

XXVIII^e Congrès de l'APLIUT

« Normes et normalisation »

02 juin 2006 - Table ronde sur les certifications

L'objectif de cette Table ronde, consacrée aux certifications, était d'offrir une vue aussi complète que possible d'un certain nombre de « produits » disponibles sur le marché, pour faciliter la comparaison entre eux (et avec d'autres produits), tout en cernant les besoins institutionnels des IUT. Pendant près de deux heures, les congressistes ont d'abord réfléchi sur les notions mêmes d'évaluation et de certification, avant d'écouter la présentation de six « traductions » concrètes différentes, et enfin de débattre de questions fondamentales.

Le débat s'était, en fait, ouvert la veille : pour pouvoir échanger sur des concepts tels que « note », « évaluation », « certification », « compétence partielle », « validité », « fiabilité », la modératrice, Gail Taillefer (Université de Toulouse-I), avait distribué un questionnaire anonyme, de type ouvert, pour recenser le sens donné par les participants à ces mots-clés (cf. Annexe 1). Une quarantaine de personnes a répondu—environ le tiers des inscrits—, certains après en avoir discuté en atelier de spécialité, d'autres de manière indépendante.

Un dépouillement rapide¹, projeté sur Powerpoint, a révélé des divergences importantes, suscitant des réactions de surprise : certaines de ces notions, semble-t-il, ont rarement fait l'objet de discussion entre collègues... Les réponses étaient plutôt divergentes, par exemple, sur la notion d'évaluation : si 24 % des répondants conçoivent ce terme comme un processus global, 32 % y voient de manière plus pointue un système d'attribution de notes ; 35 % considèrent un examen comme un bilan, alors que 48 % lui donnent le sens d'une évaluation de macro-compétences ; 22 % comprennent la « validité » d'un test, examen, diplôme... comme l'adéquation entre les notions évaluées et les questions posées, 31 % y voient la pertinence et la cohérence d'un test reconnues par une communauté, mais 41 % l'interprètent au sens plus restreint d'une période de temps (tel test—ou le score d'un apprenant ?—n'est « valide » que pendant X années).

L'échange aurait sans doute pu se poursuivre pendant un certain temps, mais pour pouvoir passer à la présentation des certifications, G. Taillefer a proposé des définitions opérationnelles d'après la littérature, notamment le *Cadre européen commun de référence pour les langues* (cf. Annexe 2).

Pour la deuxième partie de cette Table ronde, six certifications avaient été choisies pour représenter la diversité de langues, de formats, d'origines, de familiarité du produit et... de philosophie : par ordre alphabétique, le BULATS (Business language testing service), le CLES (Certificat de compétence en langues pour

¹ Une analyse détaillée des réponses a été faite, par la suite, par Dana Léon-Henri, doctorante à Paris-IV Sorbonne.

l'enseignement supérieur), le DCL (Diplôme de compétences en langue), le CELI et le CIC (Certificat de connaissance de la langue italienne, Certificat d'italien commercial), les certificats du Goethe Institut, le TOEIC (Test of English for international communication). Leurs représentants devaient eux-mêmes « rendre leur copie » pour la date limite, en remplissant une grille décrivant leur certification selon les mots-clés évoqués au préalable. Chaque présentation était limitée à 10 minutes, gageure qui a été tenue dans l'ensemble.

La grille remplie (cf. Annexe 3) résume les ressemblances et les différences évidentes entre les différents produits, notamment sur le plan des langues certifiées, des compétences évaluées, de formats adoptés, de la faisabilité, du coût humain et matériel, de la conformité avec le référentiel qu'est le Cadre européen commun, et surtout, de l'esprit de la certification : « vertical », capable de certifier un *continuum* de niveaux de compétence, ou « horizontal », ciblant un niveau donné. Le rythme ayant été très soutenu, l'attention de la salle était vive, et beaucoup de questions ont été soulevées, notamment pour comparer les trois certifications les plus connues (DCL, CLES, TOEIC).

Pour clore sur ces exemples précis, et en même temps porter le débat plus loin, G. Taillefer est revenue sur les réponses des congressistes à trois questions fondamentales dans tout choix de certification : quel rôle celle-ci doit-elle jouer ? Quel effet rétroactif (négatif, positif) pourrait-elle avoir dans la salle de classe ? Et enfin, dans la réalité pédagogique et politique des IUT, quelles difficultés y aurait-il pour intégrer une certification en langue au DUT des différents départements ?

A ce niveau, la réaction des congressistes a été beaucoup plus homogène, et visiblement très réfléchie, exprimant fortement le souci de pouvoir offrir à nos étudiants un bagage linguistique solide—et approprié—pour le contexte professionnel dans lequel ils auront à trouver leur place. Ainsi, la certification sert, pour 15 % des répondants, à avoir une idée réaliste de son niveau en langue, mais surtout, pour 30 %, à favoriser une intégration européenne et à rendre lisible la compétence en langue pour des non linguistes (employeurs, par exemple...). Sur le plan pédagogique, la certification est vue de manière positive par un pourcentage important des répondants : 41 % la considèrent comme une « carotte » motivante, 45 % la trouvent pertinente par rapport aux besoins professionnels. En revanche—et heureusement—le même pourcentage d'enseignants considèrent que la certification est négative si elle oriente le contenu des cours. Autrement dit en anglais, si « *testing to teach* » (évaluer, de manière formative, pour mieux cibler l'enseignement) est une vertu, « *teaching to test* » (enseigner en vue d'un examen, pas nécessairement cohérent, pertinent ou valide) n'en est pas une. Enfin, deux principales difficultés ont été signalées dans toute mise en place d'une certification : le manque de clarté dans la définition de l'objectif de ce nouvel outil et le manque d'heures pour y arriver, ainsi que le financement de l'ensemble (coûts directs supportés par qui ? coûts indirects en termes de service, d'attentes, de programmes...).

Cette Table ronde semble avoir atteint son objectif, qui était d'élargir la réflexion sur la certification en soulevant plusieurs questions, toutes interdépendantes ; les réponses émergeront... Parmi les interrogations fortes, citons l'importance d'une validation lisible et reconnue au niveau européen, et en corollaire la place des certifications purement nationales ; la question des moyens humains et matériels ; l'importance capitale du lien intrinsèque entre enseignement et certification ; l'avantage d'une certification « verticale » ou d'un modèle « horizontal », surtout faisant suite aux nouvelles certifications annoncées pour le secondaire ; l'analyse des besoins professionnels en langues et la reconnaissance d'une réalité exigeant des niveaux de performance différents dans les différentes compétences (lire, écrire, parler, comprendre) et les domaines de spécialité ; la place de la deuxième langue étrangère... ; et, enfin, la nécessaire formation professionnelle des enseignants chargés de mettre en œuvre toute certification, quelle qu'elle soit (l'appropriation du Cadre...).

Gail Taillefer, Université Toulouse I

Annexe 1²

Questionnaire distribué aux participants la veille de la table ronde

Veuillez, s'il vous plaît, nous faire part de vos perceptions : déposez cette fiche dans l'urne située l'Accueil à la fin de la journée du jeudi. Merci beaucoup !

Parler de la certification (obligatoire déjà pour certains, à venir peut-être pour d'autres) et des questions annexes suppose qu'on donne le même sens aux mots clés. Qu'entendez-vous par :

- note par rapport à évaluation par rapport à certification ?
 - contrôle par rapport à examen par rapport à test, certificat, diplôme ?
 - compétence partielle ?
 - la validité d'un test, examen, diplôme... ?
 - la fiabilité d'un test, examen, diplôme... ?
1. Quel rôle la certification doit-elle jouer ?
 2. Quel effet rétroactif dans la salle de classe (négatif, positif...) y voyez-vous ?
 3. Quelle(s) difficulté(s) pour ajouter une certification en langue au DUT dans votre département ?

² Le format des annexes 1, 2 et 3 a été modifié pour des raisons de mise en page de la revue (NDLR).

Annexe 2. Définitions³

- *Note* : résultat quantitatif ou qualitatif de performance à un moment donné ; outil de comparaison, de communication sociale ; « étiquette » (*Le Monde de l'éducation* : 36-38) ; « La signification des notes est intériorisée par les enseignants mais rarement définie » (*Cadre* : 38).
- *Evaluation* : recueil systématique d'informations (quantitatives ou qualitatives) pour pouvoir interpréter et prendre des décisions (Bachman : 22) ; plusieurs types (*Cadre* : 136, Tableau 7) : évaluer ce qu'on connaît ? ce qu'on ignore ? etc.). L'interprétation d'un score comme indication d'un niveau donné de compétence linguistique implique le fait de pouvoir inférer l'étendue de cette compétence à partir d'une performance linguistique observée, d'où nécessité de spécifier la compétence, de déterminer comment un score quantifie la performance réelle et potentielle, et de déterminer jusqu'à quel point la performance est influencée par le test lui-même (Bachman).
- *Certification* : attestation de savoir-faire caractérisant la mise en œuvre de la compétence en langue (lire, écrire, comprendre, parler, interagir) : « critérié » (calibré), étalonné.
- *Contrôle* : forme d'évaluation (*Cadre* : 132, 136).
- *Examen* : épreuve validant un apprentissage.
- *Test* : forme d'évaluation (*Cadre* : 136), procédure destinée à susciter un comportement spécifique à partir duquel on peut inférer certaines compétences d'un individu (Bachman : 20-22).
- *Certificat* : document attestant la réussite à un examen (*Petit Robert*).
- *Diplôme* : document qui confère et atteste un titre, un grade (universitaire, *Petit Robert*).
- *Compétence* partielle : activités de communication langagières (réception : lire, écouter ; production : écrire, parler ; interaction), tâches spécifiques dans un domaine, compétences générales (socio-linguistique, pragmatique, interculturelle, stratégique, *Cadre* : 105).
- *Validité* : dans quelle mesure ce qui est effectivement testé (le « *construct* », concept hypothétique) dans un contexte donné est ce qui doit être évalué, et dans quelle mesure l'information recueillie donne une image exacte de la compétence des candidats en question (*Cadre* : 132).
- Que tester ? Comment ? Chez qui ? Avec quoi ? Dans quel contexte ? (Bachman : ch. 7).
- *Validité du « construct »* (concept hypothétique) : jusqu'à quel point un test mesure uniquement la capacité considérée, de façon à pouvoir faire des inférences valides à partir du score ; validité convergente : corrélation entre 2 tests destinés à évaluer un même concept hypothétique.
- *validité du contenu* : jusqu'à quel point les tâches à faire sont pertinentes
- *validité « apparente »* : jusqu'à quel point un test paraît authentique et semble s'appliquer à ce que l'on cherche à mesurer (pour qu'un sujet accepte de s'y soumettre)

³ Voir les références bibliographiques ci-dessous.

- *validité « critériée »* : rapport entre un critérium donné (contenu ou niveau de performance spécifié) et un score censé indiquer la capacité en question.
- > *concurrente* : examiner les scores de sujets à des niveaux différents de capacité linguistique, ou d'un même sujet à différentes mesures d'une même capacité,
- > *prédictive*: déterminer jusqu'à quel point un score peut prédire un comportement linguistique futur dans une situation de communication hors-test.
- *Fiabilité* : mesure selon laquelle on retrouvera le même classement des candidats dans deux passations des mêmes épreuves ; exactitude des décisions prises en fonction d'une norme (*Cadre* : 132, Bachman : ch. 6).
- *fiabilité interne* : rapport des scores aux différentes parties d'un même test, ainsi que l'accord de deux correcteurs ou d'un seul correcteur avec lui-même, stabilité : consistance des scores dans le temps lorsque le même test est administré deux fois
- *équivalence* : consistance des scores à différentes formes d'un test (« tests parallèles »).
- Si la validité représente l'accord entre deux mesures différentes d'une même capacité chez un sujet – la qualité de l'interprétation et de l'utilisation d'un test –, la fiabilité représente l'accord entre deux mesures similaires (« tests parallèles ») – la qualité des scores eux-mêmes (Bachman : 240).

Références bibliographiques

- Cadre européen commun de référence*. 2001. Strasbourg : Conseil de l'Europe <<http://culture2.coe.int/portfolio>>.
- Bachman, L. 1990. *Fundamental Considerations in Language Testing*. Oxford : Oxford University Press.
- Monde de l'éducation*, février 2006, 20-40.
- Alderson, J. C., C. Clapham & D. Wall. 1995. *Language Test Construction and Evaluation*. Cambridge : Cambridge University Press.
- Bachman, L. & A. Palmer. 1996. *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford : Oxford University Press.
- Lussier, D. 1992. *Évaluer les apprentissages dans une approche communicative*. Paris : Col. F. Hachette.
- Monnerie-Goarin A. & R. Lescure (dir.). 1993. «Évaluation et certifications en langue étrangère ». *Le Français dans le monde* (Recherches et applications, numéro spécial), août-septembre.
- Tagliante C. (dir.). 1991. *L'évaluation*. Paris : CLE International.

Annexe 3 - Six Certifications

Titre	TOEIC	DCL	CLES	Goethe	CIC-CELI	BULATS
Genre	test	diplôme MENESR	certificat	certificat (B2, à partir de 2007)	certificats	test
Origine	Etats-Unis	France	France	Allemagne	Italie (ALTE)	Royaume-Uni, Espagne, Allemagne, France (ALTE)
Langue/s	anglais	anglais, espagnol, allemand, italien	anglais, espagnol, allemand, italien, portugais	allemand	italien	anglais, espagnol, allemand, français
But	professionnel	professionnel (être opérationnel)	académique, professionnel	langue générale	CIC: profession-nel (2 niveaux), CELI: académique (5 niveaux)	professionnel, enseignement supérieur
Résultat (quant./qual/)	10 - 999	5 degrés	qualitatif, par niveau (B1, B2.....)	admissibilité écrit (60%, 45/75), admission à l'oral 60% (15/25)	A, B, C > D, E; 100% qualificatif à l'oral, 90% à l'écrit	1 - 100 (calqué sur CECR et ALTE)
Format	QCM	Scénario	Scénario	CE, CO = QCM, textes courts; EE: rédaction lettre	QCM, scénario, rédaction, CO, analyse photos...	QCM, "trous"
Compétences	CE, CO	CE, CO, PO, PE, Communication interactive	CE, CO, PO, PE, Interaction	CE, CO, PO, PE, Interaction	CE, CO, PO, PE, Interaction	CE, CO, PO, PE, Interaction
Validité	"mesure CE, CO", valable 2 ans (recommandé)	support lié à l'analyse des besoins professionnels	adossée au CECR, valable 5 ans	orienté CECR, élaboration centralisée, standardisée	ALTE, EEIG	plusieurs processus, STFX

Fiabilité	pré-testé (+1 000), KR-20 coefficient, écart-type faible	grilles critériées, référentiels, conseil scientifique	référentiel CECR	pré-tests, contrôles permanents, suivi statistique (Cambridge)	ALTE, EEIG	pre-testing (600+), ALTE
Passation	2 heures	3 heures (modifications expérimentées)	1h30 - 3 heures selon niveau	3h30 écrit, 20 mns. oral (+20 préparation)	2h40 - 5h selon niveau	papier: 110 mns., informatisé: 35-75 mns.
Personnel	1/25 candidats	5/30 candidats (modifications expérimentées)	surveillance, correcteurs, jurys	surveillance écrit, 2 pour l'oral (+ surv. prép. oral)	2 examinateurs + 1 surv./15 candidats	papier: survs., administrateurs; examinateurs agréés EO, EC
Matériel	lecteur CD/K7	modifications en cours	1 poste MM/candidat (vidéo, audio)	lecteur CD/K7, 1 salle pour l'écrit, 2 salles pour l'oral	2 salles, magnéto.	informatisé: PC
Lieu/x	Centres ETS, chez client	centres d'examen agréés	université	Goethe-Institut, Maisons fr-all, partenaires agréés, écoles (sur RV)	CCI Nice (épreuves envoyées en Italie)	chez le client, centres de langues
Coût	31-38€ (66€prix public)	69€ publics non Ed. natle.	gratuit (coût estimé à 40€)	110€ individuel (à déterminer), tarifs groupes	167,69€	38-45€ pour chaque partie
Dates	libre	ang: 6 fois/an, autres: 2 fois/an	libre (actuellement 1 fois/an)	fixes (1-2 fois/an par centre)	juin, novembre	libre