

**Colloque itinérant sur la Pédagogie Universitaire
entre les Universités du Québec et les Universités de
la Communauté française de Belgique**

DOUBLE CHECK

**Un concept d'évaluation interactive permettant de distinguer les
capacités de compréhension et d'analyse**

D. LECLERCQ, B. DENIS, E. BOXUS, J.L. GILLES

Liège, 25/10/94
Namur, 26/10/94
Louvain-La-Neuve, 27/10/94
Bruxelles, 28/10/94

A. TROIS NIVEAUX EVALUES SUR SIX

En 1956, la parution de la taxonomie des objectifs cognitifs de BLOOM a fait l'effet d'une bombe. Il y critiquait beaucoup le système traditionnel d'évaluation qui fait la part du lion à la mesure des connaissances **de mémoire**. Il ne s'est pas contenté de la critique. Au terme de l'analyse de milliers de questions posées par des enseignants, il proposait une TAXONOMIE (classification) des objectifs cognitifs (en fait des processus mentaux à promouvoir ET à évaluer), en six grands niveaux :

La connaissance	Capacité d'évoquer de mémoire (ou de reconnaître) ¹ des faits, de règles, de lois, d'événements, de noms, de terminologies, de valeurs quantitatives, de dates, etc.
La compréhension	Tous les éléments étant présentés au sujet, celui-ci les INTERPRETE correctement (ex : traduit un graphique en mots et vice-versa), ce qui implique une certaine connaissance.
L'application	Capacité de comprendre une situation (où toutes les données sont disponibles), de se remémorer un (des) principe(s) ou/et un (des) algorithme(s) et de faire correspondre les éléments (concrets) de l'un aux composantes (abstraites) de l'autre.
L'analyse	Devant un problème présenté d'une façon inhabituelle, ou peu claire, il s'agit de réorganiser, de distinguer l'essentiel de l'accessoire, le nécessaire du suffisant, de déjouer les pièges,
La Synthèse	Capacité de produire une FORMULATION personnelle, de rédiger (souvent de façon originale, si possible créative). Cela peut être une compression ou une expression ²
L'évaluation	Capacité de juger en s'étant donné des critères personnels et d'assumer ses décisions et (selon LECLERCQ) de prendre des risques en situation d'incertitude.

Il faut hélas constater, près de quarante ans plus tard, que les enseignants n'évaluent **SYSTEMATIQUEMENT** que la moitié de ces processus mentaux, et ce parce qu'ils (elles) ne disposent pas des concepts et instruments d'évaluation ad hoc

Les enseignants sont en effet assez bien outillés pour mesurer trois des six types de processus mentaux et, comme par hasard, ce sont ces trois-là qui sont mesurés le plus systématiquement :

La connaissance (ex : Quelle est la valeur de g ? Quel est l'énoncé de la loi d'Archimède ?)

Donner une part trop grande à la connaissance, c'est s'exposer à ce que les étudiants mémorisent des choses qu'ils n'ont pas comprises ; or

On ne devrait jamais mémoriser ce qu'on n'a pas compris (COMPREHENSION)

L'application (ex : Voici un triangle. Mesurez (avec votre règle) et calculez son aire)

Le danger est ici que les étudiants se limitent aux cas classiques, exposés au cours (comme l'a montré HALLEUX, 1969) et soient désarçonnés lorsque le problème se présente de façon inhabituelle.

Or on devrait autant savoir SE POSER les questions qu'y répondre (ANALYSE)

¹On pourrait reprocher à BLOOM de ne pas distinguer la mémoire "de rappel ou d'évocation" de la mémoire "de reconnaissance", beaucoup moins exigeante. GUILFORD, dans son "modèle tridimensionnel de l'intelligence" fait cette distinction entre (re)cognition et mémoire (évocation).

²Inventer un titre pour un article est une compression (rendre plus compact, plus concis), mais dans le sens de BLOOM, la catégorie SYNTHÈSE contient aussi les digressions, qui peuvent être plus longues qu'au point de départ.

La synthèse (ex : Que pensez-vous de l'oeuvre de....)

L'examen oral et les épreuves écrites à réponse longue permettent en principe de mesurer ce type de performance, quoique la docimologie critique (PIERON, 1963 ; DEBAL, BECKERS et DELANDSHEERE, 1977, DELANDSHEERE, 1993) ont montré la difficulté d'assurer la validité et surtout la fidélité des notes attribuées. Cependant, la synthèse n'engage pas par elle-même.

Or il faut entraîner les apprenants à prendre leurs responsabilités (EVALUATION)

B. CURRICULUM CACHE ET VIGILANCE COGNITIVE

Dans leur célèbre expérience sur la séduction pédagogique qu'ils ont appelée *The Dr Fox experiment*, NAFTULIN et al. (1975) ont entraîné un acteur à faire un exposé devant plusieurs publics, certains comportant du personnel médical et paramédical, d'autres des psychologues et sociologues. L'acteur donne cours de façon brillante, mais truffe son exposé d'inepties, de comparaisons fallacieuses, de digressions sans fondement...non détectées ! Et lors du questionnaire écrit, la majorité des auditeurs se déclare enchantés de la leçon, ayant le sentiment d'avoir "appris beaucoup". Les auteurs nous mettent ainsi en garde contre le critère "de satisfaction du consommateur" comme seul et unique base de jugement de la qualité d'un enseignement.

Cet exemple illustre le fait que l'école ne nous a pas entraînés à nous méfier, à vérifier, à éviter les pièges d'un curriculum manifeste, et encore moins de ceux d'un **curriculum caché**, c'est-à-dire "ce que personne n'enseigne, mais que tout le monde apprend".

L'école, en effet, et ce de la maternelle à l'université, "véhicule" des messages implicites tels que

Message implicite	Or
Quand on vous pose une question, il faut répondre.	Il existe des situations où il faut refuser de répondre, soit parce que la question est indiscrete, soit parce qu'on n'est pas compétent, soit qu'une autre personne présente répondrait plus efficacement, soit parce qu'on a besoin de temps pour réfléchir, etc.
Quand une question est posée, il existe une réponse (et UNE SEULE)	Pour certaines questions, il n'existe pas de réponse, pour d'autres, il en existe plusieurs
Une question posée par l'autorité est forcément bien posée	Il arrive que (volontairement ou non), l'on ne fournisse pas à celui qui doit répondre assez d'information pour le faire valablement

Notre système d'enseignement n'entraîne pas à détecter ces situations omniprésentes dans la vie courante. Or, il faudrait exercer les apprenants non seulement à EXECUTER ce qu'on leur ordonne (dans les deux sens d'imposer et de mettre de l'ordre) mais aussi (et surtout) à faire preuve de VIGILANCE COGNITIVE.

Encore faut-il que cette exercice soit fréquente et systématique !

Pour ces deux raisons, nous recourrons systématiquement au principe de la sollicitation implicite (LECLERCQ, 1986) pour développer et mesurer les niveaux "compréhension" et "analyse" de la taxonomie de BLOOM, et ce au moyen de QCM à solutions générales implicites (QCM SGI)

C. CONSIGNES CONCERNANT LES SOLUTIONS GENERALES

Voici le texte qui est remis aux étudiants :

Les questions sont à choix multiple. Chaque question comporte une (et une seule) solution correcte. Cependant, certaines solutions (appelées solutions générales) font appel à votre vigilance; elles vous sont proposées sous les numéros de code suivant : 6 - 7 - 8 ou 9 et ne seront pas dactylographiées dans les QCM, bien qu'elles soient d'application pour toutes !!!

6 ou REJET = Aucune des solutions proposées n'est correcte.

7 ou TOUTES = TOUTES les solutions proposées sont correctes.

8 ou MANQUE = Il est impossible de répondre parce que de l'information (au moins une donnée) MANQUE dans l'énoncé de la question (donc pas dans le cours ni dans les connaissances actuelles sur le problème).

9 ou ABSURDITE = Une ABSURDITE dans l'énoncé rend toute la question sans objet parce qu'il y a par exemple une CONTRE-VERITE dans l'énoncé (il ne s'agit donc pas ici de dénoncer une absurdité dans les solutions !).

Attention ! La réponse 9 a priorité sur 6, 7 et 8 et, évidemment sur les réponses 1, 2 et 3, dactylographiées, etc.

Exemples :

Il faut répondre

La capitale de la France est	1. Lille 2. Lyon 3. Paris	3
La capitale de l'Italie est	1. Berlin 2. Prague 3. Tokyo	6
La Grande Bretagne comprend	1. L'Angleterre 2. L'Ecosse 3. Le Pays de Galle	7
Quel âge avait Rimbaud ?	1. 2 ans 2. 10 ans 3. 20 ans	8
En quelle année Jules César a-t-il rencontré Napoléon ?	1. 1850 2. 1915 3. 1945	9

Pour toutes les QCM qui suivent, vous êtes invité(e) à fournir une seule réponse, c'est-à-dire un des chiffres codant les solutions proposées ou les chiffres 6, 7, 8 ou 9.

D. L'INTERACTIVITE : CONDITION D'UN DIAGNOSTIC PRECIS

On constate que les évaluations sont insuffisamment diagnostiques. En particulier, les chercheurs ont jusqu'ici été incapables de distinguer la compréhension de l'analyse lorsqu'on vise à en mesurer la maîtrise. Pour deux raisons. Tout d'abord parce que ces concepts sont restés trop flous. C'est pourquoi nous définissons la compréhension comme "la capacité de traduire et/ou de paraphraser un contenu (de sa forme graphique à sa forme verbale ou numérique) sans aucun piège alors que notre attention est attirée sur la question". Pour la même raison, nous définissons l'analyse comme "la vigilance cognitive, c'est-à-dire la capacité de détecter des anomalies, des contradictions, des ambiguïtés, des manques de données, des absurdités".

La deuxième raison est un manque d'interactivité et de structuration dans l'échange questions-réponses, même oral. Il fallait donc concevoir une évaluation, qui ait les avantages

- d'une QCM ou d'une QROC (Question à Réponse Ouverte Courte), qui permettent de poser beaucoup de questions;	mais plus profonde
- des QROL (Questions à Réponse Ouverte Longue) qui, dans les examens écrits permettent de s'exprimer longuement et qui comme les examens oraux remettent "sur la voie" quand on s'était dirigé sur une fausse piste;	mais plus interactive
- de l'examen oral qui hélas ne peut se prolonger indéfiniment. En outre, on n'en sort pas avec un diagnostic précis (écrit). Enfin on sait, depuis les travaux des docimologues comme PIERON (1963), combien ces évaluations sont sujettes à des "effets" (stéréotypie, halo, séquence, contraste, sévérité, etc.).	moins coûteuse, moins subjective et plus diagnostique

C'est pourquoi le STE-ULG a développé **Double check** sur ordinateur. Chaque Double question (ou BIquestion) présente deux volets :

- la question PRIM (ou principale, de poids 2)
- la question BIS (ou sous-question, de poids 1)

Alors que la question PRIM fonctionne avec les 4 solutions implicites (6 = Aucune, 7 = Toutes, 8 = Manque, 9 = Absurdité), la question BIS, elle, ne fonctionne qu'avec 2 solutions implicites (6= Aucune, 7= Toutes).

On voit en quoi cette (bi)question constitue un approfondissement puisqu'elle permet de distinguer 4 situations. Pour en donner un exemple, imaginons une question où la réponse correcte à la question PRIM est "ABSURDITE dans l'énoncé". Voici les 4 situations dans lesquelles l'étudiant(e) peut se trouver :

Sit	A la question PRIM , il (elle)	A la question BIS , il (elle)	Diagnostic	Points
a	A bien répondu (ex : Absurde)	A bien désigné l'absurdité	Vigilance + Compréhension +	3
b	N'a pas bien répondu (n'a pas détecté l'absurdité)	Fait plus attention et détecte (enfin) l'absurdité.	Vigilance - Compréhension +	1
c	N'a pas bien répondu (n'a pas détecté l'absurdité)	Ne détecte toujours pas l'absurdité	Vigilance - Compréhension -	0
d	A bien répondu (ex : Absurde)	Ne désigne pas l'absurdité prévue par le professeur, mais une autre !	Vigilance ? Compréhension ?	2

E. PRINCIPES DE TESTING INTERACTIF

1. Le double passage par chaque question PRIM

Quand les étudiants reçoivent plusieurs dizaines de QCM sur papier, plusieurs stratégies sont possibles. L'une d'entre elles consiste à lire toutes les questions et à répondre à celles pour lesquelles on voit clairement quelle est la bonne réponse, gardant les autres pour plus tard, s'il reste du temps.

Une stratégie désastreuse consiste à les attaquer strictement dans l'ordre où elles sont imprimées et à prendre le temps (parfois long) pour répondre à chacune, au risque de n'avoir même pas lu un grand nombre de questions à la fin de l'épreuve.

Un tel manque de méthode mérite d'être sanctionné en soi.

On constate que trop d'étudiants gèrent mal leur examen

C'est, selon nous, parce qu'ils n'ont pas bénéficié d'un entraînement *ad hoc*

Lors d'épreuves par QCM sur ordinateur, l'étudiant doit pouvoir gérer son temps de la même façon (le temps qui reste est affiché en haut à droite de l'écran, avec un bip sonore lorsqu'il ne reste que 5 minutes). C'est pourquoi, dans CHECK, toutes les questions sont présentées deux fois (du moins leur partie PRIM).

Dans WINCHECK, l'étudiant peut revenir autant de fois qu'il le veut et dans l'ordre qu'il veut sur chacune des questions (partie PRIM).

2. Réponses forcées

Lors de la première présentation d'une question, l'étudiant DOIT y répondre, mais sa réponse (et le degré de certitude qui l'accompagne) pourra être modifiée lors du SECOND PASSAGE qui, lui, se déroule alors que l'étudiant a déjà eu connaissance du contenu de toutes les questions (partie PRIM). La réponse est alors définitive et l'étudiant peut connaître la réponse correcte s'il le désire (certains étudiants préfèrent ne confronter leur réponse aux solutions correctes que tout à la fin de l'épreuve)

Il doit toujours répondre parce que l'expérimentateur a besoin de données ! Le professeur veut en effet par là recueillir des informations sur les questions où l'on répond bien du premier coup,

rapidement (car le délai de chaque réponse est enregistré) et avec un degré de certitude élevé. S'il ne fournit pas une réponse ET un degré de certitude, l'étudiant ne peut consulter une autre question.

3. Les feedbacks détaillés

L'étudiant peut non seulement recevoir le feed-back général (comme dans DOUBLE CHECK), mais aussi, il peut recevoir, pour chaque solution, une explication de la raison pour laquelle elle est correcte ou incorrecte.

L'étudiant a-t-il intérêt à prendre connaissance de ces explications, donc à les demander ? DEPOVER (1987, p. 99) apporte une réponse à cette question :

"GILMAN (1968) [et SCHOEN (1972)] a montré, dans une situation d'apprentissage basée sur des QCM, que le fait d'expliquer à l'élève en quoi la réponse choisie est erronée, conduisait à des résultats supérieurs à ceux enregistrés à partir d'une procédure consistant simplement à lui indiquer la réponse correcte."

Il poursuit ainsi (p. 100) :

*"Alors que dans une procédure de **feed-back passif** la réponse correcte est fournie à l'élève après un ou plusieurs essais infructueux, un **feed-back actif** consiste à proposer une série de questions qui mènent l'élève à découvrir par lui-même la réponse correcte*

A l'occasion de deux expériences [menées par TAIT et al. (1973)], seuls les élèves les plus faibles ont amélioré significativement leur performance suite à la présentation d'un feed-back actif".

BIBLIOGRAPHIE

- ALBERTINI, J.M., *La pédagogie n'est plus ce qu'elle sera*, Paris : Le Seuil, Presses du CNRS, 1992.
- ALBERTINI, J.M., *L'économie en 200 schémas*, Paris : Les Editions Ouvrières, 1994.
- BOXUS E., *Les QCM à solutions générales au service de l'évaluation à livre ouvert*, Colloque international "Formation, évaluation, sélection par questionnaires fermés", Université de Liège - Centre de Technologie de l'Education, Liège, 1988.
- BOXUS E., CHECK, une banque de questions interactive, in WEBER et DUMONT, *Les Questionnaires Automatisables*, colloque de l'ESIEE, Marne-la-Vallée, 1992.
- BOXUS E. et al., *Principes communs pour évaluer les résultats cognitifs de la formation*, programme EUROTECNET, Commission des Communautés européennes, 1991.
- DENIS B. et JANS V., *Logiciel CHECK- Manuel d'utilisation (version provisoire)*, Université de Liège - Service de Technologie de l'Education, Liège, février 1994.
- HURARD, M., *Logiciel d'aide à la métacognition sur les compétences cognitives*, Mémoire de licence en informatique, Université de Liège, 1994.
- JANS, V., *Développer la vigilance cognitive à l'école primaire par des QCM à Solutions Générales Implicites*, Université de Liège, mémoire de licence en Sciences et Techniques de la Formation Continue, 1993-1994.
- JANS, V., *Evaluation et Métacognition : l'avis des utilisateurs de CERT en formation professionnelle*, Colloque 'QCM et Questionnaires Fermés', Marne La Vallée, décembre 1994.
- LAMBERT F., *Evaluation interactive informatisée et auto-évaluation*, mémoire de licence en Sciences de l'Education, Université de Liège, 1991-1992.
- LECLERCQ D., La fonction régulatrice de l'évaluation vue sous l'angle de l'implication de l'étudiant, in *Education*, n° 159, Bruxelles, déc. 1976, p. 65-75.
- LECLERCQ D., *La conception des questions à choix multiple*, coll. Education 2000, Labor, Bruxelles, 1986.
- LECLERCQ D., *Qualité des questions et signification des scores, avec application aux QCM*, coll. Education 2000, Labor, Bruxelles, 1987.
- LECLERCQ, D., *Auto-évaluation et connaissance partielle*, Université de Liège - Service de technologie de l'Education, notes provisoires, Liège, 1990.
- LECLERCQ D., DENIS, B., BOXUS, E., GILLES, J.L., *Double Check : Un concept d'évaluation interactive permettant de distinguer les capacités de compréhension et d'analyse*, Colloque Itinérant sur la Pédagogie Universitaire entre les Universités du Québec et de la Communauté française de Belgique, Liège, 25/10/94.
- LECLERCQ, D., Des banques de QCM-SGI (interactives ou non) au service de l'évaluation, in *Probio-Revue*, vol. 17, n°1-2, Louvain-la-Neuve, 1994c, p. 19-28.
- LECLERCQ D. et BRUNO J., *Item Banking : Interactive Testing and Self-Assessment*, NATO ASI Series, Computer and science, Serie F : vol. 112, Berlin, 1993.
- LECLERCQ D., BOXUS E., DE BROGNIEZ P., WUIDAR H., LAMBERT F., The TASTE Approach : General Implicit Solutions in Multiple Choice Questions (MCQs), Open Books exams and Interactive Testing, in LECLERCQ D. et BRUNO J., *Item Banking : Interactive Testing and Self-Assessment*, NATO ASI Series, Berlin, 1993, p. 210-232.
- LECLERCQ, D. et DENIS, B., *Multimédias et apprentissages*, Journée EUROTECNET, Liège, avril 1994.
- LECLERCQ, D. et GILLES, J-L, *GUESS : Un logiciel pour s'entraîner à l'auto-estimation de sa compétence cognitive*, Colloque " QCM et Questionnaires Fermés", Marne-La-Valée, ESIEE, décembre 1994.
- NIZET A.C., *Les QCM SGI et les Degrés de Certitude au Service d'une Evaluation Interactive et de l'Apprentissage*, travail de fin d'études, Université de Liège, Faculté de Psychologie et des Sciences de l'Education, 1992-1993.
- NYHAN B., *Promouvoir l'aptitude à l'auto-formation*, Eurotecnet, Presses interuniversitaires européennes, Bruxelles, 1991.