

Certifier la compétence numérique. Partie 2 : établissement des notes

Étienne Vandeput

► **To cite this version:**

Étienne Vandeput. Certifier la compétence numérique. Partie 2 : établissement des notes. Drot-Delange, B. ; Baron, G-L. & Bruillard, E. Sciences et technologies de l'information et de la communication (STIC) en milieu éducatif, 2013, Clermont-Ferrand, France. 2013. <edutice-00875636>

HAL Id: edutice-00875636

<https://edutice.archives-ouvertes.fr/edutice-00875636>

Submitted on 22 Oct 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Certifier la compétence numérique. Partie 2 : établissement des notes

Étienne Vandeput

evandeput@ulg.ac.be, etienne.vandeput@unige.ch

CRIFA Université de Liège – IUFE Université de Genève

Résumé. La mise en place, dès 2009, d'une plateforme d'apprentissage en ligne¹ dédiée à la maîtrise des progiciels d'usage courant, nous a conduit à réfléchir à la problématique de sa certification. Les épreuves se sont déroulées pour la troisième fois en Belgique et pour la première fois en Suisse, en mai 2013. Les enseignements tirés de ces différentes expériences nous ont permis d'identifier les difficultés liées à la manière de noter les participants. L'évaluateur se retrouve quelques fois devant des réponses qui le laissent perplexe, notamment lorsqu'un flou subsiste sur ce qui est réellement évalué, mais aussi parce que les outils permettent parfois d'obtenir des solutions par des voies détournées, voire compliquées. Comment et quoi noter dans ces cas particuliers ? Ces questions sont au cœur de l'article. Cet article constitue la poursuite de la réflexion entamée dans l'article précédent, sur les nombreuses difficultés que pose l'évaluation d'un usage maîtrisé et efficace d'un progiciel, dans le cas précis où celle-ci débouche sur un verdict sanctionnant.

Mots-clés: certification, évaluation, épreuve, consigne, note

Établir une note

Nous avons décrit dans l'article précédent² le contexte de certification dans lequel nous évoluons. Il est donc naturel, après avoir discuté de la manière de concevoir les épreuves, de se poser la question de l'évaluation au sens de la production d'un verdict, d'une sanction.

Diverses problématiques sont associées à ces opérations. Examinons-les en détail. Admettons d'emblée que dans cette partie de notre réflexion, on trouvera beaucoup plus de questions que de réponses. Nous naviguerons entre la macro-évaluation (la décision) et la micro-évaluation (note chiffrée d'un item d'épreuve).

Note globale

Envisageons d'abord la problématique du verdict.

Dans le cas spécifique où une sanction doit être prononcée, il convient d'établir la limite de la réussite. Dans les systèmes d'évaluation classiques, la décision est prise sur base d'une note résultant d'une addition de notes partielles, éventuellement pondérées. Certains repères font partie de l'inconscient collectif. 50%, 60% de la note maximum sont les plus couramment admis³. Il est commode de se reposer sur ces critères. Dans le cadre d'une pédagogie par objectifs, il convient toutefois de s'interroger sur la concordance entre cette limite et ce qu'il conviendrait d'appeler une atteinte minimum des objectifs.

Cette atteinte minimum des objectifs ne peut être définie qu'à travers une calibration correcte d'épreuves bien équilibrées. Il est possible d'établir une telle balance en traduisant des compétences très générales en compétences plus spécifiques elles-mêmes traduites en objectifs d'apprentissage (CRIFA-ULg(a) 2009-2013). En évaluant ces objectifs d'apprentissage aux travers d'objectifs opérationnels et donc d'activités plus élémentaires (CRIFA-ULg(b) 2009-2013), on se donne une

¹ Le site *visaTICE* est accessible à l'adresse <http://visatice.ulg.ac.be>

² *Certifier les compétences numériques. Partie 1 : conception des épreuves.*

³ Dans le système universitaire belge, une note de 12/20 dans un cours équivaut à une dispense d'examen pour la suite du cursus. Une note de 10/20 n'oblige pas à une représentation de l'examen, mais n'accorde pas la dispense si l'étudiant échoue en seconde session.

garantie d'équilibre. Dans ces conditions, on peut espérer qu'une note résultant d'un total fournisse une bonne mesure du niveau de maîtrise de la personne évaluée.

1	Action	Indicateurs	Val
2	Créer un classeur	Un nouveau classeur existe (≠ une évolution des deux premiers).	3
3	Enregistrer un classeur en le nommant	Il porte le nom "Population".	1
4	Copier une feuille d'un classeur dans un autre	Il contient la feuille Naissances.	3
5	Copier une feuille d'un classeur dans un autre	Il contient la feuille Décès.	3
6	Calculer une somme	Feuille Naissance: une formule correcte pour les totaux par ligne est écrite au moins une fois.	1
7	Calculer une somme	Une formule correcte des totaux de colonnes est écrite au moins une fois.	1
8	Recopier une formule	Les deux formules sont recopiées dans les cellules adjacentes.	3

Figure 1: Extrait d'une grille de correction

La figure 1 montre le début d'une liste d'objectifs opérationnels testés. La valeur attribuée à chaque objectif peut dépendre d'une répétition possible de celui-ci dans l'épreuve. Nous discutons plus loin de cette question.

Si une heuristique nous permet d'affirmer qu'un échec est « mérité » dans le sens où il traduit systématiquement de graves carences, nous avons pris l'habitude de procéder à quelques coups de sonde afin de vérifier que la « ligne rouge » se situe bien où nous l'avons placée. En d'autres termes, la production globale d'un apprenant qui échoue de peu est-elle insuffisante et pourquoi et, *a contrario*, la production d'un apprenant qui réussit de peu est-elle satisfaisante et pourquoi ? Cette vérification nous permet de valider nos grilles de correction (cf. plus loin) et de réviser la position de la ligne rouge dans le cas contraire. Cette procédure n'est pas douteuse. Elle relève simplement d'une incertitude qu'il est difficile d'éliminer totalement.

Nous pensons que si une limite doit être fixée, il faut pouvoir se donner la possibilité de la faire évoluer légèrement à la hausse ou à la baisse sur base d'une vérification sur l'ensemble des résultats.

Illustration. À titre d'exemple, lors de la certification 2013 des étudiants de premier bachelier en Psychologie et Sciences de l'Éducation à l'Université de Genève et sur un ensemble de 168 épreuves, une étudiante s'est étonnée de son échec et une autre de sa réussite. On peut se satisfaire de ce faible taux d'étonnement. De plus, dans les deux cas, l'évaluation que portaient ces deux étudiantes sur leurs prestations ne prenait pas assez en compte le critère d'adaptabilité de la production, ce qui est tout de même un critère majeur dans un contexte où toute production est censée évoluer. Inutile de dire que les résultats de ces deux étudiantes se situaient très près de la limite de réussite choisie.

Pour en terminer avec cette problématique de la réussite, on notera que certains systèmes (ASSAP 2011), jugent que des erreurs sont éliminatoires et d'autres pas. Il est à supposer qu'un apprenant qui ne fait pas la preuve de toutes ces aptitudes échoue. Dans un tel système, la discussion est reportée sur ce qui est éliminatoire et ce qui ne l'est pas.

Quoi qu'il en soit, il convient tout de même de rappeler que la première des qualités d'un dispositif de certification, c'est que le protocole d'évaluation soit aussi clair dans la tête des évaluateurs que dans celles des évalués.

Pondération

Idéalement, les épreuves devraient être de trois types :

- des épreuves théoriques dont on a montré l'intérêt (cf. 1^{re} partie) ;
- des épreuves pratiques imposées par des consignes très précises ;
- des épreuves pratiques libres autorisant une certaine créativité.

Les épreuves du troisième type sont difficilement concevables dans un contexte où les apprenants sont très nombreux. C'est le cas en ce qui nous concerne puisque l'apprentissage s'effectue en ligne et qu'il est accessible à des centaines d'apprenants. Nous avons toutefois montré dans la 1^{re} partie que des

épreuves pratiques bien conçues pouvaient éprouver les capacités d'analyse, de synthèse et de jugement de l'apprenant et cela, même si les consignes sont très précises et contraignantes.

Malgré cette réduction à deux types d'épreuves, nous restons devant la difficulté de les pondérer pour l'obtention d'une note globale. Si on considère qu'elles sont d'égale importance, la solution de la moyenne ou d'une simple addition s'impose. Mais on peut raisonner autrement et considérer qu'un apprenant compétent se doit d'allier résultat et manière, auquel cas on ne devrait attester d'une réussite que si celle-ci est effective dans les deux épreuves.

La réalité montre que quelques apprenants sont capables de réussir une épreuve théorique, composée de plusieurs questions distinctes, alors qu'ils échouent dans une épreuve pratique, sous l'emprise du stress de devoir terminer dans les temps, par exemple.

À l'inverse, des étudiants ayant une connaissance empirique de l'outil peuvent réussir de justesse une épreuve pratique (ce ne sera jamais brillamment car ils manqueront de temps), mais ils échoueront lamentablement dans les épreuves requérant rigueur et compréhension. Au final, ces cas sont rares et en plus, une moyenne des deux performances va tout de même dans le sens de l'échec.

Effets combinés

Les questions d'une épreuve théorique sont généralement indépendantes les unes des autres. C'est même souhaitable, au risque sinon, d'aller à l'encontre d'un principe d'équilibre dans la totalisation.

Lors d'une épreuve pratique, la tendance est plutôt de proposer la réalisation d'un travail complet sur un ensemble de données. Dans ce cas aussi, l'évaluation risque d'être perturbée par divers facteurs que nous allons analyser.

Une solution serait de proposer de petits exercices pratiques indépendants. Il serait toutefois dommage de ne pas pouvoir tester la capacité de l'apprenant à réaliser un travail complet du type de ceux que le progiciel concerné permet de réaliser. Les deux sont idéales, mais cela demande un investissement plus important encore et nous tenterons d'y échapper en prétendant que le mieux est parfois l'ennemi du bien.

Quels sont les paramètres auxquels il convient d'être attentif ? Nous en identifions deux qui sont importants : le facteur temps et les questions qui s'enchaînent.

Pour une épreuve constituée d'un ensemble de questions dont les réponses sont sans équivoque, on admet facilement que la note globale traduise les mêmes observations chez l'ensemble de ceux qui l'ont passée. C'est moins vrai lorsque l'épreuve consiste à réaliser un travail qui intègre divers savoir-faire. Le **temps alloué** est déterminant de deux manières. D'abord, on observe que certains étudiants s'acharnent sur une difficulté, même s'ils en sont avertis.

Illustration. Dans de l'épreuve pratique de la certification 2012 sur le tableur, l'apprenant devait travailler sur une base de données. Il devait d'abord créer une colonne supplémentaire et la faire compléter par la valeur « O » ou la valeur « N » selon qu'une condition sur la colonne de gauche était remplie ou non. Si l'insertion d'une colonne ne posait de problème, ceux qui n'avaient pas intégré le principe des traitements conditionnels s'acharnaient sur la formule à construire. Malgré les recommandations de ne pas y perdre du temps (il était possible de compléter la colonne manuellement sans perdre trop de temps), beaucoup d'étudiants ont consacré la moitié du temps attribué à essayer de trouver la bonne formule. Passer une étape perturbe nombre d'entre eux et cela, quelle que soit son importance.

L'acharnement sur une difficulté entraîne inévitablement un manque de temps pour la suite du travail que l'évalué ne peut terminer. Impossible de noter une partie de travail non fait. On mesure donc davantage une capacité à gérer son temps que la maîtrise d'un logiciel.

Pour la constitution d'une telle épreuve, on peut au moins faire la recommandation suivante : proposer des manipulations qui vont des plus simples ou plus complexes et, au minimum, ne pas proposer de manipulation trop complexe en début d'épreuve. De la sorte, on permet à chacun de faire la preuve d'un maximum de ses connaissances.

En revanche, le temps est un des rares moyens de mesurer la qualité du produit en termes de conception et les capacités d'organisation de l'utilisateur. En proposant des traitements particuliers à effectuer sur des données d'un certain volume, on oblige l'utilisateur à être organisé et à faire preuve de sa connaissance du système.

Les exemples de traitements pour lesquels la réflexion fait gagner beaucoup de temps sont faciles à imaginer. Il suffit de disposer d'un « corpus » volumineux sur lequel les évalués vont devoir travailler : plus d'une dizaine de pages si c'est un texte, quelques dizaines de diapositives si c'est une présentation ou encore quelques centaines d'enregistrements si c'est la base de données d'une feuille de calcul. Réaliser certains traitements sur ces documents est proprement impossible si le temps est compté. Voici deux exemples.

Illustration 1. Dans la figure 2, on trouve quelques consignes à propos d'une présentation qui contient plus d'une vingtaine de diapositives. Régler le problème de la mise en forme manuellement risque de consommer pas mal de temps alors qu'une solution automatique existe.

Tous les titres doivent posséder les attributs de mise en forme suivants : police Verdana, corps 40 pt, couleur brun foncé et alignement à droite.

Le titre principal (première diapositive) doit être en police Verdana, couleur brun foncé, corps 60 pt. Le sous-titre doit être en police Verdana, couleur gris, corps 32pt.

Les intitulés « Histoire » et « Informations utiles » sont en police Verdana, couleur brun foncé, corps 24 pt.

Le texte descriptif est en police Cambria, corps 18 pt.



Figure 2: Un travail qui demande peu de temps à celui qui est organisé

Illustration 2. La feuille de calcul de la figure 3 contient un peu moins de 600 enregistrements.

	A	B	C	D	E	F	G	H	I
1	Évolution par commune (situation au 15/01/2011)								
2	Code INS	Commune	Hab. au 1/1/2010	Hab. au 31/12/2010	Différence	Nalssances	Décès	Mariages	Cohabitations légales
3	41002	AALST	80311	81225	914	925	759	738	649
4	44001	AALTER	19519	19753	234	203	182	182	128
5	24001	AARSCHOT	28425	28652	227	254	270	210	206
6	11001	AARTSELAAR	14298	14326	28	107	143	98	106

Figure 3: Une longue liste

La question « *Combien de communes ont perdu plus de 100 habitants ?* » est plutôt stressante pour l'apprenant qui ne sait user correctement du filtrage.

En traitement de texte, la création d'une table des matières est une vraie gageure si le texte n'est pas balisé en titres et sous-titre ou si on ne connaît pas la procédure d'association des styles (ou autres éléments balisés) à cette table des matières.

Par souci d'honnêteté envers les apprenants, la description de la certification devrait faire mention de l'existence, dans les épreuves, de tels types d'opérations à réaliser afin qu'ils puissent s'y préparer.

Comment dénicher de tels documents ? Il n'est pas bien compliqué de se procurer des textes structurés que l'on a soi-même écrits, de textes qui peuvent être librement téléchargés ou copiés du Web et qu'on débarrasse de tout effet de mise en page avant de les baliser partiellement. Le travail de l'apprenant consiste alors à le compléter et à lui appliquer divers traitements. Pour les feuilles de calcul, les instituts nationaux de statistiques fournissent chaque année un nombre conséquent de données. Il est un peu plus compliqué de trouver des présentations dont l'utilisation n'est pas soumise aux droits d'auteur. On peut cependant trouver pas mal d'images et de textes libres de droits et s'orienter vers des présentations de type encyclopédique : les fleurs, les champignons, le chocolat, les voitures... Beaucoup de sujets s'y prêtent. Mais il est vrai que la préparation des matériaux demande un peu de temps.

Nous l'avons dit, dans une épreuve intégrée, les traitements s'enchaînent. Cet **enchaînement** n'est pas sans conséquence au niveau de l'évaluation. Ce problème n'est pas spécifique à ce type d'épreuve, mais à toutes les épreuves qui nécessitent des opérations en cascade.

On note au moins trois effets gênants :

- l'impossibilité d'évaluer les traitements dépendants d'autres traitements non réalisés ;
- la difficulté d'évaluer un traitement correct sur des données fausses issues d'un traitement précédent ;
- la difficulté d'évaluer des traitements dont les effets visibles ont disparu.

Les traitements non réalisés

C'est assez trivial de constater que si l'apprenant ne réalise pas certains traitements, il perd la possibilité d'en réaliser certains autres.

Illustration 1. Il est impossible de modifier les paramètres d'un style de mise en forme qui devait être défini plus tôt dans l'exercice. Dans ce cas précis, prévoir la modification d'un style prédéfini donne la possibilité à l'évaluateur d'au moins vérifier que l'apprenant est capable de modifier un style.

Illustration 2. Il en est de même si un graphique se base sur des données qui ont dû être construites dans une feuille de calcul. Ici aussi, on peut veiller à proposer la réalisation d'un graphique supplémentaire se basant sur des données fournies.

Les traitements incorrects

Si l'apprenant commet certaines erreurs, les données sur lesquelles il poursuit son travail sont fausses et cela rend compliquée l'évaluation des traitements suivants.

Illustration. Pour le tableur, on peut reprendre l'exemple du graphique dans le cas où les données ont été calculées, mais avec des erreurs. De même, la table des matières d'un document texte ne pourra être complète si certains « titres » ont été incorrectement balisés.

Quelles sont les parades ? Il n'y en a guère. La première attitude est de considérer que l'enchaînement d'opérations est de l'ordre de ce qui se passe dans le monde réel et donc n'évaluer que ce qui est correct au final. Une autre attitude est de valoriser la démarche plutôt que le résultat. À nouveau, il semble utile et honnête d'avertir la personne évaluée du système adopté.

Signalons encore que plus les épreuves sont variées (théorie et pratique, par exemple), plus les effets des problèmes évoqués sont dilués.

Il reste, nous l'avons déjà signalé, la possibilité de proposer au sein de l'épreuve plusieurs exercices plus courts contenant moins d'enchaînements.

Les traces

L'enchaînement des opérations a parfois un autre effet. La trace du traitement effectué n'existe plus. Il est donc difficile d'évaluer celui-ci.

Illustration. C'est le cas si, par exemple, des opérations intermédiaires ont lieu comme, créer une feuille provisoire avant de la copier dans un autre classeur. On aura du mal d'évaluer si l'apprenant est capable de copier une feuille d'un classeur dans un autre classeur.

De manière générale, dans tous les cas où cela se présente, il sera difficile de procéder autrement qu'en faisant enregistrer des versions intermédiaires, ce qui alourdit malgré tout le volume de documents à corriger.

Instrument de mesure

Une évaluation cohérente, surtout si elle débouche sur un verdict, nécessite de minimiser les différences d'appréciation entre les évaluateurs. Dans notre contexte, cela nous a conduit à deux choix importants :

- établir des grilles de correction,
- adopter des appréciations chiffrées.

Ceci est valable tant pour les épreuves théoriques que pour les épreuves pratiques. À nouveau, et sans surprise, ces deux choix nous ont placés devant d'autres difficultés que nous développons.

Le mode pédagogique choisi ne permet pas d'imaginer autre chose qu'un lien direct entre les items des épreuves et les objectifs spécifiques qu'on a souhaité faire atteindre aux apprenants.

Dans cette optique, nos **grilles de correction** sont construites de la manière suivante (figure 4) :

- la colonne des objectifs opérationnels (directement liés aux opérations devant être réalisées au cours de l'épreuve) ;
- la colonne des indicateurs ;
- la (les) colonne(s) de la note.

1	Produire et mettre en page un texte				
	<i>Partiel peut signifier incomplet, bricolé... et, en général, pas totalement conforme à ce qui a été demandé (compléter d'une simple croix X).</i>				
2	<i>Pas de mention possible dans les cases noircies.</i>				
3	Opération	Oui	Partiel	Non	Indicateurs
4	Fusionner le contenu de deux documents				Un seul document reprenant toutes les informations.
5	Renommer un document				Le document porte le nom demandé.
6	Insérer une tabulation				Présence d'une tabulation entre les deux informations citée (afficher les caractères non imprimables)
7	Utiliser la fonction rechercher-remplacer				Erreur corrigée dans l'ensemble du document (effectuer une nouvelle recherche)
8	Baliser le texte avec des styles prédéfinis				Chaque bloc de texte est mis en style selon les consignes.
					Tous les blocs de texte Titre 1 possèdent les paramètres définis dans les consignes : police, corps, couleur, i

Figure 4: Petit extrait d'une grille de correction

La grille idéale devrait reprendre, en sus, les objectifs d'apprentissage mis en relation avec les objectifs opérationnels. Dans la pratique, ce n'est toutefois pas indispensable si ce lien est préalablement établi⁴. On dispose ainsi de la preuve que l'on évalue bien sur base des objectifs d'apprentissage.

Autant que possible, comme nous l'avons signalé dans la 1^{re} partie pour les énoncés, la formulation des objectifs opérationnels doit se détacher de l'énoncé de la fonctionnalité. Toutefois, l'indicateur peut faire référence à la démarche. C'est même tout à fait souhaitable si on considère que celle-ci est plus importante que le résultat.

13	Modifier le pied-de-page				Un pied-de-page est présent sur toutes les diapositives (! zone prédéfinie).
----	--------------------------	--	--	--	--

Figure 5: Vérification de la démarche par l'utilisation de la zone de pied de page prédéfinie

Illustration 1. L'indicateur de la figure 5 demande de vérifier que le pied de page des diapositives se trouve en zone prédéfinie et ne résulte pas du copier-coller d'une zone de texte flottante.

⁴ Dans *visaTICE*, il est établi et consultable à partir de la page d'accueil.

Illustration 2. Dans une feuille de calcul, une mise en forme conditionnelle ne peut être observée qu'en vérifiant la démarche comme on demande de le faire en figure 6.

33	Définir une mise en forme conditionnelle simple	L'effet est visible et la règle existe.	3
34	Définir une mise en forme conditionnelle complexe	L'effet est visible et la règle existe.	3

Figure 6: Vérification de la démarche par l'existence de la règle

L'intérêt des notes, c'est qu'elles permettent de totaliser pour mesurer la performance. Cela implique toutefois, qu'elles soient estimées le plus unanimement possible.

Nous avons opté pour deux notes possibles : un point pour les opérations élémentaires dont l'observation ne souffre pas de contestation et trois points pour les opérations plus complexes dont on peut estimer qu'elles peuvent être partiellement réalisées. Ce dernier cas échéant, un seul point est attribué.

Illustration. En examinant l'illustration de la figure 7, on remarque que les titres peuvent être présents seulement dans certains graphiques, certains peuvent être significatifs et d'autres pas. Quant aux étiquettes de données qui doivent être exprimées en pourcentage, il n'y a pas de contestation possible.

39	Choisir un titre	Dans les trois graphiques, les titres sont présents	3
40	Choisir un titre	Dans les trois graphiques, les titres sont significatifs.	3
41	Choisir des étiquettes de données	Le deuxième graphique exprime des pourcentages.	1

Figure 7: Barèmes

Cette manière de procéder réduit considérablement les différences d'appréciation. Il reste que l'adjectif « partiel » cache parfois des niveaux que l'évaluation ne prend pas en compte.

Illustration. Ainsi dans l'énoncé de la figure 8, il est demandé d'insérer les diapositives nécessaires en fonctions d'autres informations fournies. Si on accorde les trois points en cas de présence de toutes les diapositives, on est forcé de n'accorder qu'un point, qu'il n'en manque qu'une ou qu'il n'y en ait qu'une seule ajoutée.

Le diaporama *chocolat* sur lequel vous devez travailler contient déjà 11 diapositives et son organisation est la suivante :

- une diapositive de titre ;
- cinq diapositives (certaines complétées, d'autres pas) fournissant des éléments historiques (type 1) chacune étant suivie d'une diapositive fournissant plus d'informations (type 2).

En tenant compte de la liste des marques présentes sur chaque diapositive, ajoutez les diapositives nécessaires en choisissant la bonne disposition. Compléter leur titre.

Figure 8: Ajout de diapositives

Une autre difficulté apparaît si la grille est directement liée aux consignes. Pour les besoins de l'exercice proposé, certains objectifs opérationnels se répètent. Il faut alors veiller à ce que ces répétitions ne prennent trop de poids dans l'évaluation, par exemple en attribuant moins de points aux opérations répétées.

Sur ce point, on pourrait rêver d'une grille basée sur des objectifs plus globaux et non autour des objectifs opérationnels.

Illustration. Par exemple, une note serait attribuée à l'objectif « *Recopier une formule* » et évaluerait l'ensemble des opérations de recopiage. On pourrait faire de même pour l'objectif « *Mettre en forme une cellule* ».

En procédant de la sorte, on augmente à nouveau le risque de subjectivité car la note n'est plus élémentaire et des différences d'appréciations sont possibles. Par ailleurs, il est indéniable qu'une grille construite en parallèle avec la feuille des consignes augmente le confort de l'évaluateur qui peut évaluer de petites séquences d'opérations en une fois.

La rédaction des indicateurs demande un certain soin. Sa description doit être sans équivoque pour les évaluateurs. Il arrive qu'un indicateur précis ne suffise pas à trancher.

Voici, pour terminer cette réflexion sur les appréciations chiffrées, un exemple comme il n'en existe heureusement pas trop. Il est puisé dans l'usage du tableur dont la richesse des usages implique une plus grande complexité de l'évaluation. Il met en évidence un manque de clairvoyance et/ou une mauvaise compréhension du concept de fonction qui peut ne pas être sanctionné. Que faut-il diagnostiquer dans les exemples suivants et comment évaluer ?

Illustration 1. Dans A3, on souhaite voir s'afficher la somme des contenus des deux cellules A1 et A2. Il semble que la solution la plus simple soit d'y écrire la formule =A1+A2. Mais l'expérience montre que beaucoup d'apprenants focalisent sur les fonctions. On acceptera volontiers la formule =somme(A1;A2), même si l'usage de l'opérateur est plus simple et plus approprié. Si l'évaluation le permet, on peut admettre la réponse et se contenter d'émettre un commentaire du genre « *L'utilisation de la fonction somme n'a d'intérêt que si le nombre de ses arguments est supérieur à deux. Sinon, l'opérateur doit être préféré à la fonction.* ».

De même, on préférera la formule =somme(A1:A5) à la formule =A1+A2+A3+A4+A5 pour les mêmes raisons d'économie d'écriture. Mais faut-il sanctionner cette seconde formule ?

Mais revenons au problème initial. Que penser de la formule =somme(A1:A2). On aurait aussi envie de l'accepter. On notera toutefois que cette formule risque de poser des problèmes si, dans la suite des opérations, on vient à insérer et à compléter une ou plusieurs lignes entre les lignes 1 et 2.

Illustration 2. Voici un cas vécu pour illustrer. Dans l'exercice de la figure 9, on demande de totaliser en colonne D les nombres de garçons et de filles pour chaque discipline. Ensuite, on demande d'insérer une colonne après la colonne des garçons et une autre après celles des filles et d'y calculer les pourcentages respectifs.

1	Secteurs	Filles	Garçons
2	Agronomie	536	726
3	Artistique	1750	432
4	Chimie	3200	657
5	Commerce	2005	1641
6	Économie ménagère	2207	589
7	Mécanique	800	2212
8	Paramédical	1200	600
9			

Figure 9: Adaptation de formule

Si la formule écrite en D2 est =somme(B2:C2), elle s'adaptera après l'insertion des deux colonnes, mais la formule de calcul du pourcentage mentionnera une erreur pour cause de référence circulaire (figure 10).

	A	B	C	D	E	F	G	H	I	J	K
1	Secteurs	Filles		Garçons							
2	Agronomie	536	=B2/F2	726		1262					
3	Artistique	1750		432		2182					
4	Chimie	3200		657		3857					
5	Commerce	2005		1641		3646					
6	Économie ménagère	2207		589		2796					
7	Mécanique	800		2212		3012					
8	Paramédical	1200		600							
9											

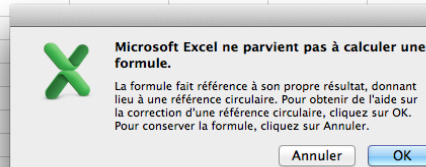


Figure 10: Référence circulaire

Illustration. Toujours sur notre question initiale et plus inquiétant, on connaît l'exemple de la formule =somme(A1+A2) qui rend perplexes les évaluateurs. Cette formule donne logiquement le bon résultat. Toutefois, elle indique une nette confusion entre opérateur et fonction.

Que penser enfin de la formule =somme(A1-A2) proposée par une étudiante pour effectuer une différence entre les contenus de deux cellules et qui, tout aussi logiquement, le fait.

Si on privilégie l'exactitude du résultat, on a moins de questions à se poser. Mais en agissant de la sorte on privilégie le produit par rapport à la démarche et à la compréhension.

Conclusion

Dans les deux parties de cet article, nous avons mené une réflexion sur la conception d'épreuves (1^{re} partie) et sur leur évaluation (2^e partie). Notre ambition s'est limitée à prouver toute la complexité de ces deux tâches. À la manière d'une analyse qualitative basée sur le principe de la théorie ancrée (Corbin en Strauss 2008), notre théorie a été construite à partir du terrain, des expériences vécues. Elle est donc nécessairement incomplète à ce stade. Dans une phase ultérieure, une comparaison avec la littérature s'impose, tant au niveau macro (Vial 2012) (De Ketele 2012) qu'au niveau micro (Leclercq, 2005). Cette comparaison permettra de développer davantage notre propre théorie.

Nous ne reprendrons que très succinctement les conclusions de la première partie. Nous retiendrons la prudence avec laquelle questions théoriques et consignes pour un exercice pratique doivent être rédigées. Nous insisterons pour que la constitution d'épreuves conduisant à une certification fasse référence à une théorie d'apprentissage. Nous recommandons une couverture maximale des objectifs et des mesures à tous les niveaux d'apprentissage. Nous suggérons des épreuves de différentes natures et, au minimum, une épreuve théorique et une épreuve pratique, l'épreuve théorique permettant de tester davantage les connaissances et la compréhension et l'épreuve pratique, la mesure des savoir-faire et l'efficacité des usages. Dans cette dernière, le temps, allié au volume d'informations à traiter, peut jouer un rôle important.

Pour ce qui concerne la manière d'évaluer, et dans l'état actuel de notre réflexion, nous préconisons un système basé sur des appréciations chiffrées. Ce choix implique la gestion d'une série de difficultés autour desquelles nous avons fait quelques recommandations. Nous retiendrons que, quels que soient les outils facilitateurs de cette évaluation, il convient de viser une uniformité des appréciations (les correcteurs pouvant être nombreux). Si un barème pour une épreuve théorique et une grille de correction pour une épreuve pratique sont des outils qui peuvent y contribuer, quelques obstacles subsistent tels la pondération des questions pour la première, la répétition des opérations, leur possible enchaînement et la qualité de leurs indicateurs pour la seconde. En outre, le but avoué d'une certification oblige à se poser des questions quant à la détermination du niveau de réussite.

On savait que, dans le contexte de la recherche d'une maîtrise des TIC, il n'est pas simple de fixer des objectifs et des contenus de formation. Cet article a voulu montrer que si cette première étape est franchie, celle de l'évaluation requiert tout autant notre attention. Il n'en est pour preuve que les choix didactiques que nous pouvons faire et qui pourraient conduire à des options dont les richesses sont très relatives : évaluer la mise en œuvre d'une fonctionnalité, évaluer l'obtention d'un résultat, évaluer une démarche.

Bibliographie

- Anderson, Lorin W., et David R. & al. Krathwohl. *A Taxonomy for Learning, Teaching, and Assessing — A Revision of Bloom's Taxonomy of Educational Objectives*. Longman: Addison Wesley, 2001.
- ASSAP. *Bienvenue sur le site U-CH*. 1 1 2011. <http://www.assap.ch/u-ch/> (accès le 6 20, 2013).
- Bloom, Benjamin. *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain*. New York: David McKay, 1956.
- Corbin, Juliet, et Anselm L. Strauss. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Third edition. Sage, 2008.
- CRIFA-ULg(a). *visaTICE: objectifs et compétences*. 1 9 2009-2013. <http://www.visatice.ulg.ac.be/pages/objectifs/> (accès le 6 20, 2013).
- CRIFA-ULg(b). *visaTICE: comprendre le traitement de l'information numérique*. 1 9 2009-2013. <http://www.visatice.ulg.ac.be/pages/objectifs/tin/> (accès le 6 20, 2013).
- De Ketele, Jean-Marie. «À la recherche de paradigmes unificateurs ou de paradigmes intégrateurs?» Dans *Modélisation de l'évaluation en éducation*, de Lucie Mottier Lopez et Gérard Figari, 195-210. Bruxelles: De Boeck, 2012.
- Leclercq, Dieudonné. *À la recherche de la triple concordance en éducation*. Liège: Les Éditions de l'Université de Liège, 2008.
- . *Édumétrie et Docimologie pour praticiens chercheurs*. Liège: Les Éditions de l'Université de Liège, 2005.
- . *La conception des Questions à Choix Multiple*. Bruxelles: Labor, 1986.
- Tyler, Ralph W. *Basic principles of curriculum and instruction*. Chicago: University of Chicago Press, 1949.
- Vandeput, Étienne. «Méthodologie d'identification des invariants du traitement de l'information numérique.» *Sciences et technologies de l'information et de la communication (STIC) en milieu éducatif : Analyse de pratiques et enjeux didactiques*. Patras: Baron, Georges-Louis; Bruillard, Éric; Komis, Vassilis, 2011. 93-107.
- Vandeput, Étienne, et Julie Henry. «visaTICE : se mesurer aux TIC et se former sous le regard d'un coach.» *Sciences et technologies de l'information et de la communication (STIC) en milieu éducatif : Analyse de pratiques et enjeux didactiques*. Patras: Baron, Georges-Louis; Bruillard, Eric; Komis, Vassilis, 2011. 141-151.
- Vial, Michel. «Les modes de pensée, organisateurs des modèles d'évaluation.» Dans *Modélisation de l'évaluation en éducation*, de Lucie Mottier Lopez et Gérard Figari, 131-146. Bruxelles: De Boeck, 2012.