : stratégies d'enseignement appropriées » - août 1995 - Université du Québec à Hull.
- GILLES, J.-L. et LECLERCQ, D., (1996). Procédures d'évaluation adaptées à des grands groupes d'étudiants universitaires - Enjeux et solutions pratiquées à la FAPSE-ULG, Actes du Symposium International sur la Rénovation Didactique en Biologie, Tunis, à paraître.
- HUNT, D. (1993). Theory and application to learning and testing. NATO ASI Series, Item Banking: Interactive Testing and Self Assessment, Berlin: Springer Verlag, 1993, Vol. 112, pp. 177-189.
- JANS, V. et LECLERCQ, D. (1996). Forum : un système d'animation et d'évaluation de grands groupes universitaires, communication au colloque de l'Association Internationale de pédagogie Universitaire (AIPU), Tunis, novembre 1996, à paraître.
- LECLERCQ, D. (1975). L'évaluation subjective de la probabilité d'exactitude des réponses en situation pédagogique. Thèse de doctorat en Sciences de l'Education, Université de Liège Institut de Psychologie et des Sciences de l'Education.
- LECLERCQ, D. (1983). Confidence marking, its use in testing. Postlethwaite, Choppin (eds.) Evaluation in Education, Oxford : Pergamon, 1982, vol. 6, 2, pp. 161-287.
- LECLERCQ, D. (1986). La conception des questions à choix multiple, Bruxelles, Ed. Labor.
- LECLERCQ, D. & al (1993). The Taste approach: General implicit solutions in MCQq, open books exams and interactive testing and self-assessment. NATO ASI Series, Item Banking: Interactive Testing and Self Assessment, Berlin: Springer Verlag, 1993, Vol. 112, pp. 210-232.
- LECLERCQ, D. & GILLES J.-L. (1994). GUESS, un logiciel pour entraîner à l'auto-estimation de sa compétence cognitive. Actes du colloque QCM et questionnaires fermés, Paris: ESIEE, 1994.
- LECLERCQ, D. & PLUNUS, G., (1996). Double Check, étude de méthodes d'innovation d'évaluations universitaires valides, formatives et efficaces, Rapport recherche-action, Université de Liège, Service de Technologie de l'Education, à paraître.
- LICHTENSTEIN, & al. (1975). Calibration of probabilities : the state of the art, decision making and change in human affairs Proceedings of the Fifth Research Conference on Subjective Probability, Utility and Decision Making, Darmstadt, 1-4 September, D. Reidel.
- PLUNUS, G. (1996). Pistes diagnostiques et procédures pour une évaluation universitaire valide, formative et efficace, Mémoire présenté pour l'obtention du grade de licencié en Sciences de l'Education, Faculté de Psychologie et des Sciences de l'Education, Université de Liège.
- SHANNON, C.E. (1951). Prediction and entropy of printed english. Bell Syst. Techn. J. 30, pp. 50-64.
- SHUFORD, E. & al (1966). Admissible probability measurement procedures. Psychometrika 31, pp. 125-145.
- VAN LENTHE, J. (1993). The development and evaluation of ELI, an interactive elicitation technique for subjective probability distributions. NATO ASI Series, Item Banking: Interactive Testing and Self Assessment, Berlin: Springer Verlag, 1993, Vol. 112, pp.
- VAN NAERSSSEN, R.F. & al (1966). Is de utiliteitscurve van examenscores een ogief ? Nederland Tijdschrift Psychologie 21(6), pp. 358-363.