

LE PROJET DE PLATE-FORME ELECTRONIQUE DE CONSTRUCTION ET DE GESTION QUALITE DE TESTS STANDARDISES *e-C&QCST*¹

JEAN-LUC GILLES¹, SYLVIE-ANNE PIETTE², PASCAL DETROZ³ ET MARC PIRSON⁴

- ¹ Chargé de recherches, directeur du Système Méthodologique d'Aide à la Réalisation de Tests (SMART) de l'Université de Liège – Belgique, chef de projet de la recherche *e-C&QCST*
- ² Coordinatrice des recherches en gestion des compétences – Unité d'Enseignement et de Recherches en Management – HEC Liège - Belgique
- ³ Coordinateur des évaluations des enseignements au Système Méthodologique d'Aide à la Réalisation de Tests (SMART) de l'Université de Liège - Belgique
- ⁴ Coordinateur de l'équipe informatique du projet *e-C&QCST* au Système Méthodologique d'Aide à la Réalisation de Tests (SMART) de l'Université de Liège - Belgique

Cette recherche est commanditée par la *Direction Générale des Technologies de la Recherche et de l'Energie (DGTRE)* de la Région Wallonne (Belgique).

Le *Système Méthodologique d'Aide à la Réalisation de Tests (SMART)* de l'Université de Liège coordonne la recherche menée en partenariat avec *l'Unité d'Enseignement et de Recherche en Ressources Humaines (UERRH)* des Hautes Etudes Commerciales de Liège.

Résumé :

La recherche en docimologie (la science qui a pour objet l'étude systématique des examens) a montré que le testing standardisé est une forme d'évaluation performante des compétences des apprenants à condition de réunir les qualités de validité, fidélité, sensibilité, diagnosticité, équité, praticabilité, communicabilité et authenticité. Lorsque c'est le cas, ces tests standardisés sont appréciés à la fois par les formateurs, les formés et les instances dirigeantes des organismes de formation.

Actuellement la construction et la gestion de la qualité des évaluations certificatives standardisées préoccupent le monde de l'enseignement supérieur confronté à une massification des effectifs ainsi que le monde de la formation professionnelle et des ressources humaines amené à réaliser régulièrement des bilans de compétences acquises en dehors du cadre scolaire. Par ailleurs, ces préoccupations relatives à la qualité des tests sont aussi à mettre en relation avec un mouvement mondial d'introduction du management de la qualité dans la sphère des activités d'enseignement et de formation.

Notre recherche vise à faciliter la création de tests standardisés, plus fiables, plus efficaces, plus rapidement mis en œuvre, en vue de satisfaire au mieux les exigences actuelles du monde de l'enseignement/formation/ressources humaines. Notre objectif est d'aboutir à une plate-forme électronique de construction et de gestion qualité de tests standardisés intitulée *e-C&QCST (electronic Construction & Quality Control in Standardized Testing)* réalisée à l'aide de logiciels libres et dans une philosophie « *Open Source* ». Les interfaces de notre logiciel intégré seront traduites en plusieurs langues et conçues en vue de fonctionner via le web, ce qui facilitera son utilisation à l'échelle mondiale.

¹ Electronic Construction & Quality Control in Standardized Testing.

1. INTRODUCTION

Le testing standardisé est une forme d'évaluation des compétences des apprenants qui, lorsqu'elle réunit les qualités de validité, fidélité, sensibilité, diagnosticité, équité, praticabilité, communicabilité et authenticité (Gilles & Leclercq, 1995 ; Castaigne, Gilles & Hansen, 2001 ; Gilles, 2002 ; Gilles et Lovinfosse, 2004), est appréciée à la fois par les formateurs, les formés et les instances dirigeantes des institutions ou organismes de formation.

Les préoccupations relatives à la qualité des évaluations standardisées des acquis des formés s'insèrent dans un mouvement qui, depuis le début des années '90, touche les institutions belges et européennes (Kaufman & al., 1995 ; Segers & Dochy, 1996). Ce mouvement participe lui-même à une tendance mondiale d'introduction du management de la qualité dans la sphère des activités d'enseignement et de formation (Ramsden, 1991 ; Nightingale & O'Neil, 1994 ; Zink & Schmidt, 1995).

Conscients de ces lignes de force actuelles, nous proposons une recherche visant la mise en place de méthodes et d'outils de gestion de la qualité des épreuves standardisées en vue de satisfaire au mieux les exigences des enseignants/formateurs et des étudiants/formés européens, voire d'autres continents. La dissémination et l'utilisation à l'échelle mondiale des résultats de notre recherche est aujourd'hui possible à l'aide d'interfaces conçues pour fonctionner via le web. C'est l'idée de notre projet *e-C&QCST*² de plate-forme électronique de construction et de gestion qualité de tests standardisés réalisée à l'aide de logiciels libres et dans une philosophie « *Open Source* ».

2. OBJECTIFS DU PROJET

Il s'agit d'abord de créer des procédures docimologiques qui permettent de maîtriser, quantifier et contrôler l'atteinte des objectifs qualité dans le cadre des étapes de construction d'une évaluation standardisée des compétences d'apprenants. Ensuite, ces procédures seront programmées. Les interfaces de la plate-forme *e-C&QCST* devront être accessibles via l'internet et seront réalisées à l'aide de logiciels libres tels que le langage de programmation PHP, le gestionnaire de bases de données MySQL et un serveur web Apache.

L'architecture de notre plate-forme électronique de construction et de gestion qualité de tests standardisés repose sur huit modules intégrés :

1. l'analyse de l'enseignement ou de la formation en vue d'en dégager les points à évaluer ;
2. la mise en forme des épreuves (conception des modalités de questionnement) ;
3. la construction de tests en mode sécurisé (banques de questions et contrôle qualité *a priori*) ;
4. l'entraînement à l'aide de « tests à blanc » en ligne ;
5. l'aide à la mise en œuvre des tests certificatifs ou formatifs permettant entre autre la passation en ligne d'examens (dans des environnements surveillés) ou le couplage avec d'autres dispositifs d'acquisition automatique des réponses tels que la lecture optique de marques ;
6. la correction des épreuves avec contrôles qualité *a posteriori* et possibilités de rectifications ;
7. les feedbacks individualisés et personnalisés accessibles en ligne pour les apprenants ;
8. la macro-régulation avec récolte et analyse en ligne des avis des acteurs.

² *electronic Construction & Quality Control in Standardized Testing*

3. JUSTIFICATION ECONOMIQUE, TECHNOLOGIQUE ET SOCIALE

Pour des raisons d'équité, relayées par des raisons de prudence juridique (afin d'éviter les recours en justice) et de massification des effectifs, les évaluations des acquis des apprenants sont de plus en plus standardisées. Si dans un pays comme les Etats-Unis les tests standardisés connaissent un tel succès, c'est notamment parce qu'ils offrent des garanties de non discrimination raciale (*culture-free test*) et de non discrimination de classe sociale (*class-free test*).

Le choix de la construction d'une plate-forme *e-C&QCST* dans une approche *Open Source* est pertinent car il permet la promotion par la dissémination via le web et via une communauté d'utilisateurs de nos résultats de recherche, de notre expertise docimologique en matière de testing standardisé des compétences ainsi que de nos services y afférents. De plus, c'est au départ de cette communauté d'utilisateurs que la plate-forme prendra vie, c'est à dire sera testée dans une perspective de long terme et continuera d'évoluer en fonction des besoins rencontrés et des avancées technologiques.

Les institutions d'enseignement supérieur et les organismes de formation sont de plus en plus confrontés à la nécessité de certifier les compétences des apprenants via des tests standardisés. Dans ce contexte, un large marché s'offre à la plate-forme *e-C&QCST* et aux sociétés de service qui assureraient un accompagnement et un suivi des utilisateurs.

Les entreprises recrutent, évaluent, recherchent un fonctionnement optimal basé, entre autres, sur la qualité de leurs ressources humaines. Les compétences des travailleurs se voient, de ce fait, évaluées selon des procédures très diversifiées. Ces dernières répondent rarement aux critères de qualité attendus dans le cadre de prises de décisions concernant pourtant un grand nombre de personnes et la rentabilité même de l'entreprise.

Il est communément admis dans les milieux industriels, politiques et académiques qu'une des clés du succès économique d'une région réside dans la qualité de la formation initiale et continuée de ses acteurs. C'est un facteur important à prendre en compte dans le cadre de la reconversion d'une partie du tissu économique actuellement en difficulté en Région wallonne.

La mesure précise et fiable des compétences et acquis des personnes insérées dans les processus de formation nous semble cruciale à divers points de vue :

- En amont de la formation l'évaluation des prérequis permet d'assurer une formation répondant aux besoins réels des apprenants et leur évite ainsi de perdre un temps précieux ;
- Pendant la formation les évaluations formatives permettent des prises de consciences chez les formés à partir de diagnostics de leurs compétences. Ces prises de conscience leur permettront de réguler leurs apprentissages ;
- A la fin de la formation, la garantie d'une certification de qualité permet de connaître les compétences et acquis réels des individus formés en vue de prises de décisions non biaisées ;
- En aval un des indicateurs de l'efficacité de la formation ou de l'enseignement résidera dans la plus-value mesurée au niveau des acquis des personnes formées.

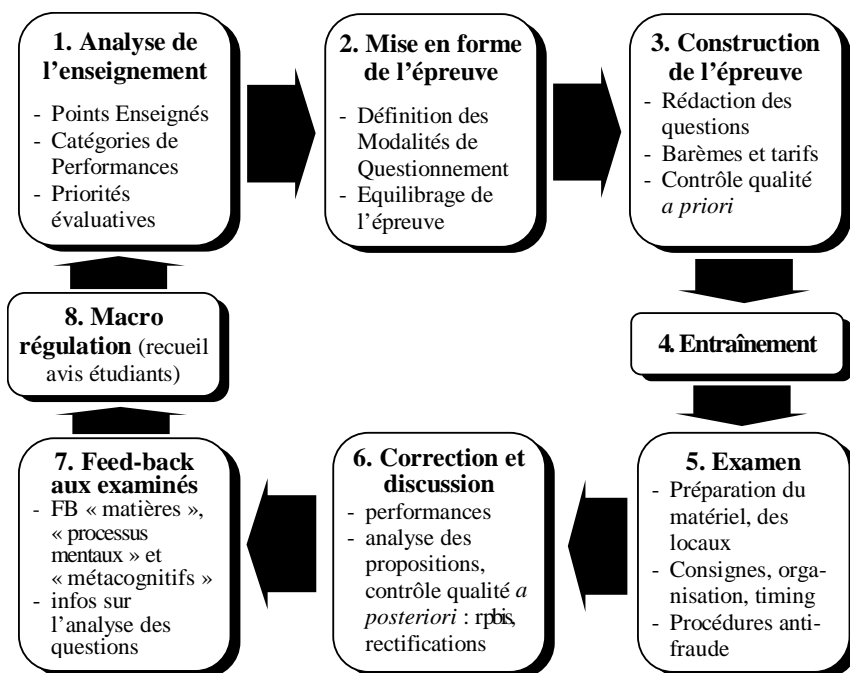
Dès lors, la non qualité dans la construction et la gestion des évaluations des acquis engendre des erreurs dans ces différentes catégories de mesure, entraînant un pilotage inefficace et des surcoûts pour les entreprises, les institutions de formation et la société en général.

4. CONTEXTE SCIENTIFIQUE DU PROJET

La recherche en docimologie et en éduométrie se doit d'offrir aux institutions d'enseignement supérieur / de formation des solutions ayant recours aux technologies de l'information et de la communication en vue de faciliter et de garantir la qualité des épreuves standardisées des acquis des apprenants.

Idéalement, ces solutions doivent s'inscrire dans un cycle SMART en huit étapes de construction et de gestion de la qualité des évaluations (Gilles et Leclercq, 1995 ; Gilles, 2002). La figure ci-contre synthétise et schématise ces huit étapes.

Parallèlement, des garanties sérieuses de qualité ne pourront être offertes qu'à la condition d'envisager des procédures et des dispositifs d'ingénierie docimologiques qui rencontrent une série de critères de qualité.



Voici la liste de ces critères de qualité proposés par le SMART (Gilles, 1998, 2002 ; Gilles et Lovinfosse 2004) :

- la validité : les scores des personnes évaluées doivent refléter ce que le formateur ou l'enseignant veut mesurer, permettre des inférences solides (validité de *construct* ou théorique) et couvrir les aspects importants des apprentissages (validité de contenu) ;
- la fidélité : une copie corrigée et classée dans une catégorie donnée doit bénéficier de la même mention si elle est corrigée par d'autres évaluateurs (concordance inter juge) ou/et à un autre moment par le même correcteur (consistance intra juge) ;
- la sensibilité : la mesure doit être précise, refléter des phénomènes subtils ;
- la diagnosticité : les résultats doivent permettre le diagnostic précis des difficultés d'apprentissage (et idéalement de leurs causes), des processus maîtrisés et de ceux qui ne le sont pas ;
- la praticabilité : la faisabilité en termes de temps, de ressources en personnel et en matériel doit être assurée ;
- l'équité : tous les formés doivent être traités de façon juste, en principe de la même manière ;
- la communicabilité : les informations non confidentielles relatives au déroulement du processus doivent être communiquées et comprises par les personnes engagées dans la passation du test ;
- l'authenticité : les questions doivent être en rapport avec des situations qui ont du sens pour les étudiants, pertinentes par rapport au monde réel et propre à favoriser les transferts dans des contextes et des pratiques qu'on peut rencontrer dans la vie.

Comme nous l'avons rappelé dans l'introduction, notre préoccupation relative à la qualité des évaluations standardisées s'insère dans mouvement mondial d'introduction du management de la qualité dans la sphère des activités de formation. Dès lors, une instrumentation de la gestion de la qualité des tests standardisés à l'aide de la plate-forme *e-C&QCST*, accessible via le web et validée scientifiquement, est susceptible d'intéresser un large marché européen d'institutions d'enseignement supérieur et d'organismes de formation confrontés à la nécessité de produire les preuves de la qualité de leurs tests, notamment lorsqu'il s'agit d'épreuves certificatives et lorsque ces tests sont liés à des audits d'accréditation ou à la gestion des ressources humaines.

5. INSTRUMENTATION DU CONTRÔLE QUALITÉ DES TESTS STANDARDISÉS

Nous proposons une recherche visant la mise en place de méthodes et d'outils de gestion de la qualité des épreuves standardisées en vue de satisfaire au mieux les exigences des enseignants/formateurs et des étudiants/formés. Notre recherche *e-C&QCST* devra aboutir à des procédures de maîtrise et de contrôle des critères de qualité décrits précédemment, ces objectifs devront se concrétiser dans les fonctionnalités des huit modules de la plate-forme. Chaque module étant lié à une étape du cycle de construction et de gestion qualité des épreuves. Nous résumons ci-dessous.

Étapes du cycle	Objectifs de la recherche	Objectifs en terme de « délivrables »
1. Analyse de l'enseignement /de la formation	Amélioration de la validité des tests à l'aide de procédures informatisées permettant d'identifier et de mettre en correspondance les points à évaluer [PE] et les catégories de compétences [CC] attendues chez les étudiants/formés. Ceci dans le but d'établir des priorités évaluatives pour les binômes [PE x CC].	Module de mise en œuvre de la démarche de validation des évaluations des acquis des étudiants/formés. L'interface web permettra de quantifier les niveaux de validité de contenu et de construct des épreuves.
2. Mise en forme de l'épreuve	Amélioration de la validité et de la fidélité des tests à l'aide de procédures informatisées permettant de définir des modalités de questionnement [MQ] adaptées aux catégories de compétences [CC] qui ont été définies à l'étape précédente. Le but étant de construire un schéma d'épreuve à partir de binômes [MQ x CC] clairement identifiés.	Module de construction via une interface web des modalités de questionnement les plus adaptées aux catégories de performances attendues chez les étudiants/formés. L'interface web fournira un tableau de bord des caractéristiques de l'épreuve et offrira des possibilités de simulations.
3. Construction de l'épreuve	Amélioration de la validité, sensibilité et diagnosticité à l'aide de l'identification des trinômes [PE x CC x MQ] construits à partir des binômes [PE x CC] et [CC x MQ] identifiés précédemment. Le but est de formaliser en fonction des priorités évaluatives chaque trinôme sélectionné en un item de testing, et ce, à l'aide de procédures informatisées.	Module sécurisé de gestion par les enseignants/formateurs via le web de leurs banques de questions construites à partir des trinômes [PE x CC x MQ]. L'interface fournira une série d'aides en vue de permettre des contrôles qualité <i>a priori</i> des questions (rappel de règles de rédaction, facilités de validation par des experts, calcul d'indices de lisibilité...).
4. Entraînement aux épreuves	Amélioration de la fidélité, de l'équité et de la communicabilité en automatisant les possibilités de « testing à blanc » de manière à familiariser les étudiants/formés avec les procédures d'examen planifiées aux étapes précédentes.	Module d'entraînement via le web des étudiants/formés aux procédures de testing. Facilités de mise en ligne d'épreuves basées sur les test des années précédentes avec correction automatique et feedbacks détaillés pour chaque utilisateur.

5. Mise en œuvre de l'examen ou de l'épreuve certificative	Amélioration de la praticabilité, de la fidélité, de l'équité et de la communicabilité par des procédures permettant une définition claire, précise et exhaustive des modalités d'utilisation du test et de ses consignes. Création de procédures anti-fraude.	Module d'encodage et de mise en page automatique des originaux des tests à partir de sélections opérées dans la banque de questions sécurisée de l'évaluateur. Des facilités de création de versions brouillées des tests seront disponibles. Procédures et logiciels permettant la passation en ligne d'un examen.
6. Correction du test	Amélioration de la fidélité par des procédures de contrôles qualité des questions ayant recours à des indices classiques (indices de cohérences internes) et nouveaux (indices spectraux). Procédures de rectification des résultats des épreuves suite aux contrôles qualité.	Module de correction en ligne des épreuves permettant à l'évaluateur d'obtenir des vues statistiques des résultats du groupe ainsi qu'un diagnostic de la qualité de ses questions. Possibilités de simulations des impacts des rectifications. Mémorisation de la qualité des questions dans la banque.
7. Feed-back aux étudiants/apprenants	Amélioration de la communicabilité et de la diagnosticité par la création de rétro-informations permettant aux étudiants / formés d'identifier clairement leurs points forts et leurs faiblesses.	Module de mise en ligne automatisée des feed-back individuels accessibles par chaque étudiant / formé via un accès personnalisé.
8. Macro-régulation	Amélioration du processus de construction de l'épreuve par des procédures de récolte et d'analyse des avis des utilisateurs. Propositions d'actions d'amélioration du cycle de construction de l'évaluation à partir des diagnostics.	Module permettant de construire un questionnaire de satisfaction en ligne pour la récolte des avis des utilisateurs à propos de la qualité de l'épreuve. Envoi automatisé des résultats et des analyses vers acteurs du processus de testing.

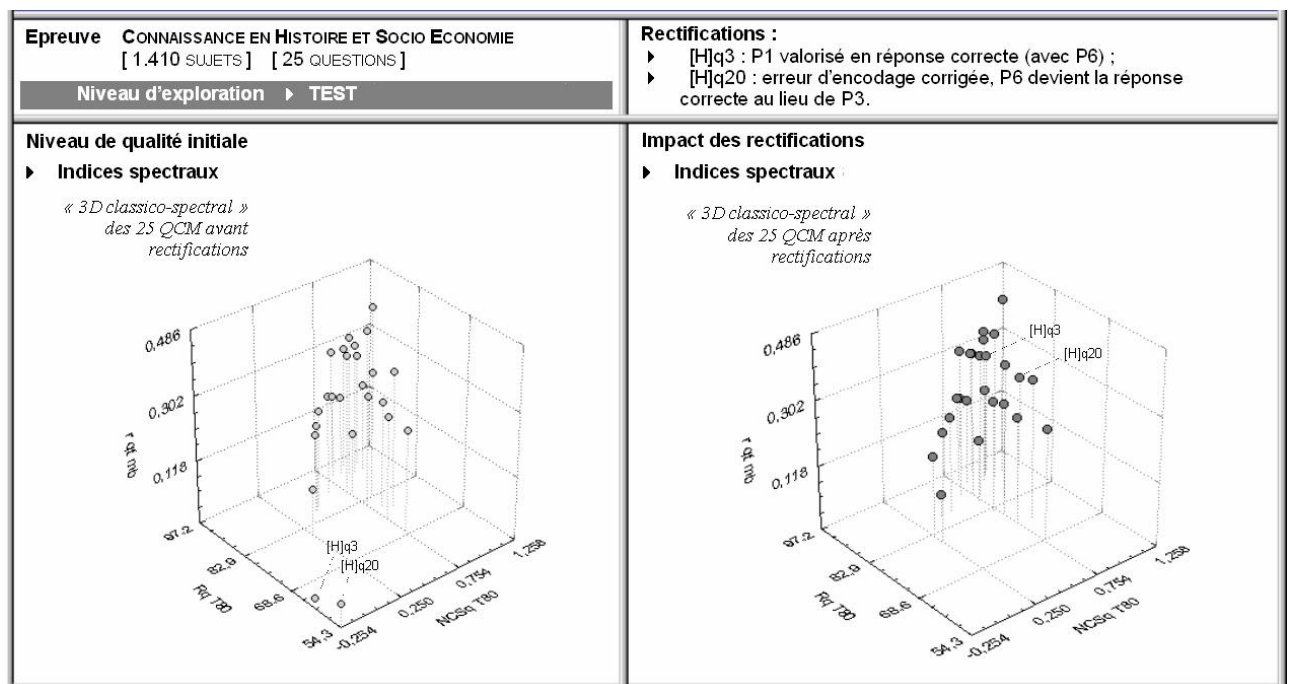
6. EXEMPLE D'INTERFACE DE CONTRÔLE QUALITÉ IMPLÉMENTABLE DANS LE MODULE « CORRECTION DU TEST »

L'exemple ci-dessous est tiré d'une recherche portant sur l'analyse des résultats de tests et sur la mise au point de nouveaux indices de qualité des questions (Gilles, 2002). Le projet d'interface ci-dessous montre les niveaux de qualité « spectrale » des 25 questions à choix multiple d'un test standardisé de connaissances en histoire et socio économie soumis à 1.410 étudiants dans le cadre de l'opération MOHICAN³ (Leclercq, 2003).

L'idée reprise dans le projet d'interface est de permettre une visualisation de l'impact de rectifications envisagées au niveau de deux questions du test : la troisième (notée [H]q3) et la vingtième (notée [H]q20). Le graphique en 3 dimensions de la partie gauche montre que ces deux questions sont situées en bas et en avant du graphique, ce qui signifie qu'elles récoltent des valeurs très faibles à trois indices de qualité. Le graphique de droite montre les positions des deux questions après quelques rectifications dans les résultats. Pour la question [H]q3, après analyse à l'aide d'autres indices et relecture par un expert, un des distracteurs a été valorisé comme réponse correcte. Au niveau de la question [H]q20, c'est une erreur d'encodage de la réponse correcte qui a été corrigée. Le graphique de droite de l'interface montre l'impact après ces rectifications sur les valeurs récoltées par les deux questions aux indices de qualité. On peut ainsi observer une amélioration de la

³ MONitoring Historique de cohortes de CANDidatures universitaires, une recherche-action qui a permis la récolte d'informations sur les acquis des étudiants qui entraînent à l'université en 1999, et ce, dans 9 institutions universitaires de la Communauté française de Belgique.

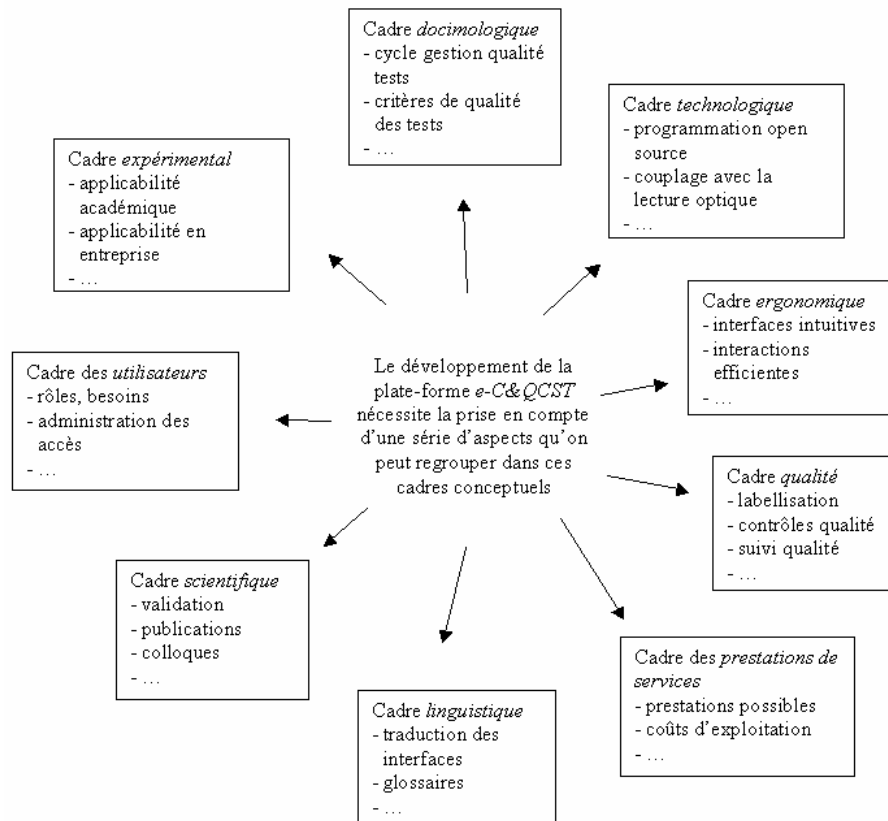
position de [H]q3 et de [H]q20 dont les points se sont déplacés en haut vers le fond du graphique, montrant ainsi une amélioration de la qualité des résultats de ces deux questions.



Selon nous, de tels instruments permettraient aux évaluateurs d'améliorer la maîtrise de la qualité de leurs épreuves.

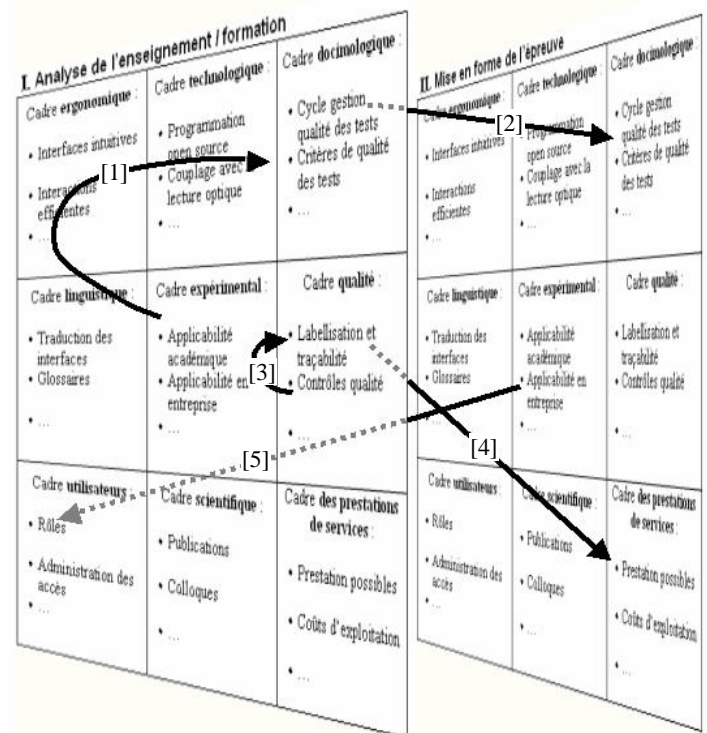
7. CADRES CONCEPTUELS ET FLUX D'INFORMATIONS PROCEDURALES

Notre recherche couvre de multiples aspects que nous devons prendre en compte. Nous regroupons ces facettes dans différents cadres conceptuels : docimologique, technologique, ergonomique, « qualité », prestations de services, linguistique, utilisateurs et expérimentation.



Dans la phase de conception des modules de la plate-forme nous prenons en compte ces différents cadres conceptuels pour chaque étape du cycle. Notre hypothèse de travail est que chacun de ces cadres peut entrer en interaction avec un autre, y compris d'une étape à l'autre.

Par exemple, [1] l'expérimentation d'une procédure prévue à l'étape I de l'analyse de l'enseignement/formation pourrait avoir des répercussions sur la production d'informations docimologiques à cette étape ainsi qu'à l'étape suivante [2]. De même, [3] des procédures de contrôle qualité pourraient avoir des effets en terme de labellisation des tests et [4] d'offre de prestations de services ultérieures liées à la valorisation des résultats de cette recherche. Il se pourrait également qu'une expérimentation à l'étape II de mise en forme de l'épreuve entraîne des modifications à l'étape précédente, notamment au niveau du cadre conceptuel des utilisateurs et plus particulièrement au niveau de la définition de leurs rôles [5]. Ces exemples montrent la complexité du travail actuellement en cours dans la phase de conception de la plate-forme.



8. VALORISATION DE LA RECHERCHE e-C&QCST DANS LA SPHÈRE DES ENTREPRISES

Nous pensons qu'une ou plusieurs sociétés de services ayant recours à notre plate-forme e-C&QCST de construction et de gestion qualité des tests pourraient voir le jour et structurer leur offre de services sur la base suivante :

- 1) L'évaluation des besoins d'entreprises ou d'institutions clientes en matière de procédures d'évaluation des performances ou des acquis des personnes formées
 - Audit du système de testing
 - Aide au choix d'une stratégie de gestion qualité de la construction de tests
 - Sensibilisation et motivation des acteurs intervenant dans la création et la passation de tests
 - Gestion du changement organisationnel
 - Aide à la définition de cahier des charges en matière d'épreuves
- 2) L'accompagnement à l'appropriation de la plate-forme de construction et de gestion qualité des tests au sein des entreprises ou des institutions de formation clientes
 - Audit de l'appropriation des outils par les formateurs / enseignants
 - Analyse coût / efficacité
 - Aide à la préparation des étapes de développement futur du système
 - Formation des utilisateurs

- 3) L'amélioration du système de construction et de gestion qualité des tests au sein des entreprises ou des institutions de formation clientes
 - Bilan docimologique et technique sur l'expérience des évaluations standardisées effectuées
 - Propositions d'améliorations tant docimologiques que techniques
 - Aide à la mise en œuvre des améliorations proposées
- 4) La création de tests pour des tiers en ayant recours à notre plate-forme électronique de construction et de gestion qualité des tests
 - Conception docimologique et technique
 - Développement des tests
 - Evaluation des performances ou des acquis par la mise en œuvre des tests

Nos premiers contacts avec la sphère des entreprises dans le cadre de cette recherche confirment notre hypothèse d'un large marché qui s'offre à la plate-forme *e-C&QCST* et aux sociétés de service qui assureraient un accompagnement et un suivi des utilisateurs. Par ailleurs, l'idée d'un label de qualité pour les évaluations standardisées (Gilles, 1998) sera proposé dans la perspective d'un déploiement d'une docimologie de qualité greffée sur les besoins des évaluateurs et des évalués. Cette perspective de labellisation semble également rencontrer une attente du monde de la formation.

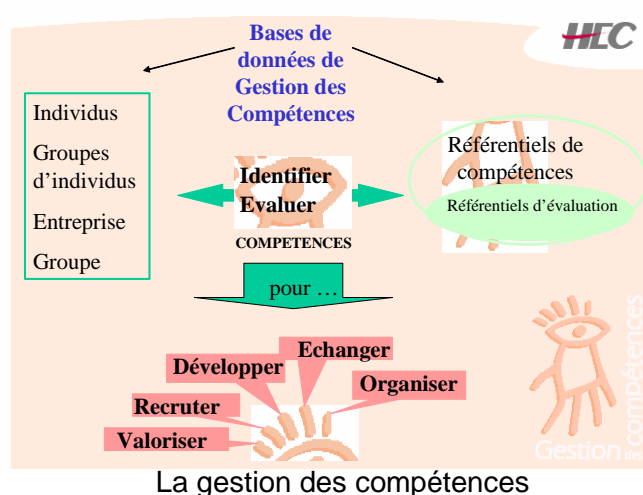
9. L'ÉVALUATION EN ENTREPRISES : ÉTAT DES LIEUX

Les entreprises évaluent ?

Une enquête sur les pratiques de GRH (Berenshot, 2002) réalisée auprès de 103 responsables ressources humaines en Belgique a démontré que la grande majorité des entreprises ont introduit la gestion des compétences dans le but « d'augmenter les performances de leurs collaborateurs et afin d'avoir un aperçu de leur potentiel. » (p.18). Ces entreprises ont également désigné le recrutement et la sélection en première position de leurs priorités essentielles, et la formation en quatrième position.

« Bien qu'il existe des différences dans le classement des priorités futures ; il est frappant d'observer que des thèmes comme la qualité des managers, les programmes de formation et de développement et la politique d'évaluation restent en pôle position au fil du temps (enquêtes de 1990, 1995 et 2002). Ces thèmes font donc incontestablement partie de ce que l'on pourrait qualifier de « quête du Graal » dans le domaine de la gestion des ressources humaines. » (p.8)

La gestion des compétences est le lien permettant la mise en cohérence de tous les aspects de la GRH. Elle est indissociable de l'évaluation : l'évaluation annuelle, la gestion des carrières, la reconnaissance, la rémunération, l'organisation du travail, la mobilité, ... » (Meignant, 2003, p.12).

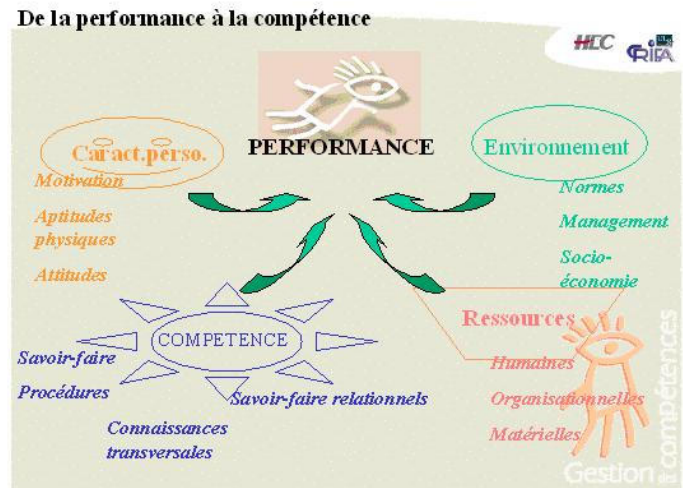


Qu'évaluent les entreprises ?

« La notion de compétence est, en entreprise, rattachée à celles de fonction et de tâches professionnelles. La compétence s'exprime au travers des performances observables et évaluables, pour autant qu'on recoure à un certain nombre de critères : indicateurs de réussite, indicateurs diagnostiques ou pronostiques, selon les enjeux de l'évaluation. ». (Piette, Orban, 2003)

Le schéma ci-contre montre les différentes composantes intervenant dans la performance.

L'évaluation d'une performance nécessite la description de critères quantitatifs et qualitatifs associé à un type d'action. Il s'agit donc de décrire le plus précisément possible les résultats attendus. Les entreprises sont surtout intéressées par ce type d'évaluation mais mélangent régulièrement les résultats attendus et les composantes sous-jacentes dans les mêmes évaluations. Les mêmes éléments sont donc évalués plusieurs fois, soit directement, soit au travers les actions qu'ils permettent d'effectuer. Par ailleurs, les résultats de ces évaluations sont peu utilisés dans une logique de développement. Ils servent surtout à définir un niveau de rémunération (prime liée à la performance) ou à établir une sorte de « reconnaissance » versus « sanction ».



Quelles sont les méthodes d'évaluation utilisées ?

Deux types d'enjeux qui conduisent à des méthodes différentes :

- L'évaluation a pour but d'identifier des compétences dans le contexte d'un positionnement salarial ou d'une promotion, d'un changement organisationnel, d'un développement de carrière,...

Dans l'enquête susmentionnée, il ressort que les entretiens (de fonctionnement, d'évaluation de prestation, de suivi et coaching) sont les méthodes de prise de décisions les plus employées. Nous sommes face à des évaluations globales complexes qui tentent d'identifier, d'après un (auto- ou hétéro-) **descriptif des actions et résultats**, un niveau de compétences des personnes. *"Les entreprises s'intéressent avant tout au savoir-faire professionnel observable en situation de travail. Ce savoir-faire n'est pas structuré par les savoirs académiques, mais par les situations de travail qui les mobilisent, et qui, le plus souvent, relèvent d'une combinatoire entre connaissances, savoir-faire, expérience professionnelles relevant de champs académiques différents"*. (Meignant, 2003, p.12)

- L'évaluation a pour but d'identifier des pré-acquis ou les résultats d'un développement de compétences en relation avec les objectifs économiques de l'entreprise et les mutations du travail. (Aubret J., Gilbert P., Pigeyre F., 2002, p.116) ou encore d'identifier les compétences dont un individu doit faire **preuve** dans le cadre d'un recrutement ou d'une mobilité.

Les outils d'évaluation de la formation sont davantage centrés sur **un comportement, une action à produire au moment de l'évaluation**. Des tests sont alors utilisés qui prennent différentes formes (tests écrits, jeu de rôle, ...). Ces procédures sont, dans le cas d'enjeu de mobilité ou recrutement, le plus souvent combinées avec des descriptifs de comportements et/ou résultats (avis du supérieur hiérarchique, Curriculum vitae, références).

10. INTÉRÊT DE L'ÉVALUATION STANDARDISÉE ACCOMPAGNÉE EN ENTREPRISE

La garantie de qualité de l'évaluation

L'application de critères de qualité est prioritaire dans certains contextes d'évaluation :

« Nous constatons que les sentiments d'équité sont surtout importants dans le cas de la mise en place d'évaluation "décisionnelle". En effet, si l'évaluation doit permettre la promotion par exemple, elle doit recourir à des critères et des procédures qui seront, au moins subjectivement, jugées comme équitables, basées sur des critères clairement établis et identiques pour tous. Dans le cadre d'une évaluation formative, l'objectif étant avant tout de favoriser des changements positifs, il peut laisser plus de place à la subjectivité, à l'auto-évaluation, aux normes spécifiques d'un groupe, d'une équipe, d'un département. » (Piette, Orban, 2003)

Aide à la construction de tests

Les entreprises n'ont pas souvent de service de formation interne, ni de personnes dans leur département RH ayant des compétences docimologiques pointues. Or, nous venons de le voir ci-dessus, les procédures d'évaluation de type décisionnel mériteraient certainement la mise en place d'un ensemble de procédures d'évaluation dont la qualité globale devrait être assurée. La plateforme, en proposant une réflexion et une procédure étape par étape, devrait permettre aux responsables d'évaluation de créer, en étant guidés, des épreuves standardisées de qualité. Ces épreuves standardisées ne répondront pas à l'ensemble de leurs besoins d'évaluation mais feront partie de l'arsenal des procédures mises en place.

La plate-forme pourra évaluer plus complètement et plus rapidement :

- les connaissances générales : connaissances théoriques pures sans aucun lien avec l'action. Elles ne permettent pas de gouverner directement l'action mais au contraire d'en constituer un éloignement préalable, qui permet la description de l'action et sa compréhension, ainsi que la création de nouvelles procédures pour faire face à des situations problèmes ou originales.
- les connaissances spécifiques à l'environnement de travail : connaissances théoriques en lien avec l'action
- les savoir-faire techniques et procédures : ce sont les comportements que le travailleur met en place pour réaliser les actions prévues dans son environnement spécifique de travail. Ils sont évaluables sur le terrain par l'observation ou l'entretien mais, dans certains cas, il serait justifié de les évaluer sur des simulateurs ou lors d'analyse de cas (entre autre pour des jeunes travailleurs et des actions ayant des conséquences potentielles graves).

Le coût lié à l'évaluation telle qu'elle est pratiquée actuellement pourra être réduit, à terme : une fois les tests créés, ils pourront être réutilisés et analysés sans frais. Le retour sur l'investissement en temps de la première épreuve sera très rapide.

11. MÉTHODOLOGIE DE TRAVAIL AVEC LES ENTREPRISES

Le but du travail avec les entreprises est d'adapter, dans la mesure du possible, le processus proposé par notre plate-forme à des pratiques existantes tout en donnant des possibilités supplémentaires liées au processus qualité.

Echantillon d'entreprises

Suite à l'analyse de l'existant en termes d'évaluation en entreprises, nous avons choisi les caractéristiques d'un échantillon à approcher afin d'analyser leurs pratiques de testing.

Quelques exemples de caractéristiques :

Comme nous l'avons dit, nous pensons que la réalisation d'évaluations nécessite un personnel formé en docimologie : la variable « taille de l'entreprise » a un impact probable sur la politique de gestion des compétences, qui elle-même en a un sur l'évaluation et sur l'existence ou non d'un service formation.

Les enjeux de l'évaluation ont aussi des conséquences sur l'importance des critères de qualité et ces enjeux sont parfois liés à la caractéristique privée / publique des organisations et au secteur (service, production).

- Relevé de « use-case » existants
Au départ d'entretiens semi-structurés et de focus group, nous allons relever les pratiques en décrivant pas à pas les actions menées par les évaluateurs lors de la construction, passation et analyse des évaluations standardisées qu'ils utilisent déjà. Une série de questions complémentaires sont prévues afin de cerner l'ensemble des problématiques telles que : les acteurs en jeu, la dimension temporelle, l'importance de la qualité du test, l'insertion du test dans d'autres processus globaux, ...
- Comparaison avec « use-case » prévus
Ces analyses vont être confrontées au processus prévu dans la plate-forme et à l'analyse menée en parallèle auprès d'enseignants.
- Analyse du possible
La discussion par rapport aux possibilités d'adaptation du processus prévu se réalisera avec l'ensemble de l'équipe, car il faut faire preuve de réalisme technologique dans le cadre du temps et du budget qui nous sont impartis. Plusieurs aller-retour entre les caractéristiques prévues et celles demandées par le terrain seront organisés.
- Apport au cahier des charges
Le cahier des charges pourra alors être rédigé puis programmé.
- Expérimentation
Il restera alors à expérimenter les différents modules auprès des entreprises et à affiner la programmation en fonction des résultats de terrain.

12. CONCLUSIONS

Rappelons que notre recherche vise à faciliter la création de tests standardisés, plus fiables, plus efficaces, plus rapidement mis en œuvre, en vue de satisfaire au mieux les exigences actuelles du monde de l'enseignement, de la formation et des ressources humaines en entreprise. Notre objectif est d'aboutir à une plate-forme électronique de construction et de gestion qualité de tests standardisés intitulée *e-C&QCST* (*electronic Construction & Quality Control in Standardized Testing*) réalisée à l'aide de logiciels libres et dans une philosophie « *Open Source* ». Dans l'optique d'une diffusion large, les interfaces de notre logiciel intégré seront traduites en plusieurs langues et conçues en vue de fonctionner via le web, ce qui facilitera son utilisation à l'échelle mondiale, notamment dans les institutions d'enseignement.

Dans les derniers points de cet article nous avons envisagé plus en détail l'utilisation de notre plate-forme dans le contexte des formations en entreprises et de la gestion des ressources humaines. Nos contacts avec la sphère des entreprises confirment notre hypothèse d'un large marché pour ce type de plate-forme. Nous y voyons des opportunités de création d'emploi dans le cadre de la reconversion d'une partie du tissu économique actuellement en difficulté en Région wallonne.

BIBLIOGRAPHIE

- Aubret, J., Gilbert, P., Pigeyre, F. (2002). Management des compétences, Dunod, Paris,
- Castaigne, J.-L., Gilles, J.-L. et Hansen, C. (2001). Application du cycle gestion qualité SMART des tests pédagogiques au cours d'Obstétrique et de Pathologie de la Reproduction des ruminants, équidés et porcins, Communication au 18^{ème} Congrès de l'Association Internationale de Pédagogie Universitaire (AIPU), *Les stratégies de réussite dans l'enseignement supérieur*, Dakar, 5-7 avril 2001.
- Gilles, J.-L. & Leclercq, D. (1995). Procédures d'évaluation adaptées à des grands groupes d'étudiants universitaires - Enjeux et solutions pratiquées à la FAPSE-ULG, in *Actes du Symposium International sur la Rénovation Didactique en Biologie*. Tunis : Université de Tunis.
- Gilles, J.-L. (1998). *Introduction d'un système qualité dans le processus de réalisation des épreuves certificatives de l'enseignement supérieur*. Communication à la demi-journée de sensibilisation à la Qualité – Université de Liège, 5 octobre 1998.
- Gilles, J.-L. (2002). Qualité spectrale des tests standardisés universitaires – Mise au point d'indices éducatifs d'analyse de la qualité spectrale des évaluations des acquis des étudiants universitaires et application aux épreuves MOHICAN check up '99, thèse de doctorat en Sciences de l'éducation. Liège : Université de Liège, Faculté de psychologie et des sciences de l'éducation de l'université de Liège.
- Gilles, J.-L. (2004). Feedbacks diagnostiques via internet : un système individualisé de communication des niveaux de performances des étudiants dans le cadre d'évaluations standardisées, in *Actes du 3^{ème} Congrès des chercheurs en éducation « (Re)trouver le plaisir d'enseigner et d'apprendre – Construire savoirs et compétences »*, Bruxelles : Administration générale de l'Enseignement et de la recherche scientifique (à paraître).
- Gilles, J.-L. et Lovinfosse, V. (2004). Utilisation du cycle SMART de gestion qualité des évaluations standardisées dans le contexte d'une Haute Ecole : regard critique en termes de validité, fidélité, sensibilité des mesures, diagnosticité, praticabilité, équité, communicabilité et authenticité, Communication au *XIV^{ème} congrès de l'Association Mondiale des Sciences de l'Education (AMSE)*. Santiago du Chili, 10-14 mai 2004.
- Jans, V. & Leclercq, D. (1999). Mesurer l'effet de l'apprentissage à l'aide de l'analyse spectrale des performances, in C. Depover & B. Noel (Ed.), *Evaluation des compétences et des processus cognitifs*. Bruxelles : De Boeck, pp. 303-317
- Kauffman, Ch., Dupont, P., Philippart, A. (1995). *Projet pilote européen pour l'évaluation de la qualité dans l'enseignement supérieur*. Bruxelles : Ministère de l'Education, de la Recherche et de la Formation
- Le Boterf, G. (1997). De la compétence à la navigation professionnelle, Les Editions d'Organisation, Paris,.
- Leclercq, D. (2003). Diagnostic cognitif et métacognitif au seuil de l'université – Le projet MOHICAN mené par les 9 universités de la Communauté française Wallonie Bruxelles. Liège : Les éditions de l'Université de Liège.
- Meignant, A. (2003). Compétence, reconnaissance, création de savoir : trois défis pour le management de la formation, *Personnel et gestion*, éd. Kluwer, n°7.
- Piette S-A, Orban, M. (2003). L'évaluation des compétences en milieu de travail, leçons de quelques expériences de terrain, Communication au colloque ADMEE, Université de Liège, septembre 2003

- Nightingale, P. & O'Neil, M. (1994). *Achieving quality learning in higher education*. London : Kogan Page
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education : the course experience questionnaire, *Studies in Higher Education*, 16, 129-150.
- Segers, M. & Dochy, F. (1996). Quality assurance in higher education : theoretical considerations and empirical evidences, *Studies in Educational Evaluation*, 1996, 22 (2), 115-137
- Zink, K.J. & Schmidt, A. (1995). Measuring universities against the european quality award criteria. *Total quality management*, 6 (5-6), 547-561.
- Berenschot Belgium (2002). La pratique de GRH, enquête, http://users.skynet.be/adp-liege/Colloque_03/Rapport_GRH_en_pratique.pdf