

ENTRAÎNEMENT A L'AUTOEVALUATION : UNE COMPARAISON FILLES/GARÇONS A L'UNIVERSITE

Jean-Luc GILLES

Faculté de Psychologie et des Sciences de l'Éducation - Université de Liège - Belgique
Tél. : + 32 - 41 - 66.20.78 - Fax : + 32 - 41 - 66.29.53 - Internet : jlgilles@vm1.ulg.ac.be

1. CONTEXTE

Comme dans la plupart des universités européennes (GIBSS, JENKINS & al., 1992), le nombre d'étudiants inscrits en 1^{ère} candidature à la Faculté de Psychologie et des Sciences de l'Éducation (FAPSE) de l'Université de Liège a considérablement augmenté ces dernières années. D'environ 200 en 1986-87, le nombre d'inscrits est passé à 400 en 1994-95. Le dispositif d'évaluation a été changé en conséquence en recourant aux questions à choix multiple, ce qui permet d'informatiser les procédures de correction des examens. Une panoplie d'outils d'évaluation assistée par ordinateur a été mise au point au Service de Technologie de l'Éducation de la FAPSE. Le logiciel CERT autorise la correction automatisée des réponses des étudiants sur feuilles traitées par lecture optique de marques. Les logiciels CHECK, DOUBLE-CHECK et WINCHECK permettent des évaluations interactives où les questions sont posées via l'écran d'un ordinateur. Tous ces logiciels gèrent le traitement des réponses accompagnées de degrés de certitude (LECLERCQ, 1975, 1983, 1990) ainsi que les QCM comportant des solutions générales implicites¹ (LECLERCQ & al., 1993).

Lors de cette année académique 1994-95 plusieurs professeurs de la FAPSE et d'autres facultés de l'Université de Liège ont eu recours à ces outils d'évaluation informatisés. Les QCM avec degrés de certitude et solutions générales implicites furent utilisés, non seulement lors des examens, mais aussi lors de « quizz » en fin de cours (les étudiants répondent à quelques questions en utilisant des feuilles spéciales, au cours suivant chaque étudiant reçoit son bilan personnel et le professeur commente les réponses).

A la suite de pionniers tels que DE FINETTI (1965), VAN NAERSSSEN (1965), SHUFORD et al. (1966), de nombreux auteurs (BRUNO, 1993; HUNT, 1993; VAN LENTHE, 1993; FABRE, 1993; LECLERCQ, 1993; DIRKZWAGER, 1993) pensent actuellement qu'il est inconcevable de traiter de la même façon l'ignorance avouée, la connaissance incomplète (réponse correcte mais peu sûre), la connaissance satisfaisante (réponse correcte et sûre), la connaissance parfaite (réponse correcte absolument certaine), la connaissance mal fondée (réponse incorrecte et très sûre) aussi appelée « fallacy » et « misconception ».

Quand on recourt aux degrés de certitude, plusieurs règles méthodologiques doivent être respectées. (1) Une consigne « probabiliste », (2) un barème de tarifs calculé selon la théorie des décisions, (3) le calcul d'indices de réalisme, (4) un entraînement à la procédure.

De nombreuses façons d'exprimer le degré de certitude ont été décrites (voir LECLERCQ, 1993, pp. 114-131). Selon la règle 1, seules celles où la consigne est probabiliste (non pas " peu sûr " mais " certitude comprise entre 25 et 50 % ") et où le barème des tarifs avantage une expression sans biais de la certitude (règle 2) sont considérées comme " Admissible Probability Measurement Procedures " par SHUFFORD & al., (1966). Voici la consigne préconisée par LECLERCQ (1983, 1993, 1995) :

¹ Les QCM avec solutions générales implicites autorisent, en plus des solutions habituellement proposées, les quatre possibilités suivantes : aucune solution proposée n'est correcte, toutes sont correctes, il manque des données dans l'énoncé pour que l'on puisse choisir UNE solution comme correcte, il y a une absurdité dans l'énoncé (à dénoncer en priorité !).

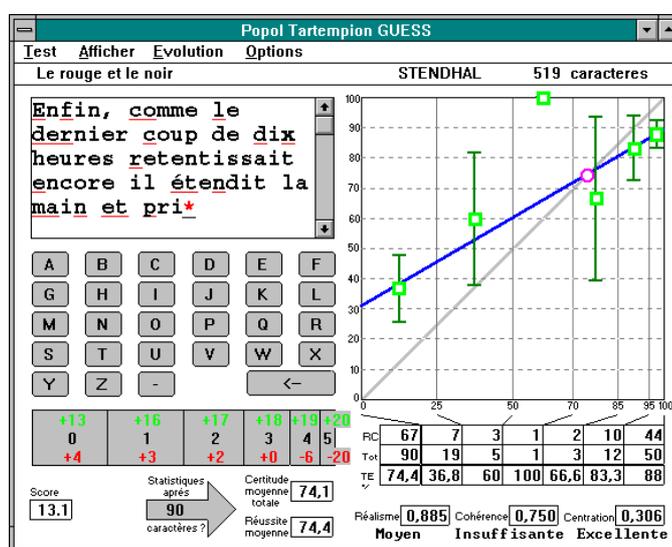
Si vous considérez que votre réponse a une probabilité d'être correcte comprise entre	Ecrivez	Vous obtiendrez les points suivants en cas de réponse	
		Correcte	Incorrecte
0 % et 25 %	0	+ 13	+ 4
25 % et 50 %	1	+ 16	+ 3
50 % et 70 %	2	+ 17	+ 2
70 % et 85 %	3	+ 18	+ 0
85 % et 95 %	4	+ 19	- 6
95 % et 100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

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 95 % et 99 % bien que la différence soit de 4 % seulement. Dans le premier cas (95 %) il a 1 chance sur 20 (1/20) de se tromper tandis que dans le second (99 %) il a 1 chance sur 100 (1/100), soit 5 fois moins. Etablir la même différence au milieu de l'échelle, par exemple entre 43 % (1/2,13) et 47 % (1/2,32), n'est pas pertinent car nous ne sommes pas capable de distinguer ces deux derniers « rapports »...

Avant d'employer les degrés de certitude tels que décrits ci-avant lors des examens, les étudiants se familiarisent avec leur utilisation (règle 4) à l'aide du logiciel GUESS.

2. L'ENTRAÎNEMENT A LA METACOGNITION PAR GUESS²

L'entraînement à l'auto-estimation de sa compétence cognitive à l'aide du programme GUESS 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 (inspiré de SHANNON, 1951 & 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. Idéalement, ces taux d'exactitude devraient se placer sur la diagonale (les valeurs attendues) et non plus haut (sous-estimation) ou plus bas (surestimation). Les taux de réussite sont encadrés de l'erreur de mesure du pourcentage. Voici l'écran tel que le voit le joueur :



² GUESS a été programmé dans le langage ToolBook par M. Michel HURARD, licencié en informatique.

3. INDICES DE METACOGNITION (règle 3)

3.1 Le réalisme

Le réalisme du joueur peut être calculé en se basant sur les différences entre les taux d'exactitude et les valeurs centrales des intervalles de probabilité qui déterminent la diagonale idéale. On est réaliste quand les taux d'exactitude obtenus pour chaque degré de certitude sont proches des valeurs de la diagonale idéale (lieu où la réalité et les prédictions devraient coïncider).

La formule du réalisme utilisée ici est :

$$R = 1 - \left(\frac{\sqrt{\frac{\sum_i (TE_i - VC_i)^2 \cdot NU_i}{NR}}}{|E|} \right)$$

avec : TE_i = Taux d'exactitude de la certitude i , VC_i = Valeur centrale, NU_i = Nombre d'utilisations, NR = Nombre total de réponses, E = écart maximum possible pour $TE_i - VC_i$ (dans le cadre des consignes préconisées par D. LECLERCQ la valeur centrale du degré de certitude le plus élevé³ autorise un écart maximum de 0,975).

Le réalisme est idéal lorsque R vaut 1 et minimal à 0.

3.2 La cohérence

La cohérence se mesure par la corrélation entre les valeurs centrales des 6 degrés de certitude et les taux de réussite des réponses fournies avec ces degrés respectifs (l'indice varie de 1 à -1). On est cohérent quand pour des certitudes élevées on obtient des taux d'exactitude élevés et qu'en même temps pour des certitudes faibles on obtient des taux d'exactitude peu élevés. On peut évidemment être cohérent avec soi-même sans être réaliste. L'indice de cohérence se calcule à l'aide de la formule de base du coefficient de BRAVAIS-PEARSON. La « droite de cohérence » (dessinée en gras sur le graphique) est la droite qui passe « au mieux » parmi les taux d'exactitude.

3.3 La centration

La centration est la différence entre la certitude moyenne et le taux d'exactitude moyen, négative si le joueur se sous-estime, positive s'il se surestime, nulle dans l'idéal.

4. DIFFERENCES METACOGNITIVES OBSERVEES ENTRE FILLES ET GARÇONS

4.1 LORS DE L'ENTRAÎNEMENT GUESS

Durant le 1er semestre 1994-1995, les étudiants de 1ère candidature de la FAPSE ont été invité à s'entraîner à l'auto-estimation de leurs compétences à l'aide de GUESS. Les données informatisées liées aux réponses fournies ont été conservées puis analysées à l'aide des indices de réalisme, cohérence et centration. Le groupe d'étudiants entraînés était composé de 119 filles et de 47 garçons. L'évolution durant l'entraînement des indices de métacognition a été étudiée à l'aide de moyennes dites « mobiles » calculées sur les données découpées en intervalles de 1/10ème des réponses fournies (le nombre de réponses fournies varie d'un étudiant à l'autre : min. = 90, max. = 424, moy. 193). Les moyennes mobiles

³ Le degré de certitude correspondant à une probabilité de réponse correcte comprise entre 95% et 100%.

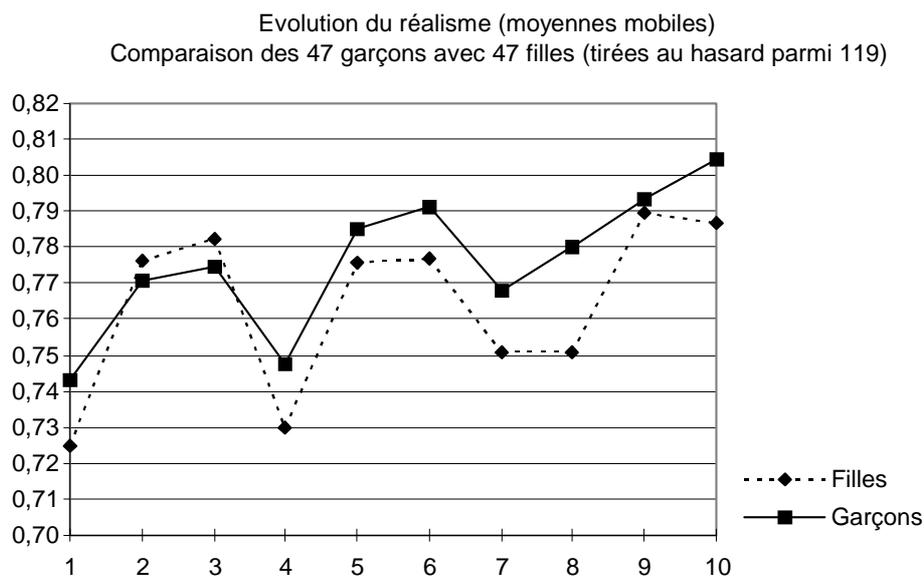
Colloque de l'AIPU « Enseignement supérieur : stratégies d'enseignement appropriées » - août 1995 - Université du Québec à Hull.

« Entraînement à l'autoévaluation : une comparaison filles/garçons à l'université »

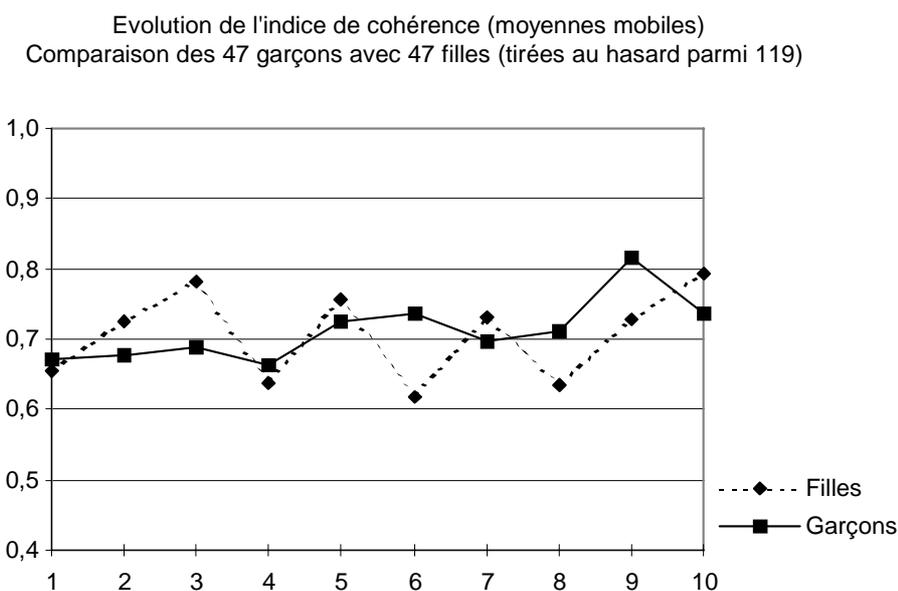
(calculées, donc, en moyenne sur une vingtaine de réponses) rendent mieux compte de l'évolution des indices de cohérence, réalisme et centration dans la durée car elles ne subissent pas l'effet cumulatif rencontré dans les moyennes classiques.

Les deux graphiques ci-après, montrent l'évolution de deux indices métacognitifs au cours de l'entraînement : le réalisme et la cohérence calculés sur les réponses de 47 filles (tirées au hasard parmi les 119) et de 47 garçons.

Rappelons qu'en ce qui concerne l'indice de réalisme, l'idéal vaut 1 et la valeur minimale 0. On observe une amélioration en cours d'entraînement. La progression n'est pas régulière, on observe des hauts et des bas dans les scores de réalisme, peut-être correspondent-ils à des stratégies différentes mises en place par les étudiants (pour améliorer leur gestion des degrés de certitude ? pour tester le système ?). Les différences entre filles et garçons semblent dues au hasard (t. de Student = -1,33).



Rappelons que la cohérence se mesure par la corrélation entre les valeurs centrales des 6 degrés de certitude et les taux de réussite pour les degrés respectifs (l'indice varie de 1 à -1). On observe, dans le groupe étudié une bonne cohérence tout au long de l'entraînement. Les différences entre filles et garçons semblent dues au hasard (t. de Student = 0,56).

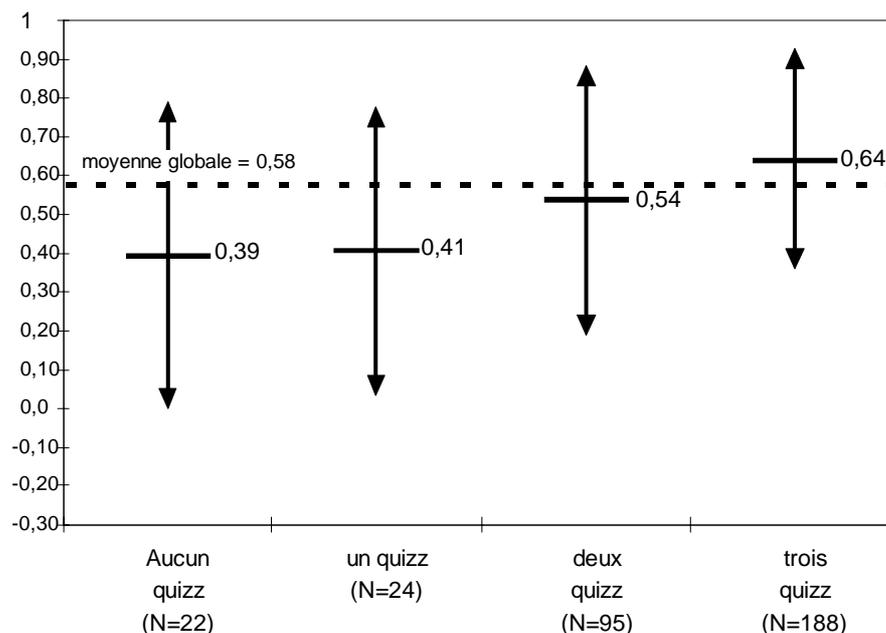


4.2 DIFFERENCES METACOGNITIVES OBSERVEES LORS D'UNE INTERROGATION PRECEDEE DE QUIZZ

Le but de l'entraînement GUESS était d'exercer les étudiants à l'auto-évaluation de leurs compétences. Le contenu de l'exercice, des extraits de textes de la littérature française, était indépendant de celui des cours proposés à la FAPSE. La consigne, compléter le texte avec une des 26 lettres de l'alphabet ou avec un espacement, était celle d'une QCM classique avec 27 solutions possibles. Pour ces deux raisons, en plus de GUESS, une autre approche fut adoptée pour familiariser les étudiants aux questions d'examen qui elles seraient à choix multiple avec Solutions Générales Implicites (SGI) et avec degrés de certitude. La méthode utilisée fut celle des quizz, c'est à dire quelques questions (entre 5 et 8) posées en fin de cours, corrigées pour le cours suivant avec feed-back personnalisé distribué à chaque étudiant. Les quizz permettent également au professeur de commenter (lors du cours suivant) les statistiques de résultats des élèves, les erreurs les plus fréquentes en réexpliquant la matière envisagée, la pertinence des questions (r.bis), etc. C'est un pas vers l'introduction d'un processus d'évaluation formative à l'université.

L'effet positif de cet entraînement à l'aide des quizz est mis en évidence dans le graphique ci-dessous qui reprend les moyennes des indices de réalisme⁴ pour quatre groupes d'étudiants ayant participé à une interrogation qui eut lieu le 8 novembre 1994, et qui avaient au préalable vécu un nombre différent (entre 0 et 3) de quizz.

Interrogation du 8/11/94 (30 QCM SGI) : moyennes des scores de réalisme en fonction des participations aux quizz



Rappelons que l'idéal de réalisme est 1. On observe le meilleur score à l'indice de réalisme dans le groupe qui avait précédemment effectué trois quizz (R = 0.64), viennent ensuite les étudiants qui en ont effectué deux (R = 0.54), puis ceux qui n'ont réalisé qu'un seul quizz (R = 0.41) et enfin ceux qui n'en ont

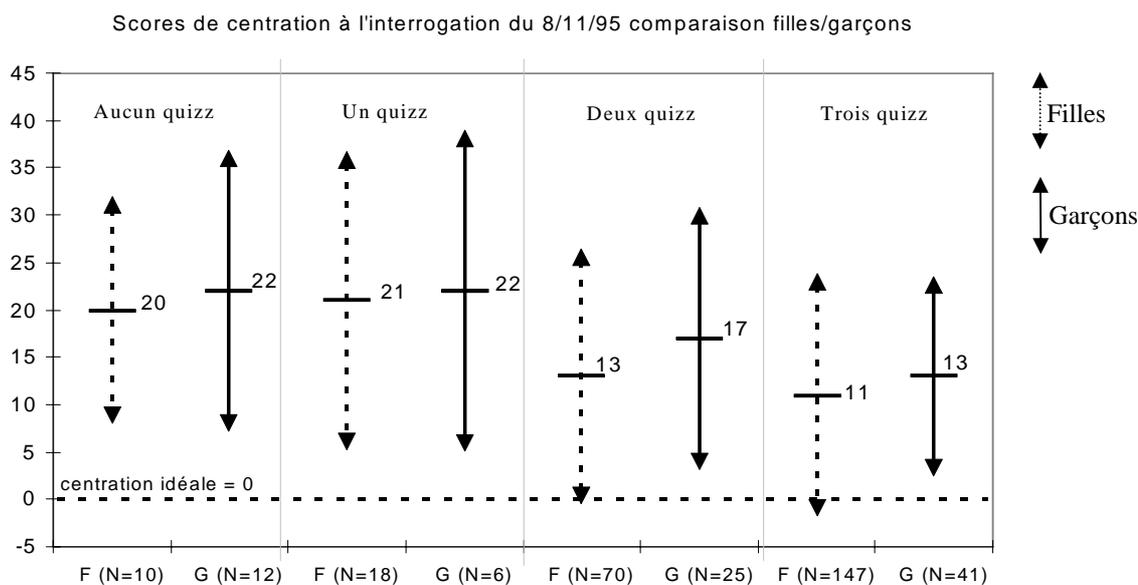
⁴ Le réalisme a été calculé ici à l'aide du logiciel CERT qui utilise la formule suivante :

$$MEM = \frac{\sum_i (|TE_i - VC_i|) \cdot NU_i}{NR} \quad \text{avec } R = 1 - (5 \text{ MEM})$$

BOXUS et al. (1991) ont établi des repères pour les valeurs obtenues à l'aide de cette formule. Le réalisme est idéal lorsqu'il vaut 1, excellent si il est supérieur à 0,97, bon entre 0,91 et 0,94, moyen entre 0,84 et 0,90, insuffisant à moins de 0,70 et minimal à -3,875.

aucun à leur actif ($R = 0.39$). Dans chaque groupe, le trait vertical représente UN écart type au-delà et en-deça de la moyenne. Plus les étudiants effectuent de quizz et plus l'écart type diminue. Ces résultats encourageants dans le cadre d'une étude exploratoire sont cependant à prendre avec prudence car nous ne sommes pas en mesure d'affirmer, dans l'état actuel de nos investigations, qu'il s'agit bien là d'un « effet quizz » et non d'un « effet assiduité » aux cours. En effet, les étudiants qui ont participé à tous les cours ont été amenés à effectuer tous les quizz, tandis que les autres, moins zélés, n'ont effectué qu'un ou deux, voire aucun quizz.

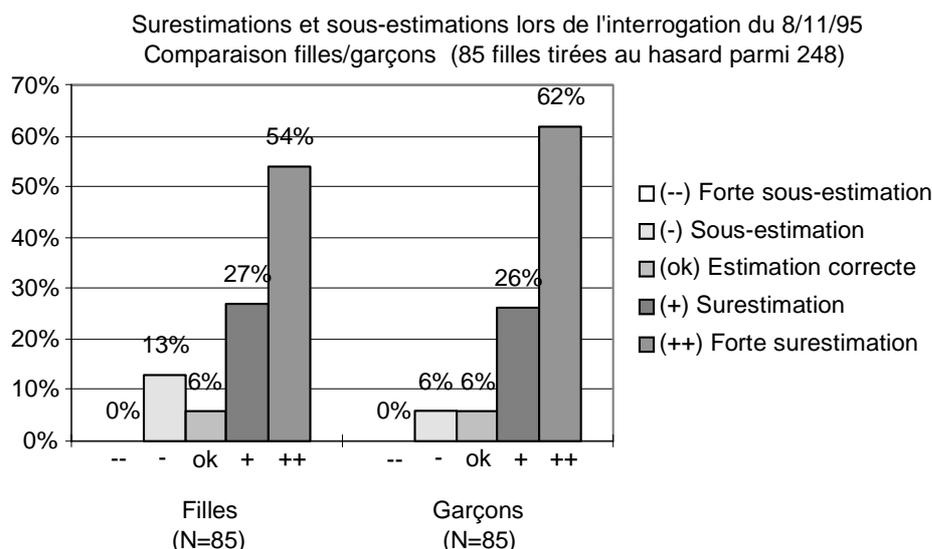
Rappelons que l'indice de centration idéal vaut 0 (différence entre la certitude moyenne et le taux d'exactitude moyen). Pour ce qui est des différences filles/garçons, le graphique ci-après montre un indice de centration à chaque fois meilleur pour le groupe des filles représenté par les traits verticaux en pointillés.



Grâce à l'indice de centration, des différences ont été observées en ce qui concerne la tendance à se sous-estimer ou à se surestimer selon que l'on fait partie du groupe des filles ou des garçons. Pour chaque étudiant, le signe de la centration permet d'établir si il y a surestimation (valeur positive) ou sous-estimation (valeur négative). Pour comparer les niveaux de sur ou sous-estimation dans les groupes des filles et des garçons, nous avons créé une échelle arbitraire à cinq degrés : « forte sous-estimation » (indice de centration au delà de -10), « sous-estimation » (de -10 jusqu'en deçà de -1), « estimation correcte » (de -1 à 1), « surestimation » (au-delà de 1 jusqu'à 10), « forte surestimation » (au-delà de 10). Les graphiques ci-après ont été établis sur base des effectifs de chaque catégorie transformés en pourcentages.

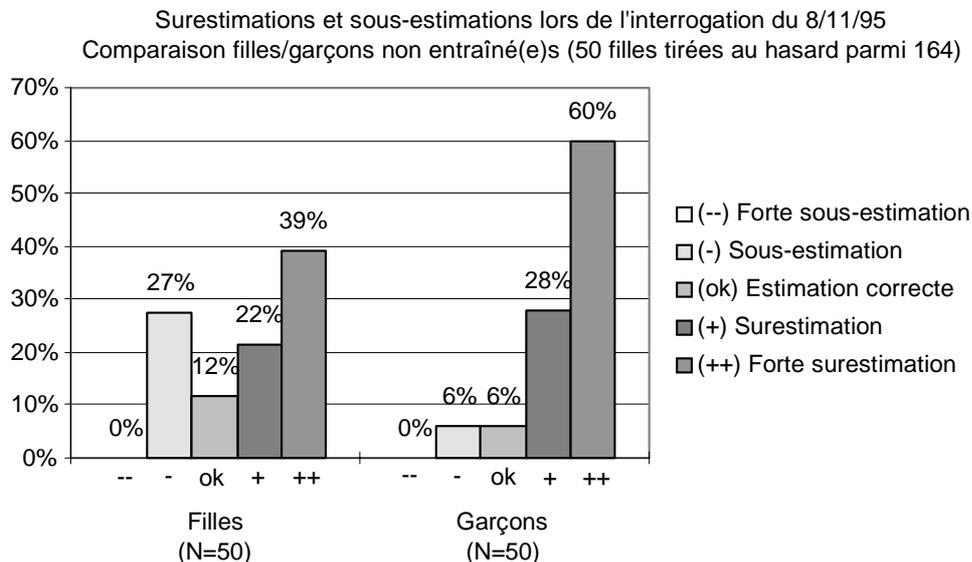
Les filles étant 3 fois plus nombreuses ($N = 248$) que les garçons ($N=85$) nous avons pour chaque comparaison tiré au hasard un nombre équivalent de filles.

On constate dans le graphique ci-après, une tendance majoritaire à la surestimation chez les filles ($27\% + 54\% = 81\%$) ainsi que chez les garçons ($26\% + 62\% = 88\%$). Cette tendance a souvent été soulignée dans la littérature (PITZ, 1974; BRUNO, 1993; DIRKZWAGER, 1993; FABRE, 1993; GATHY et DENEFF, 1993). On remarque également que 13% des filles se sous-estiment contre à peine 6% des garçons. Il faut également relever le fait qu'aucun sujet, ni chez les garçons, ni chez les filles, ne fait partie de la catégorie « Forte sous-estimation ».



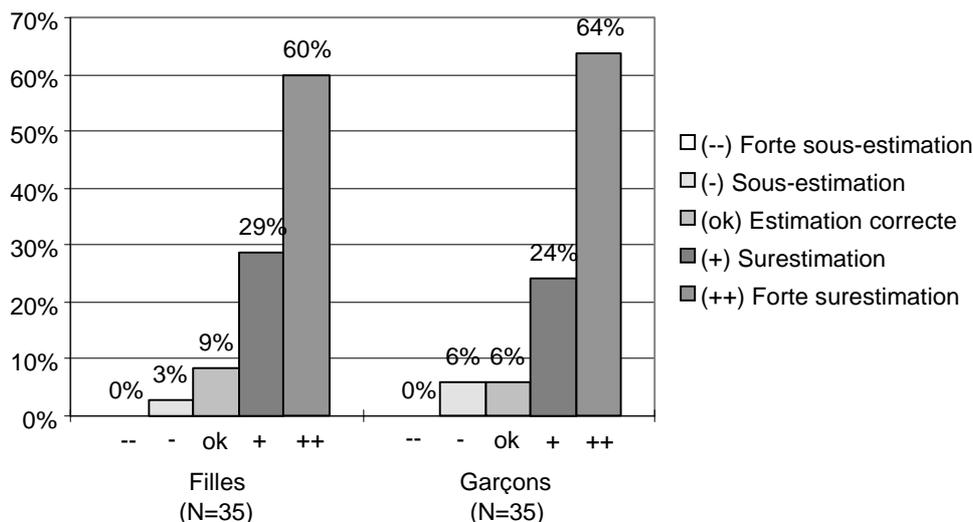
L'entraînement par GUESS influence-t-il les distributions des filles et des garçons dans les cinq catégories de sur et sous-estimation ?

La façon dont les 50 garçons non entraînés se distribuent dans le graphique ci-après est quasi identique à celle du graphique précédent. Lorsqu'on compare les filles et les garçons non entraînés, on observe des distributions assez différentes : 61% des filles (22% + 39%) se surestiment contre 88% des garçons (28% + 60%), 12% des filles s'auto-évaluent correctement contre 6% des garçons, enfin, 27% des filles se sous-estiment contre 6% des garçons seulement. Il semble donc que les filles non entraînées ont plus tendance à se sous-estimer que les garçons non entraînés.



Par contre dans le graphique ci-après des sujets entraînés avec GUESS, on observe un phénomène « d'égalisation » des distributions.

Surestimations et sous-estimations lors de l'interrogation du 8/11/95
 Comparaison filles/garçons entraîné(e)s (35 filles tirées au hasard parmi 84)



Alors que les garçons entraînés ont une distribution fort semblable à celle des non entraînés, les filles entraînées avec GUESS semblent adopter des comportements de surestimation dans des proportions presque identiques aux garçons (total surestimations filles = 89%, garçons = 88%). L'entraînement GUESS aurait-il un effet d'alignement des comportements de surestimation des filles sur ceux des garçons ? La présente étude ne nous permet pas de l'affirmer de manière catégorique car l'échantillonnage des groupes de filles et de garçons entraînés et non entraînés ne fut pas constitué aléatoirement. Quoi qu'il en soit, on eut préféré voir augmenter les pourcentages de la catégorie « Estimation correcte » chez tous les étudiants entraînés, ce n'est malheureusement pas le cas.

5. DISCUSSION

Plusieurs tendances métacognitives semblent se dégager des observations effectuées sur les groupes de filles et de garçons :

- un « effet quizz » plus profitable aux filles (elles récoltent des scores de centration à chaque fois légèrement meilleurs que ceux des garçons);
- la confirmation d'une tendance « lourde » de surestimation chez la majorité des filles et des garçons (tendance souvent décrite dans la littérature);
- des distributions de sur et sous-estimation disymétriques chez les filles et les garçons (aucun sujet ne se sous-estime fortement alors que la majorité se surestime fortement)
- une proportion de sujets non entraînés qui se surestiment moins élevée chez les filles avec chez ces dernières plus de sous-estimations que chez les garçons;
- un « effet d'alignement » de la distribution des filles sur celle des garçons dans le groupe des sujets qui se sont entraînés avec GUESS (moins de sous-estimations et plus de surestimations).

Les tendances relevées sont à prendre avec prudence car l'échantillonnage des filles et des garçons ainsi que des sujets entraînés et non-entraînés n'a pas été constitué aléatoirement. Il se pourrait que des variables cachées telles que « l'assiduité au cours » ou « zèle des étudiants » viennent parasiter ces observations. Ces variables devraient pouvoir être neutralisée par un plan d'expérience visant à démontrer « l'effet quizz » et « l'effet GUESS ».

Ces premières observations n'ont pas permis de mettre en évidence l'expression de « l'effet GUESS » attendu. En effet, dans la distribution des sujets entraînés le pourcentage des étudiants qui s'estiment correctement n'augmente pas.

6. CONCLUSIONS

Des différences métacognitives liées à l'auto-estimation de ses compétences sont observées lors d'une évaluation qui fut précédée de plusieurs quizz.

La tendance à la surestimation souvent décrite dans la littérature semble pouvoir se nuancer lorsqu'on compare les tendances à *sur* et *sous*-estimer ses compétences chez les filles et les garçons non entraînés à la gestion des degrés de certitude à l'aide du logiciel GUESS. Dans le groupe étudié, une majorité de filles non entraînées se surestime, mais proportionnellement moins que dans le groupe des garçons, et, inversement, plus de filles se sous-estiment. Est-ce là l'expression d'une plus grande « *prudence-modestie-réserve* » « féminine » et d'une trop grande « *confiance-assurance-hardiesse* » « masculine » lorsqu'il s'agit de s'autoévaluer ? Cette hypothèse est également étudiée par A. DIRKZWAGER (1995) dans le cadre d'une recherche interculturelle réunissant des chercheurs américains, australiens, espagnols, mexicains et néerlandais.

Nos observations semblent indiquer qu'un changement est possible : le groupe des filles entraînés avec GUESS s'est aligné sur celui des garçons. On eut cependant préféré observer une augmentation des estimations correctes chez les filles et chez les garçons. Un tirage aléatoire des sujets pour la constitution des différents groupes ainsi qu'un plan d'expérience rigoureux seront nécessaires pour valider les différences métacognitives observées lors de cette étude exploratoire.

7. BIBLIOGRAPHIE

- ATTNEAVE, F. (1959). Application of information theory to psychology. New York: Holt, Rinehart and Winston.
- 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 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.
- DIRKZWAGER, A. (1995). Testbet version 1.01 instructions manual. Bussum, Computers in Education, 1995.
- 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.
- GATHY, P. & DENEFF, J.-F. (1993) Self-confidence assessment during computer-assisted testing in histology. NATO ASI Series, Item Banking: Interactive Testing and Self Assessment, Berlin: Springer Verlag, 1993, Vol. 112, pp. 233-241.
- GIBBS, G., JENKINS, A. & al (1992). Teaching large classes in higher education - How to maintain quality with reduced resources, London: Kogan Page, 1992.
- GILLES, J.-L. (1995). Gender comparison of metacognition : realism in self-estimation in cognitive competency with university students, 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.
- 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. & DE BROGNIEZ P. (1990), A Fresh look on confidence marking. In Estes, Heene, Leclercq (eds.) New pathways to learning through educational technology. Proceedings of the Seventh International Conference on Technology and Education, Brussels, vol. 1, pp. 646-649.
- 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.
- PITZ, G. F. (1974). Subjective probability distributions for imperfectly known quantities. In Greg, L.W. (ed.) Knowledge and Cognition. New York: Adison Wiley, pp. 29-41.
- 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.

COMMENT SE PROCURER GUESS

Pour le logiciel GUESS : contacter Mme Brigitte DENIS, Service de Technologie de l'Education, Université de Liège, Bât. B32, 4000 Liège, Belgique. Tél. : + 32-41-66.20.72, Fax : + 32-41-66.29.53 Computers in Education