

AVANT-PROPOS

De quoi s'agit-il ? Que peut-on en faire ? Telles sont, s'agissant des résultats de la recherche COMPÉFRADE, les deux questions auxquelles nous nous attacherons à répondre dans cet avant-propos. Ces réponses pourraient tenir en quelques lignes. Nous avons préféré quelques pages, qui donnent la possibilité de situer la recherche dans la problématique de l'évaluation des compétences.

1. DE QUOI S'AGIT-IL ?

De quelques *outils d'évaluation de compétences*. De *quelques* outils destinés à l'évaluation de *quelques* compétences de lecture et d'écriture aux stades de développement que l'on peut espérer au deuxième degré de l'enseignement secondaire¹.

1.1. Compétences de lecture et d'écriture

Lecture et écriture sont parfois désignées *elles-mêmes* comme compétences. Abusivement, nous semble-t-il (DUMORTIER, 2002). En effet, toute compétence est *relative à une famille de tâches ou de situations problématiques* (ROEGIERS, 2000, p. 125-143). Une compétence est une aptitude à mobiliser les savoirs de toutes sortes qui permettent de résoudre les problèmes *typiques* d'une famille de tâches ou de situations. Est-il raisonnable de considérer que *toutes* les tâches de lecture ou d'écriture que l'on peut imposer à un élève, que *toutes* les situations de lecture ou d'écriture dans lesquelles peut se trouver un individu relèvent de la *même* famille ?

Si toutes ces tâches ou ces situations appartenaient à la même famille, il n'y aurait pas de limites à l'expertise d'un lecteur ou d'un scripteur : il serait aussi performant quoi qu'il lise ou qu'il écrive. Le cas, s'il se rencontre, est exceptionnel. La règle, c'est l'expertise plus ou moins limitée : bon ou très bon pour lire/écrire telle sorte de texte dans telle sorte de contexte ; moins bon voire très mauvais pour lire/écrire tout autre chose dans de tout autres conditions. D'où l'importance d'une réflexion sur les variétés de textes² que les élèves devraient *prioritairement* apprendre à lire et à écrire. À un moment donné du cursus et dans telle forme d'enseignement, peut-être. Dans des conditions déterminées, certainement. Et, tout aussi certainement, compte tenu d'exigences relatives au résultat de la lecture ou de l'écriture³.

¹ Les résultats de cette recherche sont bien distincts de ceux des travaux menés par la commission « Outils d'évaluation ». Ils n'en constituent pas moins une panoplie d'outils destinés à la fois à faciliter et à améliorer l'évaluation (diagnostique, formative et certificative) de certaines compétences de lecture et d'écriture.

² Cette expression, tout comme celle de « classe de textes », nous paraît préférable à celles de « type » ou de « genre » de textes, plus connotées techniquement.

³ C'est *tout au long du cursus* correspondant à la scolarisation obligatoire (du début du primaire à la fin du secondaire, donc) que devraient s'effectuer les apprentissages principaux en fait de production et de réception des variétés du discours écrit (ou oral, d'ailleurs). Cela suppose une concertation large sur ces

Mais, bien entendu, les différentes familles de tâches ou de situations de lecture et d'écriture présentent une proportion de problèmes communs. Certains de ces problèmes peuvent être qualifiés de *généraux*, car ils se retrouvent dans *toutes* les tâches ou situations. Ainsi en va-t-il, par exemple, des problèmes de traitement des mots inconnus ou de cadrage⁴, en lecture. Ainsi en va-t-il des problèmes d'orthographe ou de connexion⁵, en écriture. D'autres problèmes sont *génériques* : ils se retrouvent dans quantité de tâches ou situations apparentées, notamment par le « genre » du texte à traiter ou à produire⁶. Ainsi en va-t-il, par exemple toujours, des problèmes relatifs à l'argumentation : repérage (en lecture) ou énoncé (en écriture) d'une thèse, critique (en lecture) ou construction (en écriture) d'arguments, etc. Reconnaître l'existence de problèmes généraux et génériques, de lecture ou d'écriture, ne devrait pas porter à méconnaître les

variétés qui, selon nous, ne devraient pas différer beaucoup d'une forme d'enseignement à l'autre. En revanche, il serait souhaitable qu'à chaque étape du cursus et dans chaque forme d'enseignement, les caractéristiques du contexte de lecture ou d'écriture ainsi que les seuils de performance fassent l'objet d'une nouvelle définition.

⁴ « Par cadrage, on entend (...) l'activité sémiotique qui consiste à inférer un certain nombre d'informations à partir d'indices disposés à la périphérie du texte (dans le paratexte et le métatexte), à sa lisière (dans l'espace textuel et l'incipit), mais aussi dans le texte lui-même (les superstructures) » (CANVAT, *op. cit.*, p.115). L'activité de cadrage rapporte un texte donné à une (ou des) classes(s) connue(s), créant ainsi des attentes et ouvrant des pistes de lecture.

⁵ On désigne ainsi l'enchaînement des phrases. L'usage des connecteurs est une des conditions de la cohérence textuelle. On se gardera d'assimiler l'ensemble des connecteurs à celui des mots de liaison : les mots de liaison sont des moyens de connexion parmi d'autres (répétitions, pronominalisations, substitutions lexicales, détermination par l'article défini, le possessif, le démonstratif, ponctuation, etc.) et leur emploi n'est nullement exclusif de l'emploi d'autres moyens.

⁶ La détermination d'un genre est une entreprise ardue. Surtout s'il s'agit des genres non littéraires, relativement moins codés – pour la plupart et à l'exception notable des genres scolaires (dissertation, résumé, commentaire, synthèse, etc.) – que les genres littéraires. Si un genre textuel n'est rien d'autre qu'une classe de textes apparentés par des caractéristiques de contenu et/ou de forme, la pensée humaine, fonctionnant sur la base de perceptions discriminatives (la classe d'objets X rassemble tous les objets ayant les caractéristiques A, B, C, D, par exemple), a tôt fait de créer des classes d'objets distinctes incluses les unes dans les autres. Ainsi les romans (ou tout au moins la plupart d'entre eux) font-ils partie de la classe des récits ; les romans policiers, de la classe des romans et les romans d'énigme criminelle, de la classe des romans policiers. Admettons (en suivant GENETTE, 1977, p. 393-394) que le récit ne soit pas un genre, mais un « mode » (de l'imitation de l'action humaine, auquel s'oppose le mode dramatique). Si l'on dit que le roman est un « genre » narratif (un genre de récit), on est un peu gêné aux entournures pour dire, tout uniment, que le roman policier est un « genre » de roman, et le roman d'énigme criminelle un « genre » de roman policier. En effet, à supposer qu'un nombre X de caractéristiques communes à des objets textuels soient estimées nécessaires et suffisantes pour définir, en compréhension comme en extension, le « genre » roman, ce n'est pas le même nombre des mêmes caractéristiques qui permettront de définir le « genre » roman policier ou le « genre » roman d'énigme criminelle. Problème oiseux pour un enseignant de français ? Pas du tout ! De la compétence avérée à comprendre (et à manifester, dans les formes requises, la compréhension d') un roman d'énigme criminelle, on ne peut conclure à la compétence virtuelle de comprendre (et à manifester, dans les formes requises, la compréhension d') un roman noir (autre classe de romans incluse dans la classe roman policier) ou d'un roman fantastique (autre classe de romans incluse dans la classe roman).

spécificités d'une famille de tâches ou de situations problématiques impliquant la lecture, l'écriture ou, le plus souvent en milieu scolaire, les deux, successivement.

La détermination d'une compétence de lecture ou d'écriture suppose une analyse complète – en termes de problèmes *virtuels*⁷ – de la famille de tâches ou de situations à laquelle cette compétence est relative (DUMORTIER, 2001a).

Exemplifions cette notion de « famille ». C'est une famille de tâches de lecture-écriture que celle impliquant la compréhension d'un récit exemplaire (un récit visant à faire admettre une leçon de vie, à accréditer une morale) et la manifestation de cette compréhension sous forme de réaction à un questionnaire contenant des items à réponse choisie (lecture seule) et à réponse construite (lecture-écriture⁸). *Tous* les récits exemplaires sont *susceptibles* d'être questionnés selon les mêmes orientations et, *quelle que soit la spécificité du texte* dont l'élève doit manifester la compréhension, il *peut* avoir affaire à des questions à réponse choisie ou à des questions à réponse construite. Autre exemple : c'est une famille de tâches de lecture-écriture que celle consistant à rédiger un texte informatif, à visée didactique, d'après un plan fourni et en s'appuyant sur plusieurs documents (fournis eux aussi) pertinents au sujet imposé. *Tous* les dossiers à traiter en vue de la réalisation d'un tel travail peuvent l'être de manière analogue, et, *quelle que soit la spécificité du texte* à contenu planifié, l'élève devra respecter des conventions linguistiques, textuelles et discursives sur lesquelles le concepteur de la tâche attirera son attention.

C'est ainsi manière d'insister sur la modestie de notre entreprise. Nos outils ne sont pas propres à évaluer *la* lecture ou *l'*écriture au deuxième degré de l'enseignement secondaire. Ils favorisent l'évaluation de *certaines* compétences, sans plus.

C'est aussi manière d'attirer l'attention sur les difficultés de cette entreprise. Ce n'est décidément pas chose simple que construire un outil d'évaluation d'une compétence. Parce que ce n'est pas chose simple de cerner une compétence, une famille de tâches, une famille de situations. On marche sur la corde raide. Un faux pas à droite, on chute dans les généralités (compétence à écrire un texte argumenté... oui, mais quelle *sorte*

⁷ L'enseignant, en effet, n'a jamais aucune assurance préalable quant aux problèmes *réels* que rencontrera un apprenant *singulier*, et, moins encore, quant au degré de difficulté que présenteront, pour *cet* apprenant, ces problèmes.

⁸ Précisons que l'ensemble des tâches de lecture-écriture *inclut* l'ensemble des tâches consistant à manifester la compréhension d'un texte en répondant par écrit à un questionnaire. Dans le présent recueil, on trouvera, notamment, des exemples de tâche de lecture-écriture consistant à rédiger la fin d'une nouvelle. La réussite d'une telle tâche est conditionnée par la compréhension de la nouvelle et cette compréhension elle-même *peut* être guidée par un questionnaire (à réponses choisies et/ou construites). Mais la rédaction d'une fin implique, bien sûr, des opérations plus nombreuses et plus complexes que celles impliquées par la rédaction de réponses (correctement articulées, d'un point de vue linguistique) à des questions posées par le professeur. Autrement dit, si toute *production* d'une réponse écrite à une question écrite est une tâche relevant de la lecture-écriture, il s'en faut – et de beaucoup ! – que toute tâche de lecture-écriture consiste en la production d'une réponse écrite à une question écrite.

de texte argumenté et *destiné à qui ?*). Un faux pas à gauche, c'est dans les particularités que l'on tombe (compétence à lire *ce* conte fantastique... oui, mais les acquis de l'activité sont-ils exportables, le temps et les efforts consacrés à répondre au questionnaire de compréhension peuvent-ils être rentabilisés quand il s'agira de lire *un autre* conte fantastique?).

1.2. Évaluation

Le terme « évaluation » fait généralement penser à notes, à bulletin, à conseil de classe et il ravive, en bien des esprits, les mauvais souvenirs de maints méchants marchandages. Le terme « évaluation » traîne après lui une batterie de connotations déplaisantes et comme, de nos jours, ces casseroles-là font plus de bruit qu'à l'accoutumée, il n'est pas rare d'entendre les professeurs se plaindre : « Qu'on nous laisse donc enseigner ! », « Tout n'est pas évaluable ! », etc.

Mais, qu'il le veuille ou non, qu'il le sache ou pas, un enseignant est *essentiellement* un évaluateur⁹.

L'acte d'enseigner quelque chose à quelqu'un résulte d'une décision reposant elle-même sur le constat d'un défaut de savoir, autrement dit sur une évaluation. « Évaluer signifie confronter un ensemble d'informations à un ensemble de critères en vue de prendre une décision » (DE KETELE, 1980, *in* DE KETELE et ROEGIERS, 1993, p. 42). Imaginons que l'ensemble des élèves d'une classe fassent la preuve qu'ils ont préalablement acquis tel savoir dont l'enseignement est prévu à tel niveau d'études. Le professeur leur enseignera-t-il ce savoir-là? Il faut espérer que non. Il faut espérer qu'ayant confronté des informations relatives à l'état *réel* des connaissances des élèves aux critères de l'état *attendu* de leurs connaissances et ayant constaté la coïncidence de l'état réel et de l'état attendu, il s'abstiendra d'enseigner ce que les élèves connaissent déjà.

Au principe de l'entreprise d'enseignement, une évaluation, donc. C'est *parce que* les élèves ne savent pas qu'ils doivent apprendre. Et c'est *pour* favoriser leur apprentissage que les enseignants enseignent.

Mais cet enseignement a-t-il les résultats attendus ? Peut-on ne pas se poser la question ? Peut-on ne pas exiger les preuves de l'apprentissage ? En tout cas, la société, intéressée par la qualité des citoyens¹⁰, ne le permet

⁹ La formule peut heurter. Ajoutons donc que l'essence de la fonction d'enseignant *ne coïncide pas* avec la fonction d'évaluateur. Ce que nous voulons dire c'est qu'on ne peut pas être enseignant et ne pas être évaluateur. Bien entendu, enseigner c'est aussi – et tout aussi « essentiellement » – donner accès à des savoirs, créer des dispositifs permettant et favorisant l'appropriation de connaissances ainsi que l'acquisition de compétences. Nul n'ignore, au demeurant, que l'efficacité de ces dispositifs est en grande partie conditionnée par la qualité humaine de la relation pédagogique. C'est ainsi donner à entendre que dans l'« essence » de la fonction, il y a le respect de l'apprenant, la conviction de la valeur de ce qu'on l'aide à apprendre et le désir de faire tout ce qu'on peut pour lui fournir une aide efficace. Certains appellent ça l'amour du métier...

¹⁰ C'est cela que *toutes* les personnes impliquées dans l'entreprise pédagogique devraient garder bien en vue : la formation d'individus appelés à *exercer leur liberté*, en connaissance de cause, compte tenu d'une *situation* dans un état civil, un environnement socio-économique, une conjoncture culturelle – pas des « agents »

pas à l'institution scolaire. En fin de période, d'année, de cycle, un prononcé est obligatoire sur le succès ou l'échec des apprentissages. C'est surtout l'élève qui peut pâtir de ce prononcé. C'est *son* échec qui, le cas échéant, est proclamé. Bien que l'échec – tout comme le succès – de l'apprentissage dépende toujours, dans des proportions variables, de l'enseignement. Il faudrait parler d'échec de l'interaction pédagogique. Mais on impute l'échec au seul apprenant. C'est injuste, mais c'est ainsi. Et puisque c'est ainsi, puisque l'évaluation, lorsqu'elle a une fonction certificative, fonde une prise de décision qui détermine l'avenir du seul apprenant, ce serait, tout bonnement, faillir à l'éthique professionnelle que traiter l'évaluation en quantité négligeable.

Accorder peu de soin à l'évaluation conduit, dans la plupart des cas, à évaluer des performances que les élèves n'ont pas été *suffisamment* aguerris à réaliser. Ou à rapporter ces performances à des critères de qualité *excessifs*. Accorder peu de soin à l'évaluation conduit parfois à ceci comme à cela. Pour compenser l'une et l'autre de ces dérives – ou sans cette excuse de compensation – on additionne généralement des scores relatifs à des tâches hétérogènes, souvent très parcellaires. On obtient ainsi une moyenne autorisant le prononcé de réussite (plus rarement d'échec). Mais cette moyenne-là est dépourvue de toute vertu de certifier *des compétences*. Nous y voilà revenus.

1.3. Évaluer des compétences

Qu'est-ce que l'adoption de la pédagogie « par compétences » change à la problématique de l'évaluation ? Tout au moins, que change-t-elle à cette problématique dans les secteurs de la discipline « français » (lecture, écriture) qui nous concernent ici ? Beaucoup de choses, incontestablement, étant donné qu'un changement en entraîne (ou devrait logiquement, nonobstant les traditions scolaires, en entraîner) un autre. Nous nous en tiendrons ici à trois changements capitaux.

1.3.1. Premier changement

Une compétence de lecture ou d'écriture étant, nous venons de le voir, une aptitude à mobiliser des ressources utiles à la résolution des problèmes typiques d'une famille de tâches ou de situations de lecture ou d'écriture, on ne peut l'évaluer qu'en plaçant l'apprenant dans une situation représentative de cette famille, qu'en lui imposant une tâche exemplifiant cette famille. D'où un premier changement. Crucial.

L'évaluation de *compétences* de lecture ou d'écriture exclut, en principe, l'addition de scores obtenus pour des performances relatives à la manifestation de savoirs ou à l'exercice de savoir-faire pris distinctement, non *intégrés* dans une entreprise *complexe* de résolution de problème. Ce sont des problèmes, certes, que posent le traitement d'un mot inconnu ou l'accord d'un participe passé. Ce sont d'autres problèmes que posent le traitement d'un sous-entendu ou le choix d'un connecteur. Ce sont encore des problèmes (tout) différents que posent l'interprétation d'une histoire

ou des « consommateurs » *adaptés* à un certain (dys)fonctionnement du régime démocratique ou à un marché circonstanciel des biens tant matériels que symboliques.

imaginaire ou la construction d'un argument solide. Mais quand bien même un élève serait-il capable de résoudre, *séparément*, ces problèmes-là (et quantité d'autres), il n'aurait pas acquis une compétence – encore moins des compétences – de lecture ou d'écriture.

Pour pouvoir se prononcer quant à l'acquisition d'une *compétence* de lecture ou d'écriture, ou, plus exactement, pour pouvoir se prononcer quant à un *stade d'acquisition* d'une compétence – un stade correspondant à un seuil de performance attendu *et déterminé concrètement par une performance de référence* –, il faut imposer à l'élève une tâche de lecture ou d'écriture *complète*. Une tâche dont la complétude tient à la *variété* et à l'*interaction* des problèmes à résoudre.

Lorsqu'un professeur de français impose à ses élèves un travail de composition écrite, il est patent, en général, qu'il s'agit bien d'une *tâche complexe d'intégration* relative à l'acquisition d'une compétence. Écrire une lettre de demande, un jugement de goût motivé, un texte d'après un schéma, une synthèse de textes, etc., c'est résoudre, en même temps ou successivement, une grande quantité de problèmes *intriqués*, une grande quantité de problèmes dont les solutions sont *interactives*, si l'on peut dire. En revanche, une tâche de lecture consistant à devoir répondre à un questionnaire est moins *évidemment* une tâche complexe d'intégration. Il peut d'ailleurs arriver que ce n'en soit effectivement pas une. Il peut arriver que les questions soient tellement simples, qu'elles aient si peu à voir les unes avec les autres qu'en y répondant l'élève ne fasse rien d'autre qu'exercer *séparément* des savoirs ou des savoir-faire *limités*. Mais un bon questionnaire n'est pas une juxtaposition de questions dont l'hétérogénéité n'a d'égale que la simplicité. Un bon questionnaire se compose de questions qui *concourent* à guider la compréhension comme à faciliter la manifestation de cette dernière. Dans le meilleur des cas, les questions scandent un processus de production de sens en partie *prévisible* : elles sont relatives à des opérations cognitives que le lecteur a *nécessairement* dû accomplir. Et la réponse, *dans les formes requises*, à chacune de ces questions peut constituer une tâche dont la relative brièveté (si l'on compare avec les habituels travaux de composition écrite) va de pair avec une relative complexité.¹¹

Évaluer une compétence de lecture ou d'écriture (en voie de développement), c'est comparer une performance de lecture ou d'écriture accomplie dans des conditions précises avec une *performance étalon* conçue compte tenu, d'une part, de ces mêmes conditions, d'autre part, du niveau de développement de la compétence qu'il est raisonnable de souhaiter

¹¹ Est-il bien nécessaire d'ajouter que la qualité d'un questionnaire, jugée au moyen de la pierre de touche ci-dessus évoquée, n'est nullement une *garantie* de son efficacité en tant qu'instrument de (méta)compréhension ? Pour qu'un (bon) questionnaire soit efficace, encore faut-il que l'élève puisse l'envisager, lui, comme un instrument de (méta)compréhension. Encore faut-il qu'il *s'implique* dans la tâche de lecture. S'il s'en acquitte pour en être débarrassé et non pour apprendre à lire (comme c'est, hélas ! parfois le cas), la qualité du questionnaire ne change rien à l'affaire. Mais ceci vaut pour toute tâche scolaire. Si elle n'est pas *perçue* comme un moyen de s'approprier des connaissances et/ou comme un vecteur de savoirs et de savoir-faire constitutifs d'une compétence, il est parfaitement vain qu'elle soit, *en effet*, un tel moyen et/ou un tel vecteur.

atteint par les apprenants. On voit où est gîté le lièvre. Et l'on conçoit que l'attraper n'est pas une sinécure. Comment s'assurer d'un étalon valable ?

Si le professeur ne dispose pas d'un modèle *écrit* de la performance attendue (un corrigé de questionnaire de compréhension en lecture, un texte de référence pour un travail d'écriture), inévitablement, il évalue compte tenu d'un modèle *mental*. Et ce modèle-là, nécessairement, il le conçoit à partir de ce que *serait* sa propre performance *d'expert*. Sans doute ce qui résulte d'une telle conception est-il plus ou moins travaillé par la conscience du fait que les élèves ne sont pas, eux, des experts. Mais, si vive que soit cette conscience-là, elle s'applique à une performance *virtuelle*. Le professeur bémolise une partition vierge. Il consent, avec bon sens et bienveillance, à ce que l'élève fasse moins bien que lui n'aurait fait. Ce qui en résulte, c'est, si l'on peut dire, un étalon mou, impropre à une évaluation sérieuse.

Est-ce à dire qu'un modèle écrit de la performance attendue *suffise* à garantir le sérieux de l'évaluation ? Bien sûr que non ! La validité de l'évaluation est fonction de la valeur du modèle. Et la valeur du modèle est elle-même fonction de la détermination de *critères de réussite*. Le choix de ces critères ne devrait pas s'opérer compte tenu, principalement, de ce qu'un élève moyen (?) est censé pouvoir faire à un stade donné de la scolarisation, mais, principalement, compte tenu *du processus d'enseignement-apprentissage antérieur*. Compte tenu, autrement dit, de ce que les apprenants ont, grâce aux dispositifs conçus par l'enseignant, *réellement* appris à faire¹². Il apparaît ici que le lièvre dont nous avons parlé est en réalité un kangourou femelle, et qu'elle porte un petit marsupial dans sa poche. Quand commence – quand peut-on considérer qu'a commencé – le processus d'enseignement-apprentissage relatif à l'acquisition de telle compétence de lecture ou d'écriture ?

À cette question, ne sauraient répondre que des professeurs faisant partie d'équipes pédagogiques dont le souci principal serait la cohérence et la continuité des formations, le développement *progressif*, sur la *longue* durée, des compétences collégalement jugées prioritaires¹³. On peut douter que ces professeurs-là soient les plus nombreux. Il n'est pas douteux, en revanche, que chaque enseignant puisse, en début d'année ou en début de séquence pédagogique (au moment où il prend le relais sans trop savoir, la plupart du temps, dans quel couloir ou sur quelle piste son prédécesseur a couru), évaluer les (éventuels) résultats des (éventuelles) interactions pédagogiques relatives au développement de la compétence au

¹² Nous envisageons ici, pour simplifier les choses, que, grâce aux dispositifs en question, les résultats des apprentissages *individuels* ne diffèrent pas au point d'entraîner une trop forte hétérogénéité du groupe. C'est, nous le reconnaissons volontiers, *beaucoup* simplifier les choses...

¹³ « Collégalement jugées prioritaires » ? Et le référentiel, alors ? Le référentiel énumère les savoirs et les savoir-faire qui doivent *obligatoirement* (sigle C) ou *subsidiatement* (sigle S) figurer au nombre des ressources constitutives des compétences à développer – étant entendu qu'une compétence n'est pas une somme de ressources, mais l'aptitude à les mobiliser ensemble à bon escient. Cela étant rappelé (cf. DUMORTIER, 2001a, p.13), nous espérons que notre propos ne sent pas le souffre : les compétences sont à définir – collégalement pour bien faire – compte tenu du référentiel.

développement de laquelle il entend concourir. Il suffit pour cela qu'il dispose d'une épreuve diagnostique appropriée, ou qu'il en conçoive une.

1.3.2. Deuxième changement

Mais l'adoption de la pédagogie « par compétences » entraîne un autre changement en fait d'évaluation. Non moins important que celui qui vient d'être évoqué. Ce changement-là concerne l'évaluation formative.

« Dans une approche par compétences, on ne peut dissocier l'acquisition de connaissances des contextes dans lesquels leur utilisation prend tout son sens. Dès lors, la distinction entre situation d'apprentissage et situation d'évaluation n'apparaît guère pertinente. Toute situation est à la fois occasion d'apprentissage et d'évaluation dans la mesure où elle offre à l'élève la possibilité de porter un regard métacognitif, et donc évaluatif sur ses propres démarches et qu'elle fournit à l'enseignant ou à l'enseignante des données d'observation lui permettant de suivre le développement des compétences. » (LEGENDRE, 2001, p.16). Ça décoiffe, si on veut bien nous passer l'expression.

Plus de distinction entre « situation d'apprentissage » et « situation d'évaluation ». Traditionnellement, cette distinction est, au contraire, très nette. Élèves et professeurs font bien la différence entre les deux situations. Un temps pour l'enseignement-apprentissage des savoirs (savoir-faire, etc.), un temps pour le contrôle des résultats. Et, dans le meilleur des cas, un temps pour la « remédiation ». Qu'est-ce qui peut expliquer, qu'est-ce qui peut justifier que cette « valse à trois temps » se trouve ainsi reléguée au magasin d'antiquités ?

D'abord – et ça n'étonnera personne... pour autant que ce que nous avons dit à propos du premier changement ait été admis par tout le monde – le fait que les tâches préparatoires aux épreuves traditionnellement évaluées et ces épreuves elles-mêmes sortent du même moule. On ne demande pas aux élèves de faire tout seuls autre chose que ce qu'ils ont appris à faire sous la guidance du professeur¹⁴. D'où un effet de « nappé ». Plus de solution de continuité, plus de différence formelle – *en principe* – entre ce que les élèves font pour apprendre et pour prouver qu'ils ont appris.

Insistons sur « en principe ». En pratique, il s'avère souvent que la réalisation répétée d'une tâche – d'une tâche complexe requérant l'intégration de plusieurs savoirs et savoir-faire hétérogènes – ne suffise pas pour permettre à tous les élèves d'atteindre le seuil de performance requis. Des « exercices d'entraînement », portant sur *un* savoir ou *un* savoir-faire isolé, des « activités de structuration », portant sur la conjugaison de *quelques* savoirs ou savoir-faire isolés peuvent être indispensables à bien des apprenants. Nécessité fait loi : on ne renoncera pas à ces exercices ou à ces activités sous prétexte que la charge cognitive qu'ils impliquent est sans commune mesure avec la charge cognitive d'une tâche complexe

¹⁴ Ceci doit s'entendre compte tenu des réflexions précédentes sur le commencement du processus d'enseignement-apprentissage. Tout professeur peut – doit – tabler sur les bénéfices d'une action pédagogique antérieure... à condition de s'être assuré, au moyen d'épreuves à visée diagnostique, du caractère *effectif* de ces bénéfices.

d'intégration¹⁵. On devrait renoncer, en revanche, à (la vieille et ridicule habitude de) prendre en compte, « à la mode des marchands », les scores des apprenants relatifs aux exercices et activités susdits pour évaluer des *compétences*.

La stricte homologie entre activités relevant, traditionnellement, de l'enseignement-apprentissage et activités relevant, selon la même tradition, de l'évaluation ne suffit pas à abolir les frontières entre situation d'enseignement-apprentissage et situation d'évaluation. Si ces frontières tendent à disparaître, cela tient, comme l'affirme Marie-Françoise LEGENDRE (*Id., ibid.*), aux possibilités de métacognition qu'offrent les activités d'enseignement-apprentissage. La métacognition est la « connaissance que l'on a de ses propres processus cognitifs, de leurs produits et de tout ce qui y touche (...). La métacognition (sert), entre autres choses, à l'évaluation active (et) à la régulation (...) de ces processus (...) » (FLAVELL, 1977, *in* RAYNAL et RIEUNIER, 1998, p.226). Je puis lire ou écrire sans réfléchir à ce que je fais pour produire du sens, soit avec le texte d'un tiers, soit dans mon propre texte. Je puis lire ou écrire sans m'intéresser à ce que je fais pour comprendre ou pour donner à comprendre. Mais je puis, à l'inverse, me soucier des opérations dont résulte le « modèle mental » que j'ai formé à partir du texte d'autrui, ou des opérations dont résulte le texte que j'ai formé afin de permettre à autrui d'élaborer un « modèle mental »¹⁶. Plus je m'en soucie, plus j'exerce un contrôle sur ma lecture ou sur mon écriture, plus je me donne d'occasions de réguler l'une ou l'autre compte tenu de mon projet de communication.

Cette régulation est particulièrement importante en contexte scolaire. Dans ce contexte, l'élève *apprend* à lire et à écrire. Il est, explicitement ou non, toujours mis en demeure de *conformer* sa performance à celle que le professeur tient pour référence, celle compte tenu de laquelle le professeur évalue. La métacognition conditionne l'autoévaluation, par l'apprenant, du résultat de ses opérations cognitives. Elle *prévient* l'évaluation magistrale dans la mesure où l'élève s'est approprié les bases et les critères dont se sert le professeur. Elle favorise l'autonomisation, c'est-à-dire,

¹⁵ Par « charge cognitive », nous désignons, pour le dire simplement, la quantité de connaissances mémorisées et la quantité de relations à établir entre ces connaissances nécessaires à l'accomplissement d'une activité. Par exemple, choisir, dans une liste proposée, un moyen adéquat pour connecter deux phrases est une tâche (un exercice d'entraînement) à faible charge cognitive. Construire un texte à partir d'une série ordonnée d'informations, énoncées sous forme de phrases simples ou de groupes nominaux, en trouvant les moyens de connexion adéquats est, toujours par exemple, une tâche (activité de structuration) à charge cognitive beaucoup plus forte. Mais cette charge-là est infime si on la compare à celle que requiert la création d'un texte, laquelle suppose l'invention d'idées, la planification de ce que l'on va écrire, la textualisation (comprenant, entre bien d'autres opérations, la connexion des phrases), la révision, etc.

¹⁶ Un « modèle mental ». Dans cette acception technique, différente de celle que revêtait l'expression quand nous l'avons utilisée plus haut, à propos du modèle de la performance attendue, c'est la représentation que le lecteur construit en traitant la suite des signes constitutifs d'un texte : en échangeant les signifiants contre les signifiés appropriés, en faisant les inférences nécessaires, en opérant le tri de l'essentiel et de l'accessoire dans les propositions successives, etc.

étymologiquement, l'acquisition de la liberté de s'imposer à soi-même la loi¹⁷.

En outre, la métacognition – ou, plus exactement, la verbalisation, par l'apprenant, du souvenir de ses opérations cognitives¹⁸ – fournit à l'enseignant des informations particulièrement intéressantes pour évaluer les compétences et pour réguler l'interaction pédagogique. Si le professeur ne dispose que du *résultat* d'un processus de lecture ou d'écriture – sous la forme minimale d'un choix de réponse (Q.C.M.) ou sous la forme maximale d'un long texte –, il ne peut faire mieux que comparer (certains aspects de) ce résultat avec (les aspects correspondants d') une performance étalon, constater un nombre de points communs ou de divergences et, éventuellement, noter l'élève compte tenu de ceux-là comme de celles-ci. Chacun sait que cette composante-là de l'évaluation prend beaucoup plus de temps qu'elle n'apporte de bénéfiques. Les chances de bénéfice augmentent en proportion de ce que l'enseignant et l'apprenant parviennent à récupérer des démarches mentales dont résultent les écrits *manifestant* une compétence de lecture ou d'écriture.

1.3.3. Troisième changement

La troisième conséquence de l'adoption d'une pédagogie « par compétences » sur l'évaluation des tâches de lecture et d'écriture n'est, tout bien considéré, qu'une implication de la première (confronter les élèves à des tâches complexes d'intégration) et de la seconde (ne plus distinguer situation d'apprentissage et situation d'évaluation).

Pourquoi la traiter séparément alors ? Parce que les habitudes scolaires sont telles que les élèves, leurs parents, les supérieurs hiérarchiques et la plupart des professeurs eux-mêmes ne sont actuellement pas prêts à agir concrètement selon les deux premières conséquences. Notre « système scolaire » comporte des éléments qui déterminent l'interaction

¹⁷ On se gardera bien de croire, en lisant cela, que nous concevons l'enseignement-apprentissage de la lecture et de l'écriture comme une école de conformisme. Nous n'affirmons rien d'autre que ceci : l'enseignant faillirait à son rôle s'il n'utilisait pas des modèles – flexibles, pour bien faire – du bien lire ou du bien écrire. Il faillirait à son rôle si, excipant du droit à l'expression et s'enchantant (comme cela se trouve) de toutes les manifestations de la « créativité divergente », il laissait l'élève lire ou écrire « comme il veut ». Comme l'élève *veut*, c'est aussi comme il *peut*, compte tenu de son environnement socioculturel. Ce vouloir-pouvoir là plaît parfois beaucoup, certes, au professeur en mal de connivence culturelle. Mais l'école ne se justifie que par le projet social de permettre à tous d'acquérir des compétences que tous n'acquerraient pas si tout le soin de l'éducation reposait sur la famille et sur les proches.

¹⁸ Nous n'y faisons guère appel, à vrai dire. Le reconnaître, c'est attirer l'attention sur les limites du questionnement par écrit. Et c'est suggérer de ne pas s'en tenir à nos propositions de questions et de réponses. C'est suggérer d'interroger, aussi souvent que possible, les élèves – un ou deux, en pratique – sur les démarches mentales dont procèdent *leurs* réponses. Mais quel est le bon moment pour ce faire ? Nous doutons que ce soit celui de la correction : il est loin, alors, le souvenir de leurs démarches mentales ! Il faut intervenir *on line*, dans l'instant qui suit la production de la réponse. Et alors, adieu la possibilité d'accomplir la tâche dans les délais impartis, adieu la possibilité d'évaluer la compétence de chaque élève sur la base des performances individuelles. En pédagogie, on ne peut jamais espérer gagner sur tous les tableaux : nous suggérons donc de ne pas mettre toutes les chances de gain dans le petit panier que nous avons tressé ici.

pédagogique, qui l'empêchent d'être ce qu'elle devrait être si l'on s'en tenait, strictement, à la logique d'un enseignement-apprentissage *de compétences*. Ainsi en va-t-il, notamment, de l'organisation des cursus par année, du bulletin, des contrôles périodiques, des examens. Ces quatre éléments du « système » ont, certes, ces derniers temps, fait l'objet de révisions. Mais de révisions si timides (ou si prudentes) qu'elles n'ont abouti à rien d'autre qu'à un « placage » du terme « compétences » sur des pratiques d'évaluation conçues compte tenu de modèles pédagogiques antérieurs (la pédagogie du transfert, pour une part, la pédagogie par objectifs, pour une autre), dont on proclame l'obsolescence.

Nous ne donnons pas ainsi à entendre, bien sûr, que l'on pourrait faire l'économie d'une organisation des cursus. Nous ne donnons pas à entendre que le bulletin, les contrôles périodiques, les examens sont des éléments du système inutiles. Nous disons simplement que, *dans leur forme actuelle*, ils *parasitent* la logique des compétences¹⁹. Mais il faut bien s'en accommoder. En attendant la conception et l'institution d'une organisation moins rigide (de type modulaire). En attendant une révision plus radicale des procédures de certification et des instruments de communication des résultats des apprentissages.

Dans l'attente de cela, les professeurs sont toujours tenus de se prononcer, régulièrement, plusieurs fois par an, sur ce que l'on nomme, désormais (mais encore faudrait-il que la chose convienne au mot), les compétences des élèves. Pour ce qui nous concerne, cela signifie que, régulièrement, plusieurs fois par an, ils devraient imposer à ces derniers des tâches complexes d'intégration (en lecture comme en écriture) et les préparer, dans l'intervalle de ces épreuves, à accomplir ces tâches, compte tenu de critères de performance déterminés. Les y préparer implique les confronter à des tâches analogues, dans le cadre de l'évaluation formative.

Cette fois, ce n'est plus un lièvre qu'on débusque. C'est un autre animal à grandes oreilles, mais de la famille des pachydermes... Pour une tâche donnant lieu à une performance prise en compte dans la certification, combien de tâches préparatoires ? Tous les apprenants seront-ils obligés de les effectuer ou n'y obligera-t-on que ceux dont les performances n'atteignent pas (du premier coup, après un second essai, etc.) le niveau de performance seuil ? Quel usage fera-t-on des scores réalisés aux essais ? Quelles incidences sur le temps et le mode de correction ? Un éléphant qui surgit ? Non. Un troupeau d'éléphants ! Tellement formidable qu'on en viendrait à comprendre ceux qui veulent continuer à calculer ainsi : une dictée + un exercice de vocabulaire + une analyse de mots + une restitution de matière + une rédaction = (dans l'hypothèse d'une note moyenne) *la* compétence en français...

« Il ne faut pas, disait SARTRE, désespérer Billancourt. » Mais faut-il, pour cela, cacher la vérité ? L'adoption de la pédagogie « par compétences » a de lourdes implications. Mieux vaut le dire clairement. Et ajouter qu'il faudra du temps – beaucoup d'aide aussi, au bénéfice des enseignants de terrain – pour qu'elle produise ses effets.

¹⁹ « Comment une note pourrait-elle faire état des apprentissages après une étape de deux mois ou même de dix mois ? » (LUSSIER, 2001, p. 35).

Mais nous annonçons une troisième conséquence sur l'évaluation et le suspense a assez duré. L'évaluation dont nous allons dire un mot à présent est – on l'aura deviné puisqu'il vient d'être question de bulletin, de contrôles, d'examens et puisque nous avons déjà traité d'évaluation diagnostique et d'évaluation formative – celle à laquelle on confère une fonction *certificative*²⁰.

Si l'on certifie des compétences, il convient que l'épreuve à laquelle est soumis l'apprenant non seulement soit analogue aux tâches préparatoires, mais encore qu'elle se déroule dans des conditions analogues à celles des tâches préparatoires – qui, elles-mêmes, idéalement, ne devraient pas trop différer des conditions d'exercice d'une compétence en contexte extra-scolaire (DABENE, 1994, p.30)²¹. C'est ce que certains appellent la condition de validité « écologique » d'une épreuve. Par ailleurs, à l'avis de maints spécialistes de l'évaluation des compétences, « la démonstration d'une ou de plusieurs compétences ne peut être limitée à un moment précis comme on le ferait pour un examen terminal » (SCALLON, 2001, p. 22). Cela donne évidemment à réfléchir sur la valeur des informations prises à partir des performances accomplies dans le contexte *exceptionnel* de la certification (contrôles, examens). Et cela incite à pratiquer, à des fins certificatives (puisqu'on n'y échappe pas !), l'évaluation *continue*.

Ce n'est assurément pas une entreprise sans difficulté que la pratique d'une évaluation *continue* à fonction *certificative*. Surtout si l'on perpétue les examens, qui confèrent à certaines performances une valeur distinctive. Deux ornières sont à éviter. La première est l'addition mathématique des scores réalisés pour chaque performance relative à une même compétence. C'est le progrès de l'apprenant qui importe et la preuve qu'il donne, à un moment donné, qu'il est capable d'atteindre ou de franchir le seuil de performance minimal, dont il s'est petit à petit rapproché. La seconde ornière est la valeur *décisive* d'un examen qui, quelque précaution que l'on prenne, ne sera jamais une épreuve dont on pourra assurer la validité « écologique ». Qui ne permettra jamais, par ailleurs, et quoi que l'on fasse,

²⁰ Certains disent évaluation *sommative*. L'évaluation sommative permet de faire le bilan des apprentissages. Mais, quel que soit ce bilan, on peut prendre toutes sortes de décisions. On ne *certifie* pas nécessairement que les compétences de l'apprenant sont ou ne sont pas suffisantes pour qu'il poursuive avec fruit ses études ou pour qu'il exerce une profession.

²¹ D'où l'*urgence* d'une réflexion sur les modes de lecture et d'écriture valorisés par l'école. Ils *ne* peuvent *pas* être exactement ceux qui ont cours *extra muros*. Parce que rien ne justifierait alors la peine d'être *intra muros*. Parce que l'école a vocation de réduire les différences culturelles dépendant du hasard de la naissance dans tel ou tel milieu. Parce que l'école doit faire contrepois à des tendances (consumentistes, entre autres) que favorise la société marchande. Mais les modes de lecture et d'écriture scolaires *ne* peuvent *pas* non plus être radicalement différents de ceux qui se rencontrent hors de l'enceinte scolaire. Parce l'*enjeu* pédagogique c'est, précisément, en ce qui nous concerne surtout, l'adoption de pratiques culturelles (de lecture, d'écriture, mais aussi de parole et d'écoute) qui émancipent les jeunes des fatalités sociales. C'est cela même – cette occasion d'émanciper – que l'on peut rater ou saisir en persuadant d'adopter des pratiques *suffisamment* distinctes de celles auxquelles les jeunes recourraient « spontanément » (c'est-à-dire par la force de leur « habitus culturel »), et *suffisamment* proches de ces dernières pour ne pas leur sembler exclusivement scolaires (*cf.* pour des vues contrastées sur la question, s'agissant de la lecture, DUFAYS, GEMENNE ET LEDUR, 1996 ; ROSIER, DUPONT ET REUTER, 2000 ; DUMORTIER, 2001b.).

de se prononcer sur l'état de développement de *toutes* les compétences que l'on aura exercées. Qui le permet actuellement d'autant moins qu'on en a réduit la durée de façon drastique.

Le principe d'une évaluation certificative continue – c'est-à-dire la prise en considération prioritaire, sinon exclusive, du progrès de l'apprenant – serait, évidemment, bien plus facile à respecter s'il existait, au plus haut échelon de la prise de décision, un accord sur les compétences fondamentales de lecture et d'écriture à développer soit tout au long de la scolarité obligatoire, soit à partir de tel ou tel stade de cette dernière. Mais un tel accord *sur des compétences* (et non sur le caractère indispensable d'un chimérique savoir-lire ou savoir-écrire général) implique, insistons-y, une double et très sérieuse réflexion sur, d'une part, les familles de situation de communication où l'écriture et la lecture interviennent, sur, d'autre part, les niveaux de performance exigibles, dans chacune de ces familles de situations, à tel ou tel stade de la scolarité. Un exemple clarifiera sans doute le propos.

On s'accorderait assez facilement, sans doute, sur l'utilité de la compétence à réagir de manière critique (en manifestant un accord ou un désaccord avec les arguments avancés) aux écrits persuasifs véhiculés par les mass médias. Toute grossière qu'elle est, cette définition donne une *première* idée de la famille de situations de communication à laquelle cette compétence est relative. Une *première* idée aussi de la famille de tâches scolaires préparant les élèves à affronter ces situations. Mais quels seront les caractères types des espèces d'écrits persuasifs médiatiques que l'on utilisera, disons au premier, au deuxième et au troisième degrés de l'enseignement secondaire ? Et en quelles formes (orales, écrites) exigera-t-on qu'à ces divers niveaux d'études, les élèves puissent faire preuve de leur réception critique ? Quand, au terme d'un débat sérieux, il aura été répondu à de telles questions, l'évaluation certificative continue cessera de poser les épineux problèmes qu'elle pose aujourd'hui aux enseignants qui hésitent – et c'est tout à leur honneur – à certifier quoi que ce soit *en matière de développement des compétences*. Leurs scrupules tiennent au fait qu'on leur demande de se prononcer sur les résultats d'un apprentissage dont ils ne savent pas avec assurance quand il a commencé ou quand il aurait dû commencer. C'est dans les limites d'une année, sur la base d'un nombre infime de tâches plus ou moins pertinentes au développement de la compétence concernée qu'il leur est demandé une évaluation certificative, alors que *tous* les spécialistes de l'évaluation des compétences plaident pour une prise d'informations sur la longue – la très longue – durée...

2. QUE PEUT-ON EN FAIRE ?

Nous étant longuement attardé sur l'évaluation des compétences, nous pourrions être beaucoup plus bref sur les utilisations possibles des outils d'évaluation qui sont ici mis à la disposition des professeurs. Mais avant d'évoquer ces possibilités (usage des outils à des fins diagnostique, formative ou certificative), nous voudrions prévenir une question bien

légitime : les « modèles d'épreuves »²² du projet COMPÉFRADE sont-ils *effectivement* des outils pertinents à l'évaluation de compétences ?

2.1. Des outils pertinents ?

Nous ne sommes, bien sûr, pas les mieux placés pour répondre à cette question. Nous nous y hasardons néanmoins, sans intention de désamorcer la critique. Nous y inciterons plutôt, en en donnant l'exemple.

Ce n'est pas trahir un secret que révéler ceci : les concepteurs des outils d'évaluation ci-après n'ont généralement pas, *d'entrée de jeu*, formulé une définition *précise* des compétences que permettaient d'évaluer les outils auxquels ils ont travaillé. Ce défaut de précision pourrait passer pour vice rédhibitoire aux yeux de ceux qui estimerait, non sans raison d'ailleurs, que, pour construire un outil d'évaluation d'une compétence, la moindre des choses est d'être au clair sur la compétence à évaluer. Mais cette raison-là est théorique.

Pour nous, prime la raison pratique. Et elle s'exerce, la raison pratique, sur des objets spécifiques : des textes, qu'il s'agit de faire apprendre à lire ou à écrire – l'évaluation n'étant qu'une *composante* de l'apprentissage. Or, pour la plupart, ces textes, à comprendre ou à imiter, sont *déjà* là. Leur présence est forte. Ce sont eux qui, souvent, donnent au professeur l'idée – la première idée – d'une tâche. Les concepteurs des épreuves sont donc généralement partis d'une première idée suscitée par les textes, ou, sinon par les textes seuls, par une expérience pédagogique où des textes et des activités entraînent en conjonction. Ils se sont dit, en somme : « Que pourrions-nous faire avec ça, *dont nous disposons*, qui nous est un peu familier, pour construire une tâche pertinente à l'évaluation d'une compétence ? » Le but n'a pas été perdu de vue²³, mais, si on peut le dire ainsi, il n'a pas toujours été si clairement perçu *dès le départ* que le choix des moyens d'y arriver coulait de source. La construction des épreuves s'est souvent opérée en boucle, par ajustement progressif des items à la définition, elle aussi progressive, d'une compétence. Il me semble que ce processus circulaire, de praticien, vaut bien celui, linéaire, des théoriciens. Un pas de plus : il nous semble qu'il vaut mieux. En tout cas pour l'heure : ce n'est vraiment pas le moment de faire perdre aux professeurs une bonne partie de leurs moyens en faisant du but « compétence » une sorte d'impératif catégorique.

En fin de parcours, trouve-t-on bien des épreuves permettant l'évaluation de compétences ? Chacun a fait son possible pour que ce soit le cas. Au demeurant, comme nul n'ignore, « la plus belle fille du monde... ». Mais en donnant ce qu'elle a, la plus belle fille du monde peut décontenancer.

²² C'est l'expression que nous avons retenue pour désigner chacun de nos outils d'évaluation. « Épreuve », qui fait partie du vocabulaire usuel des enseignants, nous a paru moins rébarbatif que « tâche-problème ». Quant à « modèle », quoique le mot puisse égratigner certaines susceptibilités, il nous a paru pertinent pour attirer l'attention sur le fait que nos propositions ont un caractère à la fois reproductible et adaptable : chacun peut s'en inspirer pour faire (mieux) selon ses besoins.

²³ Lorsqu'il s'est estompé, le collectif de travail a joué son rôle : il a rappelé qu'il ne s'agissait pas de mettre au point une activité quelconque, mais de concevoir une tâche *évaluable*, l'évaluation devant porter sur une *compétence*, c'est-à-dire une aptitude à accomplir *des* tâches apparentées.

Notamment parce qu'elle ne donne pas tout tout de suite. Expliquons-nous. De certaines épreuves, prises séparément, on pourrait dire : « Ce ne sont pas là des tâches complexes d'intégration ; ce ne sont donc pas des épreuves pertinentes à l'évaluation d'une compétence. » Nous en convenons volontiers. Nous irons même plus loin : c'est une réflexion que l'on pourrait faire à propos de *la plupart* de nos épreuves. Si on les considère une par une. Ou si on n'en considère que des morceaux. Et, surtout, si l'on refuse d'envisager notre travail comme l'ébauche d'un chantier, auquel chacun est appelé à œuvrer. Soit la compétence de rédiger un texte à partir de schémas, de graphiques ou la compétence à manifester la compréhension d'une nouvelle réaliste, qui ont été, entre autres, deux de nos points de mire : qui peut croire qu'une tâche ou deux, pensées pour favoriser et faciliter l'évaluation, suffisent à la développer, partant à l'évaluer valablement ?

Nous dirions volontiers, sans le sentiment de nous avancer plus que de raison, que les concepteurs des tâches ont constamment été tirillés par deux désirs contraires : celui de construire des items intégrateurs, des items faisant appel à des connaissances hétérogènes, et celui de détailler les savoirs nécessaires pour réagir à ces items dans les formes entraînant le *satisfecit* magistral. De là une certaine boiterie des épreuves, qui peut passer pour une infirmité au regard de gardiens sévères de l'orthodoxie des compétences.

Mais qu'on réfléchisse un instant à ceci : une tâche complexe d'intégration, en lecture ou en écriture, s'énonce, la plupart du temps, en quelque mots ou en quelques phrases. Ne s'agit-il pas toujours, en somme, de manifester la compréhension d'un texte relevant d'une classe distincte, en des formes déterminées, ou de construire un texte relevant lui aussi d'une telle classe et qui soit compréhensible par un public déterminé ? Selon une perspective pédagogique, les compétences de lecture et d'écriture sont-elles autre chose que des aptitudes à prouver, dans des situations-problèmes scolaires que spécifient les valeurs particulières de la variable « texte » (à lire et/ou à écrire), de la variable « espace-temps » (en classe ; à domicile, dans tel délai) et de la variable « adjuvant » (avec ou sans outils de référence, par exemple), l'appropriation de manières de lire ou d'écrire ayant fait l'objet d'un apprentissage antérieur ? Que signifient des injonctions telles que « Commentez ce début de roman » ou « Rédigez une synthèse à partir de ces trois documents » ? Elles signifient : « Faites ça comme vous avez appris à le faire – ou comme je pense que vous l'avez appris – ou comme vous auriez dû, quoi qu'il en soit, l'avoir appris. ».

Dès lors qu'ils ne pouvaient tabler avec assurance sur les bénéfices d'un enseignement-apprentissage antérieur relatif à l'exercice des compétences qu'ils décidaient d'évaluer, les auteurs des « modèles d'épreuves » ci-après se sont estimés dans la quasi-obligation morale d'explicitier ce qui demeure souvent à l'état d'implicite dans un énoncé de tâche complexe d'intégration et de passer en revue les savoirs indispensables à la réalisation de cette tâche. Ils auraient pu opter pour une sèche énumération (pour pouvoir faire ceci, il faut s'être approprié tel, tel et tel savoirs), ils ont préféré construire des items relatifs à ces savoirs.

On pourrait considérer qu'a été ici généralisé l'attendu qui préside à l'élaboration de tout questionnaire de compréhension en lecture. Ou, plus

exactement, l'attendu qui y présiderait si le recours à cette sorte de questionnaire n'était devenu un automatisme pédagogique. Je procède au moyen d'un questionnaire attendu que mes élèves ne sont pas (encore) capables de faire part de leur compréhension sans les balises et les tréplins que constituent des questions. En procédant de la sorte, il est incontestable que je m'éloigne formellement (parfois foncièrement, si mon questionnaire manque de cohérence et si mes questions sont trop simples) de la tâche complexe d'intégration.

Le choix de procéder par questionnaire pour exercer les élèves à manifester leur compréhension des textes et pour évaluer cette compréhension fait tellement partie des habitudes que nul ne s'étonne – ni ne s'inquiète – des dérives analytiques liées au questionnement. On ne se demande pas (en tout cas nous n'entendons pas beaucoup de gens qui se demandent) si, en optant pour la communication sur la compréhension via des questionnaires, on ne court pas un trop grand risque de rompre les amarres qui rattachent les pratiques textuelles aux pontons des compétences. La sensibilité aux dérives analytiques est, en revanche, beaucoup plus vive dès qu'il s'agit de pratiques d'écriture. Sans doute parce les combats sont plus âpres, dans ce domaine-là, entre partisans des procédures pédagogiques *top-down* (celles qui partent des discours en situation pour définir les besoins langagiers des élèves) et les partisans des procédures pédagogiques *bottom-up* (celles qui partent des ressources langagières pour nantir les élèves des connaissances indispensables pour affronter les situations de discours).

C'est donc vraisemblablement sur le versant « écriture » que notre entreprise prêtera particulièrement, en certains cas tout au moins, le flanc à la critique. Nous ne pensons pas, au demeurant, qu'elle soit là plus critiquable qu'ailleurs. Mais le souci scrupuleux de pourvoir les élèves des savoirs qui conditionnent les résolutions de problèmes d'écriture impatientes vite les personnes pressées (souvent pour de fort honorables raisons, convenons-en) de mettre les élèves dans une situation de production de texte « vraiment significative ». Comment ne pas s'étonner de ce qui suit? On trouve normal de passer deux heures, parfois trois, à « lire » (communiquer la compréhension d') un texte et inopportun de consacrer deux ou trois fois plus de temps à apprendre à en écrire un. Est-on vraiment convaincu que ceci ne soit pas (beaucoup) plus difficile que cela? Nous gagerions volontiers qu'il est plus malaisé, pour la majorité des élèves, de devenir les scripteurs que l'école attend que les lecteurs qu'elle entend former. Nous le gagerions volontiers en songeant que, pour l'heure tout au moins, devenir le lecteur attendu, c'est se familiariser avec des modes de questionnement et apprendre à *réagir* à des sollicitations magistrales plus ou moins prévisibles. Devenir bon scripteur, c'est faire acquérir les ressources nécessaires à une *prise d'initiatives* beaucoup plus forte.

2.2. Pertinents à quoi ?

Même si l'on estime *excessives* certaines de nos tendances analytiques, nous ne pensons pas qu'on puisse contester que, la plupart du temps, les items constitutifs d'un « modèle d'épreuve » *concourent* bien à exercer et à permettre l'évaluation d'une compétence. Étant donné ce que nous avons dit précédemment en 1.3., on comprendra que nous ne nous appliquions pas à distinguer la vertu qu'ont ces items d'exercer une compétence et la

vertu qu'ils ont d'en rendre possible l'évaluation. On exerce une compétence en évaluant la performance qui la manifeste : c'est le principe même de l'évaluation formative, à laquelle nos outils peuvent servir.

2.2.1. L'évaluation formative

Mais si nos « modèles d'épreuves » sont pertinents à l'évaluation formative, c'est à la condition que l'utilisateur ne s'en tienne pas à l'établissement et à l'enregistrement des scores. C'est à la condition d'un *retour* sur l'épreuve, d'une *réflexion* sur les items (les plus communément) ratés ou réussis.

Certes, nous avons souvent veillé à ce que ces items favorisent la métacognition, à ce qu'ils forcent l'élève à... mettre en question sa réponse, mais ce soin est insuffisant pour que l'épreuve produise tous les effets dont elle est susceptible, en fait de formation. On sous-estime généralement beaucoup ce que nous nommerons l'inopérante docilité de fort nombreux élèves : ils font ce qu'ont leur dit de faire, parfois avec beaucoup d'application, mais plus rarement avec beaucoup d'implication personnelle dans le dispositif d'apprentissage. Ils ne deviennent pas les (co)opérateurs de la formation que l'on souhaiterait. Ils s'adonnent à des « occupations » scolaires comme ils s'adonneraient à maintes distractions : sans que cela ne change rien, sans que cela ne convoque le passé ni n'engage l'avenir. Il y a, dès lors, un risque majeur à puiser dans cette (modeste) banque d'épreuves : celui d'installer (ou de renforcer) des habitudes de « consommation », de « consumérisme » pédagogique. Les élèves consomment un produit censé propice au développement des compétences, mais sans bénéfice. Du moins sans les bénéfices escomptés en fait de compétences. Car on pourrait ne pas tenir pour rien que l'activité ait requis leur attention, canalisé leurs énergies pendant une heure ou deux.

L'efficacité de nos épreuves, en tant qu'outils d'évaluation *formative* est conditionnée non seulement par un *feed back* sur les performances, mais encore par leur utilisation *en série*, dans le cadre d'une *séquence pédagogique*. Il aurait fallu une équipe de recherche bien plus nombreuse ou un temps de travail beaucoup plus long pour mettre à la disposition des professeurs des séries plus fournies d'épreuves relatives à une même compétence. Au demeurant, il n'est pas sans avantage qu'*ostensiblement* chaque série soit pauvre : nul utilisateur de notre travail ne peut ainsi se sentir contraint de suivre un parcours que nous aurions conçu de bout en bout.

2.2.2. L'évaluation diagnostique

Il est clair qu'une épreuve estimée valide pour évaluer une compétence est valide quelle que soit la fonction assignée à l'évaluation. Qu'on diagnostique un état de développement de cette compétence censément atteint grâce à tâches antérieures, qu'on renforce ou développe cette compétence dans le cadre de l'évaluation formative ou qu'on certifie un état de développement de ladite compétence correspondant à un objectif d'apprentissage déterminé, c'est toujours à l'évaluation *de cette compétence-là* que l'épreuve est pertinente. C'est dire que le fruit de notre travail peut être utilisé à des fins de diagnostic (ou à des fins de certification : il en sera question plus loin) comme il peut l'être à des fins de formation. Mais cela ne signifie pas que chacun de nos « modèles d'épreuve » peut, dans tous les cas de figure –

dans toutes les séquences pédagogiques imaginables – être employé pour établir un diagnostic.

Sur quoi porte, en l'occurrence, le diagnostic ? Sur l'état de développement d'une compétence qui devrait, « normalement », « en principe », *avoir été* assuré par les professeurs qui se sont antérieurement attachés à développer cette compétence-là. Une évaluation diagnostique est une prise d'informations destinée à fonder en réalité les décisions concernant le début ou la poursuite d'une entreprise. Faute d'informations fiables sur ce que savent déjà les élèves et sur ce qu'ils savent déjà faire, on s'expose au risque d'exiger d'eux des performances dont ils ne sont pas capables. Si un professeur envisage une séquence sur la lecture-écriture de textes informatifs concernant des sujets culturels, par exemple, il est prudent qu'il s'assure de la solidité des bases sur lesquelles il compte bâtir. À cette fin, il soumettra ses élèves à une épreuve diagnostique censée le renseigner sur l'état de ces bases. Il ne s'agira pas d'une épreuve présentant le même nombre de difficultés ou le même degré de difficulté que celle(s) qu'idéalement il devrait déjà avoir choisie(s) pour l'évaluation certificative. Il devrait s'agir d'une épreuve comparable, sur le rapport des difficultés, à celle(s) précédemment utilisée(s) pour certifier le dernier stade de développement de la compétence envisagée. C'est cela même qui nous fait dire que si, dans l'absolu, toutes nos épreuves sont propices à l'établissement d'un diagnostic, relativement à telle ou telle séquence conçue par un enseignant singulier, elles ne sont pas toutes bonnes à cela.

Soit dit en passant, qu'on utilise nos épreuves à des fins de diagnostic ou de certification, on considérera avec beaucoup de circonspection les indications fournies, au début de chaque épreuve, sur le public destinataire. Tout le monde sait bien (même si tout le monde ne le dit pas) que, dans l'état actuel de notre enseignement, il est déraisonnable de penser qu'une épreuve destinée, par exemple, à des élèves qui entament leur quatrième année puisse être réussie, dans *n'importe quel* établissement, dans *n'importe quelle* classe, par une *confortable* majorité d'élèves. Les obstacles que surmonteront assez aisément la plupart des élèves dans la classe X d'une école Y seront infranchissables pour la plupart des élèves d'une classe correspondante, dans une autre école, ou par les élèves de la classe Z, dans la même école Y.

2.2.3. L'évaluation certificative

De même qu'elles peuvent servir à établir un diagnostic, nos épreuves peuvent servir à la certification. Qu'il nous soit toutefois permis d'insister sur ce point : nous ne les avons pas conçues pour qu'elles servent *exclusivement* à ça et nous serions fort affligés d'apprendre qu'on leur assigne cette *seule* fonction. Répétons-le, au risque de lasser : il n'est pas d'évaluation certificative acceptable, du point de vue de l'éthique professionnelle, sans une préparation très sérieuse aux épreuves imposées à des fins de certification. Or, dans notre chef, il ne fait aucun doute que nous proposons, pour chaque compétence à évaluer, des séries d'épreuves qui demandent à être complétées.

Ayant plaidé un peu plus haut (1.3.3) pour une évaluation certificative continue, nous voudrions ici saisir l'occasion de préciser trois choses. La première est que la continuité de la certification *d'une compétence*

n'entraîne pas, en principe tout au moins, la possibilité de prendre en considération, *pour certifier cette compétence*, les notes obtenues à des exercices d'entraînement ou à des activités de structuration. Or, incontestablement, certains de nos items, sinon certaines de nos épreuves tout entières, relèvent de ces exercices ou de ces activités. La seconde est que l'évaluation certificative continue ne dispense non seulement ni des premiers ni des secondes, mais qu'au surplus elle ne dispense pas de tâches complexes d'intégration imposées comme pur entraînement. La troisième précision que nous tenons à apporter, c'est que, continue ou non, l'évaluation à fonction certificative se fonde sur des tâches complexes d'intégration qui, en principe toujours, ne sont *jamaïs* plus difficiles que celles auxquelles les apprenants ont eu affaire dans un cadre d'entraînement.

En guise de non-conclusion

Les mathématiques ne sont généralement pas le fort des professeurs de français, mais le calcul élémentaire n'a pas de secret pour eux. Ainsi donc : une tâche complexe d'intégration, au titre du diagnostic + des tâches complexes d'intégration + des activités de structuration + des exercices d'entraînement, tout cela au seul titre de la préparation + des tâches complexes d'intégration, au titre de l'évaluation certificative. Dans l'hypothèse (très favorable) où les élèves ne se lasseraient pas d'autant d'insistance, combien de compétences pourrait-on développer en procédant ainsi ? Fort peu, assurément. Sauf si...

Sauf si, sachant désormais ce que l'on sait des implications de la pédagogie par compétences, on se remettait à réfléchir, ensemble, sur les possibilités qu'offre une entreprise de développement de ces compétences sur la longue durée. Mais ceci est une autre affaire, bien plus épineuse que celle qui nous a été confiée. Nous en resterons donc là, sans conclure, puisque l'affaire est...

À suivre.

J.-L. DUMORTIER (ULg)

Promoteur de la recherche

J.-L. DUFAYS (UCL)

J.-M. ROSIER (ULB)

Copromoteurs de la recherche

Olivier DEZUTTER (UCL)

Philippe ERNOTTE (ULB)

Anne JURASZ (ULg)

Hélène LEFEVRE (UCL)

Thierry ROLAND (ULg)

Chercheurs

BIBLIOGRAPHIE

- CANVAT, K., *Enseigner la littérature par les genres. Pour une approche théorique et didactique de la notion de genre littéraire*, Bruxelles, De Boeck & Larcier, 1999, coll. « Savoirs en pratique ».
- DABENE, M., « Principes et méthodes d'une évaluation didactique du savoir-lire », dans BOYER, J.-Y., DIONNE, J.-P. et RAYMOND, P., *Évaluer le savoir-lire*, Montréal, Les éditions Logiques, 1994, coll. « Théories et pratiques dans l'enseignement ».
- DE KETELE, J.-M. et ROEGIERS, X., *Méthodologie du recueil d'informations. Fondements des méthodes d'observations, de questionnaires, d'interviews et d'études de documents*, Bruxelles, De Boeck-Wesmael, 1993, coll. « Pédagogies en développement ».
- DUFAYS, J.-L., GEMENNE, L. et LEDUR, D., *Pour une lecture littéraire I*, Bruxelles, De Boeck & Larcier, 1996, coll. « Formation continuée ».
- DUMORTIER, J.-L., « Formation des enseignants de français et compétences de communication », dans *Puzzle*, n°10, juin 2001. (a)
- DUMORTIER, J.-L., *Lire le récit de fiction. Pour étayer un apprentissage : théorie et pratique*, Bruxelles, De Boeck & Larcier, 2001, coll. « Savoirs en pratique ». (b)
- DUMORTIER, J.-L., « Sur quoi les professeurs devraient-ils s'accorder pour que convergent leurs efforts lorsqu'ils fabriquent et emploient des outils d'évaluation des compétences en lecture ? », communication à l'Université d'été du CIFEN (23 août 2002), à paraître dans *Puzzle*.
- GENETTE, G., « Genres, types, modes », dans *Poétique*, n°32, 1977.
- LEGENDRE, M.-F., « Favoriser l'émergence de changements en matière d'évaluation des apprentissages », dans *Vie pédagogique*, n°120, sept-oct. 2001.
- LUSSIER, D., « Le bulletin scolaire au carrefour des pressions sociales et des mythes qui l'entourent », dans *Vie pédagogique*, n°120, sept-oct. 2001.
- ROEGIERS, X., *Une pédagogie de l'intégration. Compétences et intégration des acquis dans l'enseignement*, Bruxelles, De Boeck & Larcier, 2000, coll. « Pédagogies en développement ».
- RAYNAL, F. et RIEUNIER, A., *Pédagogie : dictionnaire des concepts clés. Apprentissage, formation, psychologie cognitive*, Paris, ESF, 1998 (1^{ère} éd. 1997).
- ROSIER, J.-M., DUPONT, D. et REUTER, Y., *S'approprier le champ littéraire. Propositions pour travailler l'institution littéraire en classe de français*, Bruxelles, De Boeck & Larcier, 2000, coll. « Savoirs en pratique ».
- SCALLON, G., « Pourquoi évaluer?... Quelle Question ! », dans *Vie pédagogique*, n°120, sept-oct. 2001.