

Pédagogies en développement

Évaluer pour former

Outils, dispositifs
et acteurs

Gilles Baillat
Jean-Marie De Ketele
Léopold Paquay
Claude Thélot
(sous la direction de)



de boeck

Sommaire

Introduction		7
<i>Gilles Baillat, Jean-Marie De Ketele, Léopold Paquay, Claude Thélot</i>		
Partie 1	LES DÉMARCHES D'ÉVALUATION : POUR QUOI? POUR QUI?	15
Chapitre 1	L'évaluation dans le système éducatif : pour quoi faire? À quelles conditions l'évaluation est-elle utile? <i>Claude Thélot</i>	17
Chapitre 2	De l'analyse à l'évaluation d'une pratique professionnelle <i>Anna Jorro, Jean-Jacques Maurice</i>	29
Chapitre 3	Que fait-on des évaluations internationales dans le système éducatif français? <i>Jean-Claude Emin</i>	43
Chapitre 4	Quelques aspects éthiques de l'évaluation <i>Bernard Rey</i>	57
Partie 2	LES OUTILS D'ÉVALUATION AU SERVICE DE L'APPRENTISSAGE	69
Chapitre 5	Conceptualiser les outils d'évaluation des apprentissages <i>Linda Allal</i>	71
Chapitre 6	Les évaluations diagnostiques en France : le diagnostic, une notion et une exploitation en classe qui évoluent depuis 1989 <i>Gérard Brézillon, Françoise Champault</i>	83

Chapitre 7	Les outils d'évaluation ouverts, ou la nécessité de clés de fermeture <i>François-Marie Gerard</i>	99
Chapitre 8	Formation, compétences et constructivisme : le référentiel d'évaluation, outil de cohérence dans les programmes de formation en santé <i>Florence Parent</i>	111
Partie 3	LES OUTILS D'ÉVALUATION AU SERVICE DE LA FORMATION DES ENSEIGNANTS ET DES FORMATEURS	141
Chapitre 9	Portfolios et entretiens de co-évaluation : des leviers de la professionnalisation des jeunes enseignants <i>Lucie Mottier Lopez, Sabine Vanhulle</i>	143
Chapitre 10	Pertinence et validité de tâches complexes dans l'évaluation de futurs enseignants <i>Pierre-François Coen, Aline Galland, Isabelle Monnard, Nicolas Perrin, Yviane Rouiller</i>	159
Chapitre 11	Évaluation des compétences et changements identitaires <i>Christine Baeriswyl, Loïc Clavier, Marlyse Pillonel</i>	177
Chapitre 12	Évaluer la formation initiale des enseignants <i>Stéphane Brau-Antony, Christine Jourdain</i>	191
Annexe	Référentiel de compétences professionnelles Professeur des Écoles stagiaires	205
Conclusion		207
Liste des figures et tableaux		209

Introduction

Gilles Baillat, Jean-Marie De Ketele, Léopold Paquay, Claude Thélot

Cet ouvrage est issu du dix-huitième colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE) qui s'est tenu à Reims les 24-26 octobre 2005. Tenu pour la première fois au sein d'un Institut Universitaire de Formation des Maîtres (IUFM), ce colloque, consacré à la question: «Comment évaluer? Outils, dispositifs et acteurs», se proposait de mettre en discussion les fonctions, mais aussi les modalités de l'évaluation, au moment où le développement de l'approche par compétences en formation renouvelle ces questions. La qualité des contributions a permis de construire cet ouvrage structuré en trois parties, consacrées chacune à une question: d'abord «à quoi l'évaluation peut-elle et doit-elle servir?», puis «quels sont les outils de l'évaluation, et pour quel apprentissage?», enfin «comment utiliser l'évaluation lors de la formation des enseignants?».

1 À QUOI L'ÉVALUATION PEUT-ELLE ET DOIT-ELLE SERVIR ?

Claude Thélot (chapitre 1) nous rappelle en premier lieu les deux types d'usages de l'évaluation: l'usage externe qui renvoie à l'information des décideurs, financeurs, parents d'élèves, opinion... et l'usage interne, auprès des acteurs du système, lorsqu'il s'agit de le réguler. Mais dans les deux cas la question reste bien celle des conditions pour obtenir des réponses effectives à ces deux types de besoin. En effet, pour que l'évaluation soit susceptible d'informer réellement la société sur l'état de l'école, deux grandes conditions doivent être réunies: la crédibilité de l'évaluation, ainsi que sa diffusion la plus large et la plus systématique possible.

Il est vrai que l'évaluation peut aussi servir à l'amélioration interne de l'école, mais il faut dans ce cas poser la question de l'autonomie des acteurs. La liberté pédagogique des enseignants appelle en effet que soit pris réellement en compte ce qui se fait effectivement dans les classes, et implique donc un regard différent de celui qui est donné par les inspections. Par ailleurs, la question de l'autonomie est aussi celle des établissements, autonomie qui reste très limitée dans certains pays. Cette double question

est cependant cruciale pour envisager une utilisation réelle des évaluations : l'évaluation des pratiques peut en effet nourrir la liberté pédagogique, mais aussi la formation des enseignants, alors que l'évaluation des établissements est indispensable pour préparer un véritable contrat d'établissement.

Si les deux regards (regard externe, auto-évaluation) sont tous deux importants, ils doivent cependant être combinés et, de toute façon, ne pas faire obstacle au regard critique vis-à-vis de l'évaluation, de ses procédures de ses résultats. Cette question du regard critique pose aussi, pour Claude Thélot, celle du rôle des cadres éducatifs qui ont pour mission de prendre en charge l'essor des usages.

La contribution d'Anna Jorro et de Jean-Jacques Maurice (chapitre 2) cible la relation entre analyse et évaluation des pratiques, deux compétences réflexives et professionnelles clés pour l'activité enseignante. Après avoir rappelé leurs caractéristiques, les auteurs présentent deux démarches de recherche conduites en didactique du français et en ergonomie cognitive. La première recherche présente quatre analyseurs de l'activité de l'enseignant en classe : le discours de l'enseignant, la transmission des savoirs, la manière dont l'enseignant communique avec les élèves et apprécie leur travail, la manière de réagir aux événements imprévisibles et de s'adapter en situation. La seconde recherche décrit l'expérience de l'enseignant dans une activité de préparation de situations d'apprentissages et montre que les enseignants expérimentés sont en mesure de réaliser des anticipations du niveau de difficulté des tâches proposées aux élèves et d'agir en conséquence, sans pouvoir pour autant satisfaire complètement à la fois les exigences du groupe classe et celles de tous les élèves.

Les deux auteurs insistent par ailleurs sur les tensions, méthodologiques et éthiques, induites par la problématique des relations entre pratiques observées et efficacité de l'enseignant et invitent à une évaluation plurielle pour en maîtriser mieux les effets. Cette évaluation plurielle pourrait prendre en compte trois registres : l'action prescrite, qui renvoie aux gestes structurants l'activité ; l'action effective, qui renvoie à l'agir professionnel de l'enseignant ; l'action réfléchie qui renvoie à l'expérience vécue.

La question posée par Jean-Claude Emin (chapitre 3) : « Que fait-on des évaluations internationales dans le système éducatif français ? » n'est pas destinée à faire l'objet d'une lecture ironique. Pour l'auteur en effet, et contrairement à une idée répandue, la France prend bien en compte ces évaluations et en particulier celles produites par PISA. Par exemple, les analyses conduites à la suite des premiers résultats de PISA ont aidé à prendre la décision de mettre en place le « socle commun », défiant en cela la tradition française solidement établie des disciplines. Tout en mettant en garde vis-à-vis de classements globaux moyens entre pays, qui dissimulent les disparités, l'auteur rappelle par exemple que la question du redoublement

des élèves pointée par PISA en ce qui concerne l'école française a effectivement conduit les responsables à promouvoir des mesures pour prévenir cette pratique.

L'auteur situe enfin l'avenir de PISA, du point de vue des responsables français, en référence à trois questions. La première est celle de la prise en compte des contextes : elle est trop faible actuellement dans ce type d'évaluation, au motif qu'elle se heurte à la question de l'extrême complexité (et donc de la multitude de variables à envisager) de l'objet. La seconde concerne la place de la France dans ces enquêtes internationales, place actuellement insuffisante du fait du faible engagement français dans PISA depuis l'origine. Jean-Claude Emin affirme cependant la volonté de la Direction de l'Évaluation et de la Prospective de s'y impliquer davantage à l'avenir. La troisième enfin pose la question de la combinaison entre évaluation internationale et évaluations nationales, combinaison souhaitée par les autorités françaises qui affirment par là même leur souhait de ne pas se contenter des actuels « compromis internationaux » préluant aux choix de PISA.

La première partie de l'ouvrage se clôt par les réflexions introduites par Bernard Rey sur quelques aspects éthiques de l'évaluation (chapitre 4). En effet, si l'évaluation ne peut exister sans valeur, l'auteur rappelle que ces dernières ne sont plus que très rarement reconnues comme universelles, ce qui nous oblige à réinterroger l'acte d'évaluer. Bernard Rey insiste notamment sur les pièges cachés au sein de l'évaluation collégiale qui dilue la dimension individuelle comme l'évaluation des collectifs (classes, établissements...) et occulte le fait qu'au bout du compte, ce sont bien des individus qui existent. Autre piège : celui de « l'évaluation technique » qui divise le travail et répartit les rôles au risque de faire perdre à l'évaluateur le sentiment d'être concerné par l'autre qui est sujet de l'évaluation, de le faire tomber dans illusion de la mesure objective.

Si l'évaluation peut impliquer un dialogue entre pairs, il ne faut cependant pas oublier qu'elle instaure une asymétrie. Il faut donc rompre avec l'empathie qui mettrait aux prises deux sujets, au profit d'un travail portant sur les actes à évaluer. Cette nouvelle prise en compte de l'évalué doit pouvoir éventuellement aboutir à sa contestation de l'évaluateur, à la mise en cause de sa subjectivité.

Bernard Rey pose enfin la question de l'éthique de la formation. Il est en effet indéniable que les transformations récentes, telles que les référentiels, sont importantes et effectivement positives, ne serait-ce que parce qu'elles permettent de mieux objectiver. Mais il ne suffit pas que l'élève sache : il faut encore qu'il sache mobiliser à bon escient. Or il est patent que cet élève est souvent victime d'une « rupture de contrat », dans la mesure où l'évaluation ne porte pas toujours sur des éléments pour lesquels il a été entraîné. Au

total, l'évaluateur a tout à gagner à être conscient que son travail n'est pas vraiment objectif, et à s'astreindre à distinguer son jugement et son pouvoir.

2 LES OUTILS AU SERVICE DE L'APPRENTISSAGE

Les outils servent-ils bien l'apprentissage? Telle est la question de cette seconde partie. Mais quels outils? Pour quels apprentissages? Et donc quels outils pour quels apprentissages?

Comme l'affirme Linda Allal au départ de sa contribution (chapitre 5), le concept d'outil est bien plus polysémique qu'il n'y paraît. Il est d'ailleurs souvent associé, voire assimilé à tort ou à raison, à d'autres concepts proches, tels les instruments, l'outillage, les supports, les dispositifs... Ce constat pousse Linda Allal à proposer un modèle de conceptualisation des outils d'évaluation, appelé DISA, croisant deux dimensions binaires. La première distingue le contexte de l'évaluation et l'outillage qui existe dans ce contexte ; la seconde différencie ce qui est instauré par l'institution et ce qui est réellement investi par les acteurs. Pour bien illustrer l'intérêt de ce modèle, l'auteur l'applique à deux exemples d'application du portfolio : le premier dans le cadre de l'enseignement primaire, le second dans la formation initiale des enseignants du primaire. Dans les contextes où elles ont été mises en œuvre, ces deux applications poursuivent, comme c'est assez souvent le cas dans la réalité, deux fonctions conjointes communes, formative d'une part, mais aussi certificative (par une démarche sommative) d'autre part. Tel qu'utilisé ici comme support d'une évaluation interactive entre enseignant, parents et élèves dans le premier cas, entre formateur et étudiant dans le second cas, l'intérêt du portfolio est de confronter les traces des apprentissages réalisés aux apprentissages à réaliser et d'agir sur les représentations qu'en ont les acteurs mis en présence.

La contribution de Gérard Brézillon et Françoise Champault (chapitre 6) fait état d'une longue expérience de la création et de l'utilisation d'outils d'évaluation diagnostique en France. Cette contribution est intéressante à plus d'un titre. En premier lieu, il s'agit au départ d'une initiative prise par les autorités du ministère de l'Éducation nationale : il s'agit d'outils mis à la disposition des enseignants, mais conçus par des enseignants, des inspecteurs et même des élèves, devant cependant respecter un cahier des charges commun à toutes les équipes de conception et d'expérimentation. En second lieu, les outils sont créés pour remplir deux fonctions possibles : une fonction d'orientation des apprentissages au début de l'année scolaire et une fonction de régulation des apprentissages en cours d'année. En ce qui concerne la fonction d'orientation, il s'agit d'une épreuve commune obligatoire passée selon un protocole standardisé dans le but, non pas de classer les élèves par niveaux de performances, mais par groupes de besoins. En ce qui concerne la fonction de régulation, les enseignants ont la possibilité de puiser libre-

ment, et donc sans aucune obligation, dans une large banque de données structurée en liaison étroite avec l'organisation du programme. Si celui-ci est centré sur le développement des compétences, il faut noter que le cadre conceptuel français n'est pas en tout point équivalent aux cadres conceptuels d'autres pays. Ainsi les compétences s'inscrivent dans des «champs», chacun d'entre eux nécessite la mobilisation de «capacités»; les compétences se déclinent en «composantes» qui s'évaluent au moyen d'un éventail d'items. Chaque item est conçu de telle façon qu'il puisse fournir une information diagnostique précise. Les outils n'ont donc pas la prétention d'évaluer la maîtrise des compétences, mais de mettre en évidence des causes ou des freins susceptibles de bloquer leur apprentissage.

Partant du principe que l'on ne peut évaluer des compétences par un ensemble de questions de sélection indépendantes, mais par des questions ouvertes ou de production qui supposent une véritable résolution de problème ou l'effectuation d'une tâche complexe, François-Marie Gerard (chapitre 7) tente de répondre aux objections qui ont présidé à leur rejet en fournissant ce qu'il appelle des «clés de fermeture» pour une utilisation pertinente des questions ouvertes dans le cadre d'apprentissages centrés sur le développement de compétences. Il propose six clés de fermeture qu'il illustre à partir de trois types d'outils: (i) le portfolio, (ii) l'évaluation des compétences à travers des situations complexes et (iii) l'évaluation de projets, de programmes ou de systèmes.

La contribution de Florence Parent (chapitre 8) s'inscrit dans le champ de la formation des professionnels de la santé. Il s'agit d'une tentative audacieuse dans le sens où elle aborde le problème de l'évaluation (certificative et formative ou de régulation) comme partie intégrante d'un processus large de co-construction avec tous les acteurs concernés (depuis les autorités ministérielles jusqu'aux praticiens des divers terrains), depuis la phase de conception des différents types de référentiels (de compétences, de formation et d'évaluation) avec le souci d'assurer une cohérence systémique jusqu'aux outils d'évaluation à mettre en œuvre sur le terrain. Intéressante par ces aspects, cette contribution est aussi intéressante à d'autres titres, notamment le fait qu'elle repose sur un cadre conceptuel non strictement équivalent au cadre présenté dans la contribution française; ainsi, elle introduit des concepts comme ceux de «paliers de compétences», de «familles de situations», d'«objectif continué d'intégration». Un souci particulier est accordé au comment s'assurer que la maîtrise d'une compétence dans une situation donnée est transférée dans d'autres situations; c'est là une question importante.

S'il est vrai que la compréhension se développe à travers la comparaison de l'objet d'étude dans des contextes différents et avec des finalités différentes, la lecture de ces quatre contributions permettra à chaque lecteur de faire progresser sa compréhension du bon usage des outils d'évaluation dans le contexte qui est le sien.

3 LES OUTILS D'ÉVALUATION AU SERVICE DE LA FORMATION DES ENSEIGNANTS ET DES FORMATEURS

Dans leur contribution sur l'usage des portfolios en formation d'enseignants, Lucie Mottier Lopez et Sabine Vanhulle (chapitre 9) examinent cet outil sous deux angles distincts : comme outil visant l'appropriation des savoirs, et comme outil de médiation, en particulier au moment des entretiens de co-évaluation. Après avoir rappelé leurs cadres théoriques, les auteurs présentent les deux dimensions de cet instrument : la fonction d'archivage des compétences construites, mais aussi celle de témoignage et de reflet de ce qui est en train de se construire

Inscrit dans une didactique des savoirs professionnels, le portfolio pose la question de l'articulation entre apprentissage, évaluation et développement professionnel. Dans cette perspective, l'alternance devient le moyen de produire les écrits témoignant d'une réflexivité en construction.

La contribution de Lucie Mottier Lopez et Sabine Vanhulle est illustrée par la présentation de l'analyse d'une étude de cas d'un portfolio ayant donné lieu à un entretien de co-évaluation. Cette étude montre en particulier l'intérêt du dispositif pour aller au-delà de l'évaluation certificative, dimension à laquelle le portfolio est souvent uniquement ramené.

Dans leur contribution du chapitre 10, Pierre-François Coen, Aline Galland, Isabelle Monnard, Nicolas Perrin et Yviane Rouiller s'attachent à la redoutable question de la pertinence et de la validité de tâches complexes dans l'évaluation des futurs enseignants. S'interrogeant sur la relation existant entre ces tâches complexes et les situations réelles auxquelles elles font référence, les auteurs rappellent en premier lieu les différentes manières d'aborder la pertinence des tâches complexes : la première interroge la relation entre formation initiale et pratiques réelles ; la seconde examine la relation entre pratiques professionnelles et dispositifs de formation ; la troisième cible le point de vue des formés.

Le chapitre décrit ensuite les résultats d'une recherche réalisée à l'aide d'entretiens menés auprès d'une douzaine d'enseignants auxquels six tâches complexes ont été proposées. Les résultats mis en évidence amènent les auteurs à s'interroger sur la pertinence des notions de pertinence et de validité. De plus, les résultats montrent l'importance de la question du référentiel implicite qui fonde les pratiques enseignantes, ce qui conduit les auteurs à mettre l'accent sur la nécessaire distinction entre référentiel de métier et référentiel de formation.

La pratique de l'évaluation des compétences n'est pas sans effets sur l'évalué mais aussi sur l'évaluateur. Christine Baeriswyl, Loïc Clavier et Marlyse Pillonel (chapitre 11) se proposent d'illustrer ce changement identitaire à travers la restitution de résultats issus de deux enquêtes.

Réalisée au sein d'une Haute École suisse en soins infirmiers, actuellement soumise à un renouvellement de son modèle de formation, la première enquête montre certes que le changement du dispositif peut s'opérer sans rencontrer de résistances a priori. Mais, si la dualisation des lieux de formation entraîne une modification des positionnements, en particulier celui de l'enseignant, dans le processus de formation, l'ouverture spatio-temporelle du processus de formation ne suffit pas à changer les anciens rôles et à en introduire de nouveaux. Les déverrouillages majeurs à opérer dans ce sens sont en effet difficiles à réaliser concrètement.

La seconde enquête présente, dans le cadre d'une recherche en cours dans un IUFM, les changements identitaires se produisant chez l'évalué à l'occasion de l'évaluation des compétences dans un dispositif de formation professionnelle. Après avoir rappelé les pratiques d'évaluation en IUFM, les auteurs situent leurs repères conceptuels, avant d'insister sur la confiance nécessaire entre le tuteur et le stagiaire évalué ; cette confiance implique des «arrangements locaux» avec l'évaluation. Ce sentiment de confiance se nourrit de dimensions personnelles et de sentiments de compétences mis en exergue par l'auto-évaluation, mais aussi de dimensions sociales qui visent, elles, une évaluation catégorisante. Les auteurs concluent à l'influence de l'évaluation sur une partie de la construction identitaire, aussi bien chez l'évaluateur qui passe du statut d'expert à celui de médiateur, que chez l'évalué pour qui la régulation strictement académique laisse place à la régulation de l'expérience professionnelle.

Dans le chapitre 12, Stéphane Brau-Antony et Christine Jourdain abordent la question de l'évaluation de la formation initiale des enseignants en France, par les enseignants stagiaires en formation. Conçue en partie comme une réponse aux critiques formulées contre la formation actuelle des enseignants, l'outil d'évaluation présenté dans cette contribution vise deux enjeux : celui de la prise en compte du référentiel de métier introduit récemment en France et celui de la régulation de la formation.

Les résultats présentés ici sont issus d'enquêtes réalisées à l'aide d'un questionnaire construit autour de quatre domaines de compétences, mais aussi du sentiment énoncé par les enseignants stagiaires sur l'origine de la construction de leurs compétences. Ces résultats mettent en évidence l'importance des stages et celle des maîtres de stages («maîtres formateurs» en France) pour la formation des compétences professionnelles des professeurs des écoles. Le travail personnel ainsi que le mémoire professionnel sont cependant reconnus comme des contributions importantes pour ces compétences. Les futurs professeurs de collège et de lycée valorisent eux aussi les stages mais insistent également sur le rôle des formateurs universitaires, en particulier dans le champ des compétences relevant de l'éthique professionnelle.

Les auteurs concluent finalement à la cohérence du point de vue des enseignants stagiaires avec l'enjeu principal de leur formation à cette étape de leur cursus : l'accent mis sur le terrain correspond en effet à la nécessité fortement ressentie de se doter de « compétences pour la classe ». Cette priorité n'occulte cependant pas les autres apports de la formation, également reconnus par les stagiaires interrogés comme complémentaires dans le cadre de leur formation initiale.

Au terme de ces contributions multiples et différenciées, le lecteur ne peut que se rendre à l'évidence : s'interroger sur les outils et les dispositifs d'évaluation implique nécessairement de se poser les questions fondamentales du sens des apprentissages et de la formation, sans oublier que ceux-ci sont toujours situés.

Chapitre 1

L'évaluation dans le système éducatif : pour quoi faire ? À quelles conditions l'évaluation est-elle utile ?

Claude Thélot

Dans ses diverses acceptions, l'évaluation est présente, et même très présente, dans tout système éducatif. Non pas que chacune des formes le soit, mais en les regroupant toutes, cette fonction présente une ampleur et un volume très substantiels. Car il y a de nombreuses et diverses acceptions, d'importance très variable. La notation des élèves, les examens (organisation et passage), l'inspection des enseignants sont trois formes traditionnelles qui prennent beaucoup de temps ; les examens ont d'ailleurs pris beaucoup d'importance ces deux dernières décennies avec la massification et la démocratisation des études. D'autres formes d'évaluation, plus récentes, prennent moins de place et de temps, et sont parfois insuffisamment développées : l'évaluation des élèves (indépendamment de leur notation), l'évaluation des enseignements, l'évaluation des établissements scolaires et universitaires, l'évaluation de tel ou tel espace ou organisation éducatif (une région, etc.) et, plus généralement, l'évaluation du système éducatif pris dans son ensemble.

Toutes ces formes d'évaluation, traditionnelles ou plus récentes, très ou peu développées, ne me semblent avoir de justification profonde qu'en fonction de leur usage : selon cette conception instrumentale de l'évaluation, elle n'a pas sa justification en elle-même, elle n'est qu'un instrument au service de la politique éducative. D'où, dans cette perspective, le caractère crucial de l'interrogation : à quoi l'évaluation peut-elle et doit-elle servir ?

Cette question, qui n'est pas nouvelle peut trouver la réponse suivante. L'évaluation a ou, au moins, doit viser à avoir deux grandes utilités. Une **utilité externe**, destinée à informer la société (décideurs, parents d'élèves, employeurs, etc.) sur l'état du service éducatif, notamment sur sa qualité, ses résultats, mais aussi son coût et son fonctionnement. La **seconde utilité est interne**, c'est-à-dire informe les acteurs du système (enseignants, élèves, administrateurs, etc.) au moins sur les mêmes éléments, ce qui les aide à réfléchir sur leurs actions et l'organisation (du système, de l'établissement, etc.) et à les infléchir pour les améliorer. Selon sa forme précise (notation, examen, évaluation des enseignements, etc.) l'évaluation remplit chacune de ces deux utilités avec une pondération différente : parfois plutôt utile de façon externe (exemples : examens, publications sur le système pris dans son ensemble), parfois plutôt de façon interne (exemple : inspection des enseignants).

C'est dans le cadre de cette exigence – satisfaire correctement à au moins l'une des deux utilités – que se situe cet article. Non pas pour réexaminer la réponse donnée à la question : « À quoi sert l'évaluation ? », qui sera considérée comme pertinente, mais pour en tirer quelques conséquences ; plus précisément quelques conditions, sur la politique d'évaluation et ses effets, ses procédures, son organisation, ses outils, pour que cette réponse soit effective. Dans cet examen, toutes les formes d'évaluation ne seront pas citées ou présentes à l'esprit avec la même ampleur, mais il serait facile d'adapter certaines des réflexions proposées à celles peu traitées.

L'article est constitué de deux parties, consacrée chacune à l'une des deux utilités : d'abord qu'implique le souci de satisfaire à l'utilité externe (1), puis qu'implique celui d'être utile en interne (2)¹.

1 CONDITIONS POUR QUE L'ÉVALUATION INFORME RÉELLEMENT LA SOCIÉTÉ SUR L'ÉTAT DE L'ÉCOLE

Si l'on veut réellement informer la société sur l'état du système éducatif, et même plus, contribuer à structurer le débat public, d'une part, et produire une évaluation utile, c'est-à-dire dont se servent effectivement les acteurs de

1. Pour avoir une description récente de l'évaluation dans le système éducatif français, on se reportera au rapport que Claude Pair a établi pour le Haut Conseil de l'évaluation de l'école (HCÉé) : *Forces et faiblesses de l'évaluation du système éducatif en France*, Rapport du HCÉé, n° 3, décembre 2001, Ministère de l'Éducation nationale.

la société auxquels cette information est destinée, d'autre part, deux grandes conditions semblent devoir être satisfaites par l'évaluation : elle doit d'abord **être crédible** (1.1), elle doit ensuite faire l'objet d'une **grande diffusion** (1.2).

1.1 Être crédible

La crédibilité de l'évaluation est capitale, et il faut bien reconnaître qu'elle est difficile à satisfaire, et que souvent elle n'est pas satisfaite. Quelques exemples élémentaires peuvent être fournis ici. Ainsi, les notes doivent-elles être crédibles aux yeux des parents (mais aussi des élèves, ce qui montre que l'exigence de crédibilité n'est pas nécessaire pour le seul usage externe de l'évaluation), en particulier entre disciplines ou entre professeurs. Ainsi, les évaluations sur le niveau et les acquis des élèves doivent-elles l'être sous peine d'être non seulement ignorées mais même contestées de l'opinion et des décideurs : le débat public risquerait d'en être non pas enrichi mais perturbé et biaisé. Ainsi encore l'évaluation d'un collège doit-elle être crédible aux yeux des parents d'élèves et des élus locaux dès lors qu'elle entend se substituer à la rumeur, et parfois s'opposer à elle.

Comment accroître (ou, selon le cas, créer) la crédibilité des évaluations pour qu'elles servent réellement en dehors de la sphère éducative, pour les personnes et institutions à qui elle est destinée ? On peut proposer de s'attacher à trois critères : les organes qui évaluent doivent être multiples et indépendants (et perçus comme tels) ; la méthodologie de l'évaluation doit être transparente (et souvent cette transparence est une conséquence nécessaire de l'indépendance) ; enfin, l'évaluation doit s'efforcer de se « rapprocher » des usagers potentiels.

Il ne saurait y avoir de monopole de l'évaluation, de l'évaluateur : la **multiplicité** des évaluateurs, des instances est indispensable pour rendre crédible cette démarche et ses résultats. Et chacun d'eux doit être **indépendant** (ou impartial, adjectif qui convient mieux à certaines formes d'évaluation). Chaque organe, chaque évaluation peut reposer en partie sur des éléments communs aux autres (des données, des segments de méthode, etc.), et même le doit pour éviter un coût excessif de la fonction prise dans son ensemble. Mais elles ne sauraient s'identifier les unes aux autres, une source de la crédibilité reposant précisément sur le fait que des conclusions communes ou proches soient tirées à travers (et en dépit de) la variété des points de vue.

Contrairement à une opinion commune, l'indépendance, nécessaire, ne requiert pas d'être extérieur à l'État, ou au système éducatif. Les organes d'évaluation peuvent être soit externes soit internes à l'administration centrale (à la condition expresse qu'ils ne soient pas les principaux responsables de la politique ou de l'état qu'ils vont évaluer), publics ou privés, dépendant de

la recherche ou extérieurs à elle, l'important n'est pas là. Car l'indépendance ne se mesure pas d'abord en termes de position ou d'extériorité. Elle s'apprécie à l'aune de trois facteurs : la liberté de définir son programme de travail (d'évaluation), la mise en œuvre d'une méthodologie scientifique le moins contestable possible (de façon être de grands professionnels de l'évaluation), la liberté de rendre publiques les évaluations. Là encore il n'est pas difficile de citer des exemples où ces critères ne sont pas absolument satisfaits, ou, mieux, d'illustrer les conséquences qu'ils pourraient ou devraient avoir. Par exemple, au titre du « professionnalisme », il faudrait, beaucoup plus qu'aujourd'hui, former les futurs enseignants à évaluer et à noter les élèves (au lieu de se contenter, comme trop souvent, de mettre en relief le caractère conventionnel et dépendant de circonstances des évaluations et des notes). Ou encore, les indicateurs sur lesquels l'évaluation d'une école primaire ou d'un collège reposera, au moins partiellement, devront être réfléchis puis rendus opérationnels avec et par des experts et des praticiens.

La méthodologie d'évaluation ne doit pas seulement être la meilleure possible, elle doit aussi, et c'est le deuxième critère de la crédibilité, être **connue**. Les médias, les décideurs ne sauraient faire confiance aveuglément à une évaluation. À nouveau les illustrations de ce principe sont faciles à proposer. En voici une, particulièrement importante : si l'on veut préciser à l'intention des employeurs potentiels futurs ce que représentent les examens, quelles compétences ils garantissent que les élèves qui les ont passés avec succès maîtrisent vraiment, il faut faire connaître et diffuser largement le type d'épreuves sur lesquels ils s'appuient. Sans doute, dans le cas présent, faut-il d'ailleurs faire plus que rendre transparentes ces modalités : elles doivent être définies, avec les partenaires sociaux eux-mêmes, c'est-à-dire les futurs utilisateurs, dans la perspective des compétences dont on souhaite que le diplôme garantisse la maîtrise ; cette exigence est capitale lors des examens professionnels. Autre illustration de l'exigence de transparence : il est nécessaire que les exercices sur lesquels se fondent les « évaluations-bilans », celles qui sont destinées à mesurer l'état des acquis des élèves, soient connus : c'est ainsi qu'ils seront crédibles. Il est vrai que cette exigence est difficile à satisfaire complètement, car elle entre en contradiction avec une autre règle de méthode, le souci de reproduire de loin en loin les mêmes exercices, pour mesurer les évolutions au cours du temps. Car, dès lors que les exercices sont connus, ce souci de les répéter dans le temps nourrit un risque de bachotage. Mais ceci est une autre question sur laquelle je ne m'étends pas.

Enfin, le troisième critère renforçant beaucoup la crédibilité des évaluations consiste à se **rapprocher** le plus possible des préoccupations des personnes et organisations susceptibles de les utiliser. Ceci n'est pas très souvent fait car dans les évaluations c'est souvent des critères et des modalités internes ou techniques qui sont retenus. Donnons deux conséquences d'un tel « rapprochement ». D'abord à propos des résultats et des types de l'évalua-

tion. Il ne faut sûrement pas s'en tenir à des moyennes, ou des agrégats très « macros » : se rapprocher de l'utilisateur éventuel consisterait à développer beaucoup les analyses en termes de dispersion (exemples : fournir non seulement la taille moyenne des classes, ou le niveau moyen des élèves, mais la proportion de classes très légères ou très lourdes, et le niveau des élèves faibles ou forts, etc.), et même à développer des résultats individuels : c'est bien cette école primaire-ci, cette classe ou ce professeur-là qu'il faut évaluer, et publiquement, ce qui impose bien entendu de sérieuses précautions et soucis méthodologiques (cf. ci-dessous). Ensuite à propos de ce qu'on choisit d'évaluer. Les évaluations des élèves sont souvent, lorsqu'elles sont conduites dans le système éducatif, structurées à partir des exigences des programmes : on cherche à apprécier à un niveau donné ce que les élèves maîtrisent parmi ce qu'ils sont censés maîtriser, c'est-à-dire parmi ce qu'indiquent les programmes du niveau. Ceci, qui est parfaitement légitime, éloigne cependant du parent d'élève, du décideur, des médias, etc., bref de la société. Car à l'extérieur c'est plutôt aux exigences de cette dernière, actuelles ou passées, qu'il faudrait répondre plutôt qu'à celles des programmes. Rendre crédibles les évaluations (et les examens) des élèves devrait conduire à beaucoup développer, peut-être même au détriment des évaluations actuelles bâties sur les programmes, des évaluations construites à partir des compétences exigées dans nos sociétés et/ou présentes à l'esprit et dans le débat publics.

1.2 Diffuser et rendre compréhensible

La seconde grande condition pour accroître l'utilité externe des évaluations conduites dans le système éducatif consiste à les rendre systématiquement publiques et à les **diffuser au maximum**. Il y a là une grande politique de diffusion (ou, si l'on veut, de transparence, là aussi, mais sur les résultats de l'évaluation, et pas seulement sur ses outils et ses méthodes) à conduire et à développer. Elle offre par rapport aux pratiques ou aux tentations fréquentes en la matière deux ruptures.

D'abord au niveau général, il est absolument nécessaire que les responsables politiques donnent leur accord *a priori* à une telle diffusion, et accordent les moyens nécessaires pour qu'elle ait lieu. En ce sens, le ministre ou, lorsqu'il s'agit d'évaluations interministérielles, le Gouvernement dans son ensemble doit non seulement accepter l'idée que son domaine de responsabilité soit évalué, mais qu'il le soit publiquement. Cette reconnaissance fait partie d'une démocratie moderne, où les citoyens et les usagers veulent et même exigent des services publics de qualité, mais il faut reconnaître qu'une telle position requiert de la part de l'homme politique de premier plan un certain courage (elle n'est d'ailleurs pas toujours assumée).

Sur un autre plan, il faut bannir, en raison même de cette condition, l'idée d'évaluation purement interne, qui n'aurait pas d'écho public. Cette

position est à prendre par les instances d'évaluation elles-mêmes (y compris dans tous les domaines et pas seulement l'éducation) : toute évaluation, sauf cas exceptionnel nommément délimité, doit être **publique**, c'est-à-dire que le citoyen doit pouvoir y accéder (et les possibilités créées par Internet rendent cette exigence accessible). La confidentialité doit devenir l'exception, alors que jusqu'à maintenant c'était l'accessibilité qui était l'exception. Retournement complet de perspective, accordé à la démocratie actuelle.

Il faut enfin tirer une dernière conséquence du choix fait en faveur de la diffusion : dès lors que le principe politique est admis, la diffusion à grande échelle des évaluations doit s'accompagner d'une pédagogie pour mettre leurs enseignements à la portée des citoyens, des utilisateurs potentiels. D'où le souci de **vulgariser** l'évaluation de l'École (ses méthodes comme ses conclusions), en en présentant les conclusions de façon claire et accessible ; d'où aussi une éducation en profondeur des utilisateurs potentiels (journalistes, politiques, parents d'élèves, etc.) à la compréhension de ce qui est ainsi diffusé. Car il ne suffit pas de vouloir diffuser. Il faut en outre créer les conditions auxquelles ce qu'on rend accessible sera reçu, compris, donnera lieu à discussion, sera utilisable, utilisé. (On retrouvera d'ailleurs cet effort d'éducation à propos des acteurs de l'École, dès lors qu'en interne cette fois l'on souhaitera que les résultats de l'évaluation soient réellement utilisés.) Les modalités de cet accompagnement pédagogique du public potentiel à la fois national et local (y compris, par exemple, sur un bassin d'éducation, ou auprès de l'ensemble des partenaires d'un collège ou d'un lycée) sont évidemment diverses. Il nécessite une ferme volonté et d'importants moyens.

2 CONDITIONS POUR QUE L'ÉVALUATION SOIT UN LEVIER D'AMÉLIORATION INTERNE DE L'ÉCOLE

Ici, la problématique de départ est un peu différente, même si on en tirera des conséquences dont certaines sont communes à celles qui viennent d'être tirées. Car en interne, l'évaluation est inséparable de l'autonomie. Plaider pour que l'évaluation et ses résultats aient des conséquences en interne, c'est-à-dire aide les acteurs du système éducatif à améliorer leurs pratiques au service d'une École plus efficace et plus juste, cela n'a de sens que si les acteurs disposent d'une certaine autonomie. L'évaluation est dans cette perspective la contrepartie naturelle, et même nécessaire, de l'autonomie. À l'inverse, si l'on plaide pour un accroissement de l'autonomie, il faut dans le même mouvement mettre en place les conditions d'une évaluation efficace et dont on tirera des conséquences concrètes. Il y a alors ici deux sujets principaux, auxquels sont rattachées les trois priorités que devrait avoir la politique d'évaluation à usage interne.

Premier sujet, la **liberté pédagogique des enseignants**. Cette liberté est souhaitable, nécessaire, même si elle ne peut être regardée comme illi-

mitée : tout enseignant, tout éducateur est libre de son action, sous réserve qu'elle conduise au progrès des élèves². Elle doit être nourrie des réussites et des pratiques des autres et elle doit évidemment être évaluée. Aussi, la liberté pédagogique appelle-t-elle deux types d'évaluation : d'une part, celle des pratiques éducatives, ce qui se fait réellement dans les classes et, si on sait le dégager, ce qui, au moins dans certains contextes, «devrait» se faire pour favoriser au mieux les progrès des élèves ; d'autre part, l'évaluation-inspection individuelle des enseignants.

Le second sujet lié à l'usage interne de l'évaluation est l'**autonomie des établissements**. Dans beaucoup de pays elle est notable et, en France, nombre de rapports et d'experts en recommandent, à juste titre, l'accroissement³. Mais cela ne saurait se faire sans mettre en œuvre une grande politique d'évaluation des établissements scolaires (et universitaires). D'où la nécessité de construire une vraie évaluation des écoles primaires, des collèges, des lycées, des universités et de définir les modalités selon lesquelles on en fera usage.

En France, l'inspection des enseignants existe et mobilise des moyens non négligeables (même si elle n'est pas assez fréquente), mais elle devrait être profondément renouvelée, car elle n'est guère satisfaisante et ne sert guère à améliorer les pratiques des enseignants. L'évaluation des établissements, elle, est pratiquement inexistante et est donc entièrement à construire. Ces deux évaluations sont par nature individuelles (chaque enseignant, chaque établissement) ; elles concernent donc un nombre trop élevé de personnes ou d'unités pour être engagées sans hiérarchisation. D'où l'énoncé de critères, avec l'idée de choisir les segments sur lesquels on fera porter pour l'essentiel l'effort, parce que c'est là que le levier sera *a priori* le plus efficace, c'est-à-dire que l'utilité de l'évaluation au service d'une amélioration des pratiques (pédagogiques, éducatives, de direction et d'animation) sera la plus grande. Par exemple, on pourra convenir qu'il faut commencer par évaluer-inspecter, avec des modalités renouvelées, les enseignants débutants, ceux qui le demandent, ceux qui seraient signalés comme particulièrement défaillants, ceux enfin de certains niveaux (en particulier en début de cycle : CP, sixième). De façon analogue, on pourrait commencer par évaluer une

2. Jamais, cette liberté n'avait été en France l'objet d'un article de loi. La loi d'orientation et de programme d'avril 2005 la consacre et, même, la renforce en quelque sorte en lui donnant le pas sur les recommandations possibles du futur conseil pédagogique de l'établissement (que la loi crée) : «La liberté pédagogique de l'enseignant s'exerce dans le respect des programmes et des instructions du ministre chargé de l'Éducation nationale et dans le cadre du projet d'école ou d'établissement avec le conseil et sous le contrôle des membres des corps d'inspection. Le conseil pédagogique prévu à l'article L.412-5 ne peut porter atteinte à cette liberté. (...)».

3. Encore récemment, par exemple, le rapport de la Commission du débat national sur l'avenir de l'École : «Pour la réussite de tous les élèves», Paris, Sceren-CNDP-La Documentation française, octobre 2004.

partie des écoles élémentaires, une partie des collèges, parties spécifiées à partir de critères précis (par exemple les établissements de notoriété « mauvais », ou, au contraire, « bons », ou dans des contextes « difficiles »), l'évaluation des lycées intervenant alors dans un second temps.

Indépendamment de la hiérarchisation des terrains sur lesquels renouveler et/ou développer l'évaluation des personnes et des unités, il est nécessaire, pour que cette politique soit réellement utile, de préciser deux choses importantes : quel type d'utilité envisage-t-on pour ces évaluations ? Selon quelle méthodologie va-t-on les rénover et/ou les construire ? Répondre d'abord à la première question permet ensuite de mieux traiter la seconde, c'est-à-dire de dégager les conditions auxquelles ces trois priorités, l'évaluation des pratiques, celle des enseignants, celle des établissements, auront l'utilité qui leur aura été assignée.

2.1 Quelle utilisation des évaluations des pratiques, des évaluations des enseignants et des évaluations des établissements ?

L'évaluation des pratiques éducatives devrait avoir deux usages. D'abord servir de support, d'aide à la liberté pédagogique des enseignants, c'est-à-dire être diffusée à tous pour leur faire connaître les pratiques des autres et, éventuellement, les pratiques les plus efficaces et, par là, favoriser leurs propres décisions, c'est-à-dire alimenter leur liberté. En ce sens, le Haut Conseil de l'évaluation de l'école avait, dans son avis sur le sujet, écarté l'idée de « modèle universel de bonnes pratiques », mais avait parlé de « canevas de progressions pédagogiques qu'ils [les enseignants] puissent adapter à la diversité des élèves »⁴. Canevas sur lequel, donc, les enseignants broderont, ce qui signifie que la liberté pédagogique, loin d'être bridée, sera nourrie. L'évaluation des pratiques éducatives doit en second lieu servir dans le cadre de la formation des enseignants, initiale et continue : elle devrait être un des principaux outils des IUFM.

L'inspection-évaluation des enseignants, profondément renouvelée, devrait, elle aussi, avoir plusieurs utilités : contrôler la qualité du système éducatif, reconnaître et valoriser l'effort, la réussite et le mérite des enseignants et en tirer les conséquences en matière de ressources humaines. Affirmer le principe de cette dernière utilité est déjà central (et constituerait, au moins en France, un changement profond), mais il faut sans doute préciser le type de conséquences envisagées. Elles devraient se déployer sur deux plans. D'abord sur celui du conseil, de la formation (continue), du soutien à l'enseignant, de manière à l'aider à progresser dans sa pratique au

4. « L'évaluation des pratiques enseignantes dans les premier et second degrés », Avis du HCéé n° 7, janvier-février 2003, Ministère de l'Éducation nationale.

service des progrès des élèves ; ensuite sur celui de son évolution de carrière : mobilité (y compris en proposant d'autres fonctions aux enseignants qui ne veulent ou ne peuvent plus l'être, en développant les reconversions, les secondes carrières) et progression salariale ou statutaire, non pas à chaque étape de cette carrière, mais aux plus importantes (exemple : changement de corps, ou de certains grades). Dans les hiérarchisations évoquées ci-dessus, retenir d'abord les débutants, les enseignants ayant de grandes difficultés, ceux qui demanderont une évaluation permettra de concentrer ces conséquences sur ceux où elles devraient être le plus nécessaire.

L'évaluation des établissements doit être directement utilisée dans le cadre de la régulation de leur autonomie. Le rapport de la Commission du débat national sur l'avenir de l'École a plaidé pour que cette régulation prenne désormais (après le relatif échec de l'idée de projet) la forme d'un contrat⁵. Et c'est dans ce cadre (contrat pluriannuel, et avec trois partenaires – l'établissement, l'État et l' élu territorial) que l'évaluation de l'établissement doit servir : pour préparer le contrat, évaluation diagnostique d'une part ; pour apprécier en quoi il a été rempli et en tirer les conséquences à son terme, évaluation *a posteriori* d'autre part.

2.2 Conditions pour que les évaluations de pratiques, celles d'enseignants, celles d'établissements servent vraiment

Quelles sont les conditions principales qu'il faudrait remplir pour que les trois évaluations précédentes puissent être utilisées selon les lignes évoquées à l'instant ?

Des conditions de **qualité méthodologique et de transparence**, analogues à celles sur lesquelles on a insisté dans la première partie, sont ici aussi essentielles. L'évaluation des pratiques éducatives par exemple, requiert d'élaborer une méthodologie commune et connue qui permettent à tous ceux qui les étudient (chercheurs, inspecteurs, évaluateurs, etc.) de parler le même langage (cadres et nomenclature d'activités, méthodes d'observation ou de recueil, etc.) et de présenter leurs résultats de façon homogène pour permettre la confrontation, la synthèse le progrès et, au bout du compte, l'utilisation hors de leur sphère, de leurs travaux. On en est loin, et rien n'est plus urgent. Dans un esprit voisin, les critères d'évaluation des enseignants sont trop peu partagés, publics et transparents, de sorte que la rénovation de cette fonction doit s'accompagner, là aussi, d'une élaboration plus commune et d'une diffusion plus grande des éléments sur lesquels l'évaluation des enseignants sera fondée. Sans être évidemment le seul, un des

5. « Pour la réussite de tous les élèves », *op.cit.* Dans ce rapport figurent par ailleurs de nombreuses notations sur l'évaluation des établissements, auxquelles on peut renvoyer. On pourra aussi se reporter à l'avis n°4 du HCcé, « L'évaluation des lycées », janvier 2002.

critères de cette évaluation rénovée devrait être étroitement lié au progrès (et non au niveau) des élèves de l'enseignant.

Au-delà de ces deux exigences de qualité méthodologique et de transparence, les évaluations, dès lors qu'elles ont un caractère individuel, c'est le cas de l'inspection-évaluation des enseignants et de l'évaluation des établissements, devraient pour être utiles, respecter deux grandes conditions.

D'abord, elles ne sauraient reposer uniquement sur leur qualité méthodologique, car elles ne sont pas qu'une question d'expertise. Elles requièrent la **participation de l'évalué**, et cela sous trois formes possibles. La définition du processus et des règles d'évaluation elle-même doit faire l'objet de discussions avec les représentants des personnels et des unités. En second lieu, l'évaluation ne saurait certes être purement interne (si elle est seule, l'auto-évaluation d'un enseignant ou d'un établissement est à bannir) : l'évaluation est et doit toujours être un regard extérieur. Mais précisément, elle ne doit pas être qu'extérieure, et c'est à une combinaison entre regard extérieur et auto-évaluation qu'il faut se livrer. Par exemple, l'évaluation des établissements devra mêler une équipe d'évaluateurs externe (où des « usagers » de l'École soient présents) et une analyse par l'établissement lui-même de son état. C'est ainsi que les constats pourront utilement nourrir un contrat destiné, sur plusieurs années, à remédier aux difficultés constatées. La troisième forme de participation de l'évalué à l'évaluation est classique, mais doit être bien organisée dans les deux cas étudiés ici : il s'agit du « contradictoire », c'est-à-dire de la possibilité pour l'évalué de commenter ou de contester les conclusions de l'évaluateur.

La seconde condition se présente, au rebours de la précédente, sous une forme interrogative : **faut-il associer dès le début développement de l'évaluation et son usage, ou bien convient-il de les séparer dans le temps**, en deux phases, d'abord le développement de l'évaluation, puis l'introduction de son utilisation ? La politique suivie depuis une quinzaine d'années dans le système éducatif français a consisté à disjoindre ces deux aspects. On a cherché d'abord à développer et à améliorer l'évaluation (des élèves, des établissements, du système dans certaines de ses parties ou dans son ensemble), pensant qu'il était nécessaire et sage de faire naître ainsi une culture d'évaluation. Et cela avant d'envisager de se servir des résultats de cette dernière. À l'inverse, certaines voix s'élèvent parfois pour soutenir le point de vue contraire : il vaudrait mieux, selon elles, notamment pour convaincre les personnes et les établissements de l'intérêt de l'évaluation et de l'intérêt d'y participer, lier dès le départ évaluation et usage⁶. Il est certain que la stratégie en deux temps a permis un réel développement de l'évalua-

6. Par exemple, Jacques Dejean dans son rapport au HCcéé sur l'évaluation des enseignements universitaires : « L'évaluation de l'enseignement dans les universités françaises », Rapport du HCcéé, n° 5, février 2002, Ministère de l'Éducation nationale.

tion, mais en contrepartie l'usage de cette dernière est de fait très limité, et même insuffisant, et on ne parvient pas vraiment à l'accroître. Non seulement les évaluations existantes ne sont guère utilisées (il faut mettre évidemment à part, ici, ces deux évaluations particulières que sont les notes et les examens des élèves, qui, elles, le sont, et parfois trop), mais en outre on n'a pas assez réfléchi aux façons de les utiliser. Car l'usage des évaluations, celles des élèves (lors des évaluations diagnostiques de masse), des personnes et des établissements notamment, ne saurait ni être automatique, ni inexistant : il faut définir et expérimenter des « usages souples » (et certaines des réflexions plus haut en définissent des possibles). À l'inverse, lier dès le début évaluation et usages bien définis risque de tuer dans l'œuf le développement de cette forme de régulation des systèmes éducatifs. Au total, la question semble ouverte. Peut-être faut-il lui fournir une réponse au cas par cas, en faisant, là aussi, l'objet d'une consultation non seulement des experts mais des personnes et unités impliquées.

3 UNE REMARQUE POUR CONCLURE

Il est utile, pour conclure ce bref panorama, de mettre l'accent sur un dernier point, conséquence ou condition, du nécessaire développement de l'usage des évaluations dans nos systèmes éducatifs. Il porte sur les personnes qui pourraient porter une telle orientation, c'est-à-dire les **cadres éducatifs**, en entendant par cadre éducatif trois catégories de personnes : les chefs d'établissement, les inspecteurs, les cadres administratifs. Ce sont eux qui devraient, peut-être contribuer au développement des évaluations, en tout cas favoriser l'essor de leurs usages, soit qu'ils les utilisent eux-mêmes, soit qu'ils facilitent leur utilisation par les enseignants ou même l'extérieur (parents d'élèves, par exemple). Aussi les perspectives tracées dans ce texte impliquent-elles certainement une grande politique de l'encadrement éducatif : définition de fonction, formation, recrutement, évaluation.

En un sens, rien n'est plus urgent que de s'appliquer à satisfaire aux conditions pour que l'évaluation, sous toutes ses formes, soit utilisée. Car dès lors qu'elle est conçue pour servir, à la fois de façon externe et de façon interne, l'évaluation est, au moins potentiellement, une très grande force de régulation du système éducatif, peut-être la seule force de régulation aujourd'hui dans un espace public marqué par des acteurs éduqués et exigeants.

Chapitre 2

De l'analyse à l'évaluation d'une pratique professionnelle

Anna Jorro, Jean-Jacques Maurice

L'activité professionnelle des enseignants fait l'objet de démarches d'analyse et d'évaluation aussi bien dans l'exercice de la pratique professionnelle que sur le terrain de la formation. Si l'analyse est perçue comme une pratique nécessaire, l'évaluation suscite toujours des réticences. Le déficit de perception qui existe entre les deux démarches est tel qu'il paraît nécessaire de montrer les relations complémentaires qui existent entre l'analyse et l'évaluation dans les pratiques quotidiennes des enseignants. Dans cet article, nous présenterons deux recherches qui rendent compte des pratiques enseignantes à partir de «grains d'analyse» différents. La spécificité de chaque recherche n'entraîne pas de divergence au regard de leurs résultats, bien au contraire. Évaluer une pratique est une mission délicate qui, au-delà de l'observation traditionnelle de l'action de l'enseignant en contexte, doit s'appuyer sur des analyses issues de protocoles de recherche pluriels.

1 DISTINGUER DEUX PROCESSUS MAJEURS

Les professionnels de l'éducation et de la formation mobilisent dans l'action des opérations de pensée qui les aident à agir, réagir, ajuster leur pratique. Deux processus majeurs sont mobilisés par les professionnels confirmés : savoir analyser le contexte et l'action en train de se faire, et prolonger

l'analyse de l'action par une évaluation afin d'anticiper son développement. Au-delà de leur dimension cognitive, l'analyse et l'évaluation constituent deux compétences professionnelles indispensables pour tenir compte de la complexité des situations professionnelles. Il nous semble important de distinguer ces deux compétences parce qu'elles se complètent et, plus encore, parce qu'elles apparaissent comme des compétences déterminantes dans l'activité professionnelle.

L'analyse est une opération capitale de la pensée, permettant une identification des paramètres de l'action qu'elle peut décomposer de façon rigoureuse afin d'en comprendre les ressorts. En ce sens, l'analyse relève d'une visée herméneutique. De son côté, l'évaluation opère une confrontation entre les éléments distingués lors de l'analyse et le système d'exigences mobilisé (un référentiel d'évaluation). À partir de cette confrontation, l'évaluation interroge la qualité de l'action (sens, pertinence, efficience...) et après un processus de problématisation (enjeux hiérarchisés...) permet de définir des ajustements. Une visée praxéologique est sous-tendue dans l'évaluation. La première compétence participe à un processus d'intelligibilité de la pratique, la seconde s'inscrit dans un processus d'optimisation de cette même pratique.

Ces deux compétences réflexives sont inégalement mobilisées dans la formation des enseignants (Jorro, 2004; Campanale, 2007). Pourtant le processus réflexif suppose un basculement de l'analyse vers la régulation de l'action. La réflexivité n'est pas seulement le lieu de la compréhension mais aussi celui de l'action régulée par anticipation (Schön, 1994). Ce qui nous conduit à relever l'interrelation entre la visée compréhensive de l'action et la visée praxéologique qui interroge le devenir de l'action, son développement.

2 QUELS « GRAINS D'ANALYSE » DE L'ACTIVITÉ PROFESSIONNELLE ?

L'activité professionnelle ne serait pas toujours directement lisible. Vouloir observer, décrire, relater une pratique relève d'une intention de compréhension. Une partie de l'activité est observable (le sujet se déplace, parle...); une autre partie peut être inaccessible à l'observateur (le sujet s'appuie sur ses souvenirs, raisonne...). De plus, certaines activités cognitives n'émergent pas à la conscience du sujet lui-même, il ne peut utiliser sa réflexivité, ses capacités métacognitives. Aucune pratique professionnelle n'échappe à la difficulté de la mise en mots; ce qui se fait parfois de façon spontanée en contexte. Ainsi, les habiletés manuelles ou intellectuelles du praticien expérimenté sont peu communicables au novice. L'enseignant qui serait en situation d'évoquer son

activité professionnelle devant un pair, un chercheur, un évaluateur serait confronté à la difficulté de l'explicitation de l'action (Maurice, 2006a).

À la suite de nombreux travaux (Postic, 1977 ; Altet, 1994 ; Tardif et Lessard, 1999 ; Perrenoud, 1999 ; Paquay, 1994, 2005...), nous présenterons deux démarches de recherche conduites d'une part en didactique du français (Jorro, 2002, 2003, 2005) et d'autre part en ergonomie cognitive (Maurice, 1996, 2006).

2.1 Des analyseurs pour comprendre les détails de l'activité

La première démarche a été développée dans le cadre de la didactique du français (Jorro, 2002, 2003), puis a été expérimentée dans des situations de mathématiques (Jorro, 2005). Le fil conducteur de ces recherches réside dans la construction d'une matrice de l'agir professoral (figure 2.1) qui rendrait compte de l'activité de l'enseignant en classe. Nous avons dégagé quatre analyseurs qui prennent appui sur

- le discours de l'enseignant, à travers «les gestes langagiers» ;
- la transmission des savoirs, avec «les gestes de mise en scène des savoirs» ;
- la manière dont l'enseignant communique avec les élèves et apprécie leur travail, avec «les gestes éthiques» ;
- la manière de réagir aux événements imprévisibles et de s'adapter en situation, avec «les gestes d'ajustement».

Ces analyseurs sont décrits sous la forme de gestes professionnels parce qu'ils rendent compte d'une action symbolique et corporelle. La corporéité de l'agir enseignant est pour nous une dimension fondamentale ; l'activité enseignante n'est pas neutre, elle signifie au-delà de l'acte. Les gestes professionnels portent des systèmes de valeurs qui infléchissent le moindre acte, le moindre regard, la moindre parole appréciative (Jorro, 2006a). Dans le schéma présenté dans la figure 2.1, les gestes professionnels peuvent être appréhendés comme des invariants de l'agir du professeur et être mobilisés pour une analyse de la pratique enseignante.

Les analyseurs distingués conduisent à caractériser le plus finement possible l'activité de l'enseignant lorsqu'il s'adresse à sa classe, lorsqu'il introduit une nouvelle leçon, lorsqu'il apporte des modifications dans le cours de l'action, lorsqu'il accompagne et conseille certains élèves... L'analyse nécessite la mise à plat d'indicateurs qui seront spécifiques pour chaque geste. Aussi, il importe de préciser les éléments descriptifs qui permettent de dégager des indicateurs ouvrant sur une analyse à grains fins.

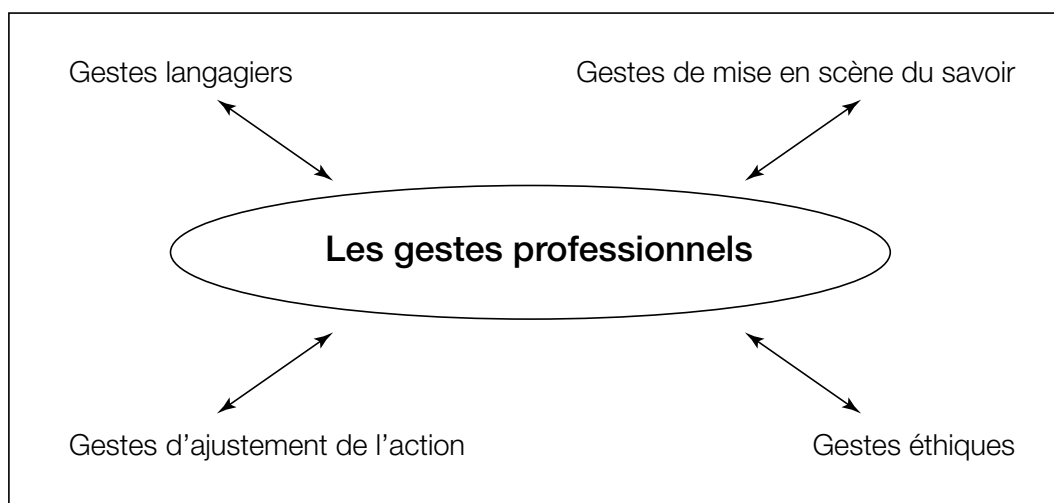


FIGURE 2.1. Matrice de l'agir du professionnel

Les gestes langagiers

Ces gestes importent particulièrement, ils sont le premier outil de travail du professeur (Bucheton *et al.*, 2005). Ils permettent notamment d'observer/analyser la posture énonciative de l'enseignant devant la classe. Les accents didactiques qui ponctuent le discours de l'enseignant orientent les élèves vers des phases de transition d'une leçon à l'autre, vers des moments de structuration et d'institutionnalisation des objets de savoir (reprise d'un concept, clarification d'une notion), vers des temps de retour au calme. Le langage du professeur peut apparaître commun ou spécialisé, et dans ce dernier cas manifester une exigence, si ce n'est une vigilance épistémologique, de la discipline travaillée. Le lexique mobilisé tout autant que la tonalité langagière ont des effets auprès des élèves. Il en est de même de la posture d'écoute, de négociation de sens...

Les gestes de mise en scène du savoir

Ces gestes permettent au professeur de relier l'activité intellectuelle des élèves aux enjeux didactiques poursuivis. Les savoirs et les modalités de traitement de ces savoirs en situation d'enseignement importent particulièrement. Les gestes de désignation d'un objet de savoir, de monstration, d'utilisation d'artefacts, de vérification des traces écrites, d'institutionnalisation du savoir sont censés orienter l'activité des élèves. La construction de la pensée abstraite, le maniement des concepts par la classe constituent l'arrière fond de l'activité du professeur.

Les gestes d'ajustement de l'action

Ces gestes relèvent de la capacité à intervenir sur le déroulement de l'activité. L'intervention de l'enseignant est multiple, elle peut porter sur le rythme de

l'action (anticipation, accélération, actualisation), sur la modification d'une consigne, sur sa reprise...sur l'invention d'une stratégie nouvelle, sur la prise en compte d'une demande émanant de la classe... Si les gestes d'ajustement sont destinés à la bonne marche de la séquence d'enseignement-apprentissage, ils peuvent concerner directement le professeur et relever de gestes d'appui et de réassurance dès que le professeur s'entoure d'aides (fiches de préparation, trame de cours, prise de notes).

Les gestes éthiques

Ces gestes témoignent du type de relation instaurée entre élèves et professeurs selon les formats de la communication et de l'appréciation scolaires. Avec le format de la communication, il s'agit de regarder comment le professeur occupe le devant de la scène ou reste en retrait afin de favoriser l'activité de la classe et ainsi promouvoir le développement de l'autonomie des élèves. En particulier, la manière dont l'enseignant incite les élèves à entrer dans une activité d'étude ou, *a contrario*, génère des rapports de domination, de captation nous renseigne sur son ethos professionnel. Les gestes éthiques supposent des gestes de retenue dans le dialogue : le silence, la tolérance à l'égard des réflexions sont des indicateurs d'une attitude d'écoute et de respect... Par ailleurs, le format de l'appréciation scolaire tend à mettre en évidence le positionnement de juge ou d'ami critique incarné par le professeur. La posture éthique du professeur peut ouvrir sur une relation d'accompagnement et de conseil qui est bien loin de la figure du maître absolu.

2.2 Les savoir-faire développés dans la conception des tâches

La seconde démarche (Maurice, 1996a, 2006a) tente de décrire l'expérience de l'enseignant dans une activité de préparation de situations d'apprentissages, en particulier les savoir faire qu'il a pu développer du point de vue de la conception des tâches.

Les premiers résultats concernent l'anticipation du niveau de difficulté des tâches proposées aux élèves. Les enseignants expérimentés maîtrisent, dès la lecture d'une tâche destinée à leurs élèves, le taux d'échec qu'elle provoquera, ils ne peuvent cependant anticiper les procédures que les élèves activeront (Maurice, 1996a et b). Les enseignants parviennent ainsi à réduire l'incertitude du lendemain et commencent à agir sur ce lendemain au moment où ils préparent leur classe. Ces invariants opératoires ont pu être identifiés grâce à un modèle¹ utilisant des probabilités, ils ne peuvent apparaître avec des procédures d'observation traditionnelles. Si, dès la lecture d'une tâche destinée à ses élèves, le maître peut anticiper ses effets

1. Il n'est pas possible, dans le cadre de ce texte, d'explicitier en détail cette méthodologie (modèle utilisant des probabilités : le théorème de Bayes). Le lecteur qui le souhaiterait peut consulter les publications suivantes : Maurice, 1996a et b.

sur les performances des élèves, alors il est possible d'affirmer que les tâches dépassent le statut d'outil utile pour enseigner : elles deviennent des « instruments » au sens de Rabardel (1993). À ces tâches, déjà utilisées les années précédentes, sont associés des schèmes. Les tâches prolongent la cognition de l'enseignant, lui permettent de lire ce qui risque de se passer le lendemain, il pense la conduite de sa classe à travers les tâches qu'il propose... mais cette lecture du lendemain semble ne viser que des performances. C'est donc bien une activité cognitive de haut niveau : anticipations, inférences... qui peut être considérée comme une « *conceptualisation en acte* » un « *théorème en acte* »² au sens de Vergnaud (1990). La pensée de l'enseignant est alors étroitement liée aux tâches qu'il choisit : il leur délèguerait une partie de son pouvoir. Mais ces dimensions sont peu conscientisées.

D'autres résultats (Maurice, 2006b) mettent en relation le pouvoir d'anticipation de l'enseignant et ses conséquences sur chaque élève. Pour ce faire, à partir du film du déroulement d'une séance de lecture au cours préparatoire, les tâches choisies par l'enseignant (par exemple le texte support d'apprentissage) sont analysées dans le but d'élaborer des tests. Ces tests, proposés le lendemain aux élèves, en protocole individuel, visent l'évaluation du « potentiel d'un sujet à qui on a prescrit cette tâche ». Cette méthodologie quantifie, de 0 à 1, la **Distance à la Performance Attendue** (DPA) de chaque élève (distance entre ce qu'il sait faire et ce qu'il fallait savoir faire).

Le traitement des données révèle que les tâches proposées restent majoritairement « calées » sur la zone proximale de développement des mêmes élèves. En effet, les élèves progressent au cours de l'année, mais le curseur du niveau de difficulté des tâches progresse également. L'enseignant semble contraint, par la rapide progression de certains élèves, à faire évoluer ses exigences didactiques vers un niveau de difficulté acceptable par le groupe classe. La nécessité de la progression dans le programme officiel ne serait pas la seule contrainte. Un certain niveau de difficulté garantit probablement le bon déroulement de la classe, maintient la crédibilité de l'enseignant, assure l'enrôlement des élèves. Il est fort probable que les élèves les plus faibles se voient proposer des tâches situées en dehors de leur zone proximale de développement et bénéficient très peu de la médiation assurée par le maître. La DPA confirme cette hypothèse : pour les élèves les plus faibles, les performances immédiates produites dans le contexte de la classe sont meilleures que celles révélées deux jours

2. Il n'est pas question de dire que les enseignants font évoluer leurs anticipations en utilisant le théorème de Bayes. Le théorème n'est qu'une modélisation de leurs inférences en action, une analogie fonctionnelle. Ils réalisent, sur le plan fonctionnel uniquement, une performance équivalente au théorème.

après par la DPA. La médiation a masqué leur distance effective ; les performances apparaissant sur leurs cahiers ne pouvant être reproduites deux jours après, il est possible d'affirmer que cette médiation a amélioré la performance en classe, sans provoquer un réel apprentissage.

Il est important de rappeler que ces tâches observées sont des supports pour apprendre et n'ont pas pour fonction l'évaluation des élèves (sachant que la discrimination des performances des élèves peut être une des fonctions de l'évaluation). Or les DPA montrent que les élèves d'une classe risquent ainsi de passer une grande partie de leur temps sur des tâches « support d'apprentissage » qui les maintiennent sur le même statut.

Lorsque nous mesurons les performances de ces élèves en début et en fin d'année, nous constatons que les élèves faibles et forts en début d'année gardent leurs statuts respectifs en fin d'année, mais ce n'est pas le cas des élèves moyens. L'enjeu est important pour les élèves moyens en début d'année, qui peuvent, en fin d'année, avoir basculé vers un statut d'élève faible ou fort.

De plus la médiation de l'enseignant, très utile pour certains, deviendrait peu efficace lorsque la tâche proposée est en dehors de la zone proximale de l'élève. L'émergence d'une performance chez tel élève ne nécessite pas la même médiation que chez tel autre. À partir d'un certain seuil, augmenter l'aide nécessaire à un élève aurait pour effet de réviser à la baisse son activité cognitive, mais légitimerait le « passage à l'étape suivante ».

Des enseignants informés sur ces différents aspects de leur expérience professionnelle trouveraient ainsi des « repères pour l'action » de second degré, autorisant une prise de distance sur leur pratique. Ces deux démarches de recherche pourraient être utiles pour donner à voir aux praticiens des dimensions (peu accessibles) de leur pratique. Elles seraient une aide à la prise de conscience, un complément à leur potentiel réflexif pendant l'action. Cependant, cette information donnée aux enseignants n'a pas le statut de formation à court ou moyen terme, il n'y a pas de procédure à offrir pour que les pratiques soient différentes : c'est la raison de la formule « repères de second degré ». La transformation ne pourrait venir que de l'individu, ce ne serait pas des repères pour agir, mais des repères autorisant la genèse d'autres formes d'action.

3 LES ENJEUX DE L'ÉVALUATION DES PRATIQUES PROFESSIONNELLES

Évaluer une activité est une mission particulièrement délicate. D'ailleurs, les programmes de recherche actuellement financés le prouvent car ils portent fréquemment sur la problématique des *relations entre pratiques observées et efficacité de l'enseignant*.

L'évaluation de l'activité professionnelle emprunte à des pratiques hétérogènes dues à l'absence de stabilisation de modèles de l'agir enseignant. Les pratiques évaluatives demeurent le fait des praticiens ou d'équipes mais ne font pas l'objet d'un affichage. Aussi, la question de l'évaluation reste entière.

L'évaluateur est tenu de faire interagir plusieurs composantes d'une pratique, il ne peut s'arrêter sur une facette du métier. L'intrication des savoirs théoriques, des savoirs d'action (Barbier, 1996), des dynamiques identitaires, des imaginaires et des systèmes de valeurs, des dimensions psychoaffectives, suppose une évaluation plurielle, tenant compte de ces éléments en interaction. La tâche de l'évaluateur apparaît donc très délicate. Le professionnel compétent mobilise des savoirs hétérogènes pour répondre aux situations et sait ajuster son activité. Il ne s'agit pas seulement de valider l'appropriation de compétences mais de faire valoir un ensemble de processus puisant dans les différents éléments constitutifs de l'activité. Serait-ce donc une tâche impossible ? Non, si l'on répond à un certain nombre d'exigences méthodologiques et éthiques.

3.1 Exigences méthodologiques

Du point de vue méthodologique, l'évaluation d'une pratique peut prendre appui sur :

- l'observation ponctuelle (très limitée dans le temps et très contextualisée) d'une pratique effective et du comportement des élèves ;
- l'analyse des traces témoignant du passé de la classe (habitudes des élèves, traces écrites³, résultats d'évaluation...);
- la prise en compte des arguments de l'enseignant et de ses propos pendant l'entretien avec l'observateur ou l'évaluateur ;
- l'identification de son rôle au sein de l'équipe éducative et de l'établissement ;
- le dialogue réflexif avec un pair.

La mise en œuvre de l'évaluation suppose une modalité opératoire précise qui implique un croisement de regards : tout d'abord, une analyse de l'activité professionnelle telle qu'elle s'est déroulée en contexte, puis la mobilisation d'un système d'exigences qui permettra un processus de confrontation

3. Traces écrites concernant l'enseignant ou les élèves.

entre la réalité d'une pratique et le souhaitable de cette même activité. Toute la question est de rendre visible le référentiel de l'évaluation. La mise au jour d'un référentiel est donc le fait de pratiques locales, d'établissement de formation, de circonscription, d'équipes d'IPR... Il reste que le référentiel détermine la nature du questionnement portant sur les critères et les jeux interprétatifs qui traversent l'activité évaluative. Car il s'agit bien de savoir si la confrontation avec le référentiel débouche sur le maintien et la valorisation de l'action, ou sur des décisions d'ajustements. Cette confrontation est importante d'un triple point de vue :

- elle revêt un caractère contradictoire, le réel étant toujours en décalage avec le souhaitable, il suscitera un questionnement sur la nature plus ou moins inédite, plus ou moins préoccupante, de la différence identifiée ;
- elle intègre une dimension heuristique dans le fait de se demander si l'activité revêt des potentialités non exploitées, si l'acteur peut mobiliser autrement ses compétences et ses gestes professionnels ;
- elle se caractérise par sa visée praxéologique dans les décisions d'ajustement qui peuvent être envisagées.

Ce processus de confrontation constitue une forme de vigilance pour le praticien qui cherche à agir de façon congruente avec le contexte et en pertinence avec la représentation qu'il se fait d'une action de qualité ou encore qui cherche à se rapprocher au plus prêt des valeurs qui l'animent.

L'acte évaluatif repose en partie sur des opérations d'inférence. En effet, la mise en évidence d'une compétence professionnelle ne peut se résumer au constat de l'exécution pertinente d'une tâche spécifique à l'instant t. C'est donc une inférence qui permet de décontextualiser l'ensemble des observables situés et datés afin de les traduire en termes de compétences. Une inférence permet d'obtenir plus d'informations que celles produites par l'observation, elle permet de généraliser, de catégoriser, de déduire... Mais une inférence peut aussi être fragile car elle n'est pas toujours explicitable, elle relève parfois du « ressenti », de « la lecture entre les lignes ». Elle ne s'appuie pas obligatoirement sur une rationalité formalisable. Il ne s'agit pas d'affirmer que les inférences moins rationnelles sont le fruit d'une maladresse de l'évaluateur d'enseignants, ou d'un déficit instrumental global des institutions chargées d'évaluer. Ces inférences moins rationnelles seraient le complément indispensable aux inférences rationnelles parce que l'évaluation des compétences professionnelles d'un enseignant est spécifiquement complexe. En effet, l'évaluation de la pratique d'un ébéniste serait plus aisée (alors que les compétences nécessaires dans ce métier sont d'un haut niveau de spécialisation et relèvent d'habiletés cognitives et manuelles coûteuses à acquérir). L'ébéniste lui-même et son évaluateur peuvent constater le résultat immédiat de l'action de ce professionnel mais également les associer aux résultats antérieurs encore observables et capitalisables : le « produit » de son action

est directement lisible, la décontextualisation de ses compétences s'appuie sur un large corpus d'observables, le lien de causalité entre l'action et le produit n'est pas contestable. L'enseignant et son évaluateur ne peuvent constater aussi aisément les résultats immédiats d'une action (malgré l'apport incontestable des évaluations des élèves qui contribuent à limiter le problème évoqué). Ils ne peuvent conclure aussi facilement en termes de liens de causalité fiables entre action pédagogique ou didactique et « produit » de cette action (l'apprentissage chez l'élève par exemple).

3.2 Exigences éthiques

Du point de vue éthique, le rapport que l'évaluateur entretient avec l'évaluation mérite d'être questionné. Sa conception de l'acte évaluatif, considéré comme outil de contrôle ou d'optimisation d'une pratique à travers des pistes d'action, comme démarche d'accompagnement du développement professionnel ou comme acte gestionnaire sur la carrière de l'enseignant, comme processus réflexif ou comme acte de mesure des écarts à la norme (parfois implicite lorsque le référentiel n'est pas discuté) déterminent un positionnement de l'évaluateur. Les paramètres qui viennent d'être énoncés sont donc déterminants pour penser l'évaluation d'une activité professionnelle. Trop souvent, le sentiment d'une épée de Damoclès en suspens a été exprimé, trop souvent des regards vécus comme trop critiques ont été dénoncés. La juste mesure de l'acte évaluatif est donc attendue.

Nous faisons l'hypothèse qu'une triple réflexion conditionne une attitude éthique et implique une élucidation :

- du rapport que les évaluateurs entretiennent avec un système d'exigences. Autrement dit, il s'agit de comprendre les jeux possibles autour de la norme et des critères d'évaluation (logique de conformité, logique d'usage);
- du rapport tissé avec le processus de confrontation avec un référentiel. Le questionnement sur les critères, la problématisation et donc la hiérarchisation des priorités liées à l'exercice de l'activité, supposent la connaissance des marges de progression dans le développement professionnel;
- du rapport assumé ou non avec le processus de régulation. Il s'agit d'aider l'enseignant à envisager des ajustements en évitant de le laisser dans la solitude de la prise de décision.

4 UNE RECHERCHE EN COURS

Nos travaux (Jorro, 2002, 2006b, 2007) nous conduisent à privilégier une évaluation plurielle et à faire valoir, de façon interdépendante, trois focales évaluatives :

- la première relative à l'**action prescrite** et à sa planification s'intéresse aux gestes structurant l'activité : choix des contenus d'enseignement, conception d'une séquence, conception de situations d'apprentissage, conception de l'évaluation... ;
- la seconde se centre sur l'**action effective** et ses transformations. Elle met en évidence la matrice de l'agir enseignant que nous avons présentée précédemment à travers les qualités énonciatives du professionnel, la pertinence des gestes de mise en scène des savoirs et l'efficacité des gestes d'ajustement en situation d'exécution ;
- la troisième porte sur l'**action réfléchie**, et sur la manière dont le professionnel revient sur l'expérience vécue. Les gestes éthiques sont alors mobilisés. Cette troisième approche suppose une capacité de l'enseignant à faire retour sur ses différents investissements professionnels dans différents cadres : celui de la classe, celui de la communauté de praticiens, celui de la communauté de la profession, celui de l'institution dans laquelle il travaille.

Dans la figure 2.2, figurent ces trois focales qui peuvent être différemment mobilisées en situation d'évaluation.

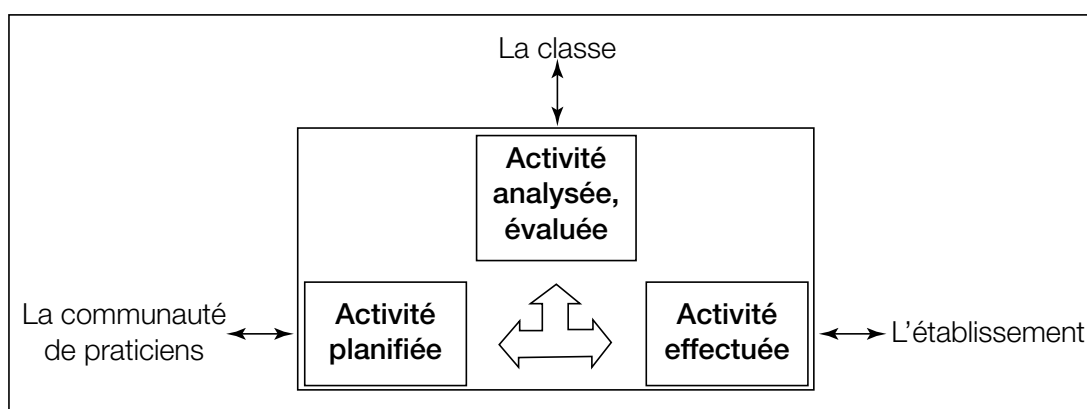


FIGURE 2.2. Trois focales pour évaluer l'activité enseignante

- L'évaluation de la planification de l'activité relève de la structuration de la séquence et permet d'apprécier les connaissances théoriques et opératoires de l'enseignant. Sa maîtrise des contenus disciplinaires, la connaissance du curriculum et sa déclinaison en tâches simples et complexes, sa connaissance des enjeux et des processus d'apprentissage, la pertinence dans la définition de la consigne, sa connaissance de la distance à la performance attendue sont autant d'indicateurs à spécifier.
- L'évaluation de la mise en œuvre de l'activité permet de porter un regard critique sur l'opportunité en situation, la qualité intrinsèque, l'efficacité, la pertinence des gestes professionnels énoncés plus haut.

- L'évaluation de l'activité réflexive de l'enseignant prend appui sur les compétences de retour de l'enseignant sur son activité, et notamment sur la lecture de l'activité en termes de points d'appui et de points d'amélioration mais aussi sur la réflexion construite autour des événements inédits et sur les engagements socioprofessionnels et organisationnels.

Ces trois dimensions n'épuisent aucunement la complexité d'une pratique mais tentent de rendre visibles des éléments d'appréciation de la pratique professionnelle. Les ajustements peuvent concerner ces différentes dimensions. L'amplitude des ajustements révélant alors la capacité de réactivité et d'adaptabilité du professionnel. Cette pluri-directionnalité de la régulation est possible si l'analyse est approfondie et si le professionnel a conscience des différents niveaux d'analyse et d'évaluation d'une pratique. Mais ces trois focales seront aussi différemment mobilisées par l'enseignant selon le positionnement de l'évaluateur. En effet, les évaluateurs peuvent mettre l'accent sur tel ou tel aspect des activités des enseignants lors de leurs évaluations, ceci reflétant diverses conceptions du développement professionnel. Ces conceptions détermineront une manière d'évaluer. Il importe aujourd'hui de les afficher au grand jour tant elles sont susceptibles de varier selon le positionnement professionnel de l'évaluateur. Par exemple, un inspecteur apprécie-t-il l'activité d'un enseignant de la même manière qu'un chef d'établissement ?

5 CONCLUSION

Les métiers de l'éducation se déroulent sur le vif, ils supposent une double compétence d'analyse et d'évaluation afin d'agir en situation de façon pertinente et efficiente. La professionnalité du praticien résiderait dans ce double regard, appréhendant plusieurs dimensions de sa situation en repérant les points d'appui, en modifiant certains paramètres de sorte que la situation effective se développe dans les meilleures conditions. Ce travail incessant d'analyse et d'évaluation relève de l'agenda de l'enseignant agissant dans l'incertitude et décidant dans l'urgence (Perrenoud, 1996) Il s'inscrit également dans une approche de la pratique enseignante ouverte à la question de la responsabilité professionnelle puisqu'en matière d'éducation, nous sommes «obligés devant l'enfant» et, par conséquent, devant l'élève (Cifali, 2005). L'analyse et l'évaluation d'une pratique impliquent un rapport assumé à l'éthique professionnelle, dialogue continu entre l'action planifiée – réalisée-projetée. Le travail réflexif place le praticien dans un processus complexe : processus de distanciation nécessaire à l'analyse, processus d'engagement requis par l'évaluation.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Altet, M. (1994). *La formation professionnelle des enseignants; analyses de pratiques et situations pédagogiques*. Paris, PUF.
- Barbier, J.-M. (1996). *Savoirs théoriques, savoirs d'action*. Paris, PUF.
- Bucheton, D., Bronner, A., Broussal, D., Jorro, A., Larguier, M. (2005). Les pratiques langagières des enseignants : des savoirs professionnels inédits en formation, *Repères*, n° 30, p. 33-53.
- Campanale, F. (2007). Analyse réflexive et auto-évaluation dans la formation des enseignants : quelles relations ? In Jorro, A. (éd). *Évaluation et développement professionnel*. Paris, L'Harmattan, pp. 57-77.
- Cifali, M. (2005). L'envers et l'endroit d'une « obligation de résultats ». In Lessard, C. et Meirieu, P. (éds). *L'obligation de résultats en éducation*, Paris-Bruxelles, De Boeck, pp. 243-256.
- Jorro, A. (2002). *Professionaliser le métier d'enseignant*. Paris, ESF.
- Jorro, A. et Croce-Spinelli, H. (2003). Gestes du professeur et activité interprétative de lecteurs de cycle 3. Colloque CERFEE-IRSA. Montpellier.
- Jorro, A. (2004). Réflexivité et auto-évaluation dans les pratiques enseignantes, *Mesure et évaluation en éducation*, 27, 2, pp. 33-47.
- Jorro, A. (2005). Le corps parlant de l'enseignant. Entente, malentendus, négociations. In *Actes du congrès international de l'AIRDF*, Québec.
- Jorro, A. (2006 a). L'agir professionnel de l'enseignant. Conférence au séminaire de recherche du Centre de Recherche sur la formation, Paris, CNAM.
- Jorro, A. (2006 b). L'ethos de l'évaluateur : entre imaginaires et postures. In Figari, G. et Mottier, L. (éds). *Recherche sur l'évaluation en éducation*. Paris, L'Harmattan, pp. 76-75.
- Jorro, A. (2007). *Évaluation et développement professionnel*. Paris, L'Harmattan.
- Maurice, J.-J. (1996a). Une connaissance de l'élève dépendante des contraintes de l'action, *Revue Française de Pédagogie*, 114, pp. 85-96.
- Maurice, J.-J. (1996b). Problèmes multiplicatifs : l'expérience de l'enseignant, l'action effective de l'élève, *Recherche en Didactique des Mathématiques*, 16/3(48), pp. 323-348.
- Maurice, J.-J. (2006a). L'expérience de l'enseignant : une réflexivité limitée, *Formations et pratiques d'enseignement en questions*, 3, 53-67.
- Maurice, J.-J. (2006b). Des contraintes du métier d'enseignant participent-elles à la construction de son expérience ? In Gentaz, E., Dessus P. (éds),

- Apprendre et enseigner à l'école : sciences cognitives et éducation.* Paris, Dunod, pp. 219-234.
- Paquay, L. (1994). Vers un référentiel de compétences professionnelles de l'enseignant, *Recherche et formation*, 16, pp. 7-33.
- Paquay, L. (2005). *L'évaluation des enseignants. Tensions et enjeux.* Paris, L'Harmattan.
- Perrenoud, P. (1996). *Enseigner : agir dans l'urgence, décider dans l'incertitude. Savoirs et compétences dans un métier complexe.* Paris, ESF.
- Perrenoud, P. (1999). *Dix nouvelles compétences pour enseigner.* Paris, ESF.
- Postic, M. (1977). *Observation et formation des enseignants.* Paris, PUF.
- Rabardel, P. (1993). Représentations dans des situations d'activités instrumentées. In Weill-Fassina, A., Rabardel, P., Dubois, D. (éds), *Représentations pour l'action.* Toulouse, Octarès éditions, pp. 113-137.
- Schön, D.A. (1994). *Le praticien réflexif. À la recherche du savoir caché dans l'agir professionnel.* Montréal, Éditions Logiques.
- Tardif, M., Lessard, C. (1999). *Le travail enseignant au quotidien, Expérience, interactions humaines et dilemmes professionnels.* Bruxelles, De Boeck.
- Vergnaud, G. (1990). La théorie des champs conceptuels, *Recherche en Didactique des Mathématiques*, 10, n° 23, pp. 133-170.

Chapitre 3

Que fait-on des évaluations internationales dans le système éducatif français ?

Jean-Claude Emin

Il a été souvent dit – notamment à l’OCDE – que la France n’utilisait pas les évaluations internationales – en particulier PISA – pour piloter son système éducatif. Je voudrais profiter de l’occasion qui m’est offerte ici de prolonger les débats qui ont eu lieu au colloque de l’ADMEE, lors de la table ronde sur «Quels usages nationaux pour les outils internationaux, l’exemple de PISA» pour revenir sur cette question.

Je m’efforcerai d’abord de montrer que la France prend effectivement en compte nombre d’enseignements tirés de l’analyse des évaluations internationales lors des réflexions et des décisions concernant les évolutions de son système éducatif ; mais elle le fait généralement en confrontant ces analyses à celles tirées d’autres travaux. L’instance qui, en France, a été chargée de faire des propositions en matière d’évaluation du système éducatif, le Haut Conseil de l’évaluation de l’école¹, s’est d’ailleurs penchée sur la question

1. Le Haut Conseil de l’évaluation de l’école, créé par le ministre de l’Éducation nationale en fin 2000 était une instance indépendante de l’administration chargée non pas d’évaluer l’École, mais d’évaluer les évaluations de l’École –quels qu’en soient les auteurs– et de donner un avis sur les méthodes utilisées, de les expertiser et de faire toute recommandation en la matière. Il comprenait trente-cinq membres appartenant à

de la position de la France à l'égard des évaluations internationales, et cet article empruntera largement à l'avis que le Haut Conseil a rendu à ce sujet², ainsi qu'aux travaux de la Direction de l'évaluation, de la prospective et de la performance du ministère français de l'Éducation nationale.

1 QU'A-T-ON TIRÉ DE PISA POUR LE SYSTÈME ÉDUCATIF FRANÇAIS ?

Une fois passées les premières réactions médiatiques, on s'est relativement peu attaché, en France, au classement des pays participants et au rang de la France dans ce classement (ce qui a sans doute été facilité par le fait que les performances de ses élèves sont, selon les domaines évalués, dans la moyenne ou supérieures à la moyenne des pays de l'OCDE !). On s'est surtout livré, progressivement, à une analyse fouillée des performances relatives des élèves français, à une réflexion sur les objectifs de notre enseignement au regard de ce que PISA se propose de mesurer, et à l'examen, à la lumière des données recueillies par PISA, de problèmes importants rencontrés par notre système éducatif. On peut évoquer plusieurs points à ce sujet.

1.1 Un questionnement des objectifs de l'enseignement français

Les performances de nos élèves – comme celles des élèves des autres pays – sont généralement très variables selon les activités proposées. Un tel constat incite à la prudence quant à une interprétation trop globale des classements qui ne constituent en somme qu'une moyenne de résultats disparates et non une mesure unidimensionnelle robuste d'une compétence générale.

trois catégories :

- tout d'abord, des représentants de la « société civile » : député, sénateur, élus régional, départemental et local, ainsi que des représentants des principales organisations d'employeurs et de salariés ;
- ensuite, des représentants des acteurs, partenaires et usagers de l'École : parents d'élèves, personnels, élèves et étudiants, associations complémentaires de l'enseignement public ;
- enfin des personnalités qualifiées, parmi lesquelles figuraient des spécialistes étrangers ayant conduit des travaux dans leurs pays et au plan international.

Les résultats de ses travaux – dont il arrêtait le programme en toute indépendance – et ses propositions étaient rendus publics, notamment sous forme d'avis.

En avril 2005, la création d'un Haut Conseil de l'Éducation par la Loi d'orientation pour l'avenir de l'École a entraîné la suppression du Haut Conseil de l'évaluation de l'école.

2. « La France et les évaluations internationales des acquis des élèves », Avis n° 16 du Haut Conseil de l'évaluation de l'école, mai 2005, disponible sur : <http://cisad.adc.education.fr/hcee>.

Il est ainsi tout à fait éclairant d'essayer de cerner ce que peuvent être nos points forts et nos points faibles relativement aux autres pays.

Par exemple, s'agissant de la compréhension de l'écrit, ces points forts et ces points faibles sont très liés à la conception de l'enseignement de la lecture qui est la nôtre. En effet, la France se positionne plutôt au-dessus de la moyenne internationale dans les compétences «s'informer» (qui renvoie à la capacité des élèves à *prélever des informations dans un texte*) et «interpréter» (qui apprécie leur capacité à *dégager du sens et à établir des inférences à partir de l'écrit*). En revanche, les élèves français obtiennent un score inférieur à la moyenne internationale dans la compétence «réagir» (qui veut mesurer leur capacité à *mettre le texte en relation avec leurs connaissances, leurs idées et leurs expériences*).

Nous nous situons ainsi au-dessus de la moyenne internationale pour les deux compétences qui sont le plus travaillées dans notre enseignement de la lecture. Ce sont d'ailleurs celles qui sont traditionnellement présentes dans nos évaluations nationales. Quant à la compétence «réagir», elle est peu présente dans nos pratiques, notamment en fin de collège. En revanche, cette compétence est beaucoup plus largement pratiquée dans nombre de pays anglo-saxons.

Ceci fait apparaître que le rendement du système éducatif français n'est sans doute pas si mauvais que cela – tout au moins relativement – par rapport aux objectifs qui lui sont assignés en matière d'enseignement de la lecture. Il a d'ailleurs toujours été dit que le protocole de PISA ne pouvait rendre compte des objectifs des programmes français. Le Haut Conseil de l'évaluation de l'école a ainsi pu souligner que dans les enquêtes internationales «*la volonté d'apprécier des compétences ou des comportements préparant de façon générale, à la vie professionnelle et sociale dans le monde d'aujourd'hui, remplace de plus en plus le souci de vérifier le degré relatif d'atteinte des objectifs affichés par les programmes scolaires des pays qui y participent. Deux raisons sont invoquées en ce sens: d'un point de vue technique, les programmes sont différents d'un pays à l'autre, ce qui rend les comparaisons difficiles; d'un point de vue plus politique, voire idéologique, l'évaluation de compétences considérées comme de portée universelle apparaît plus pertinente que celle des savoirs scolaires, pour comparer les prestations des systèmes d'enseignement, tout au moins à la fin de la scolarité obligatoire*»³.

En tout cas, c'est parce que l'on a procédé à des analyses fines des réponses des élèves aux items de PISA, et que l'on a confronté le protocole de PISA aux attentes affichées par les programmes d'enseignement français, que l'on a pu ainsi poser des questions quant au contenu de notre enseignement et à ses objectifs, notamment en fin de scolarité obligatoire.

3. Ibid.

1.2 Une contribution à la définition d'un socle commun de compétences et de connaissances pour la fin de la scolarité obligatoire

Il apparaît ainsi une certaine continuité entre des analyses comme celles évoquées ci-dessus et les réflexions qui ont conduit à donner à la scolarité obligatoire française, l'objectif de doter tous les élèves d'au moins un socle commun de compétences et de connaissances, dont la définition est ainsi donnée par la Loi d'orientation pour l'avenir de l'École d'avril 2005 : *«La scolarité obligatoire doit au moins garantir à chaque élève les moyens nécessaires à l'acquisition d'un socle commun constitué d'un ensemble de connaissances et de compétences qu'il est indispensable de maîtriser pour accomplir avec succès sa scolarité, poursuivre sa formation, construire son avenir personnel et professionnel et réussir sa vie en société»*⁴. La loi reprend ainsi une proposition de la Commission nationale du débat sur l'avenir de l'École, présidée par Claude Thélot : *«La scolarité obligatoire qui se poursuit jusqu'à l'âge de 16 ans devrait garantir à tous l'acquisition d'un ensemble de savoirs, de savoir-faire et de savoir-être. Ce socle commun ne constitue pas la totalité de ce qui est enseigné à l'école et au collège, mais doit contenir ce qui est indispensable pour réussir sa vie au XXI^e siècle»*⁵.

C'est là pour notre système éducatif une évolution importante qui s'inscrit dans un mouvement que connaissent plusieurs pays voisins du nôtre, qui définissent des compétences clefs ou des «standards» de fin de scolarité obligatoire.

En effet, la conception du socle commun rappelée ci-dessus est fort éloignée de la conception traditionnelle de nos programmes dont les objectifs sont – plus nettement encore au collège qu'à l'école – plutôt définis en termes de connaissances dans les différents champs disciplinaires, qu'en termes de compétences ou de capacités à mettre en œuvre ces connaissances. *«Les différentes instances qui ont été successivement ou simultanément associées à la définition des objectifs et des contenus des programmes... ne sont pas parvenues à faire émerger un socle accessible à tous les élèves... Les trois volumes qui décrivent ce que l'on doit apprendre à l'école maternelle, à l'école élémentaire et au collège ne décrivent pas un tel socle, mais la somme de ce qu'il est possible d'apprendre au cours de la scolarité obligatoire»*⁶.

4. Article L.122-1-1 du Code de l'éducation.

5. «Pour la réussite de tous les élèves», Rapport de la Commission du débat national sur l'avenir de l'École présidée par Claude Thélot, Paris, La Documentation française et le Sceren-CNDP, 2004.

6. *Ibid.*

Pour en revenir à PISA, il faut noter que le décret qui présente la conception et le contenu du socle commun précise que la définition du socle «se réfère... aux évaluations internationales, notamment au Programme international pour le suivi des acquis des élèves (PISA) qui propose une mesure comparée des connaissances et des compétences nécessaires tout au long de la vie»⁷. C'est très certainement la première fois qu'un texte définissant des objectifs de notre École fait une référence explicite à une évaluation standardisée, au surplus internationale, des acquis des élèves.

Mais, en l'occurrence, ce sont moins les résultats de la France, relativement aux autres pays, qui ont nourri les réflexions des diverses instances qui ont contribué à définir ce socle commun que la conception même de PISA qui se proposait d'évaluer les compétences des jeunes de 15 ans ainsi définies: «On ne peut attendre des jeunes de 15 ans qu'ils aient appris tout ce dont ils auront besoin de savoir une fois adultes. Ils doivent posséder de solides connaissances fondamentales... Ils doivent aussi comprendre des processus élémentaires et des principes fondamentaux et les appliquer avec souplesse dans différentes situations pour pouvoir poursuivre leur apprentissage... et mettre leurs acquis en pratique dans le monde réel. C'est pourquoi PISA, plutôt que d'évaluer l'acquisition de connaissances spécifiques, évalue l'aptitude à réaliser des tâches qui s'inscrivent dans des situations de la vie réelle et qui dépendent d'une compréhension approfondie de concepts fondamentaux»⁸.

1.3 Un questionnement de pratiques d'enseignement

L'évaluation PISA, tout comme l'ensemble des évaluations internationales, fait largement appel aux questionnaires à choix multiples, format peu pratiqué en France dans le cadre scolaire. On s'attendrait à constater, concernant la France, une faiblesse relative des performances due au peu d'entraînement des élèves à ce format de réponse. Parallèlement, et parce que l'écrit est, en principe, un des axes essentiels de notre enseignement, on pourrait s'attendre à une relative réussite des élèves français dans les formats exigeant une réponse construite.

Or les résultats obtenus à cette évaluation, contrairement à ces attentes, font apparaître une certaine aisance dans les réponses aux questionnaires à choix multiples, perceptible dans les performances obtenues aux exercices qui y recourent le plus, et une faiblesse dans le domaine de la réponse écrite qui se traduit par un taux de non-réponse parmi les plus forts de tous les pays. Les questions appelant une réponse longue connaissent en France de

7. Décret n° 2006-830 du 11 juillet 2006 relatif au socle commun de connaissances et de compétences.

8. *Connaissances et compétences, des atouts pour la vie. Premiers résultats de PISA 2000*, OCDE 2001.

très forts taux d'abstention. C'est particulièrement net en compréhension de l'écrit, mais ce constat concerne l'ensemble des domaines évalués.

Il est évidemment difficile de démêler, dans les difficultés constatées, ce qui relève de la compréhension, de la capacité à répondre et, enfin, de la production d'écrit elle-même. Plusieurs hypothèses, qui renvoient à nos pratiques d'enseignement et qui rejoignent des constats faits à partir de nos évaluations nationales, peuvent néanmoins être formulées.

L'une porte sur l'apprentissage de l'écrit : le très fort taux de non-réponse peut laisser supposer que c'est peut-être dans le domaine de l'expression écrite, plus que dans celui de la lecture proprement dite, que nos élèves éprouvent le plus de difficultés. Parce qu'ils sont mal à l'aise dans le domaine de l'écrit, et que celui-ci leur demande des efforts, les élèves français choisiraient souvent de ne pas répondre du tout. En effet, à score égal, dans d'autres pays de l'OCDE, les élèves, même faibles, répondent aux questions qui leur sont posées.

On obtient donc ce résultat, qui pourra être jugé paradoxal, mais qui interpelle à juste titre l'enseignement français : une enquête qui porte sur la *compréhension* de l'écrit nous indique expressément que nos élèves ont des difficultés en *expression* écrite. On peut en effet s'interroger sur les apprentissages réels entrepris dans le domaine de la production de texte ; même si l'écrit demeure l'axe essentiel de l'enseignement français, peut-être sa maîtrise est-elle trop souvent supposée par les enseignants, et les compétences qui permettent d'asseoir cette maîtrise ne sont-elles pas assez travaillées ? Or l'écrit doit faire l'objet d'apprentissages successifs impliquant la réécriture, la reprise et l'amélioration de l'expression première.

Si la volonté de faire de l'écriture une pratique quotidienne s'exprime dans les programmes, on peut penser qu'elle demeure dans la réalité relativement rare et souvent conçue comme une forme aboutie sur laquelle on ne saurait revenir. Ainsi, l'écrit considéré comme un « tout » est-il peut-être évalué comme tel dans le cadre de la classe, sans souci de distinguer et de travailler les différents savoir-faire à mettre en œuvre.

L'autre hypothèse concerne la relation du maître et de l'élève et son influence possible sur la réticence à exprimer la réponse, et plus généralement le contexte de l'apprentissage. Le fait que les élèves français aient manifestement des réticences à prendre le risque de donner une réponse fausse, alors que les élèves des autres pays assument ce risque, est très certainement révélateur de leur crainte de voir stigmatiser leurs erreurs. On retrouve ici une caractéristique majeure de notre enseignement qui, souvent, ne s'appuie pas sur les erreurs des élèves, mais considère ces erreurs comme des fautes et les sanctionne comme telles.

1.4 Des interrogations renforcées sur la pratique du redoublement

PISA a également agi en France comme un révélateur d'une question qui n'était pas nouvelle, mais que les responsables du système éducatif et plus largement l'opinion publique avaient tendance à sous-estimer : le redoublement. PISA s'intéresse en effet à une génération d'élèves, ceux qui ont quinze ans, à quelque niveau qu'ils soient dans les systèmes éducatifs, et non pas, comme la plupart des enquêtes comparatives internationales précédentes, aux élèves d'un niveau scolaire donné. Ceci permet de mettre en relation les résultats d'une génération d'élèves en fin de scolarité obligatoire dans les différents pays, appréciés à l'aune de PISA, et la pratique de ces pays en matière de promotion d'un niveau à l'autre, au cours de la scolarité obligatoire⁹.

« Il [en] ressort que les pays qui pratiquent le passage automatique affichent, globalement, de meilleurs résultats que les pays qui ont recours au redoublement. Bien évidemment, il s'agit plus d'une tendance sur un échantillon de pays que d'un résultat très tranché. Le simple fait que le Royaume-Uni et la Finlande aient de meilleurs résultats dans chacun des trois domaines évalués par PISA et, qu'à l'inverse, l'Allemagne et la Communauté française de Belgique se situent partout au-dessous de la moyenne internationale ne suffit pas à condamner le redoublement : de nombreux facteurs entrent en jeu dans le niveau moyen de performances des élèves d'un pays (économiques, sociaux, culturels, etc.). De plus, la pratique du passage automatique s'inscrit dans un ensemble plus global. Elle implique, de fait, une organisation différente du système scolaire, organisation qui elle-même peut sensiblement varier entre deux pays adeptes du passage automatique, comme par exemple, la Finlande et le Royaume-Uni. Ainsi, il est impossible d'affirmer que les compétences des élèves anglais seraient moins bonnes si une partie d'entre eux avait redoublé. De la même manière, on ne peut pas se déterminer sur l'évolution des compétences des élèves français si l'on supprimait brusquement le redoublement.

[Mais cela permet] cependant d'affirmer, et c'est là un résultat important, que le redoublement n'est pas la réponse ad hoc à l'échec scolaire. [...] un modèle « alternatif »... est fourni par les pays qui favorisent le passage automatique et dans lesquels on n'observe pas plus d'élèves en difficulté qu'ailleurs, sinon moins.

9. Les lignes qui suivent sont extraites d'une étude réalisée par Olivier Cosnefroy et Thierry Rocher, chargés d'études à la DEPP. Elle a été publiée dans le Dossier Éducation et Formations, n° 166, de mai 2005. Le redoublement au cours de la scolarité obligatoire : nouvelles analyses, mêmes constats, Paris, MENESR-DEPP.

[Il en ressort] également que le redoublement n'est pas la meilleure façon de gérer l'hétérogénéité des élèves. Si nous nous référons à l'écart type de la distribution des performances à PISA comme indicateur de dispersion, nous observons des valeurs comparables dans les pays qui pratiquent le passage automatique. Ainsi, les inégalités de performances ne sont pas moins élevées dans les pays qui pratiquent le redoublement.»

Ces conclusions sont d'autant plus importantes que le redoublement reste très pratiqué, voire très « populaire », dans le système éducatif français, comme le montre le tableau 3.1 : plus de quatre élèves sur dix y sont en retard à 15 ans.

Tableau 3.1. PISA 2003, répartition des élèves de 15 ans selon le niveau scolaire¹⁰

	Classe fréquentée	Répartition
Élèves en avance	1 ^{re} générale et technologique	2,2
Élèves à l'heure	2 ^{de} générale et technologique	49,6
	2 ^{de} professionnelle	7,4
Élèves en retard	3 ^e générale	26,8
	3 ^e autre (SEGPA, technologique, insertion)	7,7
	4 ^e	5,2
	Autre	1,1
Ensemble		100,0

Les résultats des élèves français à PISA font nettement ressortir les écarts de performance des élèves en retard. Ils le font d'autant plus que la France a adjoint à l'échantillon des élèves de 15 ans de PISA 2003, deux échantillons supplémentaires d'élèves : l'un d'élèves de 14 ans, à l'heure en 3^e, l'autre d'élèves de 16 ans, en retard en 2^{de} générale et technologique. La position de l'ensemble de ces élèves selon leur âge et leur position dans les différents niveaux du système scolaire est éclairante.

10. Les élèves interrogés au titre de PISA 2003 sont nés en 1987 ; ils avaient 15 ans révolus lors de l'enquête, en mai 2003. Selon la norme d'âge du système éducatif français, les élèves « à l'heure » doivent être en classe de seconde à cet âge. Pour la quasi-totalité de ceux qui n'y sont pas encore parvenus, leur retard est dû à un, voire à plusieurs redoublements.

La figure 3.1 situe ces positions en référence au « palmarès » international de PISA 2003, en compréhension de l'écrit, étant précisé que les données ne sont pas sensiblement différentes pour les autres domaines évalués. Globalement, les élèves français se situent à la moyenne des pays participants. Mais l'examen des différences de performances selon la classe fréquentée à 15 ans est frappant : les élèves français à l'heure à 15 ans et scolarisés en 2^{de} générale et technologique obtiennent d'excellents résultats, supérieurs même à ceux de la Finlande. À l'inverse, les élèves ayant un an de retard se situent dans le bas de l'échelle, juste avant la Russie. Les élèves encore en 4^e (deux ans de retard) sont, quant à eux, à un niveau inférieur à celui du Mexique. Les élèves en 2^{de} professionnelle à 15 ans forment une population particulière : ils ne sont pas représentatifs des élèves de 2^{de} professionnelle car la grande majorité des entrants dans cette classe a généralement au moins un an de retard. Ces élèves obtiennent des résultats inférieurs à la moyenne internationale et relativement proches de ceux des élèves ayant un an de retard.

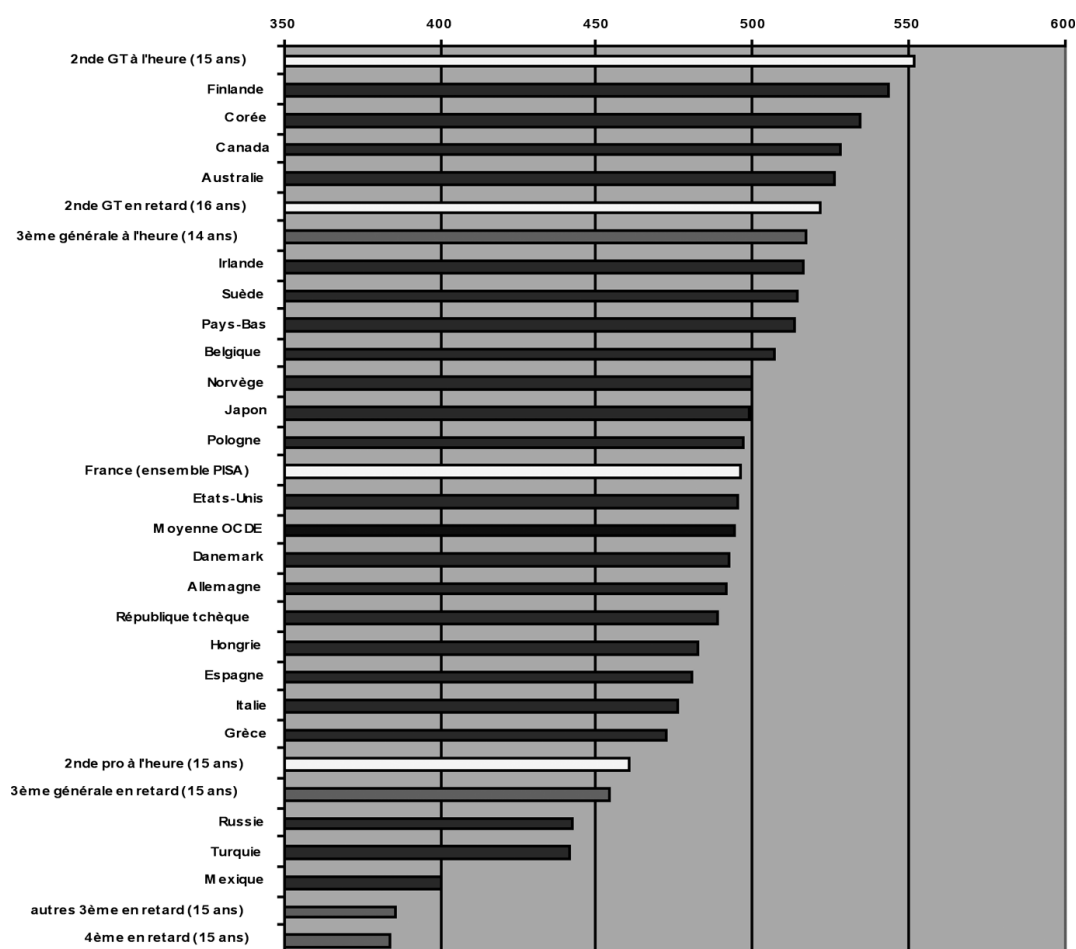


FIGURE 3.1. Situation des élèves français selon le retard scolaire (échelle internationale de PISA 2003, compréhension de l'écrit)

Comme l'indiquent les auteurs de l'étude citée plus haute : *« Ces différences amènent deux remarques. La première consiste à s'interroger sur la signification de la moyenne globale française tant l'écart entre ces sous-groupes d'élèves est important. La seconde correspond au fait que, dans cette enquête, les différences selon le retard scolaire sont amplifiées puisque les élèves de 3^e, par exemple, n'ont pas bénéficié du programme de la 2^{de} générale et technologique. »*

Toutefois, l'examen des résultats des échantillons supplémentaires des élèves de 3^e et de 2^{de} générale et technologique ne conduit pas à attribuer cet écart de performance à l'action pédagogique : l'écart est relativement faible entre les performances des élèves de 14 ans, à l'heure en 3^e générale et celles de leurs camarades d'un an de plus, à l'heure en 2^{de} générale et technologique, surtout si on le rapproche de celui qui sépare les performances respectives des élèves à l'heure et en retard à 15 ans.

L'examen de ces derniers résultats nous confirme plutôt que le protocole de PISA permet de mesurer les compétences générales des jeunes en fin de scolarité obligatoire, comme rappelé plus haut. Il ne permet pas vraiment de « capter » les connaissances et les compétences de niveau supérieur que l'enseignement de la seconde générale et technologique se propose de faire acquérir aux lycéens.

En tout état de cause, ces conclusions ont permis d'étayer plus encore le fait que le redoublement ne constitue pas un remède efficace à la difficulté scolaire. Elles ont contribué à inciter les responsables du système éducatif à promouvoir des mesures visant à prévenir ces redoublements, notamment les programmes personnalisés de réussite éducative (PPRE) institués par la Loi d'orientation pour l'avenir de l'école d'avril 2005¹¹.

2 MIEUX UTILISER PISA ?

Ainsi, contrairement à ce que l'on a parfois pu entendre, le système éducatif français met à profit les résultats de PISA de diverses manières comme il le fait des autres enquêtes internationales auxquelles il participe, notamment celles de l'IEA¹². Mais, il lui faut se poser la question du rendement qu'il peut tirer de ces enquêtes dont le coût est particulièrement élevé, aussi bien en termes financiers qu'en moyens humains et matériels de tous ordres et en mobilisation des établissements et des élèves qui font partie des échantillons.

11. Art. L.311-3-1 du Code de l'éducation : « À tout moment de la scolarité obligatoire, lorsqu'il apparaît qu'un élève risque de ne pas maîtriser les connaissances et les compétences indispensables à la fin d'un cycle, le directeur d'école ou le chef d'établissement propose aux parents ou au responsable légal de l'élève de mettre conjointement en place un programme personnalisé de réussite éducative. »

12. *International Association for the Evaluation of Educational Achievement.*

Ceci d'autant plus qu'un nouveau cycle de PISA est engagé jusqu'en 2015. Trois questions méritent d'être examinées à ce titre.

2.1 **Rendre PISA plus pertinent au regard des décisions qui incombent aux responsables nationaux des systèmes éducatifs**

Cette première question concerne PISA lui-même. C'est, comme on l'a vu plus haut avec l'exemple du redoublement, par la mise en rapport des performances des élèves des différents pays avec des caractéristiques de leurs systèmes éducatifs que l'on peut déceler des caractéristiques de ces systèmes qui méritent d'être réexaminées. S'il est hors de question, à mon sens, d'envisager d'aller vers un système éducatif « universel » dont PISA serait la pierre de touche, il est incontestable que les systèmes éducatifs des pays de l'OCDE rencontrent un certain nombre de problèmes communs, et ont intérêt à confronter les solutions qu'ils leur apportent. De ce point de vue, le comparatisme international ne fournit pas de solutions toutes faites, mais il permet, en se confrontant aux autres, de poser les questions qui incitent à rechercher des améliorations.

Dans cet esprit, l'exploitation des données de PISA peut s'avérer quelque peu frustrante, non pas tant parce que la mesure de la performance des élèves des différents pays serait discutable¹³, mais surtout parce qu'elle n'est pas assortie de données sur les caractéristiques des **différents systèmes** éducatifs, qui soient suffisamment précises et organisées de façon pertinente pour permettre de cerner les facteurs sur lesquels les responsables pourraient agir avec le plus d'efficacité. Le pouvoir explicatif de PISA s'en trouve limité.

C'est ainsi que l'on a pu dire, à propos de PISA, qu'« *un écueil classique des comparaisons internationales est d'accorder trop d'importance à un élément isolé de son contexte. C'est le cas quand l'OCDE rapproche, en suggérant, malgré des précautions rhétoriques, une relation de cause à effet, les bonnes performances de certains pays et le fait que les établissements y jouissent d'une certaine autonomie, alors que seule une corrél-*

13. On sait qu'elle l'est et le Haut Conseil de l'évaluation de l'école a pu souligner que « malgré les progrès considérables accomplis au fur et à mesure du développement des enquêtes comparatives internationales, il n'est pas possible de garantir que leurs résultats ne sont pas exempts de biais de nature diverse, d'autant plus que la volonté d'étendre les mêmes programmes d'enquêtes, considérés comme ayant valeur universelle, à des pays de plus en plus nombreux élargit l'éventail des niveaux de développement, des cultures, des langues, etc., dont il faut tenir compte » et qu'« il faut être conscient que la mesure des compétences des élèves (et des adultes) n'a pas la simplicité, la robustesse et l'universalité du système métrique. Dans le domaine de... la mesure en éducation, les mesures sont des constructions qui résultent de choix politiques et méthodologiques » (Avis déjà cité du Haut Conseil : « La France et les évaluations internationales des acquis des élèves »).

lation est observée et que nombre d'autres caractéristiques des pays performants et/ou dont les établissements sont autonomes, ne sont pas prises en compte. Il faut donc insister sur la nécessité de ne pas conclure trop vite à une relation causale entre tel facteur et tel phénomène, la démarche la plus sûre étant sans doute, en construisant des typologies, de rechercher quels sont les facteurs qui jouent de concert»¹⁴.

L'exemple de l'autonomie des établissements, concept particulièrement flou si on ne précise pas le contenu organisationnel de cette autonomie, est manifeste. De même, on pressent que les systèmes éducatifs qui ont une organisation continue de leur scolarité obligatoire, sans rupture marquée ni orientation des élèves dans des filières distinctes au cours de celle-ci, ont plutôt de meilleurs résultats que ceux qui marquent une rupture ou qui orientent les élèves à la fin ou peu après la fin de l'école primaire (CITE 1). Mais, dans l'un et l'autre cas, ce sont surtout pour l'instant des pressentiments ou des affirmations dont la démonstration doit être appuyée sur une meilleure connaissance comparative, d'une part, des modes d'organisations de l'autonomie des établissements et, d'autre part, des modalités de parcours de la scolarité obligatoire, dans les différents pays de l'OCDE.

Les progrès les plus décisifs à accomplir aujourd'hui, dans la collecte de données organisée dans le cadre des cycles de PISA, ne sont donc sans doute pas du côté de la mesure des performances, mais bien plutôt du côté de cette face quelque peu cachée ou en tout cas peu valorisée – et sans doute pas assez travaillée – que sont les «questionnaires de contexte» qui doivent renseigner sur les conditions d'enseignement des élèves, sur leur degré d'engagement dans les tâches qui leur sont proposées, etc., et surtout sur les caractéristiques «malléables» des systèmes éducatifs, c'est-à-dire celles sur lesquelles les responsables de ces systèmes peuvent agir : organisation des cursus, des procédures d'orientation, des modes d'évaluation, des processus et des niveaux de décision, etc.

Il est significatif, à ce sujet, de constater que nombre de chercheurs, qui se proposent de mettre en relation les modes d'organisation et de fonctionnement des systèmes éducatifs et les différents aspects des performances de leurs élèves, doivent consacrer de grands efforts à la constitution de bases de données comparatives que ne fournissent, ni les actuels questionnaires de contexte de PISA, ni les indicateurs internationaux de l'OCDE¹⁵.

14. *Organisation scolaire et inégalités sociales de performance : les enseignements de l'enquête PISA*, Marie Duru-Bellat, Nathalie Mons et Bruno Suchaut, in *Education et formations*, n°70, décembre 2004, MENESR-DEPP, PARIS.

15. Voir par exemple l'article cité dans la note précédente, et «Modèles pluriels d'école unique, efficacité et équité : l'éclairage de la comparaison des pays de l'OCDE», la communication de Nathalie Mons, LES-Université Pierre Mendès-France-Grenoble II, au colloque «Construction/déconstruction du collège unique : les enjeux de l'école moyenne», Paris VIII-AECSE-IUFM de Créteil, 26-28 octobre 2005.

C'est pourquoi la France a insisté, et insiste encore, pour que des données détaillées sur les systèmes éducatifs, dont de premières recherches ont pu montrer qu'elles pouvaient être pertinentes, soient recueillies, lors de la collecte de données de PISA 2009.

2.2 Participer plus activement à la conception de PISA, et plus généralement des enquêtes internationales

C'est à juste titre que le Haut Conseil de l'évaluation de l'école a pu souligner le *«faible engagement politique et scientifique français dans la conception des programmes d'enquêtes, aussi bien à l'IEA qu'à l'OCDE. Les responsables de la politique éducative et pédagogique ont laissé les experts et les techniciens décider de la participation aux enquêtes et de l'investissement dans leur conception et leur réalisation.»*

2.3 Insister pour enrichir PISA et le rendre plus pertinent

Insister pour enrichir PISA et le rendre plus pertinent, comme on vient de l'évoquer, suppose que l'on se dote d'une expertise qui rende possible une participation active à sa conception. Tel est bien le parti pris par la Direction de l'évaluation et de la prospective du ministère de l'Éducation nationale : elle avait suscité la création d'un consortium qui a répondu, sans succès, à l'appel d'offres que l'OCDE avait lancé en 1997 pour la réalisation de la première phase de PISA (2000) ; aujourd'hui, cette même direction est partie prenante du consortium qui, après avoir été sélectionné à l'issue d'un nouvel appel d'offres, a en charge la réalisation des «questionnaires de contexte» de PISA 2009.

3 COMBINER ÉVALUATIONS INTERNATIONALES ET ÉVALUATIONS NATIONALES

Ce point, sur lequel on conclura, est certainement le plus important. Pour l'évoquer, on se référera une fois de plus à l'avis du Haut Conseil de l'évaluation de l'école sur la France et les évaluations internationales des acquis des élèves : *«Le Haut Conseil juge que les enquêtes internationales comparatives des acquis des élèves constituent un “regard” extérieur qui doit impérativement être associé au “regard” intérieur apporté par les évaluations et études nationales. Elles doivent donc avoir toute leur place dans le dispositif d'évaluation et de pilotage de notre système éducatif, sans que cette place soit exclusive car, à défaut d'évaluations nationales directement liées aux objectifs propres de notre système éducatif, le risque pourrait exister de voir ériger les enquêtes internationales en modèle éducatif universel, voire, en caricaturant quelque peu, de faire de bons*

résultats à ces enquêtes la finalité principale ou unique du système éducatif.»

Il est en effet essentiel que, dans chaque pays, les enquêtes comparatives internationales soient réalisées et analysées au sein d'un dispositif d'ensemble de pilotage du système éducatif.

C'est une orientation que s'est donnée la France, puisque la Direction de l'évaluation, de la prospective et de la performance du ministère de l'Éducation nationale a engagé un programme d'évaluations-bilans nationales, qui verra se succéder chaque année, par période de six ans, des évaluations fondées sur les objectifs des programmes nationaux français dans les principaux domaines d'enseignement, et ceci à deux moments clefs de la scolarité, la fin de l'école primaire (CITE 1) et la fin du collège (CITE 2)¹⁶.

On se donne ainsi les moyens de suivre les évolutions des acquis des élèves au regard des objectifs que la société française assigne à son École, tout en se réservant la possibilité de revoir ces objectifs à l'aune des comparaisons internationales. On se donne aussi la possibilité de réinterroger l'organisation de notre système éducatif, compte tenu non seulement des comparaisons internationales mais aussi de nos objectifs propres, qui peuvent – et doivent – ne pas se limiter au « compromis international » que constitue le protocole de PISA.

16. Ont ainsi été réalisées :

- en 2003, des évaluations de *la maîtrise du langage et de la langue française en fin d'école primaire*, note d'évaluation n°04.10, par Fabienne Gibert, Jacqueline Levasseur et Jean Marc Pastor, et *des compétences générales des élèves en fin de collège*, note d'évaluation n°04.09, par Laurence Dauphin et Bruno Trosseille, octobre 2004 ;
- en 2004, des *évaluations des compétences en langues vivantes des élèves en fin d'école et en fin de collège*, notes d'évaluation n° 05-10, 05.09, 05-08, 05-07 et 05-06, de septembre 2005, par Sylvie Beuzon, Monique Levy, Annick Marquis, Evelyne Sol et Bruno Trosseille ;
- en 2005, une *approche des attitudes à l'égard de la vie en société des élèves de fin d'école et de fin de collège*, par Séverine Dos Santos, Fabienne Gibert et Saïd Yacoub, note d'évaluation n°06.02, d'août 2006, MENESR-DEPP.

Suivront :

- en 2006, des évaluations des compétences en histoire-géographie-éducation civique ;
- en 2007, des évaluations des compétences en sciences, et
- en 2008, en mathématiques.

Puis le cycle reprendra.

Chapitre 4

Quelques aspects éthiques de l'évaluation

Bernard Rey

Évaluer les produits d'une activité humaine n'est pas un acte neutre. Chacun, qu'il soit évaluateur ou évalué, sent bien que cela engage des considérations éthiques. Mais pour ne pas en rester à une inquiétude diffuse, il convient de se demander pourquoi il en est ainsi.

D'une part, toute évaluation engage une relation à l'autre : des humains émettent des jugements sur les actions d'autres humains. Même quand l'évaluation est organisée de manière très indirecte, même quand elle s'effectue selon des dispositifs qui paraissent laisser peu de place à l'initiative des sujets humains, il y a bien finalement des évaluateurs et des évalués, des personnes qui participent à la construction d'un avis et d'autres personnes sur les actes desquels cet avis est porté. Par là s'établit, même momentanément, une asymétrie. D'où l'inévitable question : qu'est-ce qui donne autorité à quelqu'un pour émettre un jugement sur les actes de quelqu'un d'autre ? Quelle est la légitimité de ses jugements ?

D'autre part, toute évaluation est, comme l'étymologie le rappelle, une référence à une valeur. C'est relativement à un ordre de préférable que l'évaluateur situe la réalisation de l'évalué. Il y a donc, dans le fait même d'évaluer, un geste profondément éthique, en ce qu'il affirme que tous les actes humains ne se valent pas et que certains peuvent être tenus comme meilleurs que d'autres.

Mais on sait bien également que la détermination de la valeur ou des valeurs auxquelles on va se référer pour juger, pose problème. C'est une banalité de dire que nous vivons dans une société dans laquelle il n'y a plus de consensus sur les valeurs. Bien entendu, il est toujours possible d'instituer, pour la société toute entière ou pour telle ou telle institution (par exemple l'école), des normes. Mais celles-ci, si elles sont incontestables d'un point de vue juridique ou réglementaire, sont toujours susceptibles d'être discutées et contestées d'un point de vue éthique.

Plus généralement, la prétention à fonder les valeurs sur un ordre des choses ou sur l'interprétation de la volonté divine s'est aujourd'hui dissipée. Nous ne pouvons considérer les valeurs que comme des constructions humaines, toujours sujettes à discussion, à délibération et à révision. C'est en ce sens que l'éthique est au cœur des procédures d'évaluation.

Du même coup, l'évaluation est un acte par lequel un humain convoque un autre humain à partager les valeurs au nom desquelles il évalue. Quelles sont les conditions qui peuvent rendre légitime cette convocation ? Nous voudrions tenter, dans les lignes qui suivent, de formuler quelques-uns de problèmes que suscite cet acte. Nous commencerons par décrire comment les dispositifs techniques de l'évaluation risquent toujours d'en occulter les enjeux éthiques (1). Nous évoquerons ensuite les contradictions que peut provoquer l'évaluation dans la relation à l'autre (2). Un troisième moment sera consacré aux problèmes éthiques plus spécifiques que pose l'évaluation dès lors qu'elle concerne des personnes en formation (élèves, étudiants, stagiaires, etc.) (3).

1 LES PIÈGES DE LA TECHNICITÉ

D'un point de vue éthique, on s'attend à ce que l'évaluateur ait pleine conscience de son statut et des risques qui y sont attachés : qu'il n'oublie pas que c'est sur les actes d'autrui qu'il émet un jugement et qu'il le fait selon une valeur susceptible d'être mise en débat.

Cette exigence n'est pas si facile à satisfaire qu'il ne le semble. La déresponsabilisation de celui qui évalue est toujours possible. Il peut avoir le sentiment que ce n'est pas vraiment lui qui évalue et que son jugement, loin d'être prononcé au nom d'une valeur, est un simple constat. Or ce sentiment fallacieux est favorisé par les dispositifs organisationnels et instrumentaux qui, dans beaucoup de situations, servent de support technique à l'évaluation.

C'est le cas par exemple lorsque l'évaluation émane d'une pluralité de personnes. Ainsi le dispositif du « conseil de classe », dans l'enseignement secondaire, permet à chaque membre, enseignant ou chef d'établissement, de considérer que la décision prise à propos d'un élève ne provient pas de

lui. Nous ne disons pas que chaque participant à un conseil de classe adopte systématiquement cette attitude ; nous disons que la situation d'évaluation collective autorise une dilution de la responsabilité. Il en va de même de toute instance collégiale et le phénomène est évidemment renforcé par le caractère confidentiel de la délibération.

Mais il existe encore une autre manière de diluer la responsabilité de l'évaluateur. Elle tient non plus à la pluralité des acteurs de l'évaluation, mais à la pluralité des évalués. Il en va ainsi chaque fois qu'on évalue une organisation, un système ou une institution : évaluer un établissement, un corps d'enseignants, un système éducatif, donne l'impression qu'on évalue une structure impersonnelle. Cette dépersonnalisation apparente de ce qui fait l'objet d'un jugement peut donner le sentiment d'une moindre responsabilité, puisque ce n'est pas les hommes qu'on juge, mais la structure dans laquelle ils agissent.

En réalité, ce sont bien des personnes qui sont évaluées, car ce sont des personnes qui opèrent au sein des structures, qui ont participé à leur mise en place ou qui assurent leur maintien. À titre d'exemple, les enquêtes internationales sur la performance des élèves dans différentes disciplines à un âge donné, donnent l'impression de ne porter de jugement sur personne. En fait, il n'en est rien : en bout de ligne, il y a des élèves, des enseignants, des inspecteurs et des responsables de divers types. Ce simple rappel engage du coup à s'interroger sur les valeurs qui sous-tendent les épreuves.

Ce que nous venons d'évoquer, ce sont des cas d'oubli plus ou moins involontaire de la condition d'évaluateur, par une sorte d'affaiblissement de la vigilance éthique. Mais on peut aussi vouloir réduire volontairement la responsabilité de l'évaluateur, pour des raisons qui ne sont pas dénuées de légitimité, comme on va le voir.

Dès lors que l'évaluation se situe dans le cadre d'une institution ou d'une organisation d'une certaine ampleur, les actes successifs qui sont techniquement indispensables sont assurés par une multiplicité d'instances : une instance détermine les règles et les critères de l'évaluation, une autre incarne ces règles dans des indicateurs, dans des tests ou des épreuves ; d'autres acteurs encore constatent les performances des évalués à ces tests ou épreuves ou bien confrontent leurs actes aux indicateurs ; enfin une autre instance en tirera les conséquences.

Or cette démultiplication des gestes d'évaluation constitue une garantie : garantie d'une situation de droit en laquelle toute la chaîne des décisions évaluatrices n'est pas rassemblée en une seule main, garantie qu'un seul individu en position autocratique ne décidera pas des principes, des tests, des performances et des conséquences. Il y a là l'équivalent, dans le domaine de l'évaluation, de la séparation des pouvoirs dans l'ordre politique.

Mais en même temps, cette répartition des actes constitutifs de l'évaluation ouvre la possibilité que chaque acteur ne se juge pas responsable de l'évaluation globale, estime qu'il a fait ce que, réglementairement, il devait faire et que ce qui est décidé en amont ou en aval ne le concerne pas. La nécessaire séparation des actes d'évaluation rend possible une tranquille indifférence à ses effets.

La littérature consacrée à l'évaluation scolaire fait parfois la différence entre la mesure de la performance et la décision. Il y aurait donc deux acteurs ou du moins deux postures différentes (éventuellement assumées successivement par la même personne) : un acte technique de confrontation d'une performance à des critères préétablis et d'autre part la décision qu'on en tire (décision didactique, certificative, organisationnelle, financière, politique, etc.). Une telle distinction peut faire croire à celui qui effectue la mesure que sa responsabilité n'est pas engagée.

Mais il existe encore une autre manière d'échapper à la responsabilité du jugement de personnes sur des personnes, c'est de faire passer l'évaluation pour une mesure ou un constat.

D'une certaine manière, l'abondante littérature consacrée à la technologie de l'évaluation consiste à tenter de ramener un jugement selon une valeur à n'être qu'une mesure objective. Bien évidemment, ces efforts ne sont nullement dépourvus d'intérêt. L'enjeu est, là encore, d'échapper à l'arbitraire et de trouver les conditions d'un jugement exact et juste. Mais il y aurait une erreur conceptuelle fondamentale à ramener l'évaluation à une mesure.

Bien entendu, il est toujours possible d'opérer sur la réalité, quelle qu'elle soit, des mesures. Mais les dimensions que l'on mesure et que l'on prend en compte sont l'effet du choix d'un sujet. Si un élève fait une erreur de calcul dans un problème de mathématiques, cette erreur est bien un fait objectif que l'on constate ; mais c'est l'évaluateur qui choisit de le considérer comme un fait à prendre en compte et de lui attribuer plus ou moins d'importance.

Quand il s'agit d'évaluer une organisation, un système ou la pratique d'un corps professionnel, la tentation est forte de faire comme s'il s'agissait d'un simple constat de dysfonctionnements qui seraient objectifs. L'évaluateur peut se poser comme le technicien d'une mécanique sociale sur laquelle il opérerait un diagnostic. Or là encore, il y a un processus d'occultation des valeurs : les dysfonctionnements n'existent pas en eux-mêmes ; ils ne peuvent être qualifiés tels que par référence à ce qu'on juge être le bon fonctionnement ou le fonctionnement « normal ». Ce fonctionnement satisfaisant est considéré comme tel par référence à une valeur. On peut s'en rendre compte en remarquant qu'une institution peut être considérée comme défailante si on la réfère à une valeur d'équité alors qu'elle sera considérée comme satisfaisante si on la confronte à une norme d'efficacité.

Évaluer, c'est émettre un jugement qui n'est pas un pur constat, d'une part en ce qu'il aura une incidence sur d'autres humains, d'autre part en ce qu'il n'est jamais relevé pur d'une réalité dont les caractères seraient déjà là : il est toujours jugement en fonction d'une valeur.

Voilà donc le premier problème : une évaluation a besoin d'objectivité ; elle doit évacuer l'arbitraire, le jugement que prononcerait l'évaluateur selon sa subjectivité ou son intérêt personnel. D'où l'importance de dispositifs qui spécialisent les rôles, qui fournissent des instruments à l'appréciation, qui, autant qu'il est possible, ramènent l'exercice du jugement à des mesures, des constats et des algorithmes. Mais cet effort indispensable peut servir à occulter que c'est toujours selon une normativité humaine que l'on juge.

2 LE RAPPORT À AUTRUI DANS L'ÉVALUATION

Parce que l'évaluation est confrontation à un préférable, et que celui-ci n'a jamais rien d'incontestable, le débat doit être ouvert sur le choix des valeurs. Il importe tout particulièrement qu'il soit rendu possible entre l'évaluateur et l'évalué. Pour que ce soit un véritable débat et non pas une simple déclaration du premier au second, il est indispensable qu'il soit conçu comme un échange de sujet à sujet, à égalité, sans hiérarchie préétablie.

De plus, le but ultime de l'évaluation n'est pas tant de mettre en cause ou de critiquer tel ou tel aspect des prestations de l'évalué. Il est plutôt de faire partager par l'évalué des considérations sur ce qui est souhaitable ou sur un idéal vers lequel on veut tendre. L'évaluation ne trouve donc son achèvement que dans un dialogue entre pairs.

Mais l'évaluation n'est possible que parce que les rôles d'évaluateur et d'évalué sont distincts, ce qui implique, nous l'avons vu, une asymétrie, laquelle ne prépare évidemment pas à l'établissement d'une parité.

En outre la fonction d'évaluateur appelle l'adoption d'une posture particulière vis-à-vis de l'évalué. Il doit se mettre en position d'extériorité et d'objectivité vis-à-vis de l'autre. S'il veut apprécier aussi objectivement que possible l'action de l'évalué et les effets de celle-ci, il doit suspendre la relation de sujet à sujet qu'il aurait spontanément avec l'évalué, relation par laquelle on réfère les actes d'autrui à des raisons qu'on lui attribue, que l'on partage ou, du moins, que l'on peut comprendre. Si je comprends les motifs de l'autre, puis-je encore l'évaluer ?

L'évaluation n'a de sens et d'intérêt que parce qu'elle ouvre, à propos d'un acte, sur un jugement distinct de celui qu'a émis spontanément son auteur au cœur de l'action. L'acte d'évaluation suppose donc la rupture, au moins provisoire, de la relation d'empathie, mais tout aussi bien d'antipathie. Non seulement il y a suspension de l'aspect affectif de la relation, qu'il soit positif ou négatif, mais aussi neutralisation du sens partagé. Saisir l'autre

et ses actes avec objectivité, c'est les constituer en objets, objets d'observations, de description, d'appréciation en fonction de critères extérieurs à leur auteur.

Certes, cette objectivation de l'autre semble pouvoir être évitée grâce à la distinction entre la personne et ses actes. On peut poser, en effet, que l'évaluateur doit prendre une posture objectivante à l'égard des actes de l'évalué, mais qu'il doit conserver, vis-à-vis de l'évalué lui-même, une posture de communication de sujet à sujet. Bien entendu, il arrive qu'il en soit ainsi. Mais souvent, dans le mouvement de l'évaluation, l'évaluateur glisse insensiblement du jugement sur la performance de l'autre au jugement sur l'autre.

Ce glissement est particulièrement fréquent lorsque l'évaluation est effectuée dans le cadre d'une formation. Car d'une certaine manière, la formation est la mise en œuvre d'un processus par lequel on s'attend à ce que le formé se modifie dans un sens qu'on juge favorable. Si l'on constate un écart entre ce que fait aujourd'hui le formé et ce qu'on souhaite qu'il sache faire, la tentation est forte d'expliquer cet écart par ce qu'il est, par les mécanismes psychologiques ou sociologiques, voire biologiques, qui l'habitent. L'autre n'est plus alors saisi comme sujet, mais comme objet à étudier. Évaluer, c'est bien aussi expliquer pourquoi ce qu'on évalue est tel qu'on l'observe.

Dans l'évaluation en formation, cette mise en explication de l'autre peut prendre des formes diverses. Elle peut se ramener à une explication fixiste : ainsi en va-t-il des explications qui nient l'éducabilité de l'autre ou ce qu'on appelait autrefois sa « perfectibilité ». Mais elle peut aussi inscrire l'autre dans un statut évolutif. C'est le cas de l'interprétation des difficultés scolaires comme effets de carences culturelles.

Cette tentation de muer l'autre en objet est d'autant plus grande que cette objectivation de l'autre ne prend pas l'allure d'une condamnation. Elle est au contraire le plus souvent bienveillante. Elle conduit en effet, précisément parce qu'on prétend saisir les causes des insuffisances de l'autre, à ne pas lui en faire grief. Le jugement moral est remplacé par une théorisation scientifique, ce qui a deux conséquences, également dommageables du point de vue éthique.

D'abord, l'évaluation, lorsqu'elle prend cette direction, peut déboucher sur une réduction des exigences du formateur vis-à-vis de l'individu en formation. Le confort que cette position peut apporter à ce dernier ne compense évidemment pas la sous-estimation qu'elle implique.

Ensuite, la substitution de l'explication scientifique au jugement en fonction d'une valeur modifie la nature de la relation entre l'évaluateur et l'évalué. On passe d'un message à l'autre à un message sur l'autre. On ne s'adresse plus à l'autre. On explique ce qu'il est et ce qu'il fait. L'évaluation, dans cette dérive, semble devoir passer par la connaissance scientifique de

l'autre, laquelle se donne comme une garantie de justesse et donc de justice. Mais, ce faisant, elle conduit à épingle l'autre à sa singularité, plutôt que de l'inviter à agir en fonction d'une valeur.

Mais de même que l'évaluateur, dans la dynamique objectivante qui est la sienne, peut déchoir l'évalué de sa qualité de sujet, l'inverse peut tout aussi bien se produire.

Il peut arriver que l'évalué ne considère pas le jugement de l'évaluateur sur lui comme légitime. Il impute le contenu de ce jugement à des caractéristiques subjectives de l'évaluateur, à ses émotions, à l'effort qu'il ferait pour défendre ses intérêts. C'est ainsi que des adolescents peuvent estimer que l'évaluation de leur travail par le professeur est l'effet de la mauvaise volonté de celui-ci ou de son hostilité. «Si j'ai obtenu une note si basse, c'est parce que le professeur ne m'aime pas», avec différentes variantes possibles : «c'est parce qu'il a voulu se venger» ou bien : «c'est parce que j'ai exprimé des idées avec lesquelles il n'est pas d'accord». Le jugement de l'évaluateur est ainsi rapporté à sa personnalité, à ses caractéristiques psychologiques, sociales ou autres.

On en trouve des exemples dans d'autres situations que la relation professeur-élève. Par exemple, dans l'évaluation du personnel par une instance hiérarchique, le soupçon est toujours possible de considérer le jugement comme un effet de l'exploitation patronale, de la volonté de tirer le maximum des employés, de l'intention de licencier, de modifier les conditions de travail.

Globalement, l'évalué peut juger que le message d'évaluation qui lui est adressé n'est pas l'effet de la mise en œuvre d'une valeur, mais plutôt l'effet d'un rapport de force ou d'un rapport de pouvoir.

Chez l'évalué, une telle attitude correspond au refus de prendre en compte le statut de sujet de l'évaluateur, sa capacité à rester libre de son jugement. Mais bien entendu, la première condition pour éviter de telles dérives est que l'évaluateur ne donne aucune prise à un tel soupçon, c'est-à-dire que l'évaluation s'établisse selon une valeur universalisable et non pas au service direct ou indirect de son intérêt particulier.

Évaluer, c'est affirmer une valeur et la faire valoir auprès d'un humain que l'on juge accessible à cette valeur. Être évalué, c'est recevoir un message sur soi-même et penser qu'il peut être fondé et inspiré par une valeur. La relation d'évaluation exige un respect mutuel de l'évaluateur et de l'évalué qui se reconnaissent comme des sujets, c'est-à-dire comme des êtres accessibles aux valeurs.

3 LES PIÈGES DU CONTRAT

Bien entendu, dans l'univers de la formation, il existe des évaluations qui n'ont rien de spécifiques, par exemple celles qui concernent les professionnels et les organisations qui interviennent dans ce champ : on peut évaluer les enseignants, les établissements scolaires, les administrations scolaires, les systèmes éducatifs, les curricula, les innovations pédagogiques, etc.

Mais dès lors que cet ouvrage a pour titre *Évaluer pour former*, on fera une place particulière à l'évaluation des personnes en formation (élèves, étudiants, stagiaires, personnes en formation continue, etc.), car, dans leur cas, l'éthique de l'évaluation exige de prendre en compte l'éthique de la formation.

Pour répondre à l'exigence, décrite au paragraphe précédent, du dialogue entre l'évaluateur et l'évalué, les travaux d'évaluation scolaire ont, depuis trente ans, conduit à des avancées majeures. Il s'est agi, pour l'essentiel, de rendre aussi explicites que possible les principes de l'évaluation : des référentiels ont été construits, précisant les objectifs attendus ; les critères d'évaluation ont été définis ; des indicateurs permettent de les appliquer en tentant de réduire la part de subjectivité de l'évaluateur. Il ne s'agit pas seulement de rendre l'évaluation plus exacte et plus juste, mais aussi de communiquer à l'évalué ce qu'on attend de lui et de l'inciter ainsi à partager les finalités de la formation dont il est l'objet. Non seulement ces efforts répondent à l'exigence de dialogue entre l'évaluateur et l'évalué, dont nous parlions au paragraphe précédent, mais ils rejoignent les impératifs de la formation, en ce qu'ils permettent de mieux entraîner les élèves aux tâches qu'on attendra d'eux lors des évaluations et à mieux centrer ainsi le processus d'apprentissage sur les objectifs visés.

Ce travail apporte des bénéfices incontestables et, notamment, les suivants :

- Il contribue à convaincre les élèves que le jugement sur leur performance ne dépend pas de l'arbitraire du maître, ni de sa subjectivité ni de la qualité de la relation qu'ils ont avec lui.
- Il garantit que c'est bien la performance de l'élève que l'on évalue et non pas sa personne.
- En connaissant mieux les critères d'évaluation et en étant entraînés à les satisfaire, les élèves arrivent à comprendre que le succès scolaire ne dépend pas de qualités mystérieuses, mais de la mise en œuvre systématique de certaines procédures. L'implicite des tâches scolaires qui peut être socialement discriminatoire s'en trouve réduit.

Ainsi les problèmes que nous avons posés dans les paragraphes précédents semblent trouver leur solution, au sein de la formation, dans un effort de transparence.

Pourtant, si l'on veut satisfaire aux exigences éthiques de l'évaluation et de la formation, il semble qu'on ne puisse s'en tenir là. En effet, est-ce que la mission de l'école est achevée, lorsque l'élève se montre capable d'accomplir une tâche à laquelle on l'a entraîné ? Il semble qu'en général l'ambition de l'école soit plus haute. Ce que l'enseignant espère, ce n'est pas seulement que l'élève sache répéter une action après entraînement, mais qu'il sache en plus déterminer par lui-même quand il y a lieu d'accomplir cette action. La capacité à déterminer par soi-même la pertinence d'une procédure est au centre de l'enseignement.

L'enjeu est d'échapper au dressage. L'école a pour but la conquête par l'élève d'une autonomie intellectuelle. Il ne doit pas se contenter de faire « ce que le maître a dit de faire » ; on veut le rendre capable de réfléchir par lui-même à ce qui convient le mieux dans chaque situation nouvelle. Il y a là une exigence éthique de l'enseignement.

Même si la notion de compétence n'est pas dénuée de confusion, elle a le mérite, en surgissant dans le champ éducatif depuis une vingtaine d'années, de mettre sur le devant de la scène cette question essentielle de l'autonomie intellectuelle de l'élève. En effet, savoir accomplir, quand on en reçoit la consigne, une tâche à laquelle on a été entraîné, cela peut être considéré comme une procédure de base, mais non pas tout à fait comme une compétence. Car une véritable compétence implique plutôt que l'individu soit capable, face à une situation inédite, de choisir par lui-même les procédures qui conviennent, parmi celles qu'il maîtrise.

Si l'on examine les référentiels de compétences dont se sont dotés les systèmes éducatifs ces dernières années, c'est bien ce type de compétence qu'on y trouve le plus souvent. Une compétence implique la saisie du bon escient.

Or cette exigence même entraîne à son tour un problème éthique de l'évaluation : puisque ce qu'on attend de l'élève, c'est qu'il détermine par lui-même ce qu'il convient de faire dans des situations nouvelles, on est amené à l'évaluer sur des tâches auxquelles il n'a pas été entraîné. Il a pu être entraîné aux procédures nécessaires à ces tâches, mais c'est à lui qu'il revient de décider de l'opportunité de mobiliser telle ou telle d'entre elles. Il y a ainsi ce que Brousseau appellerait une rupture du contrat didactique.

Ainsi l'exigence éthique, lorsqu'elle concerne l'évaluation en formation, se développe selon une contradiction : évaluer les élèves à l'école sur des tâches stéréotypées auxquelles on les a entraînés ne répondrait pas à l'impératif de former des êtres intellectuellement autonomes. Mais évaluer

les élèves sur l'affrontement à des tâches nouvelles, c'est rompre le contrat d'enseignement.

Ce problème, dont nous n'avancerons pas de solution, fait apparaître en tout cas l'insuffisance de la notion de contrat pour satisfaire aux exigences éthiques d'une formation. On pourrait penser en effet que l'on a satisfait aux impératifs de transparence et de stricte justice, dès lors que les protagonistes de la relation de formation et d'évaluation se sont engagés à des obligations mutuelles explicites : « je m'engage à t'enseigner les réponses aux questions que je te poserai lors de l'évaluation » dirait l'enseignant, ce à quoi l'élève répondrait « je m'engage à donner à tes questions les réponses que tu m'auras enseignées ».

Mais, nous l'avons vu, un tel contrat risquerait de transformer l'enseignement en dressage ; il ne serait pas formateur, au sens où il n'appellerait pas l'élève à dépasser ce qu'on lui a apporté et à se dépasser lui-même pour devenir autre que ce qu'il était au départ.

En outre, l'établissement d'un contrat dans le processus de formation et d'évaluation aurait d'autres effets négatifs :

- L'idée de contrat risquerait d'accréditer le sentiment que la relation de formation est de nature juridique. Il s'agirait pour l'élève ou l'étudiant, non pas au premier chef d'apprendre, mais de se conformer à des clauses formelles établies par l'institution de formation, avec, bien entendu, la possibilité de plaintes et de procès.
- L'évaluation risquerait alors d'apparaître comme un processus de rétribution des tâches. Le diplôme serait le salaire attribué pour avoir accompli les tâches scolaires. Or les tâches sont destinées à l'apprentissage ; leur intérêt n'est pas qu'elles soient accomplies pour elles-mêmes, mais qu'elles contribuent à ce que l'élève apprenne et se transforme. Elles n'ont d'intérêt que par cette finalité.

CONCLUSION

L'évaluation est par nature une entreprise éthiquement risquée. On prend le risque d'émettre un jugement sur les actes de l'autre. On pense pouvoir le faire au nom d'une valeur qu'on affirme universelle. Or cette valeur est une création humaine et son universalité est toujours incertaine. Par suite l'interrogation et la discussion collectives sont toujours à recommencer sur les principes qui président à un dispositif évaluatif. Les valeurs selon lesquelles on évalue sont toujours en débat. La clarté, la publicité et la transparence de ce débat sont une condition indispensable (et peut-être la seule concevable) d'une validité éthique de l'évaluation.

Cela se traduit en plusieurs exigences qu'on pourrait formuler comme une liste d'impératifs à destination de l'évaluateur :

- L'évaluateur ne peut pas céder à l'illusion qu'il est l'exécutant docile de règles qui lui seraient extérieures. Prendre une part, même très partielle, à un processus évaluatif, c'est avoir la responsabilité de l'ensemble. Nier cette responsabilité conduirait à faire de l'évaluation une mécanique sans acteur ni valeur.
- L'évaluateur ne peut céder à l'illusion que l'évaluation est un constat ou une mesure. Même dans ses dispositifs les plus techniques et les plus instrumentés, elle se fait au nom de valeurs. Toute évaluation est affirmation d'un préférable.
- Si l'évaluation appelle la distance avec l'évalué, elle exige aussi, pour répondre à sa destination, le dialogue avec lui. La question est donc de trouver ce qu'on pourrait appeler la bonne distance, celle qui marie extériorité et possibilité de compréhension réciproque et dialogue. L'évaluation, une fois encore, n'est pas un constat, mais d'abord et surtout un message à l'autre et un appel à partager une valeur.
- Pour que l'évalué puisse prendre au sérieux la valeur à laquelle l'évaluateur l'invite, celui-ci doit émettre un jugement qui soit indépendant du pouvoir dont il dispose.
- L'évaluation ne vaut pas tellement par l'écart qu'elle constate entre les faits et la norme que par la direction qu'elle propose aux actes humains. En formation, elle vaut comme appel à l'individu afin qu'il dépasse ce qu'il est actuellement et pose des valeurs.

Au total, l'évaluation est un acte humainement lourd. Mais il ne l'est pas tant parce que des humains jugent les actes d'autres humains, ni même parce qu'ils le font selon des valeurs incertaines. Le risque éthique qu'il implique tient plutôt à l'occultation de ces jugements et de ces valeurs.

Chapitre 5

Conceptualiser les outils d'évaluation des apprentissages

Linda Allal

Ce chapitre a pour but de proposer un modèle de conceptualisation des outils d'évaluation des apprentissages en situation scolaire et en formation.¹ Après avoir évoqué le caractère polysémique du concept d'outil, nous exposerons la structure d'un cadre conceptuel – DISA – qui permet de situer les outils dans leur contexte et d'analyser leur traduction en activités évaluatives. Ce modèle sera illustré par deux exemples d'évaluations interactives basées sur des portfolios. En conclusion, nous évoquerons quelques apports possibles du modèle DISA pour la formation et la recherche.

1. Ce chapitre reprend et développe une présentation faite au colloque de l'ADMEE-Europe à Reims en 2005. Je remercie Lucie Mottier Lopez pour ses remarques au sujet d'une version antérieure du texte. Certaines de ses remarques ont soulevé des questions théoriques intéressantes qui ne peuvent pas être résolues dans le cadre de ce chapitre, mais le chantier de réflexion reste ouvert.

1 OUTILS, UN CONCEPT POLYSEMIOUE

Il n'est pas simple de cerner la signification du mot «outil» dans le champ de l'éducation et dans celui de la psychologie des apprentissages. Les termes «outil» et «instrument» sont souvent employés de façon interchangeable, même si certaines nuances restent attachées à leurs usages historiques. Ainsi, l'instrument de mesure du chercheur est souvent vu comme plus complexe, plus perfectionné (comme l'instrument de musique, de chirurgie) que l'outil d'évaluation fabriqué par l'enseignant (assimilable à l'outil de l'ouvrier, de l'artisan).

Les concepts d'instrument et d'outil occupent une place centrale dans l'œuvre de Vygotsky. Dans son essai sur la méthode instrumentale en psychologie (Vygotsky, 1930/1985, p. 39), celui-ci cite une liste très vaste et assez hétéroclite d'«instruments psychologiques» : le langage, les diverses formes de comptage et de calcul, les moyens mnémotechniques, les symboles algébriques, les œuvres d'art, l'écriture, les schémas, les diagrammes, les cartes, les plans, tous les signes possibles, etc.

En analysant la conception des outils dans la théorie de Vygotsky, Moro et Schneuwly (1997) ont montré que la dimension instrumentale, dominante dans les premiers écrits, est toujours associée à une dimension sémiotique, qui ressort davantage dans les textes ultérieurs. Une conception des outils réunissant ces deux dimensions semble recouvrir la totalité de la production culturelle humaine. Faut-il considérer dès lors que, dans une situation d'enseignement, *tout* est outil ?

Vygotsky (1930/1985) a défini plusieurs caractéristiques clés des outils ou instruments intégrés dans la conduite d'activités humaines : (1) ils sont des «élaborations artificielles» créées par l'homme ; (2) ils sont «sociaux par nature» plutôt que des productions individuelles ; (3) ils sont destinés au «contrôle des processus du comportement» humain et sont aptes à modifier «le déroulement et la structure des fonctions psychiques» (*ibid.*). En plus de la fonction de contrôle ou de régulation exercée par les outils, nous pouvons ajouter leur capacité à amplifier, à étendre ou à optimiser l'action de l'utilisateur, c'est-à-dire à lui permettre des actions plus complexes que celles qu'il entreprendrait sans un outil. Ainsi, un outil ou instrument d'évaluation, telle qu'une grille de critères d'appréciation, peut contribuer à la fois à contrôler l'action de l'évaluateur (focaliser son attention sur les aspects pertinents de l'objet à évaluer) et à amplifier sa capacité de jugement (suggérer des attributs à prendre en compte auxquels il n'aurait pas pensé).

Dans le cadre de la psychologie cognitive contemporaine, Rabardel (1997) a précisé qu'un instrument intervenant dans le fonctionnement psychologique humain est formé de deux composantes :

- d'une part, un artefact, matériel ou symbolique, produit par le sujet ou par d'autres ;

- d'autre part, un ou des schèmes d'utilisation associées résultant d'une construction propre au sujet... ou d'une appropriation de schèmes sociaux d'utilisation... schèmes d'usage, schèmes d'activité instrumentée, schèmes d'activité collective instrumentée (Rabardel, 1997, p. 35)

Cette définition, à la différence de celle de Vygotsky, considère qu'un instrument n'est pas toujours une construction sociale mais peut être parfois une construction propre au sujet. Dans le champ de l'enseignement, les recherches sur les pratiques des professionnels (Huberman, 1983) ont montré clairement l'inventivité des individus dans la fabrication d'outils adaptés aux conditions de leur classe. Tout en reconnaissant la place centrale de cette individualité dans l'enseignement, on peut souligner – en résonance avec Vygotsky – le caractère fondamentalement social des bases de la plupart des inventions individuelles. Par exemple, lorsqu'un enseignant crée sa « propre » grille d'évaluation des productions textuelles de ses élèves, sa conception de l'outil en question est le plus souvent inspirée d'autres grilles présentes dans le matériel didactique officiel, proposées pendant un séminaire de formation continue et/ou utilisées par des collègues. Ce recours aux outils disponibles dans son environnement social est un trait essentiel du travail professionnel, le génie propre à l'individu résidant surtout dans le choix, l'agencement, l'adaptation, l'exploitation des outils socialement constitués dans sa communauté de pratique.

Dans une analyse didactique des outils présents à l'école, Dolz, Moro et Polo (2000, p. 43) ont relevé le caractère « double face » des outils employés tantôt comme outil d'enseignement, tantôt comme outil d'apprentissage. Par exemple, une grille de contrôle qui figure dans le matériel didactique pour la production d'un certain genre textuel peut être utilisée par l'élève pour planifier et réguler sa démarche de composition, puis par l'enseignant pour évaluer les textes définitifs rendus par les élèves. Dans les termes de Rabardel (1997), on dirait qu'un même artefact (la grille de contrôle) est associé à des schèmes d'utilisation différents, chez l'enseignant et chez l'élève. Plane et Schnewwly (2000, p. 9) proposent de considérer la présence d'un artefact matériel stabilisé comme critère pour distinguer les outils d'enseignement « au sens strict » des outils didactiques « au sens large », qui seraient, eux, des ensembles de pratiques reproductibles mais toujours réinventées, telles que les démarches constitutives d'une séquence didactique. Nous avons un souci analogue dans ce chapitre dans le sens que nous voulons situer les outils d'évaluation (au sens strict) dans un cadre plus large. Mais au lieu d'étendre (sans limites?) le terme outil, nous avons choisi de définir un ensemble de *concepts différenciés* qui s'articulent dans le modèle DISA que nous présentons maintenant.

2 UN MODÈLE DE CONCEPTUALISATION DES OUTILS D'ÉVALUATION

Le cadre conceptuel que nous proposons est basé sur le croisement de deux dimensions binaires, tel que représenté dans le tableau 5.1. La première dimension (verticale) a trait à la distinction entre le *contexte* de l'évaluation et l'*outillage* qui est présent dans ce contexte. Nous utilisons le terme «outillage» car il s'agit en général de plusieurs outils ou instruments conçus pour fonctionner de façon articulée. La seconde dimension (horizontale) correspond à une distinction entre ce qui est *instauré par l'institution* (le système scolaire, l'établissement de formation, le responsable d'une unité de formation) et la manière dont ce cadre est *investi par les acteurs* (apprenants, enseignants, formateurs, etc.). Examinons plus en détail chaque cellule du schéma.

Sur le plan du contexte de l'évaluation, l'institution instaure un *dispositif d'évaluation* régi par une spécification de finalités (ex., objectifs à évaluer, fonctions de l'évaluation) et de procédures (règles, repères temps/espace) qui existent indépendamment des personnes. Lorsqu'un dispositif d'évaluation est investi par des acteurs (élèves, enseignants), il se transforme en *situation d'évaluation*. Chaque dispositif mis en place par l'institution peut se traduire dans une multitude de situations d'évaluation qui varient selon la dynamique des rapports entre les acteurs concernés et selon leurs rapports aux objets à évaluer. Les acteurs interprètent les règles du dispositif ; à travers leurs transactions et interactions, ils co-élaborent des significations qui reflètent des valeurs partagées et/ou des valeurs personnelles en tension. La distinction entre dispositif et situation est proche de la différenciation faite par Lave (1988), dans le cadre de ses travaux sur la cognition située, entre deux aspects du contexte : *arena* et *setting*. L'*arena* est constitué des normes et pratiques codifiées d'une institution ou d'une communauté, tandis que le *setting* est constamment recréé par les activités des individus qui fréquentent une institution, par les interactions entre les membres d'une communauté.

Prenons un exemple provenant d'une recherche sur le jugement professionnel des enseignants en matière d'évaluation (Allal et Hohl, sous presse). Notre étude a montré qu'en appliquant le dispositif de notation défini par le système scolaire, les enseignants créent différentes situations de communication qui contribuent au processus d'attribution des notes et/ou qui accompagnent les notes transmises : échanges avec les élèves au sujet de leur travail scolaire, discussions entre collègues de situations d'élèves en difficulté, entretiens avec les parents d'élèves, commentaires dans le livret visant à nuancer ou à renforcer le message de la note. Ainsi, deux élèves avec une même note dans le livret scolaire vivent en fait des situations d'évaluation différentes ; selon la nature des interactions au sujet de la note, celle-ci peut être comprise comme un signe d'encouragement ou comme un signe d'avertissement.

Tableau 5.1. Le modèle DISA : un cadre conceptuel pour les outils d'évaluation des apprentissages

	Instauré par l'institution	Investi par les acteurs
Contexte d'évaluation	Dispositif d'évaluation : finalités, procédures, règles, repères temps/espace	Situation d'évaluation : dynamiques interactives, significations co-élaborées
Outillage d'évaluation	Instrument d'évaluation : artefact(s) et schème(s) d'utilisation associé(s)	Activité évaluative instrumentée : médiation de l'évaluation par l'instrument

Examinons maintenant la seconde ligne du schéma : l'outillage de l'évaluation. Le terme *instrument d'évaluation* est utilisé pour l'outillage qui est instauré par l'institution. Il comprend, au sens de Rabardel, un ou plusieurs artefacts et les schèmes d'utilisation y associés. Les schèmes d'utilisation sont parfois définis par des consignes, parfois instituées, en tant que démarches « attendues », par la formation offerte aux enseignants ou par la socialisation professionnelle. Lorsque l'utilisation de l'instrument est investie par des acteurs, dans une situation d'évaluation donnée, nous parlons d'*activité évaluative instrumentée*. Il s'agit de l'activité d'évaluation menée par les acteurs à l'aide de la médiation offerte par l'instrument. Cette activité est forcément variable selon les acteurs et les instruments concernés. Par exemple, dans une discussion entre l'enseignant et un élève du passage de ce dernier à l'école secondaire, leur activité évaluative (c'est-à-dire leurs appréciations respectives des compétences acquises par l'élève) va non seulement varier mais être profondément marquée par le type d'instrumentation invoqué : les notes attribuées par l'enseignant pour les travaux effectués en classe, ou les résultats de l'élève aux épreuves cantonales communes à toutes les classes du système. En outre, l'activité évaluative peut comprendre des transformations des schèmes d'utilisation associées initialement aux artefacts évaluatifs. Par exemple, l'enseignant peut interpréter les résultats aux épreuves cantonales d'une autre manière que celle prévue par les concepteurs des épreuves.

Le cadre proposé dans le tableau 5.1 peut se résumer par une abréviation – DISA – qui aide à rappeler les quatre composantes du modèle : Dispositif-Instrument-Situation-Activité instrumentée.

3 EXEMPLES ILLUSTRANT LE MODÈLE DISA

Afin d'illustrer le modèle DISA, nous présenterons deux exemples d'évaluations interactives ayant pour support un portfolio. Le premier exemple se situe dans le cadre de l'école primaire, le second dans la formation initiale des enseignants primaires du canton de Genève. Le terme portfolio correspond,

dans les deux cas, à une collection finalisée de documents portant sur les apprentissages réalisés pendant une période donnée ; l'apprenant est impliqué dans le choix des documents et dans l'élaboration d'éléments de réflexion personnelle sur ses apprentissages. Dans les deux cas, le portfolio sert de support instrumental à la conduite d'entretiens d'évaluation ayant à la fois une fonction formative (favoriser la réflexion métacognitive de l'apprenant sur la progression de ses apprentissages) et une fonction sommative (établissement d'un bilan des progrès accomplis au terme d'une période de formation). Une discussion plus complète des diverses conceptions, fonctions et utilisations du portfolio est présentée, pour le contexte scolaire, dans Allal, Wegmuller, Bonaiti-Dugerdil et Cochet Kaeser (1998) et, pour le contexte de la formation professionnelle des enseignants, dans le chapitre 9 de cet ouvrage (Mottier Lopez et Vanhulle).

3.1 Exemple d'une évaluation par portfolio à l'école primaire

Une analyse succincte du premier exemple selon les catégories du modèle DISA est présentée dans le tableau 5.2. L'exemple provient d'une recherche effectuée en collaboration avec deux enseignantes titulaires de classes dans une même école primaire (Allal *et al.*, 1998). Sur le plan du contexte de l'évaluation, un même *dispositif* a été instauré dans les deux classes, dont l'une de 4^e année (élèves de 9-10 ans) et l'autre de 6^e primaire (élèves de 11-12 ans). Le système scolaire a fourni des instructions et des formations encourageant le corps enseignant à constituer, en interaction avec leurs élèves, des portfolios qui pouvaient accompagner et illustrer les bilans trimestriels transmis aux parents des élèves. Par ailleurs, les directives officielles rendaient obligatoire au moins un entretien par an avec les parents de chaque élève. Ces directives n'imposaient toutefois ni la constitution de portfolios, ni leur usage lors des entretiens avec les parents. Il s'agit donc d'un cas où le dispositif d'évaluation – entretiens basés sur un portfolio – a été choisi et instauré par les enseignantes elles-mêmes.

L'outillage de ce dispositif était composé de deux sortes d'*instruments d'évaluation*. Pour la constitution du portfolio, les enseignantes ont formulé des règles concernant les types de documents que l'élève devait réunir : un formulaire de présentation générale à insérer au début du portfolio ; huit activités à choisir dans différentes branches scolaires (ex., une activité de «français communication», une activité en «environnement» (géographie, histoire, sciences), etc.) ; une brève description de deux «projets» que l'élève souhaiterait réaliser au cours du trimestre suivant. Pour la conduite des entretiens, les enseignants ont opté pour un scénario tripartite réunissant l'enseignante, l'élève et ses parents². Des consignes ont été formulées à

2. Les deux parents ont été invités aux entretiens et, dans 15 cas sur 20, sont venus ensemble, y compris quelques cas de parents divorcés. Dans 4 cas, un seul parent est venu et dans un cas, un parent était accompagné de la grand-mère de l'enfant.

l'intention des élèves pour la préparation et la réalisation des entretiens ; il était prévu notamment qu'à la suite d'une courte introduction par l'enseignante précisant les buts de l'entretien, l'élève ferait la présentation de son portfolio à ses parents. Les règles et les consignes établies par les enseignantes constituent les artefacts de l'instrumentation de l'évaluation, tandis que les schèmes d'utilisation ont émergé de manière un peu différente dans chaque classe à travers les interactions de l'enseignante avec sa classe pendant la mise en place de la démarche. Les portfolios effectivement produits sont des instruments comportant une certaine standardisation de structure mais une grande variation de contenu.

Tableau 5.2. Exemple d'évaluation par portfolio dans le cadre d'entretiens tripartites à l'école primaire (basé sur Allal et al., 1998)

	Instauré par l'institution	Investi par les acteurs
Contexte d'évaluation	Dispositif d'évaluation École primaire (classes de 4 ^e , 6 ^e années). Instructions et formations relatives au portfolio. Obligation d'un entretien par an avec les parents des élèves. Choix des enseignantes d'une procédure d'entretiens sur la base d'un portfolio.	Situation d'évaluation Entretiens tripartites (enseignante-élève-parents) sur la base du portfolio de l'élève. Dynamiques interactives de qualité variable. Problèmes posés par le non partage de la signification de la situation d'entretien.
Outillage d'évaluation	Instruments d'évaluation Portfolio : règles de constitution (formulaire de présentation + 8 documents + 2 projets). Entretiens tripartites : scénario prévu, consignes.	Activité évaluative instrumentée Exemplification permettant une compréhension approfondie. Présentation de difficultés pouvant poser problème, selon l'origine des familles.

Examinons maintenant comment le dispositif d'évaluation et son instrumentation ont été investis par les acteurs concernés. Nos remarques s'appuient sur les enregistrements effectués de 20 entretiens (10 par classe). Les données ont montré que le dispositif de base (entretien tripartite basé sur un portfolio) s'est traduit dans une diversité de *situations d'évaluation*. La majorité des entretiens se sont déroulés de manière constructive, avec des variations liées à la qualité de la dynamique interactive entre les partenaires. Plusieurs facteurs pouvaient être source de difficultés lors des entretiens. Il n'a pas toujours été possible de construire une compréhension partagée de la finalité et de la signification de l'entretien : certains parents sont restés sur leur réserve avec une attitude défensive comme s'ils se sentaient eux-mêmes objets de l'évaluation ; quelques parents ont accaparé la parole, parlé d'eux-

mêmes, de leur propre scolarité, en ignorant la tentative de leur enfant de présenter son portfolio et de parler de ses apprentissages.

En examinant plus particulièrement le rôle du portfolio dans la médiation de *l'activité évaluative instrumentée*, plusieurs bénéfices ont été constatés. Le fait que l'élève présente et exemplifie ses apprentissages scolaires à travers le portfolio a amené plusieurs parents à dire qu'ils avaient découvert des compétences qu'ils ignoraient chez leur enfant, qui ne transparaissent pas dans les bilans écrits habituels, ni dans les discussions informelles à la maison. Un aspect de l'instrumentation s'est révélé toutefois problématique. Pour les domaines du «français-technique» (conjugaison, grammaire, orthographe, vocabulaire) et des «mathématiques», l'élève devait choisir pour son portfolio une activité qui a été bien réussie et une activité où des difficultés ont été rencontrées. La présentation des activités qui ont posé des difficultés a mis plusieurs élèves et leurs parents mal à l'aise. Un père, de niveau socioprofessionnel très modeste, s'est focalisé sur les difficultés présentées par son fils pour formuler des remarques très négatives à l'égard de l'avenir scolaire de son enfant. Une élève, dont la culture familiale était assez éloignée de celle de l'école genevoise, est tombée en larmes, et sa mère l'a suivie. Ces incidents ont montré qu'il conviendrait d'envisager des adaptations de l'instrument portfolio selon la connaissance de l'enseignante de la situation familiale de l'élève.

3.2 Exemple d'une évaluation par portfolio dans la formation des enseignants

Ce deuxième exemple provient d'une étude réalisée par Mottier Lopez (Mottier Lopez et Wegmuller, 2005; Mottier Lopez et Tessaro, 2007; Mottier Lopez et Vanhulle dans ce livre). Notre analyse de l'exemple en termes des composantes du modèle DISA est présentée dans le tableau 5.3. Le *dispositif d'évaluation* a été instauré par le programme et le règlement de la formation initiale des enseignants primaires à la Faculté de psychologie et des sciences de l'éducation de l'université de Genève. Dans la dernière année de formation, tous les étudiants suivent un séminaire d'intégration ayant pour but d'assurer des articulations entre leurs expériences dans les stages en responsabilité et les apports conceptuels des enseignements suivis à l'université. Le travail de l'étudiant pour le séminaire d'intégration donne lieu à une évaluation sommative en fin d'année; une note satisfaisante est nécessaire pour l'obtention du diplôme certifiant le cursus d'études. Le formateur dispose toutefois d'une grande liberté pour fixer les modalités de l'évaluation.

Dans le cas du séminaire d'intégration étudié par Mottier Lopez et ses collègues, le formateur a demandé aux étudiants de constituer un portfolio et a organisé un entretien de co-évaluation avec chaque étudiant en fin d'année.

Ces choix se sont concrétisés dans les *instruments d'évaluation* élaborés par le formateur. Pour le portfolio, les règles de constitution fournis aux étudiants prévoyaient 14 documents dont :

- les formulaires d'évaluation que l'étudiant a reçus lors de ses stages ;
- des documents élaborés par l'étudiant ou récoltés (ex., travail d'un élève) pendant ses stages ;
- des textes d'auto-évaluation et de réflexion métacognitive élaborés pendant le séminaire ou à d'autres moments de la formation ;
- un texte final sur le thème «Ma progression en prise de responsabilité dans des situations d'enseignement».

Pour l'entretien de co-évaluation, la consigne donnée aux étudiants était formulée comme suit : «Relevez ce que vous pensez avoir construit concernant vos compétences professionnelles au terme de votre formation initiale et ce qu'il est souhaitable d'anticiper dans votre future formation continue». Selon les règles du jeu fixées, l'étudiant devait débiter l'entretien avec le choix d'un thème et en exprimant des éléments d'auto-évaluation de sa progression ; le formateur intervenait ensuite, posait des questions, exprimait ses appréciations ; un dialogue s'engageait alors et devait permettre à l'étudiant de relier ses expériences de formation à ses anticipations de sa vie professionnelle future. La note de certification pour le séminaire d'intégration était attribuée au portfolio sur la base de critères fournis aux étudiants ; il a été annoncé que la qualité des interventions pendant l'entretien pouvait relever la note mais ne pouvait pas la baisser.

Tableau 5.3. Exemple d'évaluation par portfolio dans le cadre d'entretiens de coévaluation en formation des enseignants (basé sur Mottier Lopez et Tessaro, 2007)

	Instauré par l'institution	Investi par les acteurs
Contexte d'évaluation	Dispositif d'évaluation Formation initiale, enseignants primaires. Séminaire d'intégration reliant stages et apports académiques. Note de certification en fin d'année. Choix du formateur de conduire des entretiens sur la base de portfolio.	Situation d'évaluation Entretiens de co-évaluation entre le formateur et l'étudiante sur la base du portfolio. Dynamiques interactives de nature variable : choix des thèmes à aborder. De nouveaux questionnements émergent.
Ouillage d'évaluation	Instruments d'évaluation Portfolio : règles de constitution (14 documents). Entretien de co-évaluation : consigne, règles du jeu annoncées.	Activité évaluative instrumentée Médiation instrumentale par le portfolio : variable mais plus fréquente de la part du formateur que des étudiantes.

Une analyse du fonctionnement des entretiens de co-évaluation a été réalisée par Mottier Lopez et ses collègues sur la base d'enregistrements des entretiens menés avec sept étudiantes. Les données montrent qu'un même dispositif a donné lieu à des *situations d'évaluation* assez différentes sur le plan du choix des thèmes abordés et sur celui de la dynamique des échanges entre le formateur et l'étudiante. Les analyses et interprétations ont montré aussi des points communs aux sept entretiens ; il a été relevé en particulier que les entretiens ont conduit à « de nouveaux questionnements et à l'expression de nouvelles évaluations et interprétations de la progression de l'étudiante dans les différents stages, avec l'émergence d'aspects non directement mis en évidence dans le portfolio » (Mottier Lopez et Tessaro, 2007, p. 13).

L'analyse effectuée par Mottier Lopez et Tessaro (2007) a porté aussi sur la nature de *l'activité évaluative instrumentée* par le portfolio pendant l'entretien. Leur analyse a examiné notamment le nombre de références au portfolio et son usage dans les interventions de l'étudiante et dans celles du formateur. Les résultats ont mis en évidence deux caractéristiques de la « médiation instrumentale » offerte par le portfolio : (1) l'intensité de cette médiation est assez variable (le nombre de références au portfolio varie de 12 à 34 selon l'entretien) ; (2) elle intervient plus fréquemment dans les prises de parole du formateur que dans celles des étudiantes. Plusieurs pistes d'interprétation sont proposées pour comprendre les significations données au portfolio par les différents acteurs.

4 ÉLÉMENTS DE CONCLUSION

Les deux exemples présentés succinctement ici avaient pour but d'illustrer les quatre composantes – Dispositif, Instrument, Situation, Activité instrumentée – du modèle DISA. L'utilité potentielle de ce modèle est de nous aider à mieux différencier des entités qui sont souvent emmêlés de manière confuse sous la rubrique « outils d'évaluation ». Le cadre conceptuel que nous proposons peut être utile, à notre avis, dans plusieurs démarches de formation et de recherche. Il peut faciliter et éclairer l'analyse des pratiques d'évaluation par la distinction opérée entre les composantes d'ordre différent. Il peut offrir un cadre pour la comparaison des pratiques d'évaluation : en quoi sont-elles différentes ? quelles composantes sont différentes ? lesquelles sont semblables ? Il peut servir de référence pour la planification, la construction et la mise en œuvre de nouvelles pratiques d'évaluation. On pourrait également envisager son extension plus générale à l'analyse des pratiques d'enseignement, car dans chaque enseignement on trouve un dispositif et son instrumentation, ainsi que leur transposition en situations et en activités d'enseignement/apprentissage instrumentées.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Allal, L. et Hohl, J.-M. (sous presse). Le rôle du jugement professionnel dans les adaptations et les modalités de communication de l'évaluation. In Lafortune L. et Allal L. (éds), *Jugement professionnel en évaluation : pratiques enseignantes au Québec et à Genève*. Montréal, Presses de l'Université du Québec.
- Allal, L., Wegmuller, E., Bonaiti-Dugerdil, S. et Cochet Kaeser, F., (1998). Le portfolio dans la dynamique de l'entretien tripartite, *Mesure et évaluation en éducation*, 20, 5-32.
- Dolz, J., Moro, C. et Pollo, A. (2000). Le débat régulé : de quelques outils et de leurs usages dans l'apprentissage, *Repères*, 2, 39-60.
- Huberman, M. (1983). Répertoires, recettes et vie de classe : comment les enseignants utilisent l'information, *Éducation et recherche*, 5(2), 157-177.
- Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge, Cambridge University Press.
- Moro, C. et Schneuwly, B. (1997). Introduction. In Moro, C., Schneuwly B. et Brossard M. (éds), *Outils et signes, perspectives actuelles de la théorie de Vygotski*. Paris, Peter Lang, pp. 2-7.
- Mottier Lopez, L. et Tessaro, W. (2007). Entretiens de co-évaluation instrumentés par un portfolio : pour quelle réinterprétation de la progression des compétences professionnelles ? Publication des actes du 19^e colloque international de l'ADMEE-Europe, Luxembourg. <http://jemacs.uni.lu/index.php/JEMACS/article/view/19/16>.
- Mottier Lopez, L. et Vanhulle, S. (2008). Portfolios et entretiens de co-évaluation : des leviers de la professionnalisation des jeunes enseignants. In (ajouter les indications concernant ce livre).
- Mottier Lopez, L. et Wegmuller, E. (2005). «Le portfolio en formation initiale d'enseignants : étude d'entretiens de co-évaluation et des médiations instrumentales de l'outil». Communication au 18^e colloque international de l'ADMEE-Europe, Reims, France.
- Plane, S. et Schneuwly, B. (2000). Regards sur les outils de l'enseignement du français : un premier repérage, *Repères*, 22, 3-18.
- Rabardel, P. (1997). Activités avec instruments et dynamique cognitive du sujet. In Moro, C., Schneuwly B. et Brossard M. (éds), *Outils et signes, perspectives actuelles de la théorie de Vygotski*. Paris, Peter Lang, pp. 35-49.
- Vygotsky, L.S. (1930/1985). La méthode instrumentale en psychologie. In Schneuwly B. et Bronckart J.-P. (éds), *Vygotsky aujourd'hui*. Neuchâtel, Delachaux et Niestlé, pp. 39-49.

Chapitre 6

Les évaluations diagnostiques en France : le diagnostic, une notion et une exploitation en classe qui évoluent depuis 1989

Gérard Brézillon, Françoise Champault

On se propose dans cet article de présenter la conception et les usages des évaluations diagnostiques, particulièrement développées en France.

Les finalités de ces évaluations sont distinctes de celles des évaluations-bilans, telles les évaluations comparatives internationales. *«Les évaluations diagnostiques ont pour objectif de connaître les acquis d'un élève ou d'une classe, à différents moments, afin de mettre en place des mesures pédagogiques adaptées, alors que les évaluations-bilans visent à recueillir des informations d'ensemble sur les acquis des élèves à partir d'un échantillon représentatif, en vue d'établir un bilan, d'étudier une évolution, etc.»* (avis n° 2 du Haut Conseil de l'évaluation de l'école, «Apprécier et certifier les acquis des élèves en fin de collège : diplôme et évaluations-bilans», juin 2001).

Les premières évaluations diagnostiques ont été mises en place à la suite de la promulgation de la Loi d'orientation sur l'éducation de 1989. Celle-ci demandant que soit bien prise en compte l'hétérogénéité des classes

et des rythmes d'apprentissage des élèves, il est apparu nécessaire de créer des évaluations permettant aux enseignants d'apprécier les compétences de leurs élèves pour leur proposer des progressions adaptées. C'est ainsi qu'ont été mises en place dès la rentrée scolaire de 1989, les évaluations diagnostiques de rentrée en CE2 et en 6^e. Puis a été progressivement élaborée une banque d'outils offerte en libre service aux enseignants et destinée à les aider à formuler un diagnostic sur leurs élèves. Cet article présente successivement ces deux «évaluations diagnostiques», d'abord les évaluations de masse de rentrée (1), puis la banque (2).

1 LES ÉVALUATIONS DE MASSE DE RENTRÉE

1.1 Les évaluations diagnostiques de rentrée par protocole

Outils professionnels pour les enseignants et les équipes éducatives, ces évaluations sont des aides à la décision pédagogique. Les exercices des protocoles permettent de révéler – individuellement – les points faibles qui sont des freins, voire des obstacles, aux apprentissages présents et futurs de chaque élève et les points forts sur lesquels l'enseignant peut s'appuyer pour assurer la suite de la progression de l'élève.

Les évaluations de rentrée sont constituées de protocoles standardisés qui se présentent sous forme de cahiers: cahier de passation avec des commentaires pour les enseignants et cahier de l'élève. Les consignes précises données aux enseignants permettent d'homogénéiser les conditions de passation ainsi que les corrections.

1.2 Les tableaux de compétence

Pour chaque discipline et chaque niveau, des tableaux de compétences structurent les protocoles (cf. tableau 6.1). Chaque item est lié à un triplet: «champ», «compétence», «composante». Le champ est un regroupement d'items qui correspondent à un ensemble d'activités particulières proposées dans les exercices. La (les) compétence(s) est (sont) définie(s) par rapport aux programmes. Elle(s) constitue(nt) l'échelon intermédiaire entre le champ et la composante. Elle(s) peut (peuvent) être évaluée(s) à travers plusieurs composantes qui renvoient à la description de la tâche à effectuer par l'élève.

Tableau 6.1. Exemple de tableau de compétence : Français CE2 (extrait)

FRANÇAIS CE2 extrait									
CHAMPS	COMPETENCES	COMPOSANTES	Le hérisson Items 1 à 37		Anatole change d'école Items 38 à 62		Benjamin et son papa géant Items 63 à 93		
			Exercices	Items	Exercices	Items	Exercices	Items	
	Comprendre les informations explicites d'un texte littéraire ou d'un texte documentaire approprié à l'âge et à la culture des élèves.	- comprendre les informations d'un texte et répondre à des questions « vrai - faux » - comprendre les informations d'un texte et répondre à une question ouverte	1	1, 2, 3		39, 41, 42, 43, 44	29	65, 66, 67, 68	
	Trouver dans un texte documentaire imprimé ou sur un site de la toile les réponses à des questions simples.	- trouver des informations dans un texte et répondre à des questions à choix multiples - trouver des informations dans un texte et répondre à des questions ouvertes	2 et 3	4, 5			27 et 28	63, 64	
	Dégager le thème d'un texte littéraire (de qui ou de quoi parle -t-il?)	- donner un titre en choisissant parmi plusieurs propositions - identifier les personnages d'un récit - identifier le lieu d'un récit	4	6					
	Lire et comprendre seul les consignes ordinaires de l'activité scolaire.	- appliquer une consigne - trouver une consigne			12 14	38 40	35 36	80 81	
	Comprendre en le lisant silencieusement un texte court (petite nouvelle, extrait...) de complexité adaptée à l'âge et à la culture des élèves en s'appuyant sur un traitement correct des substituts des noms, des connecteurs, des formes verbales, de la ponctuation... et en faisant les inférences nécessaires.	- donner son avis - justifier sa réponse en s'appuyant sur le texte - procéder à un traitement simple de l'information - retrouver le référent d'un substitut	6 6	8 9		45, 46, 47			
	Déchiffrer un mot que l'on ne connaît pas.	- reconnaître dans une liste de mots proches visuellement, un mot rare lu par l'enseignant	9	17, 18, 19, 20		49, 50			
	Identifier instantanément la plupart des mots courts (jusqu'à 4 ou 5 lettres) et les mots longs les plus fréquents.	- sélectionner dans une liste de mots fréquents, celui qui ne contient pas le même son	11	33, 34, 35, 36, 37					
	Écrire de manière autonome un texte d'au moins cinq lignes (narratif ou explicatif) répondant à des consignes claires, en gérant correctement les problèmes de syntaxe et de lexique.	- écrire un texte d'au moins cinq lignes - produire au moins deux phrases sémantiquement correctes - tenir compte du sujet imposé dans la consigne - utiliser de manière cohérente les temps du récit			26	57 58 59, 60 62			

1.3 La constitution de groupes de besoin indépendamment des scores moyens globaux

L'analyse des résultats aux évaluations permet de constituer des groupes de besoin à partir de critères objectifs, quel que soit le niveau de l'élève concerné. La formation de chaque groupe de besoin est faite en fonction de la typologie des erreurs recensées dans un champ et non au regard des scores globaux obtenus. C'est ainsi que l'on pourra trouver dans un même groupe de besoin des élèves ayant obtenu des scores globaux de réussite différents.

1.3.1 *Groupe de niveau / groupe de besoin*

Voici les résultats de six élèves d'une classe de CE2 à l'évaluation de français en CE2 :

Axel – Cédric	79 %
Louise – Charlotte	65 %
Arnaud – Nicolas	49 %

Si pour chacun de ces six élèves, on note les numéros des items auxquels ils ont échoué, on obtient (tableau 6.2) :

Tableau 6.2. Résultats de six élèves d'une classe de CE2 à l'évaluation de français en CE2 (extrait)

	Axel Cédric		Louise Charlotte		Arnaud Nicolas	
Groupe de niveau – Lecture verticale					14	14
				15	15	15
		16		16	16	16
			17	17	17	
		20	20		19	19
	21		21		20	
	24		23	23		
			24			
			25		25	25
			26	26	26	26
					27	27
					28	
	29		29		29	29
	31		30		30	30
	33				31	31
	34			33	32	32
		35		34	33	33
			35		34	34
				36		
				38		
	41		40	40		40
			41	41	41	41
			42	42	42	42
					43	
			45		45	
			46		46	
			47		47	47
			48		48	48
	50	50	49		49	49
	51		50	50	50	50
			51			51
	53	53	52		53	53
				53		54
		86				
	88				87	
			89		88	
						89

Groupe de **besoin** – Lecture horizontale

Note de lecture : « Lecture verticale », on prend en compte uniquement le score moyen global de chaque élève. – « Lecture horizontale », on s'intéresse aux scores obtenus aux exercices relevant de chaque notion.

On constate que : 1) les élèves, quel que soit leur score global, ont échoué à de nombreux items ; 2) une « lecture verticale » qui ne considère que ce score global, ne renseigne pas sur la nature des difficultés rencontrées par chacun des élèves.

L'intérêt des évaluations diagnostiques est donc de permettre à l'enseignant de déterminer la nature de la difficulté qui freine l'apprentissage. Comme il n'est pas possible, ni même souhaitable, de tenter de remédier à tous les problèmes, l'enseignant doit se fixer des priorités. Pour cela, il visera, d'abord, les compétences dont la maîtrise est attendue à la fin du cycle.

Si nous reprenons l'exemple ci-dessus, on constate que Louise, Arnaud et Nicolas échouent dans un grand nombre d'items relatifs à l'exercice concernant le passage du présent au futur, compétence dont la maîtrise est attendue à la fin du cycle 2. Il va donc falloir créer un groupe de besoin sur ce sujet pour ces trois élèves (tableau 6.3).

Tableau 6.3. Groupe de besoins pour trois élèves d'une classe de CE2 à l'évaluation de français en CE2 (extrait)

Axel	Cédric	Louise	Charlotte	Arnaud	Nicolas
		45		45	
		46		46	
		47		47	47
		48		48	48
		49		49	49
50	50	50	50	50	50
51		51			51

Groupe de besoin pour le futur pour Louise, Arnaud et Nicolas

1.3.2 Le diagnostic est affiné (rôle des codages intermédiaires)

Dans les évaluations diagnostiques, chaque exercice donne lieu à un ou plusieurs items, et la réponse des élèves est codée de façon fine pour, précisément, affiner le diagnostic. Voici un exemple de cette pratique (figure 6.1) :

*Lis le premier texte qui parle de la petite souris. Le deuxième texte raconte la même histoire à un autre moment. Tu dois changer les verbes soulignés. Attention au mot ajouté au début du deuxième texte. C'est : « **Demain** ».*

La souris va chez ses amies,
les autres souris. Ensemble, elles creusent un
nouveau tunnel. Elles montent,
glissent comme sur un toboggan.
La souris s'amuse tellement qu'elle
oublie de construire sa chambre.

Extrait de *Comment la souris reçoit une pierre sur la tête et découvre le monde*, Etienne Delessert, Folio benjamin

Demain, la souris _____ chez ses amies,
les autres souris. Ensemble, elles _____ un
nouveau tunnel. Elles _____,
_____ comme sur un toboggan.
La souris _____ tellement qu'elle
oubliera de construire sa chambre.

FIGURE 6.1. Extrait d'un exercice de l'évaluation en CE2

Le codage des réponses utilisé pour chaque verbe à transformer permet de distinguer les différents degrés de maîtrise du futur.

Réponse exacte : le verbe est mis au futur	code 1
Autre réponse exacte : le verbe est mis au futur proche	code 2
Réponse partiellement exacte : terminaison mal orthographiée	code 4
Réponse erronée : emploi d'un autre temps que le présent ou le futur	code 8
Autres réponses	code 9
Absence de réponse	code 0

Analyse

Les codes 1, 2 et 4 permettent de repérer les élèves qui savent mettre un verbe au futur et, parmi eux, ceux qui n'en maîtrisent pas les terminaisons (code 4). Pour ces derniers, on peut envisager un travail en orthographe. En revanche, les élèves qui utilisent par exemple un temps du passé (code 8) relèvent de situations pédagogiques très différentes. Pour eux, il est souhaitable de mettre en place un travail sur la chronologie, les déictiques... en français ou dans une autre discipline. Pour les élèves relevant des codes 9 et 0, il convient de faire une analyse spécifique de leurs erreurs (code 9) et de s'entretenir avec eux pour comprendre les raisons de leurs non-réponses (code 0) afin de mettre en place ensuite des situations pédagogiques adaptées.

1.4 Comment lier un résultat global et le profil d'un élève ?

Un bon score global peut cacher un déficit dans un champ particulier (compensé par un fort taux de réussite dans un autre champ). Ce déficit peut être révélateur de difficultés ponctuelles ou de problèmes latents, moins évidents à déceler, susceptibles de constituer un frein dans la suite du cursus. Seule l'analyse de la production individuelle permet de ne pas en rester au stade du constat. Voici l'exemple d'Amandine et de Morgane.

1.4.1 Morgane et Amandine (évaluation CE2 en mathématiques)

Morgane et Amandine obtiennent des scores globaux de réussite identiques avec des distributions de codes comparables. En revanche, les résultats par champ montrent des disparités (figure 6.2).

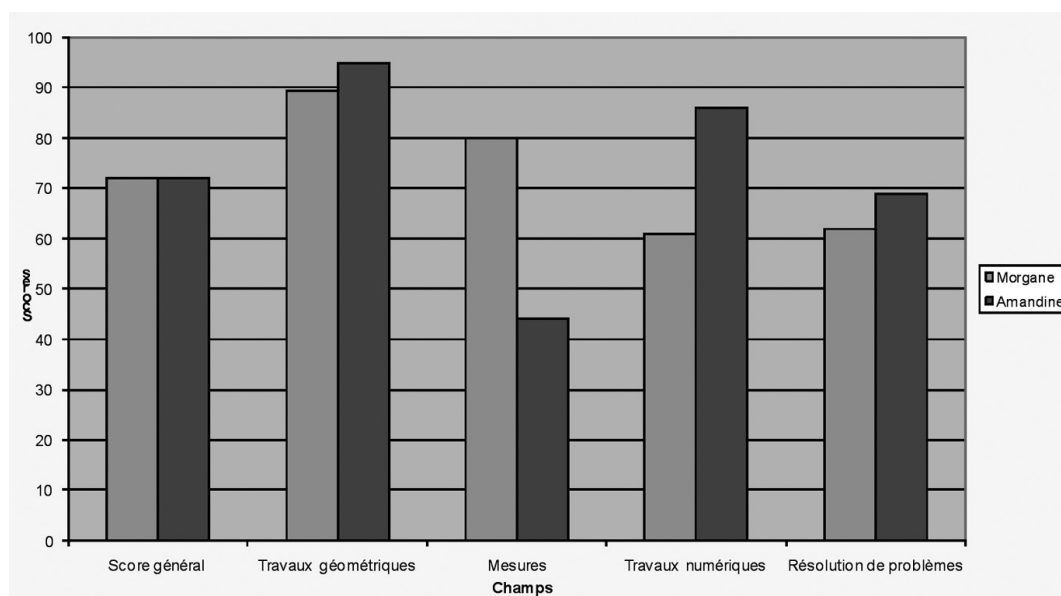


FIGURE 6.2. Les scores de Morgane et d'Amandine (CE2 en mathématiques)

Si on compare les résultats par champ, on constate qu'en «Travaux numériques» et en «Mesures» les scores de réussite présentent des différences significatives. Les scores par champ ne renseignent absolument pas sur les atouts et difficultés de ces deux élèves. Morgane a plus de codes 0 et moins de codes 9 qu'Amandine. Le nombre d'items auxquels elles ont échoué étant très faible, on ne peut guère en tirer de conclusion. Toutefois, si on regarde la répartition des différents codes, on constate que les codes 0 se trouvent concentrés en «mesures» pour Amandine, en «travaux numériques» et en «résolution de problèmes» pour Morgane, ce qui conforte l'impression dégagee par la comparaison des scores de réussite sur ces champs.

1.4.2 L'analyse des difficultés

L'examen des codes d'erreurs et le retour au cahier montrent que :

Morgane

- maîtrise la technique opératoire de l'addition à l'écrit ;
- éprouve des difficultés en calcul mental (erreur de plus ou moins 1, ou non réponse) ;
- traite les «situations multiplicatives» par une procédure d'additions itérées, stratégie qui peut se justifier dans certains cas, mais qui va se révéler coûteuse avec des nombres plus grands ;
- commet, en soustraction, des erreurs ou ne répond pas, que ce soit en technique opératoire ou en résolution de problème (l'opération est posée et non calculée).

Malgré un score général plutôt bon, la nature des erreurs de Morgane met donc en évidence une certaine fragilité pour laquelle il est nécessaire que l'enseignant fasse des investigations complémentaires pour éviter que ces points faibles ne s'aggravent et ne compromettent à plus ou moins long terme la suite du cursus.

Amandine

- n'a pas de connaissance des «familles» des unités de mesure. Elle ne sait pas associer les unités usuelles de masse, de longueur et de volume (ou capacité) à ce qu'elles représentent et choisir correctement entre différentes unités correspondant à une même grandeur. Par exemple, elle peut : «lancer une balle lestée de 200 m» ; «soulever une caisse qui pesait 8 cm» ;
- maîtrise bien le calcul mental ainsi que l'addition avec ou sans retenue. On peut repérer cette compétence dans différents contextes. Elle obtient de bons résultats en calcul mental, en technique opératoire et en résolution de problème ;
- n'éprouve pas de difficulté avec les autres opérations ;
- met en œuvre des stratégies personnelles efficaces de résolution de problème.

La nature des erreurs d'Amandine met en évidence des difficultés bien circonscrites au champ des mesures qui peuvent être traitées sans investigation complémentaire.

Morgane et Amandine ont des besoins très différents. Pour Amandine, ils sont précis, pour Morgane ils sont plus diffus. Les «lacunes» d'Amandine semblent (à ce stade de l'analyse et de la scolarité, il faut être prudent) moins potentiellement graves pour la suite de la scolarité que celles de Morgane.

2 LA BANQUE D'OUTILS D'ÉVALUATION DIAGNOSTIQUE (WWW.BANQOUTILS.EDUCATION.GOUV.FR)

2.1 Des protocoles de rentrée aux outils de la banque

Dès 1990, une *Cellule d'élaboration d'une banque d'outils d'évaluation* a été créée dans la Direction de l'évaluation et de la prospective pour «produire des exercices d'évaluation que les enseignants pourraient utiliser, selon les besoins, sans aucun caractère d'obligation, en les intégrant à leurs pratiques pédagogiques, aux moments de l'année qui leur paraîtraient les plus opportuns. À chacun de choisir, parmi les exercices proposés, ceux qui lui paraissent les plus adéquats en fonction des profils et des besoins des élèves de sa classe». Les exercices sont alors élaborés par des groupes nationaux, sur le principe d'un groupe pour une discipline à un niveau donné.

Entre 1992 et 1998, la *banque d'outils* a été publiée sous forme papier : elle couvrait le français et les mathématiques dans les trois cycles de l'école et pour le collège, les disciplines modulaires des classes de seconde de lycées. En 1995, des disquettes *banqitem* (mathématiques collège) ont initié la diffusion par fichiers informatiques.

Le site Internet *banqoutils* a été créé en 1997 avec la mise en ligne d'exercices de mathématiques du niveau collège. Il s'est enrichi avec, soit des créations, soit des adaptations d'exercices diffusés précédemment sous un autre format. Issus de projets différents (protocoles de rentrée CE2, 6^e et 2^{de} ; évaluation en grande section/cours préparatoire, sciences et technologie au collège), la présentation et l'esprit des outils ne sont pas homogènes : la *banque* est une juxtaposition de plusieurs bases d'exercices et non pas une seule base structurée et fonctionnelle.

En 2001, le ministère décide de développer la *banque d'outils* : mettre à disposition des outils dans toutes les disciplines et à tous les niveaux, créer a posteriori la cohérence de la banque, informer et former les professeurs, utilisateurs de ces outils. La production des outils est alors confiée à des équipes académiques.

2.2 Le diagnostic dans le contexte de la banque d'outils

2.2.1 Des différences entre évaluation par protocole et évaluation banque d'outils

La parenté entre protocole de rentrée et outils de la banque est certaine. Mais en quoi diffèrent ces évaluations ?

L'une est obligatoire, concerne seulement le français et les mathématiques pour les élèves de CE2 et 6^e et est conçue pour être passée en début d'année scolaire. L'autre, non obligatoire, intéresse potentiellement toutes les disciplines et toutes les classes et peut être mobilisée par les enseignants à tout moment de l'année. Pour la première, le professeur ne choisit ni les compétences, ni les champs disciplinaires qui sont évalués, un cahier lui est fourni. Pour la seconde, il sélectionne sur le site les outils correspondant aux compétences qu'il veut tester, au problème rencontré par ses élèves qu'il souhaite éclairer. Dans les deux cas (évaluation collective classe entière), l'enseignant pourra constituer des groupes de besoins. Si le professeur a procédé avec des outils *banque* à une évaluation individuelle ou à un petit groupe, il pourra proposer des activités personnalisées.

2.2.2 Des outils diagnostiques : principales caractéristiques

Les outils proposés, ciblés, permettent d'aborder de façon analytique les erreurs des élèves et les freins aux apprentissages afin d'approcher les causes des difficultés ou les contextes précis qui posent problème à certains. Ils sont diagnostiques parce qu'ils proposent, au-delà d'un simple constat, des *pourquoi* ou des *comment* aux difficultés exprimées par les élèves. L'outil est conçu à partir des hypothèses émises sur les causes probables de non-réussite et associées à un codage. Pour préciser ces éléments d'analyse, chaque outil comprend toujours une partie explicative pour le professeur indissociable de la partie destinée à la passation de l'élève.

La structure de l'outil, les questions ou tâches proposées aux élèves peuvent surprendre, car pour atteindre leur objectif, elles peuvent s'éloigner de la forme d'un exercice d'entraînement, d'application ou d'un contrôle de fin de séquence. Elles ne respectent pas systématiquement une unité de thème ou de contenu disciplinaire et une impression de morcellement peut être ressentie. Il s'agit bien de s'assurer :

- que les questions ou tâches ciblent bien les causes ou contextes du problème abordé,
- que les mêmes causes provoquent des résultats identiques,
- de ne faire varier qu'un seul élément entre les tâches ou questions que l'on rapproche,
- de l'indépendance des questions ou tâches successives proposées aux élèves.

Ces exigences de conception se doublent de contraintes d'expérimentation, ce qui explique que ces outils se positionnent à mi-chemin entre les évaluations pratiquées dans les classes et les banques d'items ou batteries de tests utilisées par les psychologues et psychométriciens : ils garantissent une évaluation valide et objective s'ils sont utilisés dans les conditions préconisées.

2.2.3 *Les compétences retenues pour la conception d'outils Banque*

Nous partons du principe que les compétences, sollicitées ou pas dans les enseignements, s'acquièrent et s'enrichissent tout au long de la scolarité et au-delà. Selon la variété et la fréquence des situations où elles auront été mobilisées, certaines seront plus développées que d'autres. Les activités scolaires demandent aux élèves de réaliser des tâches complexes qui mettent en inter-relation des savoirs, des compétences, des représentations. Évaluer isolément une compétence n'est donc pas facile car sa mise en œuvre dépend notamment :

- des savoirs à mobiliser (supposés connus) ;
- du contexte de mise en œuvre (les conditions dans lesquelles la compétence est évaluée ; il n'est pas neutre) ;
- de la compréhension par l'élève de la tâche à effectuer ;
- d'une mobilisation intellectuelle et/ou psychologique, variable presque toujours occultée.

C'est pourquoi nous recommandons aux concepteurs de commencer par analyser la tâche demandée et de questionner les élèves sur leurs procédures pour repérer ce qui peut être source de difficultés.

La lecture des programmes aide à comprendre comment les compétences sont présentées au cours de la scolarité. Selon le moment, une même compétence peut être une « méga-compétence » à acquérir ou une « mini-compétence » à utiliser (par exemple, multiplier). Croiser cette lecture entre plusieurs disciplines conduit à repérer des compétences communes (mais à des degrés de maîtrise différents au même âge), des compétences qui semblent identiques mais sont nommées différemment et des compétences spécifiques. Toutes sont potentiellement sources de difficultés pour les élèves. Cependant, n'ont été retenues pour la production d'outils que celles reconnues comme posant un problème pour la suite des apprentissages, plus précisément celles :

- pour lesquelles les difficultés sont repérables par des erreurs que les élèves commettent ;
- qui peuvent être des freins aux apprentissages ultérieurs parce que les élèves ont développé des stratégies de mise en œuvre ou s'appuient sur des représentations ou des procédures inadéquates. Cette attitude inadaptée ne devient apparente que quand la situation proposée la met en défaut.

2.2.4 *Tableau commun des compétences : un tableau d'indexation des outils*

L'entrée par les compétences, alliée à des travaux pluridisciplinaires associant tous les niveaux d'enseignement, a été choisie pour créer *a posteriori* la cohérence de la *Banque*. La lecture des travaux de recherche et des programmes scolaires montre la polysémie et la diversité d'approche des compétences. Nous avons privilégié un sens très large, la compétence permettant d'expliquer que l'action a été efficace : l'élève est compétent dans le contexte donné s'il effectue avec réussite ce qui est attendu de lui dans cette situation précise. La construction d'un cadre de travail partagé par toutes les disciplines et niveaux a donc été réalisée de façon pragmatique : général pour donner de la cohérence à la *banque d'outils*, mais ouvert pour accueillir les spécificités des niveaux et des disciplines. Le *tableau commun des compétences* qui a été défini n'est donc pas le résultat d'une élaboration conceptuelle mais bien une typologie d'indexation. C'est un document daté systématiquement car il évolue périodiquement (tableau 6.4).

2.3 Construction des outils

Pour favoriser l'ancrage de nouvelles pratiques professionnelles chez les enseignants en s'appuyant sur un réseau de proximité de personnes qui connaissent les outils et leur domaine d'application, la création des outils repose sur des enseignants et des inspecteurs issus du terrain réunis en équipes académiques de production «cycle 2» ou bien «cycle 3 – collège – lycées»¹. Les concepteurs, d'âges et de parcours professionnels diversifiés, témoignent d'ailleurs des inflexions de certaines de leurs pratiques de classe induites par leurs travaux de production d'outils.

Les équipes élaborent des outils dans un champ disciplinaire précis. Toutes les phases de production, expérimentation comprise, se déroulent localement. Chaque équipe reste libre de travailler sur les compétences qui l'intéressent ou les difficultés d'apprentissage qu'elle souhaite éclairer. Elle se réunit régulièrement pour construire des outils d'évaluation en respectant un cahier des charges. Ce dernier, rédigé au niveau national, est une pièce incontournable qui a été conçue comme un guide de référence. Des fiches décrivent la forme et le contenu des différentes rubriques de l'outil, d'autres proposent une méthode d'élaboration en huit phases. Les concepteurs peuvent s'adjoindre quelques enseignants volontaires pour les aider à mettre au point les outils et à en tester les versions successives.

1. Composition type d'une équipe de production académique :
 - «cycle 2» : 6 concepteurs, des testeurs et 1 inspecteur du premier degré
 - «cycle 3 – collège – LEGT – LP» : 6 concepteurs et des testeurs issus à parité d'école, de collège, de LEGT et de LEP (par exemple pour les concepteurs, 2 écoles, 2 collèges, 1 LEGT et 1 LP), 1 IEN 1^{er} degré, 1 IA-IPR et 1 IEN-ET/EG.

Tableau 6.4. Tableau commun des compétences

Banque d'outils d'aide à l'évaluation diagnostique des compétences des élèves

Cahier des charges - 14 octobre 2005

DOCUMENT DE TRAVAIL

ANNEXE 1 : INDEXATION DES OUTILS

Tableau commun des compétences⁶

	1- CAPACITÉ	2- COMPÉTENCES	3- COMPOSANTES
S'INFORMER	PRÉLEVER (l'information demandée, qu'elle soit explicite ou implicite) [IP]	- Repérer, identifier des informations orales - Repérer, identifier des informations écrites - Repérer, identifier des informations à partir du réel ou de représentations du réel (ni écrit, ni oral).	
	ORGANISER (l'information, à partir de critères qui lui sont donnés) [IG]	- Trier, classer - Hiérarchiser, ordonner - Mettre en relation (comparer, associer, ...)	
RÉALISER	CHOISIR (l'élève se donne les critères de sélection) [RS]	- Sélectionner, choisir, en autonomie, une ou des informations pertinentes - Sélectionner, choisir, dans une liste, une stratégie, une démarche	
	CONCEVOIR (processus intellectuel) [RV]	- Formuler une ou des questions - Proposer des solutions possibles - Proposer des stratégies, démarches, processus, - Proposer un cadre d'action, de résolution, - Construire un protocole	
	EXÉCUTER (la stratégie a été donnée par des consignes ou élaborée précédemment par l'élève) [RX]	- Respecter les contraintes, les codes spécifiques à la discipline - Accomplir des gestes techniques, produire (qualité d'exécution) - Appliquer une démarche, suivre un protocole - Reformuler un argument, une justification	
	APPRÉCIER (à tous moments de son travail) [RA]	- Faire des essais, tester, expérimenter - Contrôler, vérifier - Critiquer - Valider - Tirer des conclusions	

⁶ À ce niveau de l'arborescence, les premières colonnes sont communes à toutes les disciplines. Le terme de "compétence" est à comprendre au sens large de "famille de compétences". Les dénominations *capacité*, *compétence* et *composantes* ont été conservées car elles sont utilisées dans la base informatique du site *Banqoutils*. Les formulations portées dans ce document sont donc celles à inscrire dans le cartouche de chacun des outils.

À partir de la dernière colonne, grisée, le libellé des composantes peut varier d'une discipline à l'autre, d'un niveau de difficulté à l'autre (Se référer aux différents tableaux disciplinaires de compétences). Certaines d'entre elles peuvent même n'exister que dans une seule discipline.

Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche — Direction de l'évaluation et de la prospective - DEP C1

- 27 -

Les élèves et les enseignants qui expérimentent les outils participent aussi à la création des outils. Leur collaboration est capitale pour s'assurer de la qualité de la production car il s'agit d'une passation avec un public varié (élèves en nombre qui n'ont pas testé l'outil en cours de conception, des sites diversifiés en taille et public scolaire, des enseignants qui ne connaissent pas les outils). Le respect du protocole d'expérimentation (dont l'observation stricte des conditions et des consignes de passation, et des consignes de correction), permet de garantir la validité des résultats obtenus. Il n'est pas demandé de constituer un échantillon représentatif, ni d'effectuer une analyse statistique des prestations des élèves. Mais l'équipe académique doit s'assurer de la validité de l'expérimentation, tant par l'étude des prestations des élèves que dans les aspects matériels de son déroulement.

2.4 L'utilisation des outils de la banque dans les classes

Bien que la *banque d'outils* soit nationale et d'origine institutionnelle, les enseignants n'ont pas obligation de l'utiliser. L'espace de liberté pédagogique des professeurs est respecté mais ceci explique pourquoi la *banque d'outils diagnostiques* demeure encore inconnue par la plupart d'entre eux.

Un pic de connexions au site *Banqoutils* est constaté autour de la rentrée scolaire, la consultation est importante jusqu'en novembre, période propice à la mise en place des progressions pédagogiques. L'utilisation réelle des outils en classe reste une information mal connue. Dans une étude récente sur les pratiques en collège², environ 20 % de l'échantillon des professeurs, toutes disciplines confondues, déclarent utiliser quelquefois, souvent ou toujours les outils de la banque. Ce résultat, qui peut sembler faible, est toutefois prometteur puisque le diagnostic *Banque* est un complément aux autres types d'évaluations pratiquées en classe.

Les professeurs ont souvent une relation ambivalente avec leur mission d'évaluation dans ses différentes dimensions (régulation, orientation, validation), d'autant plus qu'ils ont le souci «de ne pas faire n'importe quoi» en raison des enjeux qui y sont liés. Ils doivent répondre à la demande de repère de la part de l'institution mais aussi des familles : ce sont les contrôles, les examens et des notes qui permettent de décider de l'avenir scolaire des élèves (normes scolaires et sociales). La plupart de ces évaluations, parce qu'elles sont élaborées pour des groupes d'élèves et à partir d'objectifs généraux de classe, permettent des constats globaux (réussite ou échec) sur l'épreuve proposée. Pour se soucier davantage de chaque élève, il faut faire de l'éva-

2. Braxmeyer N., Guillaume J.-C. et Lévy J.-F., (2005), *Les pratiques d'évaluation des enseignants en collège*, Paris, Ministère de l'Éducation nationale, de l'enseignement supérieur et de la recherche – Direction de l'évaluation et de la prospective, collection «Les dossiers» n° 160.

luation/régulation des apprentissages pour chacun, ce qui est un aspect de l'individualisation. La difficulté pour le professeur est de rendre compatible la progression individuelle optimale de chaque élève et la progression collective optimale de toute la classe. Une même situation d'évaluation sera plus efficace pour l'un ou l'autre type de régulation poursuivi.

C'est probablement parce que les outils de la *Banque* conduisent à une investigation, à une analyse de la tâche et de la mise en œuvre de la compétence dans différents contextes, parfois éloignés du programme, qu'ils surprennent les enseignants car ce sont des pratiques qui ne s'intègrent pas facilement au fonctionnement collectif de la classe. C'est aussi parce que ces outils sont analytiques qu'ils demandent aux utilisateurs de comprendre le changement de perspective : il ne s'agit pas, comme en apprentissage, de guider l'élève dans la résolution des situations proposées, de l'aider à donner du sens à l'activité présentée. Il faut rechercher comme un détective pourquoi ces situations peuvent lui poser problème, vérifier si certains types d'erreurs ou difficultés sont commis : les situations sont construites pour vérifier dans quels cas l'élève échoue.

Deux constats peuvent être établis à partir des remarques transmises par les utilisateurs du site. D'abord, un diagnostic qui cherche à comprendre des pourquoi plausibles aux difficultés des élèves n'est pas la préoccupation principale. Des réponses en termes de comment résoudre la difficulté sont attendues. Nombreux sont ceux qui cherchent des remédiations sur le site.

En second lieu, utiliser tel quel quelque chose conçu par d'autres est difficile : adapter, ré-assembler les situations est demandé. La cohérence interne de chaque outil, tout comme le lien constitutif entre l'analyse et la forme du questionnement ne sont pas toujours perçus par les enseignants.

Ceci confirme un décalage avec les attentes exprimées et un besoin d'appropriation des outils proposés. Des dispositifs d'information et de formation à destination des utilisateurs sont indispensables mais restent encore insuffisamment développés.

Chapitre 7

Les outils d'évaluation ouverts, ou la nécessité de clés de fermeture

François-Marie Gerard

Les outils d'évaluation ouverts sont aussi appelés «outils ou items de production» ou encore «items à réponses construites». Ils s'opposent aux «outils de sélection» ou «items à réponses choisies». Si les textes qui font mention de cette distinction, présentent souvent les items de sélection avant les items de production, il faut bien avouer qu'historiquement, les outils ouverts sont les plus anciens. Depuis que l'homme essaie d'évaluer les acquis d'apprenants – c'est-à-dire vraisemblablement depuis que l'homme existe! –, il utilise des outils de production en demandant à l'apprenant de produire un comportement ou un discours qui témoigne de sa maîtrise.

Ce sont d'ailleurs ces questions ouvertes qui ont provoqué les premiers travaux de la docimologie, à partir de la dissertation philosophique, de la composition française, mais aussi de travaux de mathématiques. Dès 1938, Laugier et Weinberg mettront ainsi en avant qu'il faut 127 correcteurs d'une dissertation philosophique pour avoir une moyenne de notes qui ne varie plus par l'ajout d'un nouveau correcteur. Il en faut 78 pour la composition française, et 13 pour les mathématiques (de Peretti, 1993). Ces constats édifiants ont débouché sur la création des items de sélection. Ceux-ci se caractérisent par deux éléments :

- **d'une part, la réponse est fournie (parmi d'autres) à l'élève qui ne doit donc pas aller la chercher dans son répertoire cognitif mais l'identifier ;**
- **d'autre part, en corollaire, l'item permet une correction « objective » puisqu'il suffit de constater que l'élève a bien sélectionné la réponse attendue.**

Ces deux caractéristiques mettent en évidence certaines limites bien connues des deux types d'outils. En proposant des réponses, les items de sélection peuvent introduire dans le répertoire cognitif du sujet une information ou une démarche qui ne s'y trouve pas nécessairement. L'identification de la « bonne réponse » ne signifie donc pas nécessairement que cette réponse était connue du sujet. Ce problème ne se pose pas pour les items de production, mais celle-ci risque d'être parfois limitée par les exigences de communication écrite qu'elle nécessite et qui peuvent occulter les connaissances réelles.

La problématique de base des « outils de production » est aussi posée : ils débouchent sur une évaluation subjective et cette subjectivité serait à bannir, comme l'ont mis en évidence les travaux de la docimologie. La question est complexe et si plusieurs auteurs (Cardinet, 1992 ; de Peretti, 1998 ; Gerard, 2002 ; Weiss, 1986) ont montré que non seulement la subjectivité est inévitable, mais qu'elle peut même constituer une composante essentielle de l'acte d'évaluation, il n'en reste pas moins qu'un enjeu important des outils d'évaluation ouverts est d'apporter des « clés de fermeture » susceptibles de réduire la part de subjectivité et surtout d'éviter de transformer celle-ci en arbitraire.

Face à ces éléments, de nouveaux « outils de production » ont été développés ces vingt dernières années pour participer à l'évaluation. Parmi d'autres, nous conduirons notre propos sur la base de trois outils, que nous présenterons d'abord brièvement : le portfolio, l'évaluation de compétences à travers des situations complexes, l'évaluation de projets, de programmes ou de systèmes.

1 TROIS OUTILS OUVERTS

1.1 Le portfolio

S'il est bien un outil d'évaluation à propos duquel la littérature abonde depuis quelques années, c'est le portfolio (Allal, Wegmuller, Bonaiti-Dugerdil et Cochet Kaeser, 1998 ; Bélair, 2002 ; Dansereau, 2005 ; Eyssautier-Bavay, 2004 ; Forgette-Giroux et Simon, 1998 ; Mottier Lopez et Allal, 2004 ; Mottier Lopez, 2002 ; Paris et Ayres, 2000 ; Scallon, 2004).

Le portfolio est «un assemblage finalisé (*purposeful collection*) des travaux de l'élève qui démontrent ses efforts, ses progrès et ses acquisitions dans un ou plusieurs domaines. Il implique la participation de l'étudiant à la sélection des contenus, à la définition des critères de sélection et d'appréciation des travaux, ainsi que des manifestations d'autoréflexion (*self-reflection*) de la part de l'étudiant» (Northest Evaluation Association, 1990, traduit par Allal *et al.*, 1998, p. 7).

1.2 L'évaluation des compétences à travers des situations complexes

La principale caractéristique de l'approche par les compétences, désormais largement mise en œuvre dans les systèmes éducatifs francophones, est le fait de susciter la mobilisation de ressources intégrées pour résoudre des situations problèmes (De Ketele, 2000 ; Dolz et Ollagnier, 2002 ; Jonnaert, 2002 ; Le Boterf, 1994 ; Legendre, 2001 ; Perrenoud, 1997 ; Rey, 1996 ; Roegiers, 2000, 2003). Dès lors, ces compétences ne peuvent plus être évaluées par des items isolés formulés en termes de contenus et/ou d'objectifs opérationnels (Blais et Loye, 2003 ; Dauvisis, 2006 ; De Ketele et Gerard, 2005).

Dauvisis (2006) relève plusieurs outils pour l'évaluation scolaire des compétences : les livrets scolaires, des grilles analytiques, des évaluations authentiques en situation, des grilles descriptives... Mais l'outil le plus communément admis pour évaluer les compétences est la présentation d'une situation complexe relative à la famille des situations problèmes de la compétence, nécessitant de la part de l'élève une production complexe de résolution de la situation (Crahay et Detheux, 2005 ; De Ketele, 2001 ; Gerard, 2005 ; Rey, Carette, Defrance et Kahn, 2003 ; Roegiers, 2004 ; Scallon, 2004).

Les questions principales que pose ce genre d'outil sont : Comment valider les épreuves ? Comment assurer l'équivalence d'épreuves évaluant la même compétence ? Comment les corriger ?

1.3 L'évaluation de projets, de programmes ou de systèmes

L'évaluation existe aussi en dehors du champ scolaire, avec les mêmes problématiques en ce qui concerne l'élaboration des outils. L'évaluation de projets, de programmes ou de systèmes constitue un enjeu important (légitimité du projet, nécessité de faire la preuve du retour sur investissement de chaque action) mais rencontrent de nombreuses difficultés techniques, psychologiques et financières (objectivité de l'évaluation, peurs des acteurs à l'égard de l'évaluation, coûts parfois exorbitants de certains dispositifs).

Certains outils fermés existent en ce domaine (Commission européenne, 1999, 2001), mais ils ont du mal à s'adapter à la variété des situations. Aussi, à la suite des travaux de l'équipe de Stufflebeam (1974/1980), de D'Hainaut (1981) et de Kirkpatrick (1998), des outils d'évaluation ont été développés sous la forme de modèles structurant un certain nombre de critères permettant d'évaluer un projet ou un système selon toutes ses facettes et quel que soit le moment de son existence (Bourgeois et Roegiers, 1993; Gerard, 2001, 2003; Roegiers, 1997; Sall et De Ketele, 1997).

2 DES CARACTÉRISTIQUES COMMUNES, POUR DES CLÉS DE FERMETURE

Ces outils, malgré leurs différences, ont un certain nombre de caractéristiques communes, qui sont vraisemblablement liées à leur place et à leur spécificité dans l'évaluation. Certaines de ces caractéristiques conduisent à cadrer les outils, apportant ainsi des clés de fermeture qui contribuent à leur validité.

2.1 Les outils d'évaluation ouverts privilégient les trois fonctions de l'évaluation

Plus encore que pour un outil fermé, il importe de se demander à quoi sert l'évaluation réalisée avec des outils ouverts (De Ketele, 2006a) : poursuit-elle une fonction d'orientation, de régulation et/ou de certification ?

L'orientation peut être un élément crucial au niveau individuel, par exemple lorsqu'un institut de formation pour adultes utilise une épreuve sur les compétences génériques pour orienter les candidats lors de la procédure d'admission (De Ketele, 2004). Lors de la même procédure, les candidats sont invités à présenter leurs expériences en y joignant des pièces probantes, à l'instar du portfolio. L'orientation est, à un niveau plus collectif ou systémique, une fonction essentielle de l'évaluation de projets. L'évaluation *ex ante* est appelée à déterminer les objectifs, ainsi que les stratégies et les ressources nécessaires, sur la base des besoins issus du contexte et des bénéficiaires (Commission européenne, 1999).

La régulation est au cœur de l'utilisation du portfolio (Allal *et al.*, 1998; Mottier Lopez, 2002). Pour ces auteurs, le portfolio est avant tout un outil de régulation en permettant de contrôler la progression vers le but, d'assurer un retour sur l'action et de confirmer ou réorienter la trajectoire de l'action vers le but. Dans ce cas, le portfolio est souvent appelé *dossier d'apprentissage* ou *dossier de progression*. L'évaluation des compétences par situations complexes poursuit aussi avant tout une fonction de régulation. Ainsi, les « modules d'intégration » mis en place dans le cadre de la pédagogie de l'intégration (Roegiers, 2000, 2004) consiste à confronter

les élèves, tant à des fins d'apprentissage que d'évaluation, à des situations complexes à la suite d'apprentissages ponctuels. L'utilisation des critères et des indicateurs permet ensuite de dégager des stratégies personnalisées de remédiations. Enfin, les modèles d'évaluation intégrée de projets accordent une place importante à l'évaluation du processus, afin de régulariser l'ensemble du projet, notamment lors des *évaluations à mi-parcours*.

La certification est également visée par tous ces outils. Lorsque le portfolio s'appelle *dossier de présentation* ou *d'évaluation*, c'est avant tout cette fonction qui est poursuivie. De plus en plus, des épreuves d'évaluation certificative sont élaborées sur la base de situations complexes, appuyées par la construction de barèmes permettant de disposer d'une note. Les situations complexes apparaissent aussi dans les épreuves PISA et autres épreuves internationales. Enfin, les évaluations de projets accordent une grande importance à la certification, notamment quand il s'agit de décider de libérer ou non la dernière tranche budgétaire sur la base d'une évaluation finale, voire *ex post*. Cette dernière inclut l'évaluation de l'impact, essentielle pour garantir qu'un projet a servi à quelque chose, même si son interprétation est toujours difficile (Gerard, 2001).

Cette question de la fonction de l'évaluation est fondamentale pour les outils ouverts. C'est elle qui permettra de décider des caractéristiques premières de l'outil, d'orienter son élaboration et son utilisation. En quelque sorte, c'est une première clé de fermeture à apporter : savoir pourquoi et pour quoi on veut utiliser un outil.

2.2 Les outils d'évaluation ouverts deviennent des outils d'action

Cette évolution des outils d'évaluation est particulièrement sensible dans le cas du portfolio. Dossier d'apprentissage, il devient lui-même outil d'apprentissage : en créant son portfolio, l'élève est supposé apprendre mieux et plus en profondeur en le rendant conscient de ses apprentissages (Forgette-Giroux et Simon, 1998). Si cette idée généreuse, renforcée par le taux élevé de satisfaction des utilisateurs, n'a pas encore fait l'objet, à notre connaissance, de validations empiriques, il n'en va pas de même de l'utilisation de situations complexes : plusieurs études ont montré que leur utilisation dans le cadre de « modules d'intégration » permettaient aux élèves de gagner environ 15 points sur 100 (Gerard, 2005).

Comme le montre Scallon (2004), il importe sans doute de bien fixer le ou les buts qui sont poursuivis par les outils qu'on utilise, et notamment de savoir si on veut avant tout évaluer (pour préparer une décision ou une action) ou agir, sans quoi on risque de s'y perdre. Il s'agirait d'une deuxième clé de fermeture aux outils ouverts d'évaluation. Cette nécessité est peut être moins importante dans le cadre de l'évaluation de projets. Par exemple, c'est

naturellement qu'un modèle d'évaluation de projet ou de système devient outil de pilotage de ce projet ou de ce système (De Ketele, 2006b), car il fournit tous les éléments pour dégager – avant, pendant, à la fin et après – les nécessités d'action et de pilotage.

2.3 Les outils d'évaluation ouverts intègrent le sujet dans la démarche

Par définition, le portfolio fait de l'élève un acteur de son évaluation : l'élève agit dans une démarche de *co-construction de sens* (Derijcke, 2000), qui peut inclure non seulement l'enseignant et l'élève, mais aussi ses parents (Allal *et al.*, 1998). Ce rôle de l'élève acteur de son évaluation est moins présent dans le cadre des épreuves d'évaluation des compétences par situations complexes, quoique déjà en 1977, Tourneur et Bouillon faisaient état d'études qui ont montré qu'un moyen efficace pour apprendre à un élève à résoudre un problème était de lui demander de rechercher ou de construire lui-même des énoncés de problèmes.

Les outils d'évaluation de projets ont bien intégré, notamment dans le domaine de la santé, cette notion d'*évaluation participative* (Aubel, 2000) qui vise à considérer que les parties prenantes d'un projet, ses bénéficiaires, sont les meilleurs experts pour participer à la construction de sens recherchée par l'évaluation qu'on peut dès lors qualifier, à la suite de Weiss (1986) d'*évaluation intuitive et interactive objectivée*.

2.4 Les outils d'évaluation ouverts se fondent sur des critères et des indicateurs

Les spécialistes du portfolio insistent sur la détermination des critères de sélection et d'appréciation des travaux en montrant l'importance de l'implication de l'élève dans cette phase. L'évaluation de projets, même lorsqu'elle n'est pas entièrement participative, passe par une phase où les promoteurs et/ou bénéficiaires du projet définissent avec les évaluateurs les critères qui seront utilisés pour l'évaluation. Dans l'évaluation des compétences par situations complexes, la connaissance par les apprenants des critères sur lesquels ils seront évalués est essentielle, car c'est par rapport à ceux-ci qu'ils pourront optimiser leur performance. Une manière d'intégrer le sujet dans la démarche d'apprentissage relève donc du travail sur les critères.

Ceux-ci sont d'ailleurs un élément clé de l'élaboration des outils ouverts. Souvent, dans un outil fermé, le critère est implicite, non formulé. Dans l'évaluation des acquis, ce critère caché est souvent l'exactitude de la réponse (Gerard, 2005). Mais ce critère ne peut fonctionner dans les outils ouverts, tout simplement parce que la réponse correcte n'existe pas. L'ouverture de ces outils tient à la diversité des réponses, des productions

qui peuvent être apportées. Pour pouvoir limiter une subjectivité arbitraire dans l'appréciation de ces productions, il est essentiel de disposer de critères qui détermineront de manière précise les qualités qui sont attendues de cette production. Des indicateurs concrets permettront d'opérationnaliser ces critères de telle sorte que l'évaluateur sache exactement ce qu'il doit observer, la manière dont il doit gérer son observation et inférer à propos de la maîtrise de la compétence.

Les critères et les indicateurs sont une clé de fermeture supplémentaire – et fondamentale – des outils ouverts. Sans eux, ces outils ne seraient que des outils de recueil d'information. C'est parce que des critères sont définis et utilisés de manière systématique qu'ils deviennent des outils d'évaluation.

2.5 Les outils d'évaluation ouverts portent sur des dimensions multidimensionnelles

Les travaux qui peuvent apparaître dans un portfolio ou dans la résolution d'une situation complexe sont eux-mêmes complexes et flexibles, non réductibles à une seule et unique démarche ni à un seul et unique produit. Cette multidimensionnalité s'oppose au postulat de l'unidimensionnalité du trait mesuré qui fonde les techniques de la théorie classique des scores ou des théories de réponse à l'item.

Cette caractéristique n'est pas sans poser des difficultés méthodologiques et peut conduire à devoir limiter l'inférence qui est faite au départ de l'observation d'une réalité. Ainsi, Rey (2006) considère qu'il n'est pas possible d'inférer la maîtrise d'une compétence large à partir de la résolution d'une situation complexe. Il ne remet pas en cause l'utilisation de situations complexes pour évaluer une compétence, mais limite la définition de celle-ci au «fait d'être capable d'accomplir une tâche déterminée», homogène dans sa finalité technico-sociale, sans qu'on puisse inférer de la complexité d'une situation à la multidimensionnalité d'une compétence. D'autres auteurs mettent cependant l'accent sur la nécessité d'une inférence multidimensionnelle, en se fondant sur la notion de famille de situations, non limitée à une tâche déterminée (Barbier, 2006 ; Beckers, 2004 ; Roegiers, 2000, 2004).

Travailler ces questions liées à cette multidimensionnalité des réalités évaluées par les outils ouverts constitue un défi important pour l'avenir de ces outils.

2.6 Les outils d'évaluation ouverts favorisent la communication

Allal *et al.* (1998) ont bien montré combien le portfolio est un outil précieux pour favoriser la communication entre les enseignants, les élèves et les parents. Le portfolio présente l'avantage de proposer de l'information brute.

La personne qui a devant elle une production de l'élève est bien mieux renseignée sur la compétence de son auteur que par une note ou une appréciation. La discussion qui s'en suit peut dès lors se concentrer sur l'essentiel, sur le concret de la production de l'élève.

Dès 1992, Cardinet soulignait qu'on évite ainsi deux sources d'erreur majeures : celle du codage du premier correcteur et celle du décodage de celui qui doit interpréter le rapport d'évaluation, le bulletin. On évite en même temps de perdre l'information concernant tous les aspects qualitatifs du travail.

Ce problème de codage-décodage est très présent au moment de la correction d'épreuves par situations complexes et représente vraisemblablement une difficulté essentielle de ce type d'outil, nécessitant des études complémentaires (Blais et Loye, 2003 ; De Ketele et Gerard, 2005). Si le codage peut déboucher sur une note unique, non porteuse en elle-même de sens, il peut aussi s'inscrire dans une démarche plus descriptive que sommative (Roegiers, 2004 ; Scallon, 2004). Ce qui importe, c'est de pouvoir mettre en évidence quels critères sont maîtrisés ou non, et pourquoi il en est ainsi. On peut ainsi déboucher sur un tableau susceptible de favoriser la communication, pour autant qu'il reste compréhensible par des utilisateurs profanes.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Allal, L., Wegmuller, E., Bonaiti-Dugerdil, S. et Cochet Kaeser, F. (1998). Le portfolio dans la dynamique de l'entretien tripartite. *Mesure et évaluation en éducation*, 20, 5-31.
- Aubel, J. (2000). *Manuel d'évaluation participative de programme*. Beltsville Drive, Child Survival Technical Support – CSTS et Catholic Relief Services – CRS (2nd edition).
- Barbier, J.-M. (2006). Les voies nouvelles de la professionnalisation. In Y. Lenoir et M.-H. Bouillier-Oudot (éds) (2006). *Savoirs professionnels et curriculum de formation*. Laval, Les Presses de l'Université Laval, pp. 66-82.
- Beckers, J. (2004). *Qu'apporte la notion de familles de tâches à l'évaluation de compétences?* Lisbonne, Communication au colloque de l'ADMEE.
- Bélair, L.M. (2002). L'apport du portfolio dans l'évaluation des compétences. *Évaluation et formation. Questions vives*, 1 (1), 17-38.
- Blais, J.-G. et Loye, N. (2003). Une étude de l'accord et de la fidélité inter juges comparant un modèle de la théorie de la généralisabilité et un modèle de la famille de Rasch. *Actes du Colloque de la revue Mesure et*

- évaluation en éducation, lors du 71^e congrès de l'ACFAS, à Rimouski le 21 mai 2003.*
- Bourgeois E. et Roegiers X. (1993). *Une révision du modèle de Stufflebeam*. Lyon, Communication au colloque de l'ADMEE.
- Cardinet, J. (1992). L'objectivité de l'évaluation. *Formation et technologies – Revue européenne des professionnels de la formation*, n°0, 17-26.
- COMMISSION EUROPÉENNE (1999). *Évaluer les programmes socio-économiques*. Collections MEANS, 6 volumes.
- COMMISSION EUROPÉENNE (2001). *L'évaluation à la Commission européenne*. Europe Aid Office de Coopération.
- Crahay, M. et Detheux, M. (2005). L'évaluation des compétences, une entreprise impossible? (À propos de la résolution de problèmes mathématiques complexes). *Mesure et évaluation en éducation*, 28, n° 1, 57-78.
- Dansereau, V. (2005). L'utilisation du portfolio comme outil d'évaluation dans les classes préscolaires. *Mesure et évaluation en éducation*, 28, n° 1, 37-55.
- Dauvisis, M.-C. (2006). L'évaluation des compétences au risque des barèmes et des notes scolaires, *Usages sociaux de la notion de compétence*, Journée d'étude organisée par CNAM-CRF-MRPP-CEREQ-IUFM Champagne-Ardenne-ENSIETA, le 9 mars 2006 au CNAM, Paris.
- De Ketele, J.-M. (2000). En guise de synthèse: Convergences autour des compétences, in C. Bosman, F.-M. Gerard et X. Roegiers (éds). *Quel avenir pour les compétences?* Bruxelles, De Boeck, pp. 187-191.
- De Ketele, J.-M. (2001). Place de la notion de compétence dans l'évaluation des apprentissages. In G. Figari et M. Achouche (éds). *L'activité évaluative réinterrogée. Regards scolaires et socioprofessionnels*. Bruxelles, De Boeck, pp. 39-43.
- De Ketele, J.-M. (2004). *La reconnaissance et la validation des acquis de l'expérience dans le cadre de la formation d'adultes à l'Université de Louvain*. Lisbonne, Communication au colloque de l'ADMEE.
- De Ketele, J.-M. (2006a). Contrôles, examens et évaluation. In J. Beillerot et N. Mosconi (éds). *Traité des sciences et des pratiques de l'éducation*. Paris, Dunod, pp. 407-420.
- De Ketele, J.-M. (2006b). La qualité et le pilotage du système éducatif. In M. Behrens (éd.). *La Qualité en éducation. Pour réfléchir à la formation de demain*. Québec, Presses de l'Université du Québec, collection Éducation-Recherche, pp. 19-38.

- De Ketele, J.-M. et Gerard, F.-M. (2005). La validation des épreuves d'évaluation selon l'approche par les compétences. *Mesure et Évaluation en Éducation*, 28, n° 3, 1-26.
- de Peretti, A. (1993). *Controverses en éducation*, Paris, Hachette Éducation.
- de Peretti, A. (1998). Aspects objectifs et subjectifs de l'évaluation, in A. de Peretti, J. Boniface et J.-A. Legrand (1998). *Encyclopédie de l'évaluation en formation et en éducation. Guide pratique*. Paris, ESF éditeur, pp.481-484.
- Derijcke, M. (2000). La grille critériée et le portfolio à l'épreuve du suivi pédagogique. *Revue française de pédagogie*, n° 132 «Évaluation, suivi pédagogique et portfolio», 23-32.
- D'Hainaut, L. (1981). *Analyse et régulation des systèmes éducatifs*. Bruxelles, Labor-Nathan
- Dolz, J. et Ollagnier, E. (éds) (2002). *L'énigme de la compétence en éducation*. Bruxelles, De Boeck.
- Eyssautier-Bavay, C. (2004). *Le portfolio en éducation : concept et usages*. Inédit réalisé dans le cadre d'une thèse en informatique de l'Université Joseph Fourier Grenoble 1.
- Forgette-Giroux, R. et Simon, M. (1998). L'application du dossier d'apprentissage à l'université. *Mesure et évaluation en éducation*, 20, n° 3, 85-103.
- Gerard, F.-M. (2001). L'évaluation de la qualité des systèmes de formation. *Mesure et Évaluation en Éducation*, vol. 24, n° 2-3, 53-77.
- Gerard, F.-M. (2002). L'indispensable subjectivité de l'évaluation. *Antipodes*, n° 156, avril 2002, 26-34.
- Gerard, F.-M. (2003). L'évaluation de l'efficacité d'une formation. *Gestion 2000*, vol. 20, n° 3, 13-33.
- Gerard, F.-M. (2005). L'évaluation des compétences par des situations complexes. *Actes du Colloque de l'Admee-Europe*, IUFM Champagne-Ardenne, Reims, 24-26 octobre 2005.
- Jonnaert, Ph. (2002). *Compétences et socio constructivisme. Un cadre théorique*. Bruxelles, De Boeck.
- Kirkpatrick, D.L. (1998). *Evaluating Training Programs*. San Francisco, Berret-Koekler Publishers, Inc. (2nd edition).
- Laugier, H. et Weinberg, D. (1938). *Recherche sur la solidarité et l'interdépendance des aptitudes intellectuelles d'après les notes des examens écrits du baccalauréat*. Paris, Chantenay.

- Le Boterf, G. (1994). *De la compétence. Essai sur un attracteur étrange*. Paris, Éditions de l'organisation.
- Legendre, M.-F. (2001). Sens et portée de la notion de compétence dans le nouveau programme de formation. *Revue de l'AQEFLS*, 23(1), 12-30.
- Mottier Lopez, L. (2002). Interroger la pratique du portfolio en situation scolaire dans une perspective «située» de l'apprentissage. *Actes du 15^e colloque de l'ADMEE-EUROPE*, Université de Lausanne, Suisse.
- Mottier Lopez, L. et Allal, L. (2004). Le portfolio : pratiques et perspectives. *Éducateur* (numéro spécial 04), 46-51.
- Paris, S.G. et Ayres, L.R. – Traduction Aussanaire, M. (2000). *Réfléchir et devenir. Apprendre en autonomie. Des outils pour l'enseignant et l'apprenant*. Bruxelles, De Boeck.
- Perrenoud, Ph. (1997). *Construire des compétences dès l'école*. Paris, ESF.
- Rey, B. (1996). *Les compétences transversales en question*. Paris, ESF.
- Rey, B. (2006). Les compétences professionnelles et le curriculum : des réalités conciliables ? In Y. Lenoir et M.-H. Bouillier-Oudot (éds) (2006). *Savoirs professionnels et curriculum de formation*, Laval, Les Presses de l'Université Laval, pp. 83-108.
- Rey, B., Carette, V., Defrance, A. et Kahn, S. (2003). *Les compétences à l'école. Apprentissage et évaluation*. Bruxelles, De Boeck.
- Roegiers, X. (1997). *Analyser une action d'éducation ou de formation*. Bruxelles, De Boeck.
- Roegiers, X. (2000). *Une pédagogie de l'intégration*. Bruxelles, De Boeck.
- Roegiers, X. (2003). *Des situations pour intégrer les acquis*. Bruxelles, De Boeck.
- Roegiers, X. (2004). *L'école et l'évaluation*. Bruxelles, De Boeck.
- Sall, H.N. et De Ketele, J.-M., (1997). Évaluation du rendement des systèmes éducatifs : apports des concepts d'efficacité, d'efficience et d'équité. *Mesure et Évaluation en Éducation*, 19, n°3.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par les compétences*. Bruxelles, De Boeck.
- Stufflebeam D.L., Foley W.J., Gephart W.J., Guba E.G., Hammond R.L., Merriman H.O. et Provus M.M. (1980), *L'évaluation et la prise de décision en éducation*, Victoriaville (Canada), N.H.P. (1974 pour l'édition américaine).

- Tourneur, Y. et Bouillon, D. (1977). *Typologie des questions en mathématique, une étude critique*. Communication devant la Commission Internationale pour l'Étude et l'Amélioration de l'Enseignement de la Mathématique, Lausanne, juillet 1977.
- Weiss, J. (1986). La subjectivité blanchie? In J.-M. De Ketele (éd.) (1986). *L'évaluation: approche descriptive ou prescriptive?* Bruxelles, De Boeck, pp.91-105.

Chapitre 8

Formation, compétences et constructivisme : le référentiel d'évaluation, outil de cohérence dans les programmes de formation en santé

Florence Parent

1 INTRODUCTION

Le développement et la demande de plus en plus marquée de programmes de formation basés sur des approches par compétences dans le secteur de la santé (Nguyen, 2007 ; Gay et Heyrman, 2007) témoignent d'un besoin de changement dans les formations professionnalisantes de ce domaine de la formation. Bien que le rationnel qui sous-tend ce besoin soit variable, la même préoccupation anime l'ensemble des acteurs du changement : une meilleure efficacité des activités de formation relativement aux besoins des systèmes de santé, des professionnels de la santé et des populations bénéficiaires des services de santé. Derrière cette variabilité, une clarification du concept de compétence permettra de préciser les finalités des réformes envisagées et d'analyser la cohérence des dispositifs pédagogiques et d'évaluation qui sera mise en place (Jonnaert, Barette, Boufrahi et Masciotra, 2005). En effet, l'approche par compétences peut véhiculer des paradigmes épistémologiques très différents quant à la construction des savoirs. Le secteur de

la santé n'échappe certainement pas à cette possibilité de confusion, partagé entre des protagonistes du comportementalisme (empreinte et conditionnement) et d'autres ouverts au constructivisme et au socio constructivisme. Bien que la préoccupation prioritaire de l'ORE (Observatoire des réformes en éducation) soit d'assurer un maximum de cohérence quant à l'utilisation du concept de compétence dans les refontes des programmes (Masciotra, 2007), il est probable que le changement soit progressif, du moins dans le secteur de la santé. L'approche transmissive et l'approche comportementaliste y sont particulièrement ancrées, malgré des avancées de plus en plus précises en faveur d'un paradigme constructiviste dans certains contextes éducationnels (Grand'Maison *et al.*, 2007 ; Boulé et Langlois, 2007) ou de santé (World Health Organization, 1986 ; Gagnayre, 2007).

Ce chapitre met en avant l'intérêt de questionner les liens au niveau du curriculum entre les dispositifs d'évaluation et de formation dans le cadre d'un enseignement-apprentissage basé sur une approche par compétences et plus précisément d'une approche située de la compétence (Masciotra, 2007). Par ailleurs, il démontre que le développement d'une approche par compétences « intégrative » dans le secteur de la santé est une opportunité pour renforcer l'accessibilité à des ressources, méthodes et outils et pour évaluer de manière plus cohérente les objectifs éducationnels et pédagogiques pour l'ensemble de ses acteurs : apprenants, enseignants, professionnels, communautés.

2 OBJECTIF

L'objectif est de présenter, à partir d'un exemple, comment quand la porte d'entrée privilégiée est celle de l'évaluation, cela peut renforcer la cohérence d'une formation en santé basée sur une approche par compétences. Privilégier l'évaluation comme porte d'entrée est un excellent fil conducteur à celle de la cohérence globale. C'est à ce titre que l'on peut se référer à la phrase « Dis-moi comment tu évalues, je te dirai qui tu formes » (De Ketele, Chastrette, Cros, Mettelin et Thomas, 1989) dans une projection à la fois institutionnelle et de référentialisation (Figari, 2004 ; Legendre, 1993). Il s'agit de définir un cadrage formel en termes de référentiel d'évaluation et d'aller jusqu'au bout de la logique au niveau institutionnel afin de développer un cadre de référence cohérent mais aussi dynamique et actualisable.

3 CADRE CONCEPTUEL DE RÉFÉRENCE

Le concept de compétences auquel il est fait référence est celui « d'un savoir agir en situation ». On privilégiera la tendance actuelle (ORE, 2007) qui est de parler d'une approche située de la compétence ou celle plus ancienne faisant appel au concept d'une « pédagogie de l'intégration » (Roegiers, 2000) afin d'orienter les préoccupations pédagogiques sur la cohérence des finalités aux

moyens. En formation professionnelle et particulièrement dans le secteur de la santé, la globalité (ou intégration) et l'authenticité des situations devraient être au centre des apprentissages.

Les exemples dont il sera question dans cet article font clairement référence, en théorie, à une définition «située» de la compétence et à un paradigme socio constructiviste de construction des savoirs. Comme déjà précisé, ce choix est d'autant plus pertinent qu'il s'agit de contextes de formations professionnelles, dont les classes ou familles de situations (Jonnaert, 2002; Roegiers, 2000) peuvent être particulièrement précises et «donner sens» aux attendus des organisations, des professionnels et populations bénéficiaires. Par ailleurs, le constructivisme véhicule des principes d'action similaires à ceux rencontrés dans le champ de la Promotion de la Santé (Parent, 2007). Les concepts d'évaluation formative, d'interdisciplinarité, de participation, d'autonomie et de renforcement des capacités y prennent une place prépondérante. C'est ainsi que les dérives utilitaristes et les logiques de contrôle quant aux «mauvais usages» dont il pourrait être fait de l'utilisation du concept de compétence ne devraient pas y avoir une place émergente (Rozenblatt, 2000).

Une définition précise du concept de compétence est essentielle : «les changements de paradigme en perspective sont radicaux, ils ne peuvent se concrétiser que sous l'éclairage de finalités définies sans ambiguïté» (Jonnaert, Barette, Boufrahi et Masciotra, 2005). La compétence, c'est (figure 8.1) :

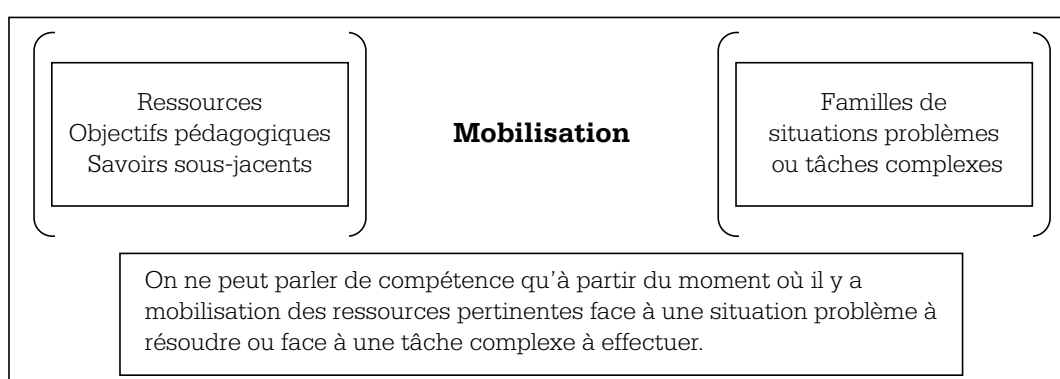


FIGURE 8.1. La compétence¹

1. Extrait de «Les bases théoriques de l'approche par compétences en lien avec les formations professionnalisantes», présentation par Jean-Marie De Ketele dans le cadre des Journées partenariales «L'approche par compétences et l'intégration des programmes verticaux dans la formation en santé» ARE@ Santé, Be-cause health et NESI, Bruxelles, septembre 2006 (www.areasante.org).

Un premier schéma conceptuel (figure 8.2), auquel il sera fait référence, permet d'observer que la compétence est reliée de manière équivalente à trois piliers du curriculum que sont les Référentiels de Compétences (RC), de Formation (RF) et d'Évaluation (RE). La définition de la compétence ne change pas, ce sont les référentiels qui s'adaptent selon leur fonction propre. De cette manière, une clarification du concept de compétence devrait guider à la fois la conception mais surtout la cohérence des programmes par compétences. Le Référentiel de Métier (RM) n'est pas repris sur ce schéma car il s'agit bien, ici, d'une analyse de la cohérence interne au programme de formation indépendamment de la question de «l'utilité systémique» (Parent, 2007) qui intègre notamment les réalités professionnelles et les normes sanitaires.

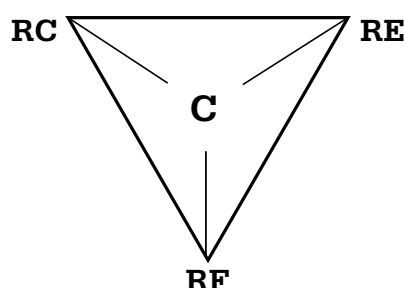


FIGURE 8.2. Compétence et piliers de la compétence

Le schéma suivant (figure 8.3), intitulé «le triangle d'adéquation», permet de mieux comprendre le concept d'«utilité systémique» pour le secteur de la formation. Le cercle qui relie les trois pôles du triangle témoigne d'un niveau de rapprochement minimum nécessaire à la question de «l'utilité systémique». Sans cela, chacun des pôles ne pourra pas interagir avec les

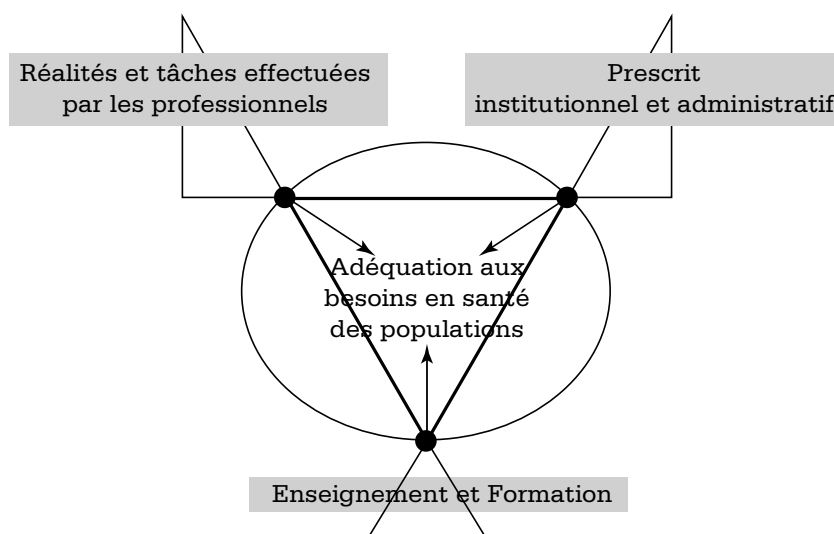


FIGURE 8.3. Le triangle d'adéquation

autres de manière harmonieuse. Le résultat étant une perte d'efficacité, voire des effets contre productifs entre les réalités professionnelles, les normes et les programmes de formation.

Ces deux schémas peuvent être adaptés à la formation continue dans la mesure où le RE est remplacé par le Bilan de Compétences (BC) et le RF par le Plan de Formation (PF). Le schéma suivant (figure 8.4) permet de mieux cerner les liens entre cette logique de référentiels, la formation initiale et la formation continue dans un contexte de valorisation du concept de «life long learning» (Parent, Coppieters et Kittel, 2007 ; Ministère de la Santé de RDC, 2006 ; Ministère de la Santé et de la Prévention médicale du Sénégal, 2007).

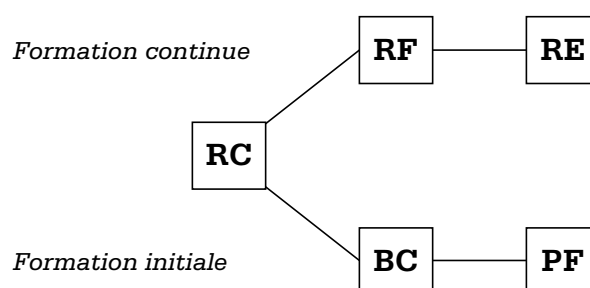


FIGURE 8.4. Le référentiel de compétences comme cadre de référence pour la formation continue et la formation initiale

Certaines adaptations sont cependant nécessaires entre un RC pour la formation initiale et celui qui sera utilisé pour la formation continue.

La progression dans la construction d'un programme «situé» par compétences dans le secteur de la santé commence par un travail de fond sur le référentiel de métier (notamment en regard du critère d'*exhaustivité*) avant de se poursuivre par une transposition didactique sous forme de référentiels de compétences, de formation et d'évaluation. La transposition didactique commence bien avec le référentiel de métier afin de renforcer l'adéquation de la formation avec les pratiques professionnelles (Perrenoud, 1998 ; Parent, 2007 ; Parent *et al.* 2006).

Il est utile de préciser que le RC reflète la vision du curriculum et devrait intégrer le cadre de référence en pédagogie (paradigme sous tendu). Le RC devrait être particulièrement bien validé et mis à jour (mécanisme de gestion de la qualité dynamique dans le temps) pour répondre à son rôle de cadre de référence pour la formation et l'évaluation. Le RC est un outil qui doit être accessible aux apprenants ; en tant que tel, il remplit bien son rôle de renforcement d'une identité professionnelle et de guidance pédagogique. Il s'agit en quelque sorte d'une clarification du contrat entre les apprenants et l'institut de formation. Le RC est constitué d'une part des compétences

transversales déclinées en Paliers de Compétences (PC)² ou capacités qui reflètent les ressources à mobiliser en situation. D'autre part, on y retrouve les Familles de Situations (FS)³ et l'Objectif Continué d'Intégration (OCI)⁴. Les compétences tout comme les FS sont déclinées par année d'étude, selon une complexité grandissante, en témoigne l'énoncé par année des FS, des OCI, des PC, des critères et indicateurs. Le RF en formation initiale et le plan de formation en formation continue devront refléter la cohérence avec le RC.

Le schéma suivant (figure 8.5) donne une place prépondérante à l'analyse de la cohérence entre les différentes activités d'un programme de formation.

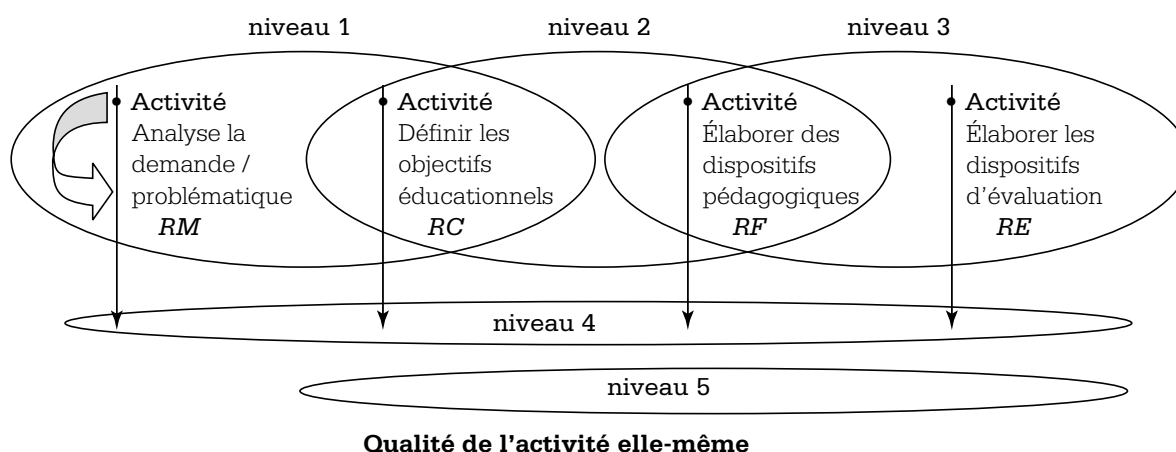


FIGURE 8.5. Analyse des niveaux d'écart dans l'action pédagogique en santé

Légende : RM = Référentiel de métier ; RC = Référentiel de compétences ; RF = Référentiel de formation ; RE = Référentiel d'évaluation

2. Soit de capacités propres à la compétence ou de ressources qui, au sein d'une même année, se juxtaposent aux autres pour former l'ensemble de la compétence pour cette année. Il devrait y avoir progression dans le niveau de complexité d'un même palier d'une année de formation à l'autre.
3. Ensemble des situations appartenant à une même catégorie de situations que l'agent de santé rencontre et qu'il est de sa responsabilité de gérer. Pour une même année de formation il peut y avoir plusieurs familles de situation selon les grandes fonctions et/ou lieux d'exercices professionnels.
4. La formalisation des compétences attendues se concrétise dans la formulation de ce que nous avons nommé l'«objectif continué d'intégration» (OCI) et ce en tous les cas dans les contextes de formation de base. L'OCI est une variante de ce que Roegiers a appelé l'objectif terminal d'intégration (OTI). Il s'agit de rendre directement visible ce qui est attendu tout au long de l'année comme macro compétence d'intégration, soit la synthèse des familles de situations.

Ce schéma donne un cadre pour l'évaluation des activités de formation prenant en compte la qualité des activités mais aussi et surtout la cohérence des activités entre elles (niveaux d'écarts). Par exemple, l'écart de niveau 3 se situe entre le RE et le RF, tandis que l'écart de niveau 4 inclut l'ensemble des niveaux d'écart qui peuvent exister entre chacune des 4 activités. Ce schéma permet de ne pas dissocier, tout en les analysant, les liens entre la mise en place d'un dispositif d'évaluation, la formation et les objectifs de formation. La simplicité de ce schéma est paradoxale à la réalité qui voit de nombreuses discordances entre les objectifs de renforcement des compétences et les dispositifs pédagogiques et d'évaluation.

Seule la cohérence du programme de formation, qui inclut le RC, le RF et le RE (niveaux 2, 3 et 5), est analysée, par la porte d'entrée de l'évaluation, dans l'exemple qui va suivre⁵.

Le processus d'élaboration des outils sur le terrain a respecté une approche résolument socio constructiviste, basée sur la participation, l'appropriation et la co-construction méthodologique (Lafortune et Deaudelin, 2002 ; Anadon, 2007).

4 L'ÉVALUATION DANS LA FORMATION INITIALE DES PROFESSIONNELS DE SANTÉ : EXEMPLE D'UN RÉFÉRENTIEL D'ÉVALUATION POUR L'ENSEIGNEMENT INFIRMIER

L'exemple qui est présenté émane d'une synthèse théorique à partir de divers contextes de mises en œuvre de formations, pour infirmier(ère)s et médecins, basée sur une logique d'APC aussi bien en formation continue (contexte du Sénégal) (Parent *et al.*, 2007) qu'en formation initiale (contextes de la RDC, du Rwanda, de la Belgique). L'exemple choisi n'est qu'une alternative parmi d'autres, mais veut rendre compte de la manière de considérer les liens entre le RC, le RF et le RE dans une perspective qui respecte les principes d'actions de l'APC, des pédagogies actives et du constructivisme.

4.1 Caractéristiques principales du programme

Le dispositif d'évaluation devra servir les besoins d'un programme de formation professionnelle caractérisé de la façon suivante :

- l'existence d'un RC décliné en compétences, Paliers de Compétences (PC) et Familles de Situations (FS) pour lesquels des critères et indicateurs de réalisation sont définis ;

5. De manière complémentaire, mais sans que cela ne soit décrit dans cet article, le même modèle est analysé au travers d'une grille composée de 8 critères : la *validité*, la *fidélité*, la *précision*, la *progression*, l'*acceptabilité*, la *faisabilité*, l'*intégration* ou la *cohérence*, la *place de l'évaluation formative*. Cette grille de critères fait référence notamment aux travaux de Tardif (Tardif, 2006 ; Scallon, 2004 ; Normand, 2007).

- la répartition des activités d'apprentissage en trois catégories dans la planification annuelle : les cours (enseignements théoriques), les situations d'intégration⁶ et les stages (enseignement pratique);
- les tableaux de spécification des PC «répartis» dans les cours, les stages et les Situations d'Intégration (SI);
- les PC et les FS deviennent les objectifs pédagogiques.

Sans pour autant changer de vocabulaire, mais afin de mieux cerner celui utilisé, certains parallélismes sont opérés. On peut considérer que les paliers de compétences sont équivalents au concept de «ressources» utilisé notamment par Tardif (Tardif, 2006). D'autres auteurs feront référence au concept de «capacités». Derrière ces «ressources» on retrouve les différentes catégories de savoirs cohérentes aux activités professionnelles ayant servi à construire, depuis le référentiel de métier, le référentiel de compétences. Il s'agit des savoirs réflexifs, cognitifs, procéduraux, sociaux et psychoaffectifs (Charlier, 2003; Parent *et al.*, 2006; Parent, 2007). En effet, dans les travaux menés avec le secteur de la santé cette catégorisation s'est avérée particulièrement cohérente aux activités professionnelles rencontrées.

4.1.1 Les outils du référentiel d'évaluation

Le RE est constitué notamment :

- d'un répertoire de grilles de critères et d'indicateurs, en provenance du RC, pour guider la récolte qualitative des données sur les terrains de stage (qui se fait par observation, entretien et/ou analyse de contenu);
- d'une banque de situations d'intégration avec leur fiche de notation individuelle, construite elle aussi sur base des critères et indicateurs du RC;
- de tableaux des appréciations par cours selon les PC qui y sont mobilisés et certifiés;
- d'un tableau de spécification des PC à mobiliser et à certifier dans chaque cours;
- d'un tableau de spécification des PC à mobiliser et à certifier dans l'enseignement pratique;

6. Ce sont des situations qui, à l'école, préparent l'apprenant à ses stages. Moments clés de la formation, ces périodes d'intégration mobilisent dans des situations qui simulent le stage, les savoirs, savoir-faire et savoir-être appris dans les cours. Une situation d'intégration est une situation complexe, contextualisée et signifiante qui engage l'apprenant à utiliser tous ses savoirs, issus de plusieurs cours et qui nécessite la réalisation d'une production clairement précisée. L'activité est orientée vers l'OCI et vers les FS. Les situations d'intégration sont préparées et encadrées par les enseignants des modules qui y entrent en jeu pendant les semaines d'intégration prévues au programme (ARE@ Santé, 2007). Il s'agit toutefois dans ces situations d'intégration de jouer sur des «habillages» différents qui remettent à chaque fois l'apprenant dans une autre situation à gérer; soit il s'agit d'éviter de n'être que dans la reproduction ou la simple application de mêmes ressources déjà mobilisées dans d'autres SI.

- d'un bulletin ;
- des règles de délibérations.

Les critères et indicateurs dont il est question dans les outils du RE sont rendus accessibles à l'apprenant à travers le RC. Seuls les tableaux de spécification et le bulletin sont repris ci-dessous. L'utilisation de l'ensemble des outils du RE est abordée aux divers endroits de la discussion.

Le choix d'un modèle d'évaluation basé sur une logique qualitative prédomine, sans exclure la possibilité de se référer à certains moments à l'utilisation d'une logique quantitative (par exemple pour répondre à la problématique du critère de *fidélité*).

Les tableaux de spécification précisent les objectifs pédagogiques et guident autant la formation que l'évaluation. En effet, il importe de former et d'évaluer sur les mêmes objets. Dans le tableau ci-dessous, la lecture verticale renvoie aux cours disciplinaires, tandis que la lecture horizontale fait apparaître les compétences et paliers de compétences. La synthèse de l'évaluation est concrétisée sur la droite du tableau, tandis qu'en bas se retrouvent le nombre d'heures par cours et le nombre de PC mobilisés ou devant être maîtrisés dans chacun des cours.

[illegible]

Noter + ou – dans les carrés blancs selon que le PC est maîtrisé ou non. La maîtrise ou non du PC est basée sur le seuil des 2/3 de l'atteinte des critères/indicateurs du PC

3. Prester les soins	1.3.1 Réaliser les soins.																								/3	/4
	1.3.2. Réaliser les activités liées au laboratoire																								/3	
	1.3.3 Réaliser activités liées à l'assainissement																								/3	
	1.3.4 Réaliser les activités de réadaptation																								/3	
4. Manager	1.4.1 Enregistrer un patient																								/3	/2
	1.4.2 Utiliser rationnellement le matériel																								/3	
	Heures de cours	15	30	45	30	15	30	15	45	30	30	30	45	90	15	45	30	15	30	30	30	30	15	90	30	42
	Nombre de PC	1	2	2	1	1	2	1	1	2	2	1	1	3	1	1	1	1	1	2	2	1	1	3	1	42

Légende : PC = Palier de compétence; A = Aucune maîtrise; P = Maîtrise partielle; S = Maîtrise satisfaisante; M = Maîtrise maximale.

Pour construire ce tableau de spécification, les balises suivantes doivent être respectées :

- En lecture horizontale, s'assurer que chacun des PC soit mobilisé au moins trois fois dans des cours différents, mais avec un maximum de 5 fois afin de rester suffisamment précis et d'éviter de « tout évaluer partout ».
- En lecture verticale, s'assurer que le nombre de PC évalués par cours soit lié, autant que possible, au nombre d'heures dans chacun des cours. Ce critère « temps » est cohérent aux dispositifs pédagogiques de l'APC qui demandent de l'espace temps pour l'intégration des pédagogies actives en classe. Dans l'exemple, un cours de 15h évalue un seul PC. Il ne devrait pas être proposé plus de 4 PC pour un cours même si le cours dépasse les 60 heures. Pour des raisons de *faisabilité et de pertinence* une flexibilité est toujours nécessaire.
- Une approche participative et représentative des acteurs du système, basée sur le consensus afin de déterminer la pertinence des croisements PC/cours.

Pour remplir ce tableau de spécification, les règles suivantes doivent être appliquées :

- L'évaluation des PC et des compétences se fait de manière binaire, maîtrise ou non maîtrise. Un gradient peut être obtenu pour la compétence sous forme de maîtrise maximale (tous les PC sont atteints), maîtrise satisfaisante (2/3 des PC sont atteints), maîtrise partielle (moins de la moitié des PC sont atteints) et aucune maîtrise (aucun PC n'est atteint). C'est au minimum la maîtrise satisfaisante qui assure la maîtrise de la compétence (le système reste binaire).
- Quel que soit le nombre de PC par cours, la décision, dans les cas de discordance se prend par consensus. Dans la mesure où des notations chiffrées auraient été maintenues pour les PC dans les cours, la moyenne peut aussi être envisagée afin de faire face au problème de *fidélité* dans le cas d'une forte discordance (Scallon, 2004).

Ce premier tableau de spécification met en exergue le PC comme unité d'évaluation pour le modèle de RE proposé.

4.1.2 Particularités de l'approche pour l'enseignement théorique

Cette manière de faire est intéressante car, en favorisant un bilan « horizontal », c'est-à-dire transversal aux cours, c'est le PC et non le cours disciplinaire qui l'emporte en termes d'objectif d'évaluation et partant de là, de formation. On favorise ainsi une optique résolument orientée sur l'acquisition et la maîtrise des ressources ou savoirs sous-jacents. Le choix d'évaluer des PC et non des cours par discipline vient renforcer l'interdisciplinarité

dans les dispositifs de formation et d'évaluation. Par ailleurs, cette manière de procéder ne donne plus le pouvoir d'une décision d'échec ou de réussite d'un cours à un seul enseignant, mais au contraire favorise le consensus et le débat qualitatif nécessaire au moment de l'évaluation. En effet, ce type de modèle respecte la complexité et l'intégration pour la construction des savoirs (PC distribués dans les cours). Une même cohérence devra être assurée au moment de l'évaluation de la maîtrise de ces savoirs.

Néanmoins, en favorisant les regards croisés sur un même objet, le PC, dans différents cours, ensuite sur une même compétence à travers différents PC et finalement sur différentes compétences, la simple sommation quantitative des notes obtenues pose des problèmes de *fidélité* et de *validité*. Le biais majeur de la notation chiffrée tient au fait que cela conduit à un modèle de notation compensatoire où des sommes d'unités ou de points associés à des PC ou des compétences compensent pour des «compétences» qui ne sont pas pour autant maîtrisées (Scallon, 2004 ; Normand, 2007). Un autre critère de qualité attendu pour un référentiel d'évaluation, d'autant plus dans un cadre de formation par compétences, est la *précision*. En effet, pour un modèle de formation basé sur ce type de tableau de spécification, il devrait être aisé de repérer les compétences et les composantes des compétences (les PC) qui posent problème afin de faciliter les mécanismes de feedback et les stratégies de remédiations. Si on procède à une quantification des notes de chacun des PC et ensuite des compétences, on est en mesure de se demander ce que signifie le score final ainsi obtenu. Ce manque de précision vient encore renforcer les problèmes d'interprétation des résultats.

En se référant à ces critères de *fidélité*, de *validité* et de *précision*, on se rend compte qu'un choix devrait être opéré entre le modèle de référentiel de formation auquel on adhère dans un cadre d'approche situationnelle des compétences ou de pédagogie de l'intégration, et l'approche qualitative en évaluation à laquelle il faudra faire face. Ce choix d'une approche qualitative du RE est d'autant plus cohérent que l'évaluation dans un cadre d'APC se base sur l'utilisation de grilles de critères et d'indicateurs et sur des méthodes de récoltes qualitatives des données déjà écrit plus haut.

Dans la mesure où le modèle d'évaluation serait maintenu dans une logique exclusivement quantitative de sommation pour les enseignements théoriques, il est sans doute préférable de s'en maintenir à un modèle classique d'évaluation disciplinaire par cours où chaque enseignant donne sa note.

Les protagonistes des approches quantitatives argumenteront la possibilité de liens PC/cours comme suffisante dans le référentiel de formation sans y voir le levier institutionnel de l'évaluation comme un facteur renforçant des changements de comportements, notamment en termes de pratiques professionnelles en interdisciplinarité. D'autres arguments font appel

		FS Hôpital			FS Maison repos			FS Domicile			FS Communauté			FS recherche			Décisions			
		SI maîtrisés	Stage	Port folio	SI maîtrisés	Stage	Port folio	SI maîtrisés	Stage	Port folio	SI maîtrisés	Stage	Port folio	SI maîtrisés	Stage	Port folio	Total PC maîtrisés	Maîtrise (+ ou non maîtrise (-) par PC	Niveau de maîtrise de la compétence	Décision A/S/P/M
Noter + ou – dans les carrés blancs selon que le PC est maîtrisé ou non. La maîtrise ou non du PC est basée sur le seuil des 2/3 de l'atteinte des critères/indicateurs du PC																				
1. Communiquer	1.1.1 Entrer en relation avec personne																/4			
	1.1.2 Informer la personne																/4			
	1.1.5 Rédiger un rapport																/3			
	1.1.6 S'intégrer à l'équipe																/3			
2. Analyser	1.2.1 Collecter les données																/3			
	1.2.2 Identifier les besoins/les problèmes																/3			
	1.2.9 Planifier les soins requis																/3			
	1.2.11 Évaluer sa démarche																/3			

Noter + ou – dans les carrés blancs selon que le PC est maîtrisé ou non. La maîtrise ou non du PC est basée sur le seuil des 2/3 de l'atteinte des critères/indicateurs du PC

[illegible]

au critère de *faisabilité* qu'il ne faut cependant pas confondre avec le critère d'*acceptabilité*. En effet, la question de la faisabilité devrait être abordée comparativement à l'utilisation d'une approche quantitative qui peut notamment être difficile à gérer dans des contextes à faibles ressources technologiques (pour l'encodage et l'élaboration de tableaux par exemple). Par ailleurs le critère d'*acceptabilité* est particulièrement complexe car il se confronte à des représentations divergentes de l'évaluation et des savoirs, provenant d'acteurs impliqués différemment dans l'institution éducative, depuis les parents d'élèves, les apprenants ou les enseignants. En effet, la question de l'*acceptabilité* est particulièrement prégnante dans le cadre de l'évaluation des enseignements théoriques suite aux représentations très ancrées de l'importance à accorder aux seuls savoirs cognitifs et à la garantie de leur «bonne évaluation» à travers une évaluation quantitative. Un troisième argument pour favoriser l'évaluation disciplinaire et quantitative dans l'enseignement théorique, toujours aussi dans une volonté de simplifier le système et de le rendre soi disant plus *faisable*, est de considérer que c'est le terrain de stage qui est le plus propice à l'évaluation des compétences. Cet argument qui est juste en soi fait fi de la nécessité de favoriser, dans le cadre d'une formation basée sur une APC, la maîtrise au préalable des ressources ou PC dans une logique similaire d'intégration. De plus, la dichotomie qui s'installe entre l'enseignement théorique et l'enseignement pratique (présenté à travers le tableau de spécification suivant, tableau 2) n'est pas favorable à l'implantation d'un changement réel, en double boucle (Argyris et Schön, 2002) faisant appel à un changement de valeurs ou de finalités et de pratiques.

Le tableau 8.2 présente de la même manière que pour l'enseignement théorique le croisement des paliers de compétences. Dans ce cas, il ne se fait plus avec les cours, mais avec les situations d'intégration, les stages et le rapport de stage ou le port folio correspondants à une même famille de situation dans l'enseignement pratique. Une lecture horizontale nous donne une appréciation de la compétence, tandis qu'une lecture verticale nous donne une appréciation de l'intégration. En effet, dans la mesure où l'unité d'évaluation n'est pas seulement le palier de compétence mais aussi la famille de situation, il devient impératif d'évaluer les lieux d'intégration de manière globale.

L'exemple proposé correspond à une année de formation type dans laquelle 5 FS sont visées. Le nombre de FS peut augmenter en fonction des années d'étude et ce n'est qu'en année ultérieure qu'on pourrait voir apparaître dans certains contextes, d'autres lieux d'exercice professionnel. Le tableau de spécification devra alors s'adapter avec des colonnes supplémentaires. En général, en formation initiale, les FS sont au nombre de 2 à 6 maximum et qui vont en se complexifiant d'année en année.

Comme pour l'enseignement théorique, il est encore plus important de favoriser dans l'enseignement pratique des regards

croisés sur un même objet (critère de *fidélité*). C'est ce qui est assuré lors du jugement de maîtrise d'un même PC au sein d'une même FS. La question de la *fidélité* se joue également en amont du tableau 2, par une décision consensuelle entre les différents observateurs toujours sur base des fiches de notations individuelles qui intègrent les critères et indicateurs du RC (en général ces observateurs devraient être au nombre de 3). Soulignons enfin que chaque PC devrait être évalué au moins dans trois familles de situations.

Par ailleurs, les circonstances spécifiques des projets ont amené à considérer le port folio comme un élément équivalent aux SI et aux stages. Cette situation entraîne cependant une certaine confusion. Les SI et les stages sont des situations professionnelles «authentiques», tandis que l'écriture du rapport de stage est une production, qui certes nécessite la mobilisation d'un certain nombre de ressources, mais ne reflète pas en soi un lieu d'exercice professionnel au quotidien. Néanmoins, le port folio et les travaux de fin d'étude sont des occasions, non seulement d'évaluer les mêmes PC en situation, mais également peuvent être considérés comme des ressources favorables au développement de pratiques réflexives et d'explicitation, facteur pouvant renforcer l'intégration globale.

4.1.3 Modéliser l'intégration

Il ne s'agit pas seulement de vérifier la maîtrise des ressources, mais également la maîtrise de la compétence (Tardif, 2006). C'est à ce niveau qu'il apparaît utile de faire un retour sur la clarification du concept de compétence et de sa définition. En effet, dans l'approche méthodologique choisie, ce qui est nommé «compétence» reflète plutôt des «macro capacités», tandis que la définition plus exacte de la compétence, en référence à l'ASC (Approche Situationnelle de la Compétence), serait celle de la FS. La FS intègre plusieurs compétences et nécessite de ce fait, la mobilisation en situation de ressources différentes. Méthodologiquement ceci paraît logique car l'émergence des compétences telles qu'elles sont actuellement nommées⁷, provient

7. *Compétence 1.* Établir une communication professionnelle: Échanger, dans le respect mutuel, de manière à établir une relation de confiance avec l'individu, la famille, la communauté et les autres organisations.

Compétence 2. Analyser une situation et développer un esprit critique: Identifier au départ des données collectées les problèmes de santé des individus, familles ou collectivités et prendre de manière participative des décisions pertinentes.

Compétence 3. Réaliser les prestations de soins et offrir des services: Dispenser les soins et offrir les services de manière globale, continue et intégrée en partenariat avec l'individu, la famille ou la communauté.

Compétence 4. Manager les ressources: mobiliser et utiliser rationnellement les ressources en vue de répondre efficacement aux besoins des individus, des familles, des communautés et des autres organisations (extrait de Parent *et al.*, 2006).

de la catégorisation des verbes d'actions en fonction des savoirs sous-jacents auxquels ils se réfèrent. Il paraît donc logique de voir différents types de savoirs mobilisés dans chacune des FS. Les alternatives pour essayer d'avoir le plus de cohérence entre le concept, sa définition et ses représentations derrière le vocabulaire qui sera utilisé, sont notamment de :

- redéfinir les compétences actuelles comme des macro-capacités, tandis que les PC sont bien des capacités. Envisager la FS comme étant le reflet le plus proche de la définition de la compétence ;
- envisager les compétences actuelles comme des compétences «transversales», tandis que les FS reflètent les compétences «spécifiques» ou d'«expertise» de la profession. Cette proposition n'est cependant pas radicale car les compétences «transversales», témoignent aussi, à travers leurs PC, d'une certaine spécificité professionnelle. Ceci, au-delà du fait que lors de la construction du RC c'est le verbe d'action et non le contexte de son application qui devrait l'emporter dans la formulation des PC. L'objectif est bien de favoriser le côté opérationnel de l'outil et de renforcer la possibilité de construire un curriculum favorable aux mécanismes de transferts des apprentissages ;
- considérer que c'est l'ensemble du RC qui doit refléter l'intégration et donc qu'il s'agit de valoriser, lors de l'évaluation, tout autant les compétences que les FS dans le but d'évaluer la maîtrise d'objets différents (ressources ou PC et intégration ou FS).

Un critère de «cohérence» est ajouté à ce tableau de spécification car l'évaluation de la mobilisation d'un nombre minimum de PC dans un même contexte n'assure pas que les liens entre ces mêmes PC mobilisés aient été faits. Par exemple, la prise en charge d'une problématique de santé peut être bien faite et en lien avec le diagnostic de maladie et pourtant l'intégration n'est pas complète car la prise en charge manque d'une approche globale. Sans une attention précise sur ce critère, les liens peuvent vite se perdre au risque de nouveau de se focaliser simplement sur chacune des composantes (à savoir chacun des PC). Il est nécessaire dans la construction des outils du RE, de placer systématiquement ce critère en amont, dans le canevas des situations d'intégration, les grilles d'observation des terrains de stage et dans celui du port folio. Dans le temps, il est probable que ce critère devienne moins prioritaire, étant intégré dans le savoir-faire en évaluation des enseignants et des apprenants.

Ensuite, c'est à partir de ces deux tableaux que le bulletin pourra être établi selon une logique d'évaluation qualitative cohérente à l'approche par compétences.

Le tableau 8.3 présente un exemple de bulletin. Celui-ci peut prendre des formes différentes, il est nécessaire de retrouver les grandes catégories ou objets évalués pour lesquels un consensus a été posé :

- les compétences «transversales» ou macro-capacités dans chacun des deux types d'enseignement ;
- les FS spécifiques au niveau d'études et le critère de cohérence.

La fusion des situations d'intégration, stages et rapports de stage, en regard des FS, est un choix qui peut varier selon les contextes.

Tableau 8.3. Exemple d'un modèle de bulletin

	Enseignement théorique				Enseignement pratique			
	A	P	S	M	A	P	S	M
Compétence 1								
Compétence 2								
Compétence 3								
Compétence 4								
					A	P	S	M
Famille de situations «Hôpital»								
Famille de situations «Maison de repos»								
Famille de situation «Domicile»								
Famille de situation «Communauté»								
Famille de situation «Recherche»								

Légende : A = Aucune maîtrise ; P = Maîtrise partielle ; S = Maîtrise satisfaisante ;
M = Maîtrise maximale

Il est proposé de noter une croix dans la colonne adéquate et de définir des critères de délibération en considérant à la fois la maîtrise des compétences dans l'enseignement théorique et pratique et de manière encore plus intégrative dans les FS. En effet, la lecture de l'intégration se fait à différents endroits et à différents niveaux selon les dispositifs pédagogiques mis en place. Dans l'enseignement théorique il s'agit en général plus d'évaluation de PC (ou de capacités/ressources/savoirs sous jacents) en y incluant bien sûr déjà un niveau d'intégration pour autant que les dispositifs pédagogiques et d'évaluation soient pensés dans cette optique. La place de l'intégration dans l'enseignement pratique est directement plus évidente mais c'est bien dans la lecture verticale de la situation (FS) que l'intégration est la plus en proximité avec l'approche située de la compétence.

Comme déjà abordé précédemment, il s'agirait dès lors, dans le référentiel d'évaluation, de parler plutôt de macro-capacités dans l'enseignement théorique et pratique (pour l'enseignement pratique, uniquement dans la lecture horizontale du «tableau de spécification des PC à mobiliser et à certifier dans l'enseignement pratique», tableau 8.2) et de n'avoir recours au terme de compétences que lorsqu'on synthétise les niveaux de maîtrise dans la FS (lecture verticale du tableau 8.2). Une correction similaire devrait prendre cours dans le RC et dans le RF (figure 8.2).

Ce type de modèle de bulletin, rempli à partir des tableaux 8.1 et 8.2, donne des perspectives en termes de *précision* et d'*évaluation formative*. Et c'est bien de cette précision rendue visible que peuvent être renforcés institutionnellement des mécanismes de remédiations (notamment en ce qui concerne le système d'organisation des secondes sessions).

De manière plus concrète, en cours d'année on peut favoriser des retours à l'étudiant (à l'occasion des activités cumulées dans tous les types d'enseignement) et l'orienter très tôt vers des dispositifs de remédiations tantôt organisés par l'institution de formation et tantôt à mettre en place par l'étudiant lui-même.

Dans tous les cas ces retours multiples en cours et fin d'année donnent à l'étudiant suffisamment de précisions de façon à ce que, dans une démarche constructive et autonome, il puisse déployer des stratégies de réussite plus efficaces. Un exemple d'application peut être clarifiant : celui de l'étudiant démontrant sa maîtrise intégrative des compétences dans la FS en communauté et pas à l'hôpital. Il s'agit alors d'aller rechercher les informations qui expliquent cette non maîtrise dans cette FS en se posant les questions suivantes : cette non maîtrise porte sur combien de macro capacités ? Une, deux, ou plus ? Que lit-on de la maîtrise des composantes de ces macro capacités de l'étudiant pour cette FS dans l'enseignement théorique, dans les stages et dans les port folio ? S'agit-il dans ces différentes occasions d'évaluation des mêmes PC non maîtrisés ? Etc.

Les différents documents ayant «nourri» le bulletin final viennent donc donner les informations utiles en vue d'une part d'explicitier l'échec mais surtout d'identifier, comme déjà dit plus haut, les remédiations possibles.

Pour diverses raisons (contrôle de l'enseignant sur ses notes, visibilité des résultats pour les parents (critère d'*acceptabilité*)) il peut être envisagé d'associer à l'évaluation des PC une évaluation par discipline de «connaissances cognitives», avec des points proportionnels au nombre d'heures de cours. Indépendamment du type de bulletin que l'on choisirait, il faudrait y faire apparaître le pourcentage global obtenu en théorie et l'associer à une fiche de notes par discipline. Les règles de délibération devraient prendre en compte ce pourcentage. L'évaluation des savoirs cognitifs peut aussi servir à

prendre des décisions en délibération dans le cas d'hésitation entre la réussite et l'échec.

Une échelle de qualification de la réussite ou de l'échec peut, selon les contextes, être envisagée afin de renforcer l'*acceptabilité* d'un système d'évaluation basé sur une logique qualitative et binaire d'appréciation de la maîtrise des compétences.

4.1.4 *Synthèse et discussion*

Les réformes de l'enseignement et de la formation, menées dans plusieurs contextes dans le secteur de la santé, permettent de mieux en mieux de préciser et de formaliser les démarches méthodologiques. Néanmoins, si l'élaboration des référentiels de compétences et de formation semble parfois bien maîtrisée et leurs utilisations bien acceptées, l'élaboration des référentiels d'évaluation pose de nombreuses interrogations fondamentales. Le changement de paradigme semble plus se situer au niveau de l'évaluation que de la formation et du cadre de référence. On réalise que même dans des institutions porteuses de changements et volontaires d'un enseignement-apprentissage basé sur les pédagogies actives ce sont les modifications des pratiques d'évaluation et d'usage du bulletin qui témoignent réellement d'un changement en profondeur (figure 8.5). Il est actuellement de plus en plus prioritaire, dans la mesure où le changement réel est voulu et que des acteurs s'engagent dans le processus, de documenter et critiquer les chemins potentiels ou déjà pris en faveur de plus de cohérence entre les référentiels de compétences, de formation et d'évaluation dans le cadre de formations basées sur la logique des compétences en santé.

La formalisation de la démarche de construction d'un RE a permis un certain nombre de constatations et de liens. Des éléments de discussion ont, par ailleurs, déjà été abordés. La synthèse et la discussion aborderont : (a) d'une part la question du choix de la porte d'entrée dans l'élaboration d'un programme de formation par compétences et la mise en œuvre du changement ; (b) d'autre part, certains éléments techniques et méthodologiques en termes d'ingénierie des compétences seront formalisés.

a) Débat sur le choix de la porte d'entrée dans l'élaboration d'un programme de formation par compétences et de mise en œuvre du changement

Il est nécessaire de séparer l'activité de construction du curriculum de celle de sa mise en œuvre malgré la nécessaire intégration.

La progression au sein des différents terrains d'action en santé dans l'élaboration d'un curriculum par compétences en passant du RM, au RC, au RF pour finalement en arriver au RE est, d'une certaine manière, logique et d'autant plus facilement comprise. Néanmoins, tout au long du processus de

construction, on observe un certain nombre d'allers retours entre chacun des référentiels. Le RF et l'écriture d'une banque de SI par année d'étude (selon l'organisation de la progression de l'enseignement) répondent aux préoccupations des enseignants de varier leurs dispositifs de formation et d'introduire les pédagogies actives en classe. Le travail sur le RF est renforcé par l'existence en réseau (colloques, écoles partenaires, centre de ressources documentaires, Internet et ouvrages en pédagogies) d'un grand nombre d'exemples de pratiques enseignantes et apprenantes qui favorisent l'autonomie.

Il apparaît cependant plus difficile d'avancer de manière cohérente avec le changement quand on en arrive à la question de l'évaluation. Les freins et les résistances peuvent émaner d'acteurs différents : enseignants de la base, inspecteurs, parents d'élèves ou apprenants. L'emprise, voir le mythe de l'évaluation quantitative, prédomine de même que le « retour sur le contrôle des savoirs purs » revient en force. La vérification de « savoirs purs » et l'approche quantitative en évaluation sont des facteurs rassurants. Les habitudes ancrées des pratiques d'évaluation sont vite mises en échec quand le choix de départ est de favoriser la mise en œuvre d'une approche par compétences.

C'est ainsi qu'on a pu observer que des avancées significatives sur le référentiel de formation peuvent être perdues face aux changements radicaux que l'évaluation exige dans un cadre d'approche situationnelle des compétences. Étant nécessairement dans des modèles de formation plus complexes et proches des réalités professionnelles, les modèles d'évaluation devront intégrer cette complexité. Ce sont les modèles d'évaluation qualitative, basés notamment sur l'utilisation de grilles critériées, de méthodes de récoltes qualitatives des données et valorisant les échanges et le consensus entre enseignants, qui sont les plus valides (Scallon, 2004). Cependant, le manque de comparaison vient fragiliser certaines avancées, et l'évolution constructive, par tâtonnement avec l'ensemble des acteurs d'un processus de changement, est plus facilement envisageable lors de l'élaboration du RC et du RF que lors de l'élaboration du RE. Plus de ressources sont nécessaires afin d'argumenter le changement de paradigme qui devient explicitement palpable à travers la décision de qualification de l'étudiant.

Néanmoins, plusieurs hypothèses complémentaires peuvent être posées, outre le manque de compréhension de l'impact du changement sur l'évaluation par l'ensemble des acteurs. Le rationnel sous-jacent à la mise en œuvre d'une APC peut se limiter d'une part à avoir une liste de compétences et capacités conforme aux problématiques et contenus de santé et de l'autre, à l'utilisation de pédagogies actives en classe. L'évaluation peut être un objet à part, visant ou non les compétences attendues, en regard seulement du RC et non du RF. La cohérence globale qui devrait inclure à la fois l'objet réel de l'évaluation dans un cadre d'ASC, et un changement dans les pratiques d'évaluation de l'institution, n'est pas la finalité poursuivie par toutes les réformes dans un cadre d'approche par compétences.

D'autre part, l'approche comportementaliste dans l'utilisation de la logique des compétences peut aisément se satisfaire d'un cadre d'évaluation basé sur un modèle plus linéaire par objectifs exclusivement. C'est là un des dangers de voir le curriculum par compétences être facilement détourné au profit d'une pédagogie par objectifs dont on connaît les limites, et ce malgré l'existence de plus en plus fréquente d'un référentiel de compétences dans les institutions de formation en santé.

Ces constats nécessitent de questionner d'une part l'importance qu'il faut accorder tout au long du processus à une clarification constante du concept de compétence (figure 8.2) et, d'autre part le choix de la porte d'entrée en faveur du changement (figure 8.5). Malgré la logique progressive de penser d'abord au cadre de référence, ensuite à la formation, pour finalement questionner l'évaluation, il serait pertinent de favoriser une double entrée par le RC et le RE. De cette façon, le travail sur les outils du RF pourra renforcer la cohérence globale en assurant le pont entre le RC et le RE et en vérifiant les liens entre évaluation et formation. Une autre alternative, complémentaire, est de s'assurer en début de construction, d'une adhésion des acteurs clés au changement de paradigme qui sera traduite dans les outils d'évaluation. Une négociation clairement établie et qui précise les tenants et aboutissants du changement est une étape à envisager selon les contextes et l'origine des protagonistes du changement.

La porte d'entrée dans le changement ne pose pas seulement la question de la meilleure manière d'aborder l'élaboration des outils du curriculum, mais aussi celle de la question sur le plan du processus. Un travail de sensibilisation et de formation aux dimensions que véhiculent les pédagogies actives en termes d'évaluation formative et de constructivisme est aussi une priorité pour la masse critique des acteurs qui vont progresser avec le changement. Dans certains contextes, ce travail préalable a permis une meilleure *appropriation* et une plus grande *acceptation*, voir volonté de mettre en place des modalités d'évaluations innovantes (Parent *et al.*, 2004).

b) Conséquences en termes d'ingénierie des compétences du RE sur le RF et le RC

Il est important de noter combien l'utilisation d'un cadre de référence en termes de RE peut devenir un facteur renforçant en termes de modifications de pratiques enseignantes et apprenantes. Dans la mesure où le cadre de référence propose un enseignement théorique orienté sur les savoirs disciplinaires, versus un enseignement théorique orienté de manière interdisciplinaire sur les savoirs pluriels et la maîtrise des ressources ou capacités (PC), on peut émettre l'hypothèse de l'existence d'un levier institutionnel, par le biais du RE, en faveur d'un changement de pratiques.

En effet, le risque est grand de revenir (ou rester) à un enseignement «traditionnel» dans la mesure où une évaluation disciplinaire serait maintenue. Seuls certains enseignants plus engagés dans les pédagogies actives et/ou le constructivisme seraient motivés par l'organisation de dispositifs pédagogiques et d'évaluation intégrateurs. Un choix d'évaluation négocié clairement au niveau de l'institution, avec l'ensemble de son corps professoral, permet de renforcer les efforts des enseignants porteurs d'une dynamique créative pour l'élaboration des dispositifs pédagogiques telle qu'une banque de données de situations d'intégrations notamment. Par ailleurs, la formation des enseignants aux dimensions de l'évaluation et de la formation dans un cadre d'ASC doit rester une priorité de toute réforme.

Une attention aux stratégies de contournement devrait systématiquement être questionnée afin d'améliorer le système dans sa globalité et de mieux comprendre les résistances. Un travail de cohérence méthodologique, entre les différents outils, est également important afin d'éviter des endroits de confusion qui pourraient renforcer les résistances ou être des freins à la mise en œuvre.

Il est en effet intéressant de noter que le processus méthodologique d'élaboration du RE renvoie naturellement à revisiter les autres référentiels dans une visée de cohérence systémique (figures 8.2 et 8.5). Les conséquences de cette cohérence sont manifestes : toute correction dans un outil nécessite de la penser dans les autres. Il y a là un effet de d'interaction, entre les différents référentiels. Cela ne pose pas de problème pour autant qu'un mécanisme de gestion de la qualité ait été mis en place afin d'accompagner le processus. À titre d'exemples de révision du RF ou du RC en regard d'une entrée nouvelle par le RE, on peut citer :

- La nécessité que le nombre de critères et le nombre de PC correspondent au seuil qu'on se serait donné pour assurer la réussite. En effet, il est plus aisé d'avoir un minimum de 3 critères et de 3 PC quand on veut appliquer un seuil de réussite au 2/3, ou de 4 critères et de 4 PC si on envisage d'appliquer un seuil à 3/4 par exemple (soit d'avoir réussi 3 unités d'évaluation sur les 4 pour considérer la réussite).
- Quand le modèle d'évaluation maintient une logique d'évaluation des PC dans différents cours, il devient moins pertinent de vouloir catégoriser dans le RF les cours selon les compétences dans une logique de module. Ceci nous amène à constater que certains passages du RF sont plus des témoins de l'ordre du processus ou de la méthodologie que d'une place opérationnelle à terme, quand le produit est finalisé.
- Une cohérence devrait se maintenir dans le choix des cours depuis la première à la dernière année de formation. À cet égard ceci est décisif dans le choix de maintenir un cours s'inscrit dans la vérification de la complexité grandissante du PC en regard des critères.

De manière générale, en termes d'ingénierie pédagogique, il est nécessaire de respecter les 8 critères définis en note de bas de page précédemment et de ne pas tendre uniquement vers le critère de *faisabilité* qui peut, en simplifiant de manière « pragmatique » les choses, répondre au besoin de certains acteurs moins imprégnés du paradigme constructiviste et de sa complexité inhérente. Il existe alors un risque de voir les avancées dans le discours (théorique) se connoter d'un modèle opérationnel beaucoup plus éloigné de la réalité d'une APC.

Il est indispensable de souligner que nous pouvons aller très loin dans de l'ingénierie de plus en plus standardisée. C'est alors que la question de la pertinence reprend le dessus et le pour « quoi » choisir l'APC. Ces allers retours sur la question du sens et des dérives de toute directive pédagogique devraient être constants et venir par là nourrir les choix des référentiels d'évaluation. Ceci est d'autant plus important qu'on sait l'impact que ces choix ont sur la formation et l'apprentissage.

4.2 Synthèse

Le référentiel d'évaluation peut être un facteur renforçant essentiel pour engager l'institution éducative dans des changements de pratiques d'évaluation, d'enseignement et d'apprentissage en cohérence au concept de compétence véhiculé.

L'enjeu est de taille car il se situe non seulement dans la complexité même des dispositifs à créer, mais aussi dans le continuum qui devrait exister pour les professionnels de la santé entre l'évaluation de leurs propres compétences et celles des patients et des communautés. À partir du moment où le lien entre évaluation, formation et compétences est mieux cerné (figure 8.5), le professionnel de la santé sera plus à même de développer des dispositifs pédagogiques adéquats dans le domaine de l'éducation thérapeutique du patient et celui de la prévention et de la promotion de la santé. En effet, l'objet de l'évaluation se situe principalement à trois niveaux :

- la formation initiale, qui vise l'apprenant comme futur professionnel de la santé ;
- la formation continuée, qui vise le professionnel de la santé ;
- l'éducation thérapeutique, l'éducation à la santé et la promotion de la santé, qui visent le patient et la communauté.

Dans la mesure où un modèle d'enseignement peut imprégner les pratiques courantes et professionnelles, le développement réel d'une ASC à travers un paradigme socio constructivisme devrait renforcer une efficacité globale des activités de formation en favorisant une approche centrée non seulement sur l'apprenant (en formation initiale) mais aussi sur le professionnel (en formation continuée) et sur la communauté et le patient (dans le cadre

de l'éducation thérapeutique du patient (Lefebvre, Pelchat, Levert, David et Gauvin-Lepage, 2007), la prévention et la promotion de la santé).

Néanmoins, malgré une définition claire de la compétence et une volonté affichée de tendre vers l'autonomie et le constructivisme, ce n'est pas pour autant que le changement soit abordé et accepté par tous les acteurs de la même façon. Notamment le risque de revenir vers une pédagogie par objectifs reste présent. C'est à ce niveau que la cohérence entre l'évaluation et les objectifs pédagogiques doit prendre une place particulière et fondamentale comme levier de changement institutionnel, pour autant que les dispositifs d'évaluation soient eux-mêmes cohérents au concept de compétence véhiculé (schéma1).

Bien qu'un référentiel d'évaluation établi de manière cohérente aux objectifs de formation et aux compétences attendues, peut représenter un levier institutionnel sérieux en faveur du changement, l'élaboration et ensuite l'acceptation (phase d'implantation) d'un tel référentiel dépendra des ressources mises dans le développement des compétences éducationnelles des enseignants et des acteurs concernés par le changement. Des travaux de fond sur les représentations, les théories de l'apprentissage, les capacités de former et d'évaluer via les pédagogies actives dans un cadre d'ASC, la mise en réseau sont autant de leviers pour la cohérence. C'est ainsi qu'il s'agit de mettre en place, selon chacun des contextes, un véritable accompagnement socio constructiviste des réformes en éducation (Lafortune et Deaudelin, 2002).

Remerciements. Ce chapitre représente un travail d'écriture en interdisciplinarité directe avec C. Lejeune et Y. Coppieters. Les remerciements vont aussi à l'ensemble des acteurs des projets et processus qui ont nourri cette synthèse et aux personnes ayant assuré une relecture, E. d'Hoop, R. Baulana, G. Kahombo, D. Lemenu et A. Santucci. Un merci tout particulier à J.-M. De Ketele pour ses révisions de l'article et son accompagnement pédagogique constant depuis le début de ce cheminement.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Anadon M. (2007). *La recherche participative*. Multiples regards. Presses de l'Université du Québec.
- ARE@ Santé. (2007). Évaluation et suivi de la mise en œuvre de la réforme infirmière en RDC. Bruxelles, www.areasante.org.
- Argyris C., Schön D.A. (2002). *Apprentissage organisationnel. Théorie, méthode, pratique*. Bruxelles, De Boeck Université.
- Boulé R., Langlois L. (2007). Trois outils pédagogiques visant l'amélioration de la compétence à résoudre des problèmes de santé. Présentation

- au 4^e Forum international francophone de pédagogie médicale, Mont-Tremblant (Québec).
- Charlier P. (2003). Élaborer une description de fonction et un profil de compétences. Séminaire de formation et de réflexion. BIEF. Louvain-la-Neuve, 2003.
- De Ketele J.-M., Chastrette M., Cros D., Mettelin P. et Thomas J. (1989). *Le Guide du Formateur*. Bruxelles, De Boeck.
- Figari. (1994). *Évaluer: quel référentiel?* Bruxelles, De Boeck.
- Gagnayre R. (2007). L'éducation thérapeutique du patient: un champ à part entière en éducation médicale. *Pédagogie Médicale*, 8, 115.
- Gay P., Heyrman J. (2007). Développer un programme de formation basé sur les compétences en médecine générale. European Academy of teachers in general practice (EURACT). Présentation au 4^e Forum international francophone de pédagogie médicale, Mont-Tremblant (Québec).
- Grand'Maison P., Gaudreault M., Schofield A., Lamarche S., Lalande R., Gagné P. (2007). Offrir la totalité de la formation médicale pré-doctorale à l'extérieur du campus universitaire: des défis pédagogiques à la hauteur des besoins des communautés. Universités de Sherbrooke et de Montréal. Présentation au 4^e Forum international francophone de pédagogie médicale, Mont-Tremblant (Québec).
- Jonnaert P., Barette J., Boufrahi S., Masciotra D. (2005). Contribution critique au développement des programmes d'études: compétences, constructivisme et interdisciplinarité. *Revue des Sciences de l'éducation*, 3, 667-96.
- Lafortune L., Deaudelin C. (2002). *Accompagnement Socioconstructiviste. Pour s'approprier une réforme en éducation*. Presses de l'Université du Québec, collection Éducation Intervention.
- Lefebvre H., Pelchat D., Levert M.-J., David C., Gauvin-Lepage J. (2007). Programme d'intervention familiale interdisciplinaire. Partenariat entre professionnels de la santé, personnes ayant un traumatisme craniocérébral et proches soignants. *Pédagogie Médicale*, 8, 91-100.
- Legendre R. (1993). *Dictionnaire actuel de l'éducation*. Paris-Montréal, Guérin/Eska.
- Masciotra D. L'agir compétent: une approche situationnelle. Consulté sur le site <http://www.ore.uqam.ca/Documentation/Masciotra/Agir%20comp%20tent%20situ%20.pdf> (23.04.2007).
- Ministère de la Santé de RDC. (2006). *Référentiel de compétences infirmières Sciences de santé du niveau secondaire*. Direction de l'enseignement des sciences de santé, Kinshasa.

- Ministère de la Santé et de la Prévention médicale du Sénégal. (2007). *Recueil d'outils de bilans de compétences et de plan de formation pour l'Infirmier Chef de Poste de santé (ICP). Contexte de supervision formative et/ou d'identification des besoins de formation*. MSPM, Dakar, 2007.
- Nguyen D.-Q. (2007). Approche basée sur les compétences dans la formation clinique : implications pour l'enseignement et l'évaluation. Université de Montréal. Présentation au 4^e Forum international francophone de pédagogie médicale, Mont-Tremblant (Québec).
- Normand S. (2007). Commentaires et observations sur la mise en place d'un dispositif d'évaluation pour un modèle de l'enseignement infirmier en République Démocratique du Congo. Université de Montréal, Bureau d'évaluation, Faculté de Médecine.
- ORE (2007). Colloque de l'observatoire des réformes en éducation. Logique de compétences et développement curriculaire : débats, perspectives et alternative pour les systèmes éducatifs. UQAM, Montréal. <http://www.ore.uqam.ca/>
- Parent F. (2007). *Déterminants éducationnels et facteurs favorables à une meilleure adéquation entre formation et compétences attendues des professionnels de la santé dans les organisations de santé en Afrique*. Étude sur la gestion et le développement des ressources humaines en santé. Thèse de Doctorat en Sciences de la Santé Publique. École de Santé Publique/ULB. D/2007/Florence Parent, Auteur-Éditeur.
- Parent F., Coppieters Y., Kittel F. (2007). *Gestion des connaissances et la formation continuée des médecins prestataires de l'ONE*. Université libre de Bruxelles (École de Santé Publique).
- Parent F, Kahombo G, Bapitani J, Garant M, Coppieters Y, Levêque A, Piette D. (2004). A model for analysis, systemic planning and strategic synthesis for health science teaching in the Democratic Republic of the Congo : a vision for action. *Human Resources for Health*, 2, 16.
- Parent F., Lemenu D., Lejeune C., Baulana R., Komba D., Kahombo G., Coppieters Y. (2006). Présentation du référentiel de compétences infirmier en République Démocratique du Congo dans le cadre du renforcement des ressources humaines en santé. *Santé Publique*, 18(3), 459-473.
- Parent F., Ndiaye M., Coppieters Y., Deme S., Sarr O., Lejeune C., Lemenu D., De Ketele J.-M. (2007). Utilisation originale de l'approche par compétences en supervision formative en Afrique subsaharienne. *Pédagogie médicale*, 8(3), 2-22.
- Perrenoud P. (1998). La transposition didactique à partir de pratiques : des savoirs aux compétences. *Revue des sciences de l'éducation* (Montréal), 3, 487-514.

- Roegiers X. (2000). *Une pédagogie de l'intégration : compétences et intégration des acquis dans l'enseignement*. Paris-Bruxelles, De Boeck.
- Rozenblatt P. et al. (2000). *Le mirage de la compétence*. Paris, Syllepse.
- Scallon G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Paris-Bruxelles, De Boeck Université.
- Tardif J. (2006). *L'évaluation des compétences. Documenter le parcours de développement*. Montréal, Chenelière Éducation.
- World Health Organization (1986). The Ottawa Charter for Health Promotion. First International Conference on Health Promotion, Ottawa, 21 November 1986.

Chapitre 9

Portfolios et entretiens de co-évaluation : des leviers de la professionnalisation des jeunes enseignants

Lucie Mottier Lopez, Sabine Vanhulle

Certains programmes de formation initiale à l'enseignement recourent désormais au portfolio en tant que pratique d'évaluation. Les objectifs fréquemment assignés à cet outil sont d'encourager des prises de conscience chez les étudiants à propos de l'émergence de leurs compétences et de leur identité professionnelles. Les documents qu'il contient attestent de leurs capacités à entrer dans des démarches réflexives et métacognitives et dans des formes de pensée critique et créatrice. Le portfolio se présente encore comme un support à l'intégration de savoirs théoriques et pratiques.

Nous l'examinons ici sous deux angles interdépendants, liés à nos démarches de formation et de recherche¹ : nous l'abordons d'abord comme un outil servant à guider l'appropriation de savoirs et la construction identitaire professionnelle. Il entre alors dans des dispositifs de formation fortement interactifs qui vont jusqu'à vouloir articuler les processus régulateurs de l'évaluation formative aux exigences de l'évaluation certificative.

1. Présentées à l'ADMEE-Europe, Reims, 2005, dans une conférence à deux voix.

Ensuite, nous analysons la fonction de médiation du portfolio dans le bilan des apprentissages tel qu'il se déroule à travers des entretiens de co-évaluation. Un exemple est exposé afin de mettre en évidence, notamment, quelques apports de l'articulation entre le portfolio et les échanges intersubjectifs. Mais avant d'entrer dans ces deux volets, voyons brièvement ce qui semble faire consensus chez plusieurs auteurs francophones à propos du portfolio.

1 LE PORTFOLIO EN FORMATION À L'ENSEIGNEMENT : QUELQUES ÉLÉMENTS DE CONSENSUS

Au-delà de références épistémologiques et théoriques diverses, comme la perspective située, le constructivisme piagétien ou l'interactionnisme social inspiré de Vygotski, la littérature anglophone et francophone relative au portfolio relie l'usage de cet outil à la nature sociale du fonctionnement cognitif. Ce qui peut varier, ce sont entre autres : la définition du «social», entre des perspectives situées très focalisées sur les contextes des situations éducatives et des perspectives davantage culturelles ou décentrées de ces situations ; les conceptions de la notion de «médiation» dans les apprentissages ; le statut attribué au «sujet» apprenant et à la notion d'«identité» ; l'importance effectivement accordée à l'interaction avec les pairs dans le développement cognitif ; la définition des compétences et de l'usage de référentiels ; les conceptions de l'accompagnement et de l'étayage *versus* l'autonomie du formé et, partant, les conceptions mêmes de l'activité évaluative.

Ainsi, la littérature du champ est loin d'être homogène. Des nuances subtiles mais aussi de réelles divergences traversent les conceptions relatives à l'usage du portfolio en formation professionnelle et au type d'évaluation qu'il soutient. Sa définition en tant qu'«instrument», notamment, diffère donc d'un programme de formation à l'autre. Nous ne nous risquons pas à analyser ces différences ici, surtout dans les limites de ces pages. Relevons plutôt quelques éléments clés qui semblent faire consensus au moins parmi divers formateurs-chercheurs francophones².

Comme le résume Bélair (2005), le portfolio se définit fréquemment comme «un outil d'auto-évaluation et d'évaluation dynamique, facilitant le dialogue, la négociation et la réflexion sur des thèmes, des objectifs ou des compétences abordées». Il «permet à l'étudiant de fournir les preuves factuelles de son niveau de compétences et de justifier la qualité de ses

2. Nous nous appuyons notamment sur les communications de formateurs-chercheurs lors du colloque de l'ADMEE-Europe, Reims, 2005 : A) Symposium consacré à l'évaluation par portfolio en formation professionnelle à l'enseignement (organisation : L. Mottier Lopez) avec L. Allal (discutante), L. Bélair, C. Fantoli, G. Hoefflin, L. Mottier Lopez, B. Noël, J. Rouiller, E. Runtz-Christan, S. Vanhulle (discutante), E. Wegmuller. B) Communications en atelier, notamment C. Vanieuwenhoven, D. Ledur et M. Labeeu ; M. Deum, C. Mattar et S. Vanhulle.

performances», de prendre conscience de ses forces et faiblesses, mais aussi, de se positionner, de «prendre parti».

Au-delà d'une diversité d'appellations (dossier d'apprentissage, dossier de développement professionnel, etc.), le portfolio en formation à l'enseignement est le plus souvent traité comme outil d'autoquestionnement, de développement d'une pratique réfléchie (*praxis*), d'intégration théorie-pratique ou encore de conscientisation, pour l'étudiant, de ses rapports propres aux savoirs, à la profession d'enseignant, à des valeurs, aux autres et à soi-même (Runtz-Christan et Rouiller, 2005). Orienté vers le développement de compétences complexes, il fait intervenir diverses formes de pensée, réflexive, critique et créative (Vanhulle et Schillings, 2004); métacognitive et ancrée dans des concepts autant que dans des pratiques (Noël, 2005); expérientielle et narrative mais aussi distanciée et offerte à l'abstraction réfléchissante; objectivante et (inter)subjective (Vanhulle, 2005).

2 UN OUVRAGE ÉVOLUTIF POUR UN PROCESSUS DE DÉVELOPPEMENT

Tout cela met en exergue la double dimension du portfolio, autant «produit» matériel réalisé en fonction de consignes déterminées, qu'élément d'un processus.

2.1 Les deux facettes de l'instrument : instrumental et sémiotique

La dimension «produit» se reflète dans la fonction d'archivage du portfolio, dans son statut d'outil de monstration – pouvant aller, selon les cas, jusqu'à la présentation des compétences acquises à des étapes déterminées. Ainsi, c'est comme produit réalisé à un moment donné qu'il vient outiller la certification à travers des entretiens de co-évaluation. Mais le portfolio fait aussi figure d'«œuvre», ou d'«ouvrage», ou d'«artéfact», achevé provisoirement aux étapes clés de la formation, toujours ouvert à des réorganisations. Il illustre les moments du processus que ces pièces jalonnent, il alimente un devenir. La première condition de réussite du recours à cet outil est, dès lors, dans le rapport de l'étudiant au portfolio comme projection concrète de sa progression et des éléments de cette progression qu'il désire donner à voir. De matériel (c'est un simple porte-document, au départ), l'instrument se fait ainsi reflet, témoignage : il montre, il signifie, c'est un outil sémiotique qui médiatise une activité prolongée d'apprentissage.

Si le portfolio soutient un parcours singulier, il sert aussi de base à la construction de savoirs et de questionnements partageables. Ses documents sont propices à l'échange avec autrui (les pairs, les formateurs) : d'après les portfolios, où en sommes-nous par rapport à telle compétence ? Pourquoi tel concept fait-il sens aujourd'hui ? Comment redéfinir tel type de problème que

plusieurs rencontrent fréquemment dans les classes? Comment sommes-nous parvenus à telle compréhension, reformulation, solution? Quelle culture de l'enseignement/apprentissage se dessine à partir de nos portfolios respectifs? Comment y sommes-nous parvenus?

2.2 Portfolio et didactique des savoirs professionnels

En d'autres termes, le portfolio-processus a pour fonction fondamentale de socialiser l'engagement des formés dans leurs apprentissages à partir du principe même de ce que Jorro (2001) nomme «l'apostrophe évaluative». Dans notre perspective, l'auto-évaluation continue, que le portfolio stimule, vient nourrir la construction collective de savoirs. En articulant ainsi l'auto questionnement et sa socialisation, le portfolio s'inscrit dans une didactique des savoirs professionnels (Vanhulle, à paraître).

Celle-ci vise à intégrer dans un dispositif qui associe étroitement apprentissage, évaluation et développement professionnel, des démarches d'appropriation qui mettent en tension les contenus mêmes de toute formation en alternance, partagée entre des cours et des stages, mais surtout oscillant entre des visions multiples et contradictoires sur l'éducation et l'enseignement.

Certains théoriciens de l'alternance en formation parlent volontiers d'une pédagogie du «tiers-temps formatif» (Vanhulle, Merhan et Ronveaux, à paraître) pour désigner la construction de soi par l'«alternant», à travers des types de savoirs, des espaces, des temporalités, des communautés de pratiques différents: ce travail consiste à «faire avec» l'hiatus partiel entre la théorie et la pratique, irréductibles l'une à l'autre, et à se construire une culture de manière autonome. C'est une traversée faite de ruptures et de tensions pour le formé, qui, à notre sens, doit être accompagnée et outillée. L'instrument portfolio fait partie des modalités possibles d'accompagnement. Sans résoudre la question des hiatus théorie-pratique et des contradictions, il sert au moins à mettre à plat les différentes sources possibles de progression, de compréhension des phénomènes rencontrés dans les stages, d'interprétation des faits éducatifs à partir de théories et de définition de sa propre approche d'enseignant en devenir.

2.3 Des écrits de formation balisés par des critères

Dans cet esprit, le portfolio doit contenir davantage que de simples bilans d'activités de stages ou autres grilles de performances attestables à travers des comportements ad hoc. Dans des dispositifs tels que ceux de Vanhulle et Schillings (2004) et Deum, Mattar et Vanhulle (2005), les consignes d'écriture des textes à verser dans le portfolio ont pour buts d'entraîner les futurs enseignants dans des démarches intellectuelles complexes.

Les textes à verser au portfolio sont dans ce cadre des écrits réflexifs de retour sur des pratiques ou sur sa trajectoire intellectuelle, de conceptualisation, d'auto-évaluation des compétences, etc. Chaque type de texte demandé est balisé par des critères utiles à sa rédaction, qui sont en même temps des critères pour l'(auto)évaluation. Ils se déclinent et se hiérarchisent avec des priorités différentes selon qu'ils servent l'évaluation formative ou l'évaluation certificative. Parmi ces critères, selon les consignes d'écriture proposées, on retrouve par exemple, l'appropriation du genre discursif spécifique, la pertinence des analyses et des exemples choisis, la mise en évidence des liaisons effectuées entre la théorie et la pratique, le positionnement dans une énonciation personnelle, etc. Ces critères soutiennent avant tout une régulation dialogique et interactive des apprentissages, à travers des réponses personnalisées, écrites ou orales, des formateurs aux portfolios.

Tous les textes ne sont pas obligatoirement montrés par l'étudiant (par exemple des descriptions très intimes d'expériences difficiles) : la consigne est de réaliser son portfolio personnel à partir d'une série de propositions d'écriture possibles, d'une part, et de ne donner à voir au formateur qu'un portfolio « officiel » qui contient des textes obligatoires bien définis (des bilans d'auto-évaluations, des comptes rendus de stages, des analyses systématiques de pratiques, un récit final de formation). On l'a dit plus haut, ces textes, comme les portfolios, sont multi fonctionnels : en termes d'apprentissage proprement dit, ils servent de points d'appui à des échanges au sein de cours et de séminaires, entre pairs dans des travaux en groupes et en séances plénières avec les formateurs ; ils servent à l'étudiant pour lui-même ; ils servent au dialogue régulateur entre formateurs et formés.

Voyons à présent plus en détail en quoi consiste ce processus d'évaluation et de régulation que le portfolio soutient.

3 LE PORTFOLIO, UN OUTIL DE MÉDIATION DES ÉVALUATIONS CONJOINTES

3.1 Évaluations conjointes

La littérature s'accorde à souligner l'effet positif des processus d'auto-évaluation dans la régulation des apprentissages³. Élaborer un portfolio offre précisément des occasions d'auto-évaluation au sens strict (e.g., instruments d'auto-évaluation insérés dans le portfolio, écrits réflexifs), mais également d'évaluations conjointes par des pratiques :

- d'évaluation mutuelle entre pairs (dite « inter-évaluation » dans certains écrits) : présentation mutuelle du portfolio, élaboration concertée de certains éléments, de projets de régulation ;

3. Voir, par exemple, Allal et Mottier Lopez (2005) pour une revue des écrits francophones sur l'évaluation formative, Black et William (1998) pour une méta-analyse des écrits en langue anglaise.

- de co-évaluation qui désigne une auto-évaluation confrontée à l'évaluation d'un expert ou de plusieurs experts. En cours d'élaboration ou une fois constitué, le portfolio sert de base empirique à des échanges de co-évaluation, y compris dans des entretiens tripartites réunissant, par exemple, un étudiant, un formateur de terrain et un formateur académique ;
- d'évaluations guidées dans des interactions collectives (échanges entre un/des formateur/s et le groupe de formation)⁴, vues comme spécialement propices à la co-construction de représentations communes, de significations, de normes relatives à la pratique du portfolio (Mottier Lopez, 2006).

L'auto-évaluation est intentionnellement confrontée à des interprétations et des appréciations d'autrui, manifestées dans des relations sociales symétriques et/ou asymétriques. Ces confrontations encouragent l'émergence de régulation par la fonction critique de la réflexion imbriquée dans les processus d'auto-évaluation (Jorro, 2004). Ailleurs, nous avons distingué quatre dimensions inter reliées pouvant relever d'une évaluation-négociée au sein d'une réflexion partagée avec autrui (Mottier Lopez, à paraître) : (1)*constater, communiquer, décrire une situation vécue*, une compétence professionnelle en construction, une connaissance, un comportement, une progression, un ressenti ; (2)*exprimer et confronter des positionnements et des valeurs* sur les objets dégagés, comprenant l'expression plus ou moins explicite d'une mise en tension entre les valeurs (inter)personnelles et celles qui sont perçues dans les différents contextes de formation expérimentés ; (3)*s'auto- et s'inter-questionner par le moyen d'une réflexion critique et par une problématisation* des objets soumis à réflexion et évaluation conjointe ; (4)*dans le but ultime de se réguler* ; une régulation qui se manifeste dans le « dire » de ce qui a déjà été effectué ou de ce qui est projeté. Le modèle de « l'auto-évaluation explicite » de Campanale (1997) argumente une *confrontation tripolaire* entre le référentiel interne d'action du sujet, les référents externes et les objets évalués. Les référents externes s'expriment par les interprétations des pairs et des experts, ainsi que par des éléments matériels, symboliques, représentationnels présents/construits dans les situations d'évaluation (consignes, grilles de critères, ouvrages de référence, modèles théoriques). Notre conception est que les évaluations conjointes, notamment celles qui impliquent des interventions d'un expert, engagent des processus de référentialisation au sens de Figari (2002), c'est-à-dire la construction et la négociation de référents communs susceptibles de changements et d'évo-

4. La typologie « auto-évaluation, évaluation mutuelle, co-évaluation » est présente dans la littérature francophone (Allal, 1999) et anglophone (Dochy, Segers et Sluysmans, 1999). Nous ajoutons, pour notre part, la catégorie d'évaluation guidée dans des interactions collectives. L'élaboration d'un portfolio implique souvent la combinaison de ces différentes formes d'évaluation.

lutions au fil des expériences confrontées et subjectivées. Mais rappelons que toute auto-évaluation aboutie, quelle que soit sa modalité, débouche, *in fine*, sur une autorégulation de l'activité individuelle (Allal, 1999).

3.2 Médiation du portfolio

Plusieurs auteurs mettent en avant que l'implication des apprenants dans l'évaluation peut être perçue comme une « prise de risque » non négligeable par certains étudiants. Paquay, Darras et Saussez (2001) explicitent les injonctions paradoxales qui sous-tendent toute forme d'auto-évaluation dans une formation certifiée : le fait, par exemple, d'inciter les étudiants à énoncer leurs forces et leurs faiblesses, à exprimer leur interprétation et appréciation alors que, finalement, c'est toujours le formateur qui décide de la certification. Afin de déjouer ces effets pervers, les auteurs proposent différentes stratégies dont celle d'*instrumenter* les démarches d'auto-évaluation. Dans le prolongement de cette idée, notre hypothèse est que l'outil portfolio peut agir sur les paradoxes situationnels en tant que système d'attentes et d'activités négociées qui prépare et participe pleinement au « contrat » de la co-évaluation, en amont et au cours des échanges asymétriques. Les traces issues des différents contextes de formation, sélectionnées pour figurer dans le portfolio, peuvent être mises en relation, analysées et problématisées par le moyen d'une écriture réflexive, par des interactions structurées entre pairs et le formateur, par des interactions collectives. Les outils d'auto-évaluation et les textes réflexifs encouragent un questionnement explicite des réussites, des difficultés plus ou moins surmontées, des régulations mises en œuvre ou projetées. C'est l'occasion de discuter collectivement ou de façon plus individualisée du droit à l'erreur et du droit à l'apprentissage, mais aussi du devoir de la régulation dans la conception d'un développement qui se déroule tout au long d'une vie professionnelle. Ce faisant, le portfolio et sa démarche d'élaboration offrent un étayage aux échanges de co-évaluation.

La recherche de Mottier Lopez et Tessaro (2006) a étudié les *média-*tions liées à l'usage d'un portfolio dans des échanges interpersonnels de co-évaluation entre un formateur académique et un étudiant. Ces médiations, produites par le portfolio qui instrumente les échanges, ont été analysées à partir des discours élaborés par les participants sur l'évaluation de la progression professionnelle au fil de stages pratiques réalisés dans la dernière année d'une formation initiale à l'enseignement. L'hypothèse soutenue par les chercheurs est que l'outil portfolio peut rendre possible l'établissement d'intercompréhensions, d'ajustements, de nouvelles interprétations et évaluations entre les acteurs qui se l'approprient. Le portfolio peut devenir un moyen de l'action pour les participants (Béguin et Rabardel, 2000).

Sans entrer dans le détail des résultats de cette recherche, la section suivante expose l'exemple d'un entretien : dans quelle mesure les échanges de co-évaluation ont-ils été propices à l'émergence de questionnements et

d'interprétations nouvelles concernant la progression initialement mise en évidence dans le portfolio ?

4 L'EXEMPLE D'UN ENTRETIEN DE CO-ÉVALUATION, CELUI DE VÉRA

L'entretien a eu lieu dans le cadre d'un « séminaire d'intégration » qui fait partie d'un programme universitaire de formation à l'enseignement primaire. L'évaluation certificative du séminaire portait sur l'élaboration d'un portfolio individuel associé à un entretien de co-évaluation. Le portfolio, en tant que « produit final », a été évalué avant l'entretien, sur la base de critères explicités aux étudiants. Il a été annoncé que l'entretien ne pouvait pas entraîner une baisse de l'appréciation attribuée au portfolio, mais qu'il pouvait la relever. La consigne pour préparer l'entretien était la suivante : *relevez ce que vous pensez avoir construit concernant vos compétences professionnelles au terme de votre formation initiale et ce qu'il est souhaitable d'anticiper dans votre future formation continue*. Dès les premiers échanges, le formateur a communiqué son évaluation du portfolio. Puis il a poursuivi par la demande rituelle : *je vous laisse démarrer avec une ou deux choses que vous avez relevées avec ce portfolio*.

Il y a des points dont je n'ai pas vraiment parlés dans le portfolio mais que je trouvais importants de quand même relever par rapport à ma progression dans l'enseignement, justifie Véra d'entrée de jeu. Ainsi, du point de vue de l'étudiante, l'entretien est une situation qui lui permettra d'aller au-delà des aspects formulés dans son portfolio. Parmi les différents objets de discussion introduits par Véra, nous rapporterons ici les échanges portant sur « la planification de l'enseignement ». Le tableau 9.1 compare les contenus du portfolio de l'étudiante et les échanges de co-évaluation traitant de cette question.

Tableau 9.1. Contenus du portfolio et échanges de co-évaluation

Évaluer sa progression concernant la planification de son enseignement	
Contenus du portfolio de Véra	Entretien de co-évaluation
1. Une grille d'auto-évaluation en début d'année 2. Une grille d'auto-évaluation de la progression au fil des stages 3. Un document de terrain : une planification effectuée pendant un stage, accompagné d'un commentaire (figure 9.1) 4. Un texte réflexif écrit en fin de formation (extrait 9.1)	– Véra introduit l'idée d'anticipation. Le formateur relance sur la distinction entre « anticipation » et « planification » notée dans le portfolio de Véra (extrait 9.2) – 72 tours de parole plus tard, Vera revient sur le thème de la « planification » mise en tension avec la question de la « souplesse » (extrait 9.3)

4.1 Contenus du portfolio

Dans le portfolio de Véra, quatre documents traitent de la planification de l'enseignement : deux grilles d'auto-évaluation, dont une dans laquelle Véra indique qu'elle se sent à l'aise concernant la « planification » mais que l'« anticipation » doit être améliorée. Aucune indication n'est fournie concernant la distinction que l'étudiante effectue entre ces deux notions. Un troisième document est une planification effectuée lors d'un stage (figure 9.1). Un commentaire est agrafé au document, justifiant notamment sa présence dans le portfolio.

Semaine du 8 au 12 novembre

→ devoirs préparés

	Lundi	Mardi Thème = Vies	Jeudi Entraide + respect des autres	Vendredi Respect mutuel
8.00-8.30	1/2 groupe TM	Proverbe	Phrase avec groupes	Sketch S. 11 + L.
8.30-9.45	géométrie	Calcul mental	Opérations	Piscine
10.00-11.30	géométrie	Allemand Projet écrit	Mathématiques Géométrie leçon 2	Correction devoirs → 1 collect New devoirs Dictée Sketch
		PDT Biplan Expo Ad. Ad. Ad.	Géographie Explication pr. Cyr. Hyp-Hop.	Sketch Dessin Dessin Dessin
13.30-14.15	Allemand Tambourin	Dessin explication, dias, projet sur papier	Gym	Sketch Dessin Dessin
14.15-15.00	PDT	Dessin explication, dias, projet sur papier		Dessin Dessin
15.15-16.00	Conseil de Classe	Gym		Dessin

Reprendre
cette di. 9. 5. 11. 12

Commentaire agrafé au document de terrain :

Ce document présente la planification d'une semaine lors d'un stage de 4 semaines. Elle rend compte des disciplines qui ont été enseignées ainsi que la complexité rencontrée chaque jour. Les diverses annotations faites à la main présentent les imprévus. Les choses à ne pas oublier etc. Ce document est très révélateur de ce qui peut se passer dans une semaine au sein d'une classe de 6P. J'ai choisi de le présenter dans le portfolio car la planification est une compétence que j'explique dans mon dernier texte.

FIGURE 9.1. Une planification effectuée pendant un stage

Enfin, un texte réflexif, écrit en fin de formation, aborde la question de la planification, comme annoncé par Véra dans le document précédent. Le thème de ce texte, imposé par le formateur, était le suivant : *Sur la base des documents précédents [du portfolio]... est-ce que je peux être considérée comme une enseignante, certes débutante, de l'école primaire?*

Extrait 9.1 : Texte réflexif de fin de formation

Afin de répondre à cette question, je vais m'appuyer sur certaines compétences qui me semblent essentielles dans le métier d'enseignant et que j'ai pu acquérir tout au long de ma formation à travers la pratique sur les terrains et la théorie à l'université [...] En ce qui concerne la planification, celle-ci était pour moi quelque chose de difficile à élaborer sur un long terme. Puis à travers l'évolution des stages, j'ai pu découvrir différentes façons de s'organiser dans une planification [...] De plus la « consolidation différenciée » (unité de formation) sur ce sujet m'a permis de développer cette compétence sur une longue période. [...] Aujourd'hui je pense que la planification n'est plus pour moi un obstacle dans mon enseignement [...] Cependant, je pense que, lors de ma première année d'enseignement, la collaboration avec mes futures collègues sera nécessaire pour consolider la planification des différents objectifs. (Texte de Véra)

On note que Véra souligne tout spécialement les apports des différents contextes expérientiels de formation qu'elle a rencontrés, une caractéristique de son discours écrit et oral comparativement à d'autres étudiants (Mottier Lopez, à paraître). Elle se projette également dans ses pratiques professionnelles futures dans une idée de consolidation de ses compétences. Son auto-évaluation reste relativement succincte, essentiellement énoncée en termes de difficulté et d'obstacle surmontés. Globalement, il ressort que le thème de la planification est présent dans le portfolio de Véra, mais sans développements particuliers, tant aux plans théorique que pratique. Véra en a conscience puisqu'elle argumente, en début d'entretien, que c'est la raison qui la motive à aborder ce point. Mais on constate aussi qu'elle choisit un objet dont elle sait que du « matériau » existe dans son portfolio.

4.2 Échanges de co-évaluation

Sans exposer évidemment l'ensemble des échanges portant sur le thème de la planification, il est intéressant de rapporter quelques extraits significatifs de l'articulation entre les apports du portfolio et de l'entretien. Comme indiqué dans le tableau 1, deux épisodes dans l'entretien sont concernés. L'extrait 2 illustre le premier épisode qui est initié par Véra lorsqu'elle énonce la valeur

qu'elle attribue au fait d'anticiper les activités didactiques – tour de parole 1 (tp1):

Extrait 9.2. Épisode 1

1. V L'anticipation, j'ai vraiment constaté pendant ces trois stages que c'était quelque chose de **nécessaire** pour chaque activité, même si c'est une petite activité, d'avoir préparé l'exercice avant. D'avoir prévu les difficultés, etc.
2. F Alors vous distinguez planification et anticipation?
3. V Ouais (légers rires).
4. F C'est quoi la différence?

V = Véra, F = formateur, gras = appuyé dans le discours.

Le formateur relance sur la distinction entre « planification » et « anticipation » (tp2, tp4) sans que cette distinction n'apparaisse explicitement dans le discours de Véra. On peut faire l'hypothèse que c'est sa connaissance du portfolio qui l'amène à formuler cette demande. Véra est incitée à expliciter la différence qu'elle conçoit entre ces deux notions, dans un discours qui ne paraît pas convenu d'avance :

Extrait 9.2. (suite)

5. V Pour moi planification c'est euh... c'est s'organiser dans ouais pour atteindre les objectifs etc. Et anticipation c'est vraiment euh... prendre le temps quand on prépare nos activités, c'est vraiment prendre le temps de faire l'exercice en tant qu'élève, de se mettre à la place de l'élève et puis de prévoir les difficultés [...] pour moi planification, moi je l'ai vu aussi avec la CD qu'on a eu, c'est plus euh... ouais c'est s'organiser dans l'enseignement.
6. F (acquiescements pendant le discours de Véra) Est-ce que vous vous êtes sentie plus à l'aise d'abord en planification de l'enseignement? et puis qu'après l'anticipation a pu être possible?
7. V Oui [...] moi j'ai assez besoin de voir d'abord comment faire pour après approfondir c'est un peu l'entonnoir je trouve la planification et puis l'anticipation [...].
8. F Je vous pose la question parce que dans l'auto-évaluation que vous avez remplie ici, il y a quelque chose de cet ordre-là.

... = interruption dans le discours, [...] paroles non rapportées dans l'extrait

On note à nouveau la médiation offerte par le portfolio : le formateur s'appuie sur sa connaissance de ce dernier pour soutenir à la fois l'auto-évaluation de Véra et sa propre interprétation de la progression de l'étudiante (tp6). Dans le tour de parole⁸, il corrobore les propos de Véra en se référant au document² du portfolio. Après quelques échanges, le formateur formule sa propre appréciation : *visiblement vous êtes déjà plus à l'aise [...] donc c'est qu'il y a une progression*. Véra confirme et, dans le prolongement, précise encore sa pensée concernant une planification qu'elle voit plutôt à long terme et une anticipation perçue, quant à elle, à court terme. La conception de l'étudiante paraît s'être affinée au cours des échanges interpersonnels.

Le thème aurait pu se clore. Véra choisit néanmoins d'y revenir 72 tours de parole plus tard :

Extrait 9.3. Épisode 2

1. V Je peux peut-être juste éclairer quand même parce que j'ai mis... (cherche dans son portfolio et montre le document 3 de terrain) la planification pendant la 6P. C'est plein d'annotations [...] c'était juste en fait pour refaire la remarque que chaque semaine il y avait la planification donc que j'établissais avec l'enseignante, mais que pour finir tout changeait de jour en jour. Parce que chaque fois, il y avait des nouveaux événements [...] enfin ce tableau il reflète bien tout ce qui peut se passer pendant une semaine dans une classe quoi. Donc il faut être aussi souple (donne l'exemple d'un événement précis)

Véra justifie de façon plus étoffée que dans son portfolio la présence du document de terrain. Mais surtout, elle nuance les propos tenus dans l'épisode 1 en affirmant la valeur qu'elle accorde aussi à l'idée d'être «souple» dans la gestion des activités didactiques. Le formateur relance immédiatement par une sollicitation d'auto-évaluation : *et puis alors comment vous vous situez par rapport à cette nécessité d'être souple?*

Extrait 9.3. (suite)

3. V Ce n'est pas quelque chose qui me déstabilise. C'est vrai que j'ai besoin encore de faire mon planning. Mais s'il y a tout un coup des imprévus [décrit une situation vécue] je ne suis plus complètement déstabilisée.
4. F À un moment donné on doit donc à la fois avoir cette souplesse mais à la fois garder à l'esprit qu'il faut passer suffisamment de temps sur la tâche [...].

5. V Mais je pense que des fois il faut savoir dire stop aussi par rapport à cette souplesse. [...] Ouais être souple mais des fois il faut faire attention justement à ne pas se laisser dépasser.
6. F C'est une question qui n'est pas évidente. Il faut faire attention à ça. Je suis d'accord.

La question difficile de l'équilibre à trouver entre souplesse et planification est abordée pour la première fois. Des positionnements sont exprimés, y compris de la part du formateur. Plus généralement, on note que les échanges de co-évaluation sont sous-tendus par l'expression de valeurs et de significations réinterrogées, offrant des occasions de nouvelles appréciations interprétatives, notamment concernant la progression mise en évidence dans le portfolio.

5 ÉLÉMENTS DE CONCLUSION

Le portfolio, orienté vers le développement de compétences complexes et faisant intervenir des formes de pensées multiples, a été conçu dans ce chapitre comme un outil authentique d'apprentissages professionnels et d'évaluation qui parvient à articuler les processus régulateurs de l'évaluation formative aux exigences de l'évaluation certificative. Nous avons souligné la double dimension du portfolio en tant que « processus » qui témoigne de la trajectoire des formés dans des dispositifs d'alternance et en tant que « produit » qui, par exemple, peut servir de base empirique à des échanges de co-évaluation. Les formes spécifiques d'évaluation associées à ces deux dimensions ont été esquissées : par le moyen de critères de production et d'évaluation des écrits réflexifs, par la distinction entre les textes personnels et ceux donnés à voir au formateur, par les régulations interactives et dialogiques fondées sur le portfolio en cours et en fin de formation par exemple.

Associer, comme nous l'avons fait, un entretien de co-évaluation au portfolio est une façon, parmi d'autres, de souligner le rôle important d'étayage du formateur ainsi que des phases d'intersubjectivité à des fins de validation des acquis, de remise en question, de formalisation, d'amélioration et surtout, pensons-nous, de construction de références et de significations partagées inscrivant les formés dans des contextes sociaux et professionnels. L'exemple de l'entretien de Véra, effectué en fin de formation initiale, est intéressant sur plusieurs points : il montre que les échanges ont permis d'aller au-delà de l'évaluation certificative du portfolio « produit », tout en soulignant les ressources offertes par ce même portfolio aux échanges interpersonnels de co-évaluation. Ceux-ci ont engagé de nouvelles interprétations et négociations de significations favorables à une projection dans les *pratiques professionnelles futures* de Véra. Des échanges *entre professionnels*, certes de

niveau d'expertise différent, ont vu le jour. Dans la littérature, l'apport formatif des évaluations conjointes est tout spécialement mis en avant en cours d'élaboration du portfolio à des fins d'ajustement des éléments qui le constituent. Notre exemple montre que des processus régulateurs peuvent également se développer dans la transition entre la fin de formation initiale et l'entrée dans le métier : portfolios et entretiens de co-évaluation font figure de leviers de la professionnalisation des jeunes enseignants dans la conception d'un développement professionnel continu. Une condition essentielle pour que l'outil s'investisse réellement de cette fonction de levier est l'engagement et la prise de risque que l'étudiant accepte de prendre et, en contrepartie, l'accompagnement assuré par le formateur qui se place lui-même dans une démarche de co-construction de significations.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Allal, L. (1999). Impliquer l'apprenant dans les processus d'évaluation : promesses et pièges de l'auto-évaluation. In C. Depover et B. Noël (éds), *L'évaluation des compétences et des processus cognitifs, modèles, pratiques et contextes*. Bruxelles, De Boeck, pp. 35-56.
- Allal, L. et Mottier Lopez, L. (2005). Formative assessment of learning : A review of publications in French. In *Formative Assessment. Improving Learning in Secondary Classrooms*. Paris, OECD-CERI Publication (What works in innovation in education), pp. 241-264.
- Béguin, P. et Rabardel, P. (2000). Designing for instrument mediated activity. *Scandinavian Journal of Information Systems*, 12, 173-190.
- Bélair, L. (2005). *Le portfolio en formation professionnelle des enseignants : parcours d'obstacles et analyse critique*. Actes du 18^e colloque international de l'ADMEE-Europe, Reims.
- Black, P. et William, D. (1998). Assessment and Classroom Learning. *Assessment in Education*, 5 (1), 7-74.
- Campanale, F. (1997). Auto-évaluation et transformation de pratiques pédagogiques. *Mesure et évaluation en éducation*, 20, 1-24.
- Dochy, F., Segers, M. et Sluijsmans, D. (1999). The use of self-, peer- and co-assessment in higher education : A review. *Studies in Higher Education*, 24, 331-350.
- Deum, M., Mattar, C. et Vanhulle, S. (2005). *Le portfolio en formation initiale : la professionnalisation au croisement de la théorie et de la pratique*. Actes du 18^e colloque international de l'ADMEE-Europe, Reims.

- Figari, G. (2001). Us et abus de la notion de référentiel. In G. Figari et M. Achouche (éds), *L'activité évaluative réinterrogée, regards scolaires et socioprofessionnels*. Bruxelles, De Boeck, pp. 310-314.
- Jorro, A. (2001). L'évaluation comme ouverture de soi. In G. Figari et M. Achouche (éds), *L'activité évaluative réinterrogée, regards scolaires et socioprofessionnels*. Bruxelles, De Boeck, pp. 83-89.
- Jorro, A. (2004). Réflexivité et auto-évaluation dans les pratiques enseignantes. *Mesure et évaluation en éducation*, 27, 33-47.
- Mottier Lopez, L. (2006). Interroger la pratique du portfolio en situation scolaire dans une perspective «située» de l'apprentissage. *Mesure et évaluation en éducation*, 29 (2), 1-21.
- Mottier Lopez, L. (à paraître). Entretiens de co-évaluation instrumentés par un portfolio : quelle zone de médiation entre les contextes de la formation théorique et pratique ? In A. Jorro (éd.), *Évaluation et développement professionnel*. Paris, L'Harmattan.
- Mottier Lopez, L. et Tessaro, W. (2006). *Entretiens de co-évaluation instrumentés par un portfolio : pour quelle réinterprétation de la progression des compétences professionnelles ?* Actes du 19^e colloque international de l'ADMEE-Europe, Luxembourg.
- Noël, B. (2005). *Le portfolio comme dossier d'apprentissage dans la formation des enseignants du secondaire*. Actes du 18^e colloque international de l'ADMEE-Europe, Reims.
- Paquay, L., Darras, E. et Saussez, F. (2001). Les représentations de l'auto-évaluation. In G. Figari et M. Achouche (éds), *L'activité évaluative réinterrogée, regards scolaires et socioprofessionnels*. Bruxelles, De Boeck, pp. 119-133.
- Runtz-Christan, E. et Rouiller, J. (2005). *Le portfolio de compétences professionnelles : moyen d'évaluation d'un cours-séminaire universitaire de formation initiale en enseignement et de ses impacts sur les conceptions et les pratiques professionnelles des apprenants*. Actes du 18^e colloque international de l'ADMEE-Europe, Reims.
- Vanhulle, S. (2005). L'écriture réflexive, une inlassable transformation sociale de soi. *Repères, revue de didactique du français*, 30, 13-31.
- Vanhulle, S. (à paraître). Côté cour, les compétences, côté jardin, la conscience ; et au milieu, l'activité formatrice. In Y. Lenoir et P. Pastré (éds), *Didactique professionnelle, didactique des savoirs professionnels, didactiques des disciplines : quelles relations pour une formation à l'enseignement ?* Toulouse, Octarès.

- Vanhulle, S. et Schillings, A. (2004). *Former des enseignants compétents en lecture-écriture: interactions sociales, écriture réflexive et portfolio*. Bruxelles, Labor.
- Vanhulle, S. Merhan, F. et Ronveaux, C. (à paraître). Texte de cadrage: Formes et enjeux de l'alternance dans les formations des enseignants et des formateurs d'adultes. *Raisons éducatives*, 11. Bruxelles, De Boeck.

Chapitre 10

Pertinence et validité de tâches complexes dans l'évaluation de futurs enseignants

Pierre-François Coen, Aline Galland,
Isabelle Monnard, Nicolas Perrin, Yviane Rouiller

INTRODUCTION :

CURRICULUM EN TERMES DE COMPÉTENCES ET ÉVALUATION

Depuis quelques années, les formations professionnelles initiales des enseignants en Suisse ont l'obligation de construire leurs curricula en termes de compétences. Cette nouvelle ingénierie remplace une structuration de la formation élaborée principalement autour de connaissances théoriques complétées par une formation pratique souvent juxtaposée. Le changement opéré a pour but de garantir, dans un contexte où la formation interroge la tradition du compagnonnage, une approche plus réflexive du métier et de la construction de compétences. Dans une perspective de professionnalisation, il s'agit de garantir que les futurs enseignants puissent agir de manière compétente en développant aussi bien des routines pour faire face à des situations habituelles que des moyens pour problématiser les situations professionnelles ou pour faire face à la diversité des contextes et à la complexification des situations d'enseignement.

Une formation construite en termes de compétences interroge de fait les modes d'évaluation. Il ne s'agit plus d'évaluer, d'une part des connaissances théoriques et, d'autre part, des actes professionnels courants. L'ambition est de pouvoir évaluer durant la formation initiale de véritables compétences au travers de tâches complexes – intégratives – au sens où l'entend Roegiers (2004). Il faut donc vérifier que les futurs enseignants savent mobiliser les ressources acquises durant la formation pour faire face aux situations complexes du terrain. Or, dans ce contexte de formation, il n'est guère possible de proposer des tâches complexes en situation naturelle, dans les classes par exemple. Il serait difficile de garantir la représentativité de ces situations et le dispositif poserait des problèmes logistiques importants pour que tous les étudiants puissent accéder à de telles tâches ; sans compter le risque de perturbation que cela pourrait induire sur l'écologie de la classe. Si la démarche adoptée consiste donc à proposer des tâches complexes, généralement hors de toute interaction pédagogique avec des élèves, la question se pose de savoir si les situations proposées ont un lien de parenté suffisant avec des situations réelles et si elles sont assez explicites quant à la nature des compétences qui sont évaluées. En d'autres termes, la pertinence d'une telle évaluation est-elle garantie ? Ou dit encore autrement, une évaluation de cette nature permet-elle de récolter des informations qui garantissent une décision adéquate aux termes de l'évaluation ?

Ce chapitre interroge la manière d'assurer la pertinence de tâches complexes utilisées dans l'évaluation des compétences de futurs enseignants. Il présente une démarche articulée autour de cette exigence qui consiste à exposer ces tâches complexes – nous en avons retenues six élaborées dans les Hautes Écoles Pédagogiques (HEP) vaudoises et fribourgeoises – à la critique de praticiens et en regard d'un cadre théorique. Cette contribution est structurée en trois parties : premièrement, nous exposons différentes manières d'aborder la question de la pertinence des tâches complexes ; deuxièmement, nous nous appuyons sur les résultats d'une recherche pour rendre compte des problèmes que pose la pertinence des tâches complexes ; troisièmement, nous discutons cette notion de pertinence et la mettons notamment en regard de la question de la validité des tâches complexes.

1 PLUSIEURS MANIÈRES D'ENVISAGER LA PERTINENCE DES TÂCHES COMPLEXES

Poser la question de la pertinence revient à vérifier si on évalue des compétences qui sont véritablement importantes pour la pratique professionnelle. En d'autres termes, est-ce que les tâches complexes proposées évaluent les mêmes compétences que celles qui sont nécessaires pour faire face aux situations d'enseignement sur le terrain ?

Pour cerner les enjeux liés à cette première question, trois perspectives peuvent être convoquées. La première interroge la relation qui existe entre la formation initiale – qui peut être un vecteur de changement – et les pratiques d’enseignement existantes. La deuxième consiste à vérifier si les pratiques professionnelles, auxquelles se réfèrent les dispositifs de formation et d’évaluation, sont bien prises en compte par ces derniers. La troisième aborde le point de vue de l’apprenant et pose la question de la didactisation des situations d’évaluation en regard des situations rencontrées sur le terrain.

1.1 La formation professionnelle comme vecteur de changement

Dans la première perspective, il est nécessaire de comprendre le rôle que peut avoir la formation professionnelle en regard des changements qui interviennent au niveau du contexte de l’enseignement. La formation initiale se doit d’être en phase avec la réalité pédagogique. Elle doit tenir compte de la complexité des situations effectivement rencontrées sur le terrain par les professionnels. Mais elle peut également se trouver en décalage parce qu’elle propose des tâches novatrices face à des pratiques d’enseignement routinisées. Aussi, proposer des tâches complexes qui se réfèrent à des pratiques peu fréquentes à ce jour peut être considéré comme pertinent dans une certaine mesure.

Cela renvoie à la question de la norme qui permet de trancher par rapport aux critères de pertinence. Est-il judicieux de se baser sur le référentiel de l’institut de formation plutôt que celui de l’activité réelle/actuelle des praticiens, cette dernière n’étant d’ailleurs pas uniforme ? Il faut reconnaître qu’un écart peut exister entre le métier tel qu’il est pratiqué sur le terrain, c’est-à-dire au sein d’une culture professionnelle, et le métier tel qu’il pourrait être pratiqué en regard de l’évolution des savoirs pédagogiques. L’enseignant se retrouve face à une situation complexe, au sein d’un réseau de contraintes et de ressources, incluant une jurisprudence et une culture professionnelle. Les ressources et les contraintes peuvent aussi bien contribuer à rendre des solutions novatrices impossibles – alors que celles-ci auraient pu déboucher sur un agir plus efficace et plus économique pour la personne – que constituer des routines très efficaces dans un système donné. Inversement, un écart peut exister entre le travail réel de l’enseignant et les transformations préconisées des pratiques pédagogiques qui ne tiennent pas forcément compte de la complexité et de l’histoire de l’ensemble du système, qui contraint de manière inadéquate l’action de l’enseignant.

1.2 Se référer à l'activité des professionnels

Pour sortir de ce dilemme, et ceci correspond à notre deuxième regard, la construction des tâches complexes proposées dans le cadre des évaluations jalonnant la formation initiale des enseignants devrait se baser sur une analyse du travail comme le préconise l'ergonomie de langue française (Guérin, Laville, Daniellou, Duraffourg et Kerguelen, 1997). Ceci peut se faire par exemple au travers de situations emblématiques, en tenant compte aussi bien des pratiques existantes que des dysfonctionnements qui y sont consécutifs. À l'heure actuelle, les formations professionnelles proposées dans les HEP fribourgeoises et vaudoises ne sont pas systématiquement construites de cette manière.

À défaut, les tâches complexes sont souvent élaborées par les formateurs à partir de leur propre expérience. Dans un contexte légèrement différent, Collins, Brown et Duguid (1989) soulignent l'importance d'un apprentissage situé, c'est-à-dire à l'aide de tâches authentiques, pour développer des compétences qui correspondent bien à celles du métier, c'est-à-dire celles qui sont effectivement mises en œuvre en situation. On peut supposer que l'expérience des formateurs peut faciliter cela; encore faut-il vérifier si la tâche proposée est réellement pertinente.

1.3 La didactisation des situations

Roegiers (2003) va dans le même sens en parlant de «situation naturelle» et de «résolution réelle». Il oppose tout d'abord situation naturelle et cas construit. Un cas construit correspond alors à une didactisation de la situation naturelle, dans le but d'appréhender un point précis de la situation naturelle. Il oppose ensuite résolution réelle et résolution simulée. Une résolution simulée correspond alors à une tâche qui sera traitée par l'étudiant mais pas utilisée effectivement dans une situation d'évaluation.

Quatre cas de figure peuvent se réaliser et être pertinents, ils n'ont alors pas la même fonction. Le premier cas de figure (situation naturelle et résolution réelle) se présente dans les formations professionnelles, dans des situations de stage où l'apprenant est actif. Le deuxième (situation naturelle et résolution simulée) représente un cas qui pose des questions bien réelles mais où l'apprenant pour des raisons de sécurité, de gaspillage ou de responsabilités ne résoudra pas réellement les problèmes. Ce sont des situations très souvent utilisées en formation, telles qu'une étude de cas. Le troisième (cas construit et résolution réelle) renvoie à des situations qui exigent la réalisation d'un projet. Le quatrième (cas construit et résolution simulée) sont souvent écartées car elles exigent un gros travail de préparation de la part des concepteurs.

D'autres auteurs comme Lasnier (2000) mettent en évidence l'élément de signifiante des tâches proposées. Celles-ci permettent ainsi la construction de liens avec une situation authentique. Roland (2005) précise en cinq points les composantes d'une tâche en situation authentique : la mise en situation la plus authentique possible, la description de la tâche, les consignes et contraintes à respecter, le contexte d'application des résultats (impact sur l'environnement), les critères d'évaluation. Si ces précisions ne relèvent pas essentiellement de l'authenticité de la tâche, les premier et quatrième critères y sont étroitement liés.

Enfin, la question de la pertinence des tâches complexes renvoie aussi à la notion d'authenticité (*authentic assessment*) ou de « naturalité » (Scallon, 2004). Cet auteur donne quelques pistes de lecture en proposant des moyens pour déceler, par exemple, des problèmes mal définis, la nature de la production attendue, la quantité de données produites et leur pertinence.

2 LE REGARD DES ENSEIGNANTS COMME MOYEN POUR VÉRIFIER LA PERTINENCE DES TÂCHES COMPLEXES

Face aux enjeux multiples liés à la conception d'une évaluation par tâches complexes, notre intention est d'interroger la pertinence de ces tâches à travers le regard de quelques enseignants expérimentés. Il s'agit alors de considérer les tâches complexes étudiées comme des révélateurs permettant d'interroger le processus de construction et la faisabilité de leur mise en œuvre. En effet, les écarts qui pourraient être mis en lumière, interpellent aussi bien les pratiques existantes que les pratiques souhaitées dans un nouveau paradigme de formation, que ces pratiques soient celles des enseignants ou des formateurs-concepteurs-évaluateurs.

Cela nous amène à nous interroger sur le fait que les enseignants « du terrain » apprécient la pertinence des tâches complexes en fonction d'un référentiel implicite lié à leurs pratiques effectives alors que les formateurs se réfèrent à un autre référentiel intégrant des innovations souhaitables. Nous y reviendrons dans la discussion.

Dans le cadre de la recherche qui sous-tend ce chapitre, une douzaine d'enseignants expérimentés ont été sollicités pour des entretiens semi-directifs d'une trentaine de minutes. Dans le but d'obtenir des regards « croisés » sur les tâches, nous avons soumis trois tâches différentes à chaque enseignant. Cela nous a permis d'avoir pour chacune d'elle en moyenne six avis centrés sur les aspects suivants :

- pertinence des tâches proposées en lien avec l'exercice du métier (selon vous, les compétences mobilisées par la tâche sont-elles pertinentes en regard de la profession ?)
- lien avec la pratique quotidienne du métier (pensez-vous vous même réaliser cette tâche au quotidien ?)

- place de la tâche dans la formation (selon vous cette tâche doit-elle se faire en début / milieu / fin de formation?)
- inventaire des ressources (capacités et compétences) nécessaires pour réaliser la tâche.

Au terme de l'entretien, nous demandions encore aux enseignants de formuler des remarques ou des commentaires généraux concernant les trois tâches qu'ils devaient analyser. Nous leur avons finalement demandé d'identifier la tâche qui selon eux était la plus représentative de leur métier. Après une retranscription exhaustive des entretiens, une analyse catégorielle a été effectuée et un codage ouvert des textes a été mené.

2.1 Tâches complexes proposées à l'examen des enseignants expérimentés

Six tâches complexes ont été présentées aux enseignants expérimentés. Ces tâches sont destinées à évaluer les compétences à différents moments de la formation (qui dure trois ans):

- La première tâche complexe (T1, 1^{re} année) conçue pour la certification d'un cours d'introduction à la recherche en sciences de l'éducation¹ consiste à esquisser une démarche de recherche sur la base d'un article scientifique et d'une thématique proposée.
- La deuxième tâche (T2, 1^{re} année) évalue des compétences dans le domaine des médias, de l'insertion dans un système éducatif et de la communication. Par groupe de 4, les étudiants doivent réaliser un film vidéo illustrant différents aspects de la profession.
- Dans la troisième tâche (T3, 1^{re} année), les étudiants doivent analyser une séquence vidéo présentant plusieurs enfants travaillant à la résolution d'une situation en mathématiques; leurs compétences à réaliser une observation objective et à analyser des expériences pratiques sont évaluées.
- La quatrième tâche (T4, 2^e année) permet d'identifier chez les étudiants leur conception de l'école. Individuellement, ceux-ci doivent choisir un concept et le développer en se référant aux apports théoriques des diffé-

1. Examiner la pertinence d'une tâche complexe liée à une activité de recherche pourrait paraître saugrenu, tant la recherche est parfois peu présente dans l'activité quotidienne des enseignants. Toutefois, il faut préciser d'une part que le référentiel de compétence des enseignants inclut cet aspect. D'autre part, il ne s'agit pas de former des chercheurs mais de proposer une formation par la recherche; l'ambition est alors de favoriser des actions telles que problématiser une situation, mobiliser des cadres d'analyse pour mieux les maîtriser, préciser les données pertinentes à récolter, interpréter des données, bref, favoriser et instrumenter un point de vue réflexif et renforcer une articulation théorie-pratique.

- rents dispositifs de formation (cours, ateliers, stages) puis confronter leur vision avec celle d'un autre étudiant.
- La cinquième tâche (T5, 2^e et 3^e année) concerne la planification et la conduite d'un projet d'enseignement-apprentissage. Ici, ce sont les compétences dans le domaine des didactiques qui sont évaluées. Les étudiants élaborent un projet qu'ils conduisent en stage. Ils préparent ensuite un dossier présentant l'analyse de leur expérience qu'ils discutent oralement.
 - Dans la sixième tâche (T6, 3^e année), une situation problématique typique de la pratique professionnelle est proposée aux étudiants qui doivent identifier les données du problème, les analyser d'un point de vue théorique et élaborer un plan d'intervention.

2.2 Analyse de la pertinence aux yeux des enseignants

Nous allons aborder, dans les lignes qui suivent, les questions évoquées plus haut. Comme nous l'avons mentionné, les résultats présentés ici s'inscrivent dans le cadre d'une recherche menée dans les deux institutions². Nous avons jugé utile pour illustrer le propos de mentionner des extraits des entretiens conduits auprès des enseignants dans le cadre de cette recherche.

2.2.1 Pertinence en lien avec l'exercice du métier

Faut-il maîtriser les tâches pour exercer correctement le métier? D'une manière très générale, on constate que les enseignants interrogés voient un lien direct entre les tâches complexes proposées aux étudiants et l'exercice du métier. Leur point de vue est assez homogène face aux différentes tâches proposées. À propos de la T1, un enseignant résume ce point de vue: *c'est effectivement des choses qu'il faut savoir faire en tant qu'enseignant, c'est une base de notre travail*; un autre ajoute à propos de la T2: *apprendre à communiquer, c'est quand même un passage obligé pour nous*; un troisième poursuit parlant de la T4: *Je pense que ça correspond à ce qu'on demande maintenant, à ce qu'on attend de nous comme enseignant. On est de plus en plus amené à faire ça dans notre métier: pouvoir argumenter de ce qu'on fait auprès des autorités, de la commission scolaire ou des parents*; concernant la T6, un dernier enseignant conclut en des termes équivalents: *je pense que ce qui est proposé, c'est vraiment le travail de l'enseignant*.

Ces appréciations très positives sont néanmoins à nuancer car elles concernent davantage la pertinence des compétences attendues que la situation dans laquelle elles sont observées. En effet, malgré une bonne acceptation globale, les enseignants estiment que les tâches (à l'exception de la T5)

2. Pour plus de précisions: <http://www.hepfr.ch> ou <http://www.hep.vd.ch>.

ont un caractère artificiel. Ils voient certes un lien entre tâches complexes et exercice de la profession mais considèrent que la didactisation des tâches les éloigne de la réalité du métier. Parlant de la T6, un enseignant s'exprime en ces termes : *je trouve chouette d'avoir l'occasion de réfléchir sur ses pratiques sans être obligé de le faire dans le feu de l'action, bon c'est quand même assez artificiel de le faire comme ça.*

Certaines tâches ne sont pas estimées comme nécessaires pour l'exercice du métier, les enseignants y voient un plus dans la formation, rien de plus. C'est typiquement le cas de la T2 (réalisation d'un film vidéo), qui ne semble pas indispensable à réaliser pour certains enseignants mais apporte un aspect nouveau dans la formation : *c'est une tâche qui me ferait peur à moi, parce que réaliser un film, je ne suis pas sûr d'en être capable seul. Je n'ai jamais fait de film, mais je pense que c'est très intéressant, c'est un plus pour l'enseignant que de savoir faire ça.*

Enfin, les tâches proposées aux étudiants qui sont moins reconnues comme faisant partie du quotidien (T1 par exemple) questionnent les enseignants et le regard sur leurs propres pratiques. Leurs réponses mettent en évidence une nette pertinence pour la profession tout en témoignant d'un manque de concrétisation au niveau des pratiques. Ainsi, parlant de la T4, un enseignant estime *faire des allées et retours vers la théorie, quelquefois, il pense que quand on est sur le terrain on oublie trop de le faire.* Se référant à la T6, un autre estime encore *qu'il faudrait s'arrêter chaque jour pour réfléchir à ce qu'on a vécu dans la journée noter les points importants qu'on a observés.* Il affirme : *nous, on ne fait peut-être pas assez ça sur le terrain.*

2.2.2 Pertinence en lien avec les pratiques réelles

Le deuxième volet de notre analyse consiste à interroger les pratiques réelles et à voir dans quelle mesure les tâches proposées les recouvrent. Là encore, nous constatons une bonne appréciation générale et les enseignants interrogés voient dans la plupart des tâches proposées des actions qu'ils exécutent au quotidien, cependant ils nuancent leur propos et ajoutent volontiers au oui posé d'emblée un *mais* qui réduit parfois l'étendue de ces pratiques.

On constate que la majorité des enseignants reconnaissent mettre en œuvre une partie ou un aspect de la tâche : *par rapport à la question qui était de lancer une mini-recherche, c'est vrai que ça ne se fait peut-être pas de manière scientifique, mais ça se fait par des discussions.* S'agissant de la T3, un enseignant affirme : *c'est clair que personnellement je ne vais pas écrire 4-5 pages sur le déroulement, c'est-à-dire qu'on connaît ses élèves, on sait des choses, c'est quand même clair ; ces propos sont complétés de la sorte : on ne fait peut-être pas les choses de manière aussi ciblée, c'est très rare que j'aie une grille d'analyse pour analyser un*

groupe, mais c'est vrai qu'on a plutôt un regard global. D'autres apportent des nuances sur les modalités d'exercer l'activité comme cet enseignant lorsqu'il parle de la T3: nous, on ne voit pas des séquences vidéo, mais par exemple dans le cadre du décroisement, on voit des enfants de l'école infantile de première et deuxième et puis on fait des tournus. Là, on n'a pas de séquences vidéo, mais on voit les enfants en situation. Se référant à la T4, ils insisteront volontiers sur l'aspect de collaboration et d'échange qu'ils mettent en pratique et où ils se reconnaissent. Ce dernier extrait résume ce point de vue: maintenant dans l'enseignement, on est tenu de travailler en équipe et je pense que c'est indispensable de développer ce partage avec les autres où on se confronte. Les enseignants ne réalisent aussi que partiellement la T6, ou en d'autres termes: est-ce que tu prends vraiment beaucoup de temps pour analyser, ça je ne suis pas vraiment sûre parce que souvent tu dois réagir en fonction d'une situation bien précise et rapidement, donc l'analyse, tu l'as fait parfois par après mais sans doute pas de façon systématique.

Parfois, le lien avec les pratiques réelles est beaucoup plus ténu. S'agissant de la T1, les enseignants disent ne pratiquer ce genre de démarche que très occasionnellement: *c'est quelque chose d'important, mais si je regarde ma pratique, ce n'est pas des choses que j'utilise, disons régulièrement. Ce n'est pas en soi étonnant dans la mesure où les compétences visées par la cette tâche sont nouvelles et correspondent à la volonté de tertiarisation de la formation et du métier. Se référant particulièrement à des lacunes d'ordre technologique (T2), ils disent carrément ne pas exécuter cette tâche comme une enseignantes lorsqu'elle fait référence à la maîtrise des nouvelles technologies: je suis sceptique sur la vidéo, est-ce que c'est vraiment une compétence que nous nous utilisons parce que vraiment c'est une des compétences accessoires au niveau des moyens informatiques. Je trouve déjà que dans notre métier y a pratiquement personne qui maîtrise ça, à part un ou deux ou trois dans le canton qui sont des passionnés. Le montage vidéo en tout cas pas.*

Tout comme pour l'exercice du métier, certaines tâches induisent le conditionnel: *moi, j'ai le sentiment que c'est quelque chose qu'on devrait faire plus. On le fera peut-être dans le cadre d'une formation continue ou alors à la maison, mais personnellement, je ne lis pas beaucoup. Arrivée à la maison après une journée de travail, je n'ai pas spécialement envie de prendre un bouquin de pédagogie et de le lire de façon approfondie. Cette posture démontre paradoxalement leur pertinence sans que leur mise en œuvre effective s'avère nécessaire.*

Lorsqu'on interroge les enseignants sur la pertinence de ces tâches en lien avec leur pratique quotidienne, certains d'entre eux font référence au contexte éducatif ou social d'aujourd'hui. Leur regard englobe dès lors les exigences de la profession qui à leurs yeux se développent encore toujours

plus. C'est particulièrement visible dans la T2 où l'on voit bien pour les enseignants en place l'enjeu important lié à l'usage et la maîtrise des nouvelles technologies. Cet enseignant ne le cache pas : *les jeunes d'aujourd'hui apprennent tellement vite sur les ordinateurs qu'on doit nous aussi maîtriser ces choses ; alors qu'au autre ajoute : je pense qu'à l'heure actuelle, quand on voit les évolutions technologiques, c'est indispensable de le faire.* C'est également visible pour la T3 qui suppose la nécessité de se documenter sur certains thèmes actuels et d'être capable d'argumenter son point de vue vis-à-vis des parents ou du monde extérieur : *il y a 25 ans, quand j'enseignais les enfants étaient presque formatés [...] aujourd'hui c'est plus pareil, ça fonctionne différemment, il faut expliquer, argumenter.* Se référant à la tâche 6, un collègue va dans le même sens : *de plus en plus, au vu des exigences de la Direction de l'instruction publique, il faut collaborer entre nous ou avec les services auxiliaires par exemple dans le cadre de projets d'école ou dans les conseils d'arrondissement.* Les enseignants perçoivent ainsi à travers les tâches proposées les nouveaux enjeux (souvent éducatifs) du métier et de l'école.

En résumé, d'une manière générale, on peut dire que les pratiques des enseignants sont relativement proches des tâches proposées bien que dans la majorité de nos interviews, les sujets interrogés disent n'effectuer qu'une partie ou qu'un aspect des tâches. Cela s'avère particulièrement clair pour la T1 (mini-recherche) dont la pertinence n'est pas remise en question, mais dont l'application n'est que très partielle. Cela se démontre également pour la T2 (faire un film) même si la pertinence (au niveau de la maîtrise des technologies) apparaît comme moins évidente. Dans d'autres cas, les enseignants considèrent que même s'ils n'effectuent pas certaines activités, ils souhaiteraient – ou devraient – le faire en particulier pour s'adapter au contexte éducatif actuel en évolution constante. Finalement, il semble que certaines tâches font partie de routines établies et qu'il ne soit donc plus nécessaire pour les enseignants de formaliser leurs pratiques ; le passage par l'écrit la verbalisation des analyses de leur enseignement appartiennent à cette catégorie.

2.2.3 La pertinence en lien avec le moment choisi pour cette tâche dans la formation

La formation dispensée à la HEP de Fribourg et Vaud est de trois ans (niveau BA européen) et les tâches complexes proposées aux étudiants sont réparties durant tout le cursus. Nous avons interrogé les enseignants de terrain pour connaître leur avis sur le moment choisi dans la formation pour présenter ces tâches. D'une manière générale, les enseignants s'accordent pour fixer ces tâches durant la formation initiale. Ils voient dans leur réalisation des opportunités de faire des choses qui plus tard ne se feront plus vraiment de façon aussi approfondie ou détaillée (la routine aidant, les choses se font

de manière plus synthétique ou automatique). Se référant à la T1 (mini-recherche) deux enseignants s'expriment ainsi : *je trouve que c'est bien de faire ça quand on n'est pas dans l'urgence du quotidien ; c'est en formation initiale que ça vaut la peine de faire ce genre de chose plutôt qu'en formation continue*. Évoquant la T3, un autre complète : *je trouve bien de faire ça en formation initiale parce que si on ne fait pas ça à ce moment-là, on ne le fait jamais*. La réalisation des tâches s'entend donc volontiers en formation initiale même si pour certains de nos sujets, il demeure un caractère artificiel d'où l'idée, pour certains d'entre eux, de prolonger ce type de réflexion lors de la formation continue. *Ça va être un peu artificiel de faire ça dans la formation initiale, donc il faudrait reprendre les choses quand il a sa classe* dit un sujet lorsqu'il parle de la T4.

Ils apportent ensuite quelques nuances sur le moment dans la formation initiale. Pour eux, certaines tâches doivent plutôt se faire en première année, d'autres plus tard dans la formation. S'agissant de la réalisation du film vidéo (T2) portant sur les représentations de la profession, un premier enseignant s'exprime en ces termes : *je crois que ça doit se faire au début de la formation, en première année, car il me semble que réfléchir sur le métier qu'on va faire c'est bien de le faire au début*. Parlant de la T3, un deuxième précise qu'il faut quand même avoir une base de connaissances théoriques et didactiques, *donc je pense que ce serait difficile de faire cette tâche en tout début de formation*. Les enseignants sont encore plus nuancés en parlant de la T6. Cette dernière propose des situations de la vie d'enseignant auxquelles il faut répondre. Certains enseignants estiment que l'expérience constitue un facteur important pour faire face à ces situations, dès lors, ils estiment qu'il est difficile de proposer des pistes pour résoudre les problèmes proposés lorsqu'on a peu d'heures de pratique. C'est le point de vue d'un enseignant qui dit : *c'est un métier où on apprend tous les jours [...] on apprend petit à petit sur le tas, par exemple les relations avec les parents [...] donc je ne vois pas comment on pourrait les préparer à ça si ce n'est en les mettant dans la pratique*.

Les avis des enseignants témoignent d'une bonne analyse et dans la plupart des cas recouvrent les choix effectués dans l'institution quant au moment prévu pour la passation des tâches complexes. Ils soulignent toutefois le problème important de l'expérience professionnelle qui ne peut évidemment pas s'acquérir en formation initiale mais qui pour eux constitue une ressource importante soit pour faire face aux problèmes quotidiens de la profession soit comme moyen de réduire le temps alloué à certaine tâche via la routinisation des tâches.

3 QUESTIONNEMENT DES NOTIONS DE PERTINENCE... ET DE VALIDITÉ

Notre propos était de clarifier comment il est possible d'évaluer la pertinence des tâches complexes proposées dans le cadre d'une évaluation par compétences. Les résultats de notre recherche montrent combien cette question de la pertinence est délicate. En conclusion de ce chapitre, il nous semble donc nécessaire de revenir sur cette notion. En effet, si notre approche à consister à interroger des enseignants expérimentés, nous devons également interroger la nature de leurs réponses. De plus, évaluer par tâches complexes nécessite également de vérifier que la compétence visée correspond bien à la compétence effectivement évaluée et d'aborder du même coup la notion de validité. Par ailleurs, il s'agit aussi de clarifier si ce sont bien des compétences qui sont évaluées au travers des tâches complexes, et pas seulement des ressources. Ces questions nous montrent l'importance de mettre en regard le critère de pertinence et celui de validité.

3.1 La notion de famille de situations

Selon Roegiers (2004), La pertinence d'une tâche tient à la pertinence des informations que l'on recueille à travers la résolution de la tâche. Dans une évaluation de compétences, cette pertinence est liée au fait que la tâche appartient à la famille de situations de la compétence. Si nous souhaitons nous assurer que ce qui est évalué permet effectivement en situation d'aboutir à un agir efficace, il faut pouvoir circonscrire la famille de situations correspondante.

Certes, le concept de famille de situations ne fait pas l'unanimité au sein de la communauté des chercheurs. Crahay (2005) estime que ce concept est peu opérationnalisé et s'interroge à son sujet. Pour Gerard (2005), l'équivalence des situations comme leur appartenance à la même famille de situations n'est pas évidente ni lors de la construction des tâches ni lors de leur validation.

Cependant, il n'est pas certain qu'une tâche complexe, telle qu'elle a été conçue, corresponde à la situation rencontrée sur le terrain, à la situation qu'on souhaitait reproduire. Le fait de construire une tâche complexe à réaliser «sur table» ou dans le cadre d'un projet réalisé en dehors d'une situation «naturelle» d'enseignement peut aboutir à la mobilisation de compétences différentes. Par exemple, l'analyse d'une situation telle que présentée dans une étude de cas nécessite des stratégies d'analyse différente de celles qui sont requises dans la pratique, cette dernière ne laissant que peu de temps à la réflexion. Cet écart entre la compétence visée et la compétence effectivement évaluée peut aussi découler du modèle de compétence en vigueur sur le terrain et dans l'institution de formation. Rogiers (2004) montre bien les

différentes manière d'appréhender la notion de compétences. Notamment, ce qui est nommé par «compétence» peut se référer à une pratique limitée ou impliquant des apprentissages importants, à une approche strictement disciplinaire ou inclure une perspective interdisciplinaire, et faire référence à une tâche contextualisée ou décontextualisée. C'est ainsi qu'un objectif général, une habileté, une attitude, un savoir-faire ou tout autre notion peut être implicitement assimilée à une compétence par les différents praticiens. Rey, Carette et Kahn (2003) distinguent quant à eux différents niveaux de compétences (élémentaires, élémentaires avec cadrage ou complexes) et démontrent par là une certaine difficulté à saisir ce concept.

De plus, la volonté de minimiser le nombre d'épreuves conduit à proposer des tâches complexes qui recouvrent plusieurs compétences. De telles situations ne sont pas toujours évidentes à trouver et peuvent également être à l'origine d'un décalage avec la compétence visée, c'est-à-dire effective sur le terrain.

3.2 La notion de ressource

Les enseignants interrogés ont éprouvé parfois des difficultés à identifier les ressources à mobiliser dans les différentes tâches. Le statut de ces dernières n'était pas vraiment clair pour eux : se trouve-t-on face à des tâches d'apprentissage ou d'évaluation ? Dans le premier cas, les enseignants mentionnaient plutôt des ressources à développer en formation, alors que dans le deuxième, ils citaient les ressources propres à leur pratique quotidienne : *[il y] avait un dilemme. Il y avait une collègue qui voyait quelles compétences l'étudiant devait développer pour faire ce travail-là alors qu'on était d'autres collègues à se dire plutôt les compétences qu'un enseignant devait avoir, quelles compétences d'enseignants on visait.*

Une comparaison pour l'ensemble des tâches montre que les enseignants mentionnent des ressources beaucoup plus nombreuses que celles annoncées : de 7 pour T5 à près d'une vingtaine pour T 6, diversité qui confirme la complexité des activités à réaliser.

Toutefois, les enseignants s'accordent assez peu sur les ressources à mobiliser, un nombre important n'étant même cité que par une seule personne. Il semble donc qu'il soit difficile de garantir que les tâches évaluent bien ce que les institutions souhaitent mesurer. Autrement dit, pour effectuer ces tâches, les enseignants utiliseraient d'autres ressources que celles sensées être mises en œuvre par les étudiants.

Par exemple dans la T2, tous s'accordent pour dire qu'elle nécessite de *comprendre et utiliser les médias et les TIC*, mais aucun n'a estimé qu'elle implique d'*assumer l'ensemble des tâches liées à sa fonction*. La nécessité de mobiliser des ressources très spécifiques paraît dominer tous les autres aspects : *dans la T2, la vidéo c'est une fin en soi.*

Le même phénomène apparaît dans T3. Ici l'observation est au centre du processus, les particularités de la situation – le travail sur les représentations des élèves – ne semblant pas nécessiter la mobilisation de ressources spécifiques.

Dans T4, ce qui paraît le plus important aux yeux des enseignants est de construire des représentations sur un thème en utilisant des sources théoriques et de les confronter avec celles d'autres personnes. Là encore, les ressources liées au contenu des situations sont peu citées. En fait, tout semble indiquer que les caractéristiques de surface des tâches soient les seules prises en compte (Crahay, 2006). Les enseignants citent la plupart des compétences visées par cette tâche, mais tous mentionnent en plus le besoin de recourir aux collègues pour résoudre de telles situations. Pourtant, cette ressource ne peut être mobilisée dans cette tâche puisqu'il est demandé aux étudiants de travailler seuls.

En résumé, si la tâche est considérée comme naturelle, les enseignants identifient les ressources qu'ils mobilisent en situation. Par contre s'ils estiment que la tâche est artificielle, ils imaginent les compétences qu'il faut mobiliser principalement de type réflexif. Les compétences/ressources mentionnées ne se réfèrent que rarement au contenu proposé (contrairement aux formateurs) puisque comme le dit un enseignants *ça change à chaque fois, même dans une situation semblable c'est pas la même chose*, donc ils ont tendance à citer les compétences en référence aux procédures à mobiliser (faire une vidéo, observer, utiliser des ressources théoriques) et non aux caractéristiques de surface (travailler à partir des représentations des élèves, assumer l'ensemble des tâches liées à sa fonction, différencier).

3.3 Pertinent... par rapport à quel référentiel ?

Les propos tenus par les enseignants du terrain démontrent également qu'ils apprécient la pertinence des tâches complexes en fonction d'un référentiel implicite lié à leurs pratiques effectives. Les résultats présentés ci-dessus montrent effectivement que les enseignants du terrain évaluent la pertinence des tâches complexes non pas en fonction du référentiel de compétences qui a fondé la construction des tâches complexes (celui de l'institution) mais en fonction de leur propre référentiel construit implicitement à partir de leurs pratiques. Dès lors se profile ici une limite à leur regard : les praticiens ne peuvent évaluer la pertinence d'une tâche que s'ils l'exécutent eux-mêmes. Le référentiel de compétences élaboré par l'institution et le référentiel implicite des praticiens peut néanmoins correspondre si la formation initiale vise à reproduire les pratiques en vigueur. Il n'en va pas de même en cas d'innovation, de changement de paradigme professionnel lorsque les pratiques désirées ne sont pas encore implantées sur le terrain.

Ceci nous amène à faire clairement la différence entre un référentiel de formation et un référentiel métier. Il est en effet nécessaire de pouvoir se référer à une situation qui constitue un passage obligé, une manière provisoire de s'y prendre, en vue d'acquérir une compétence professionnelle. Cette situation sera alors différente d'une situation qui attestera d'une professionnalité. Ceci permet notamment de distinguer des situations d'évaluation formative et d'évaluation certificative au sein d'une unité de formation ; elle permettra aussi de porter un regard sur l'ensemble des trois années de formation.

Toutefois, cette distinction entre référentiel de formation et référentiel métier ne résout pas complètement la problématique de l'innovation que nous venons d'exposer. Une pratique routinisée peut également traduire la perte de gestes professionnels exercés en formation. Ces gestes peuvent avoir été écartés progressivement sous la pression des contraintes du terrain ou des régulations effectuées au sein d'une équipe pédagogique traduisant ainsi la construction d'un genre professionnel, la négociation progressive des pratiques, la construction d'une jurisprudence au sens où l'entendent Clot et Faïta (2000). Il n'est toutefois pas souhaitable de laisser s'installer une incompréhension en ce qui concerne les pratiques de référence et la manière de s'y référer. Les différents acteurs – enseignants et formateurs³ – ont besoin de pouvoir se comprendre, de pouvoir négocier des pratiques de référence, c'est-à-dire pouvoir négocier la nature des compétences et la manière de les traduire dans des tâches complexes. C'est précisément ce que Zulauf (2006) appelle le « paradigme de la co-responsabilité » où les différents acteurs de la formation collaborent véritablement pour négocier les savoirs et les compétences à développer à la fois dans l'institution et sur le terrain.

3.4 Vers la définition de situations emblématiques ou de concepts pragmatiques

À défaut de pouvoir s'appuyer vraiment sur la notion de famille de situations, la notion de situations emblématiques semble intéressante. Certaines formations sont construites en fonction de cette référence et proposent des situations d'évaluation à partir d'une description de situations caractéristiques qui peuvent se retrouver sur le terrain.

Si cette perspective peut paraître séduisante, il nous faut bien entendre que certains enseignants questionnent fortement la notion de situations en mettant en évidence leurs aspects toujours singuliers plutôt que leurs similitudes. Faut-il y voir la traduction d'une culture professionnelle ou la conséquence d'un ancien paradigme de formation normalisant

3. Il serait également judicieux d'identifier la place que jouent les étudiants dans cette négociation...

qui ne pouvait déboucher que sur le constat de son insuffisance à caractériser des processus de formation complexe? Ou faut-il reconnaître, comme le suggère une analyse ergonomique (Durand, 2002; Rogalski, 2003) qu'enseigner équivaut à gérer un système dynamique extrêmement complexe vu le nombre des élèves (chacun ayant une dynamique propre) et vu le long temps de réponse (puisque une action d'un enseignant n'aura des effets perceptibles parfois que plusieurs semaines après l'action de l'enseignant)? Une piste théorique pourrait être de se référer à la notion de concept pragmatique, c'est-à-dire à un mélange d'indices et d'actions souhaitables à entreprendre dans certaines situations, conceptualisations qui font d'ailleurs l'objet d'une socialisation (Pastré, 1999, 2002).

RÉFÉRENCES BIBLIOGRAPHIQUES

- Clot, Y. et Faïta, D. (2000). Genre et styles en analyse du travail, concepts et méthodes. *Travailler. Revue Internationale de Psychopathologie du travail et de Psychodynamique du travail*, 4, 7-42.
- Collins, A., Brown, J. S. et Duguid, P. (1989). Cognition contextualisée et culture de l'apprendre (Situated cognition and the culture of learning), *Educational Researcher*, vol. 18 n° 1, pp. 32-42, janvier-février 1989 (traduit de l'anglais par G. Fritsch). URL : http://www.rezozero.net/articles/jsbrown_art.htm.
- Crahay, M. (2005). L'évaluation des compétences, une entreprise impossible? *Mesure et évaluation en éducation*, 28 (1), 57-78.
- Crahay, M. (2006). Dangers, incertitudes et incomplétude de la logique de la compétence en éducation. *Revue française de pédagogie*, 154, 97-110.
- Durand, M. (2002). *L'enseignement en milieu scolaire*. Paris, Presses universitaires de France.
- Gerard F.-M., (2005), Évaluer des compétences ou ne pas se tromper de cibles, *Liaisons*, 40, 7-9.
- Guérin, F., Laville, A., Daniellou, F., Duraffourg, J. et Kerguelen, A. (1997). *Comprendre le travail pour le transformer. La pratique de l'ergonomie*. Paris, ANACT.
- Lasnier, F. (2000). *Réussir la formation par compétences*. Montréal-Toronto, Guérin.
- Pastré, P. (1999). La conceptualisation dans l'action : bilan et nouvelles perspectives. *Éducation permanente*, 139, 13-35.
- Pastré, P. (2002). L'analyse du travail en didactique professionnelle. *Revue Française de Pédagogie*, 138, 9-17.

- Rey, B., Carette, V. et Kahn, S. (2003). *Les compétences à l'école. Apprentissage et évaluation*. Bruxelles, De Boeck.
- Roegiers, X. (2003). *Des situations pour intégrer les acquis scolaires*. Bruxelles, De Boeck.
- Roegiers, X. (2004). *L'école et l'évaluation. Des situations pour évaluer les compétences des élèves*. Bruxelles, De Boeck.
- Rogalski, J. (2003). Y a-t-il un pilote dans la classe? Une analyse de l'activité de l'enseignant comme gestion d'un environnement dynamique ouvert. *Recherches en Didactique des Mathématiques*, 23 (3), 343-388.
- Roland, L. (2005). Comment rendre l'évaluation orientante. Papier présenté au 4^e colloque sur l'approche orientante, Québec, Canada.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Bruxelles, De Boeck.
- Zulauf, M. (2006). Quel paradigme pour l'éducation musicale de demain? In P.-F. Coen et M. Zulauf (éds), *Entre savoirs modulés et savoir moduler: l'éducation musicale en question*. Paris, L'Harmattan, pp. 233-258.

Chapitre 11

Évaluation des compétences et changements identitaires

Christine Baeriswyl, Loïc Clavier, Marlyse Pillonel

Parce que les compétences s'inscrivent dans le social, les évaluer induit des changements chez l'évalué comme chez l'évaluateur. En fonction des situations et des contextes la pratique de l'évaluation est instituante chez l'évaluateur et constituante chez l'évalué. Nous entendons par là qu'en fonction du statut, du projet et des savoirs de l'évaluateur au sein de l'institution, évaluer va travailler son identité professionnelle que ce soit en termes d'affirmation ou en termes de déconstruction.

Ce changement identitaire à l'œuvre dans une pratique pédagogique fait l'objet de la première partie de ce texte. Il s'agit ici de questionner un dispositif d'évaluation des compétences mis en place, dans une optique de formation professionnelle en alternance. La recherche se veut exploratoire : enquête par questionnaires aux professeurs et praticiens formateurs. Elle conduit à analyser d'une part, les possibilités et les difficultés rencontrées par les formateurs (professeurs et praticiens) en ce qui concerne l'évaluation des compétences des formés et d'autre part, tente de mettre en évidence les opportunités et les difficultés identitaires vécues par les formateurs.

La récolte de données a lieu dans un contexte de Haute École spécialisée pour des étudiants en soins infirmiers, plus précisément à la Haute École de Santé de Fribourg, Suisse, où le programme de formation a démarré en octobre 2002. À cette occasion, des changements majeurs ont été apportés

dans la formation professionnelle : un programme cadre, identique pour dix sites de formation de Suisse romande, basé sur un référentiel de compétences et sur des contenus de formation.

Dans ce changement radical de la formation professionnelle, les compétences des professeurs et des praticiens-formateurs notamment dans le domaine de l'évaluation se doivent d'évoluer du fait que l'action éducative se situe dans le paradigme de la compétence.

Dans la seconde partie, nous ferons référence aux changements identitaires chez l'évalué au travers de l'évaluation des compétences dans un dispositif de formation professionnelle. Choisir un tel angle d'observation permet d'envisager l'évaluation autrement que comme un dispositif ou comme une construction instituée. L'évaluation devient un construit d'actions sociales, théâtre des influences des acteurs. Elle devient constituante (au sens de fondatrice) pour l'évalué en reliant les identités d'aspiration à l'acquisition de compétences. Ce positionnement permettra de mieux lire les stratégies des uns et des autres au sein d'une formation et d'identifier les arrangements locaux opérés par les acteurs comme autant de témoins de l'influence de l'évaluation sur les changements identitaires. Afin d'étayer notre réflexion, nous présenterons une recherche en cours à l'IUFM des Pays de Loire qui vise à préciser le rôle de la validation sur la construction de l'identité professionnelle des professeurs stagiaires.

1 ANALYSE DU VÉCU DES ACTEURS SUITE À DES CHANGEMENTS DANS LES DISPOSITIFS ET L'ÉVALUATION DE LA FORMATION PRATIQUE EN HAUTE ÉCOLE SPÉCIALISÉE

1.1 Du côté des professeurs

Avec l'introduction de praticiens formateurs pour l'accompagnement et l'évaluation de la formation pratique les professeurs ont été contraints de diminuer leur présence sur les lieux de pratique professionnelle, transférant de ce fait la responsabilité de l'évaluation des compétences professionnelles aux praticiens formateurs. Les professeurs sont obligés de chercher d'autres repères concrets sur la progression des étudiants auprès des praticiens. Seuls des échanges avec les praticiens permettent aux professeurs d'accéder aux indicateurs démontrant le niveau de maîtrise des compétences des étudiants.

Ce nouveau dispositif de formation pratique a des conséquences sur l'identité professionnelle des professeurs. Un des obstacles rencontrés est la perte de contact direct avec des situations professionnelles qui réduit les possibilités d'ancrer les cours théoriques dans une pratique professionnelle.

Une autre problématique émergeant a trait à la perte du statut d'évaluateur de la pratique professionnelle pour les professeurs, rôles pour

lesquels ils ont été amenés à développer des compétences. Des professeurs avouent vivre assez mal le fait de ne plus évaluer la pratique professionnelle, leur rôle actuel se limitant à un accompagnement et une évaluation de l'activité réflexive. Les professeurs relèvent la nécessaire complémentarité des ressources des professeurs et praticiens formateurs pour optimiser l'accompagnement de la formation pratique. L'apport spécifique des professeurs réside dans la connaissance du niveau attendu, la maîtrise des contenus théoriques, la connaissance de la démarche d'évaluation et sur les concepts sous-jacents à l'évaluation. Le regard extérieur à l'institution de soins dans ce processus d'évaluation est également un élément capital. L'apport spécifique des praticiens se trouve dans l'analyse de la situation observée, la conception professionnelle «en acte» et la connaissance du contexte de travail. En outre, le référentiel de compétences s'avère être un instrument difficile à s'approprier car le niveau d'attente est peu explicite et manque de repères concrets.

Malgré le fait que, pour les professeurs, une partie de leur activité professionnelle actuelle sur les lieux de pratique se situe dans le processus d'informations et de formation des praticiens formateurs, cette activité n'est pas identifiée comme un nouvel élément positif de leur identité.

1.2 Du côté des praticiens

Les praticiens s'investissent plus dans l'encadrement du fait de leur responsabilité accrue. Leurs nouveaux rôles liés au statut de praticien formateur, les amènent à réaliser seuls des évaluations formatives et sommatives, à clarifier et expliciter avec l'étudiant les processus d'apprentissages et à analyser leurs interventions. Dans le dispositif de formation pratique mis en place, les praticiens formateurs ont été projetés principaux responsables de l'apprentissage et de l'évaluation des étudiants en situation de pratique professionnelle. Ils bénéficient pour ce faire d'une formation post-grade de courte durée pour un tel changement identitaire.

Des difficultés relevées sont liées, d'une part, à l'utilisation du référentiel de compétences et, d'autre part à l'évaluation des compétences. Concernant le référentiel, les obstacles font référence à l'appropriation des termes utilisés dans le référentiel, et au fait que les compétences décrites ne correspondent pas toujours aux pratiques professionnelles contextuelles ou, encore, au manque de critères d'évaluation.

Les praticiens formateurs disent se sentir de plus en plus à l'aise avec le dispositif, du fait de leur formation, du réseau constitué avec d'autres praticiens formateurs et de la collaboration avec l'équipe soignante et les enseignants de l'école. Dans l'évaluation des compétences, les praticiens accordent une importance particulière à certaines attitudes de l'étudiant telles que son intégration au sein de l'équipe, son investissement dans la prise en

charge de sa formation, sa capacité à communiquer auprès de l'équipe et du patient, la pertinence de ses questions durant sa formation pratique ou encore sa capacité d'autonomie.

On peut relever une certaine ambivalence quant à la responsabilité de l'évaluation des compétences. Si le praticien formateur trouve, au sein de l'équipe soignante, une concordance liée au rôle professionnel visé et au rôle d'étudiant, cette responsabilité est volontiers assumée. Mais, parfois, le praticien se sent seul, largué pour assumer une telle responsabilité dans l'évaluation des compétences. Dans des situations limites, cette solitude peut le pousser à ne pas prendre de décision contestable pour l'étudiant, pour l'équipe soignante ou pour le site de formation.

Avec le nouveau dispositif de formation pratique, les praticiens formateurs estiment avoir modifié leurs pratiques d'évaluation : leur fonction les obligerait à développer la pratique réflexive personnelle et avec les étudiants, à déployer des méthodes d'encadrement pertinentes, à maintenir leurs connaissances professionnelles, à proposer des démarches d'évaluation plus élaborées.

1.3 Synthèse – discussion

Si les échanges entre partenaires de formation s'inscrivent officiellement dans un contexte d'alternance, il semble bien qu'ils s'insèrent dans des jeux d'acteurs en train de se constituer en réseau. Les besoins des praticiens perçus par les professeurs sont liés à l'appropriation du référentiel et du paradigme de la compétence.

Même si les praticiens formateurs ont une formation récente, leurs propos sur leur pratique de formateur démontrent une conception cohérente d'une formation par compétences. Les compétences effectives, sont le résultat de la mobilisation, de la sélection, de la coordination, de la mise en œuvre et des nombreux ajustements des ressources utiles pour le traitement des tâches dans une situation donnée ou dans une classe de situations (Jonnaert, 2002). Dans les dispositifs d'évaluation mis en place conjointement par les praticiens formateurs et les professeurs, ce sont les activités effectivement réalisées qui déterminent les compétences évaluables, car c'est uniquement en situation que l'étudiant se construit, modifie ou réfute des connaissances situées et développe des compétences situées. De ce fait, seule la situation donne les critères de réussite de la compétence ; ces critères sont issus de tous les éléments de la situation, ils sont contextualisés et ne sont pas significatifs a priori.

Un obstacle, dans le dispositif de formation, explicitement relevé par les partenaires, se situe dans l'utilisation du référentiel de compétence. Cette difficulté peut s'expliquer par le fait que ce référentiel de compétences

n'est pas un référentiel métier, rédigé suite à une analyse de l'activité des infirmières.

Dans une formation professionnelle en alternance qui vise le développement de compétences, la pratique réflexive prend une dimension clé. Car si, à travers une compétence, un sujet mobilise, sélectionne et coordonne une série de ressources (dont certaines de ses connaissances, mais aussi une série d'autres ressources qui seraient affectives, sociales et celles reliées à la situation et à ses contraintes) pour traiter efficacement une situation, encore faut-il que ce même sujet pose un regard critique sur les résultats de son activité (Perrenoud, 2002). Selon Denoyel (2005) la réflexivité, articule mise en intrigue et mise en commun; la mise en dialogue de l'expérience singulière permet aux apprenants, par la confrontation avec les formateurs de se professionnaliser. En ce sens, pour mener à bien une pratique réflexive, l'activité professionnelle observée est identifiée comme un des moyens d'articulation entre le discours et l'activité réelle.

Dans notre enquête, l'activité observée est primordiale car elle sert de support à l'auto-évaluation, procurant ainsi le substrat de base à la discussion. Elle permet également de confronter le discours à la réalité, discours pouvant poser certains problèmes tels que la cohérence avec l'agir, la prégnance du discours sur l'activité favorisant les étudiants avec de bonnes capacités langagières.

La duplication des lieux d'apprentissage détrône l'enseignement et l'enseignant de leur place centrale dans le processus de formation. Cette ouverture spatio-temporelle du processus de formation entraîne une diversification enrichissante à condition de changer les anciens rôles et d'en introduire de nouveaux. Des déverrouillages majeurs sont à opérer dans ce sens. Et les données récoltées montrent qu'il est bel et bien difficile de réaliser concrètement ces déverrouillages, même si le dispositif mis en place ne suscite pas en soi de résistances a priori. Les rôles actuels de formation et d'informations auprès des praticiens formateurs que s'attribuent les professeurs ne comblent pas encore la perte identitaire.

On peut relever ici que, contrairement aux professeurs qui vivent une perte identitaire, les praticiens formateurs, par leur nouveau statut, se trouvent enrichis de nouveaux rôles.

Les partenaires développent une posture favorable à l'alternance qui suppose un projet localement négocié, construit et partagé. Parler d'alternance nécessite de concevoir l'action éducative autour d'un réseau de partenaires co-formateurs, c'est chercher à articuler des temps et des lieux de formation différents, supposant des échanges permanents (Golhen, 2005).

2 LA VALIDATION EN IUFM : DE L'INFLUENCE DES PRATIQUES D'ÉVALUATION SUR LA CONSTRUCTION DE L'IDENTITÉ PROFESSIONNELLE DE L'ENSEIGNANT STAGIAIRE

2.1 L'évaluation en IUFM

Dans un précédent ouvrage (Clavier et Aubegny, 2004) nous avons déjà identifié la richesse et la complexité de l'évaluation des enseignants stagiaires en IUFM. Pas moins de sept rapports seront nécessaires à la commission idoine pour prononcer la validation en fin d'année. Cette profusion d'évaluations désamorce totalement, pour certains stagiaires, les enjeux de l'évaluation de leur travail. Il n'est d'ailleurs pas anodin semble-t-il que d'un point de vue sémantique on utilise souvent et d'une manière informelle le terme « rapport » alors qu'il s'agit d'une évaluation ! Il demeure dans cette démarche une visée hiérarchique (à dominante normalisante) même si les documents utilisés s'enorgueillissent du mot « évaluation », tendant ainsi à donner à penser qu'il y a un référentiel et des critères permettant une telle pratique. Sur une population de quarante stagiaires (Professeur de Lycée et Collège – PLC – et Professeur de Lycée Professionnel – PLP) étudiée sur deux années universitaires (2004 à 2006) à l'IUFM des Pays de Loire, nous avons pu remarquer que la redoutable technicité de l'évaluation mise en place nous fait perdre en efficience ce que nous y gagnons en lucidité.

Tout d'abord en centre IUFM, malgré une présentation détaillée en début d'année des modalités et enjeux de l'évaluation et plus particulièrement de la validation, les stagiaires avouent tous se sentir concernés tardivement par les procédures d'évaluation. Cette « prise de conscience » s'opère entre mi décembre pour les plus précoces et fin février pour les plus irrésolus. Cela est d'ailleurs très contingenté à la première visite sur le lieu de stage en responsabilité et à ses retentissements auprès du stagiaire. C'est à la suite de ce moment que l'activité en centre est mise en perspective par les stagiaires tant du point de vue des apprentissages à l'IUFM que du point de vue de l'évaluation propre aux activités qui y sont menées. Le corollaire de cette maturation hiératique s'incarne dans un sentiment qui revient très souvent dans les entretiens que nous avons menés avec ces stagiaires. À leur sens, la découverte de l'importance de l'activité en centre et de l'évaluation qui en résulte (un rapport écrit par le Responsable du Groupe de Référence – RGR) les amène à se sentir évalués à tout instant au travers de leur participation aux différents travaux du groupe que ce soit en tant qu'acteurs directs (présentant une séquence pédagogique) ou que participants aux débats du groupe. Ils ont l'impression que tout ce qu'ils disent peut être retenu contre eux. À ce sentiment vient s'ajouter la difficulté à comprendre les ressorts des démarches d'évaluation dont ils sont l'objet. En fonction de leur degré de compréhension ils seront très revendicatifs quant au manque de clarté et à la complexité de l'évaluation.

L'évaluation de l'activité sur le terrain est la plus observée. Toutefois, la plupart du temps, les stagiaires envisagent ces visites comme des inspections implicites pouvant mettre en porte à faux la relation élaborée avec leur tuteur. Cette crainte d'un décentrement atteint son paroxysme lorsqu'à la suite de la visite, le tuteur révisé son jugement à la lumière des conclusions du visiteur. Le stagiaire se retrouve face à une contradiction dans sa relation à son tuteur et dans une incertitude très inconfortable quant à la qualité de sa pratique. Peu de stagiaires entendus lors de ces entretiens m'ont signalé une évaluation construite à partir d'une grille de critères. Les évaluations se mènent au travers d'un dialogue entre évaluateurs et évalués sans recours explicite à un référent. La situation évoquée précédemment sans être souvent récurrente témoigne de la difficulté pour le tuteur de se poser en évaluateur. En effet, malgré un plan de formation conséquent pour l'accompagnement des Conseillers Pédagogique Tuteurs (CPT), les difficultés que nous signalons précédemment se manifestent trop souvent au goût des RGR. Ainsi, Bertrand (2004) montrait bien la difficulté du tuteur à se positionner en tant qu'évaluateur, retardant ainsi le plus longtemps possible l'échéance d'une évaluation du stagiaire. Les rencontres permettant de fouiller la pratique du stagiaire sur la base d'un dialogue sont, de loin, préférées. Cette logique est à rapprocher des réflexions d'Anne Jorro (2005) lorsqu'elle montre que dans les représentations de certains formateurs l'évaluation étant le siège de l'excellence, une dérive s'instaure vers une analyse des pratiques sans référent explicite afin de ne pas confronter le stagiaire à un profil professionnel difficilement accessible. Nous avons relevé au cours de nos entretiens de nombreuses situations de tension autour de l'évaluation du stagiaire dans son activité d'enseignant. Trois ensembles se distinguent. Un premier ensemble de stagiaires affiche de nombreuses difficultés à comprendre ce que l'on attend d'eux. Ils disent ne pas avoir mis en œuvre ce que le tuteur proposait en termes d'intervention... parce que cela ne correspondait pas à leurs représentations du métier (tant du point de vue de la pédagogie, que de la discipline). Ils ne parviennent pas à faire coïncider leurs propres modèles avec ceux attendus par l'institution.

Un second groupe de stagiaires se distingue par sa volonté de travailler à ses propres représentations. Cela se manifeste par des tensions lorsque, par exemple, le tuteur dans son évaluation manque de méthode, déstabilisant ainsi le travail de lecture de l'institution entamé par le stagiaire. Ce sont des stagiaires qui sont parfois dans une logique de reproduction de gestes professionnels et qui, pour réussir dans cette démarche, réclame une grande transparence dans l'évaluation. Enfin, le troisième ensemble de stagiaires est constitué par ceux qui très rapidement sont en confiance avec leur tuteur. Bien au-delà de la confiance, il s'agit ici d'un partage de valeurs. L'histoire du stagiaire, sa réflexion personnelle et sa capacité à communiquer lui permettent d'entrer dans une relation de compétences et de réflexivité quant à sa professionnalité avec son tuteur. On voit donc ici les conditions d'une influence sur les pratiques professionnelles des stagiaires. Cela ne

remet pas en cause la totalité des pratiques mais bien plutôt celles qui aux yeux de chaque professeur stagiaire va lui permettre d'établir un sentiment de confiance à l'autre... l'autre étant en l'état, son tuteur ou le RGR.

Le mémoire professionnel mêle deux types d'évaluation, comme j'ai pu déjà le montrer, (Clavier, 2002): une évaluation formative lors de l'accompagnement et une évaluation certificative à l'issue de la soutenance. Cependant, l'enjeu du mémoire est tel que la seconde contamine la première. De sorte que les stagiaires s'emploient (il faut ici relativiser ce qui suit en fonction de la population étudiée lors de la recherche et du degré de mise en œuvre de cette stratégie en regard de chaque stagiaire interrogé) plus ou moins à décoder la demande de l'institution au détriment d'un travail personnel guidé par la perspective d'une plus value de leurs apprentissages. La soutenance, par ailleurs, est symboliquement une mise à mort (Lecoine, 1997)... On sacrifie l'impétrant pour lui permettre de ressusciter comme pair de ses propres juges. Le passage du statut d'apprenant à celui d'impétrant est douloureux pour les stagiaires. La perspective de la soutenance est la promesse périlleuse d'une mutation quand la validation est une re-naissance.

Entre efficience et lucidité... selon nous, l'évaluation ainsi pratiquée en IUFM fait le choix de la complexité et de la technique et donc de la lucidité pour apprécier au mieux l'évolution du stagiaire. C'est bien cela qu'on attend d'un organisme préparant les stagiaires au recrutement de l'état. Si, d'un point de vue institutionnel ces choix se justifient, il n'en demeure pas moins que les acteurs sur le terrain tentent de construire ensemble le quotidien d'un métier riche et complexe. Ainsi avons-nous vu que les stagiaires donnent une place non négligeable dans leurs choix et leurs investissements à la compréhension du projet de l'institution afin de répondre dans les meilleures conditions aux enjeux de la validation. Cela influence leurs pratiques professionnelles mais aussi une transaction identitaire (Clavier, 2001) entre le statut d'apprenant et celui d'impétrant. Passage difficile de celui qui, métabolisant ses apprentissages, les convertit en compétences professionnelles. Ce passage ne peut se faire que dans la confiance, autre élément important qui ressort de notre travail de recherche; cette confiance s'établissant au sein d'arrangements locaux à propos des procédures d'évaluation.

Au cours de l'année de formation on peut assister à un glissement entre ingénierie de formation et ingénierie de compétences. Les formateurs et tuteurs passent des savoirs aux savoirs en action et à la mise en œuvre de compétences. Ce passage est accompagné d'une modification du statut du stagiaire (d'apprenant à impétrant). L'évaluation dans un tel dispositif change d'objet en se fixant sur les compétences tandis que la formation voit son influence relativisée à un terrain plus ambitieux recouvrant la professionnalisation. Une telle variation n'influence t'elle pas, dès lors, la construction de l'identité professionnelle du stagiaire?

2.2 Identité professionnelle, quelques repères conceptuels

Dans un premier temps nous nous situerons à la suite de Chaix (2000) pour considérer que les professeurs stagiaires sont dans une logique de socialisation professionnelle. Elle sollicite le modèle proposé par C. Dubar (1991) selon lequel l'identité sociale est le résultat de l'articulation de deux transactions : une transaction biographique entre « identités héritées » et « identités visées », et une transaction relationnelle entre « identité pour soi » et « identité pour autrui ». La transaction biographique est subjective tandis que la transaction relationnelle est objective. Il précise :

Cette problématique suppose de concevoir et d'analyser la transaction « objective » comme une confrontation entre des demandes et des offres d'identités possibles et non simplement comme des produits d'attributions d'identités pré-construites. [...]. Elle doit se concevoir comme une négociation véritable entre les demandeurs d'identité en situation d'ouverture de leur champ du possible et les offreurs d'identité en situation d'incertitude quant aux identités virtuelles à proposer (Dubar, 1991, p. 115).

Il devient donc nécessaire :

de repérer des mouvements qui affectent les modèles sociaux d'identification, c'est-à-dire les types identitaires pertinents. Ceux-ci ne peuvent être assimilés aux catégories sociales existantes officiellement à un moment donné qui sont toujours menacées de relative obsolescence notamment dans les périodes de crise. Ils doivent aussi et surtout être saisis à partir des identifications « réelles » des individus entre eux et pour eux-mêmes (Dubar, 1991, p. 118).

Il poursuit plus loin en s'interrogeant : « quelles sont les “identités d'aspiration” qui permettent des projections d'avenir efficaces pour l'action ? » (p. 122). Ce qui nous lance sur la piste des identités d'aspiration des CPT.

M.-L. Chaix (2000) qui est la première à avoir exploré cette piste à propos des élèves ingénieurs, envisage de compléter cette théorisation en rapprochant construction des identités et construction des compétences. Elle fait appel à Delbos et Jorion (1983) pour lier compétences et identité d'aspiration. Suite à leur étude de la construction des compétences chez les enfants de conchyliculteurs et de saliculteurs ils ont montré comment leurs compétences se construisaient sous la responsabilité du père, de l'initiation à la maîtrise, accédant ainsi aux images de maîtrise véhiculées par les aînés. Selon eux, ce qui se transmet correspond non pas à du savoir mais à du travail. Chaix reprend donc ces éléments pour les compléter en précisant que dans ce cadre là, la transmission à l'œuvre est chargée de savoir sur le travail. Les conditions d'apprentissage en situation reposent sur la capacité du jeune à anticiper sa maîtrise d'une situation afin d'acquérir les compétences indispensables à cette maîtrise. En transposant aux élèves ingénieurs elle identifie dans la relation tuteur stagiaire des formes de filiation symbolique. Dans

notre cas, l'alternance est très spécifique dans le sens où l'employeur est aussi le centre formateur. Cette singularité renvoie dès lors le RGR dans une situation quasi identique à celle du tuteur. La question de la filiation symbolique opère donc sur le terrain et à l'IUFM. En reprenant, le raisonnement de Chaix, il nous semble possible que le CPT comme le RGR travaillent auprès de leurs stagiaires respectifs, «à la rupture et à l'anticipation, à la démonstration/reconstruction identitaire» afin de les aider à construire une position d'enseignant... leur permettant d'accéder aux compétences d'enseignants. Les identités d'aspiration dont ils sont porteurs sont principalement celles issues de leurs histoires respectives et de leur pratique d'enseignant en formation.

C'est ce dernier point que développent les membres du CIPE dans un article au travers d'un modèle :

fondé sur une conception de l'identité professionnelle participant des dimensions sociales et psycho-individuelles, s'intégrant en cela à l'identité globale de la personne. Il pose que l'identité professionnelle se construit dans les conditions relationnelles impliquant pour la personne congruence (au sens humaniste du sentiment d'être soi) et en contiguïté avec l'autre (c'est-à-dire sentiment de confiance à l'autre quelque soit cet autre pair...) et passe par les mécanismes d'identification et d'identisation (Anadon, Bonchard, Gohier, Chevrier, 2001).

Ces différents éléments d'approfondissement de l'identité d'aspiration en lien avec la pratique d'enseignant en formation montrent qu'une telle construction se réalise dans un relationnel favorisant le sentiment de confiance dans une alternance entre identification et identisation et contingentée à une filiation symbolique. Face à cet axe de recherche, les travaux de V. Cohen-Scali (2000) nous renseigne sur les liens entre l'identité personnelle sociale et professionnelle.

Elle postule que les dimensions sociales et personnelles jouent un rôle «sur la construction de l'identité professionnelle: les expériences précoces, le milieu social, la formation mais aussi les aptitudes liées notamment à l'estime de soi, au sentiment de compétences et les attitudes et représentations concernant le travail» (p. 130).

La construction professionnelle du professeur stagiaire passe notamment par une identité d'aspiration animée par leur pratique d'enseignant en formation. Cette identité alterne entre identification et identisation au travers d'une recherche personnelle de congruence et d'un sentiment de confiance à l'autre (en contiguïté). Cette démarche se nourrit de dimensions personnelles autour de l'estime de soi et de sentiments de compétences mis en exergue par une auto-évaluation mais aussi de dimensions sociales dont les axes, attitudes et représentations, se traduisent par une évaluation catégorisante.

Voilà bien le bénéfice espéré d'une démarche prenant en compte la relation entre compétences et identité professionnelle. Du savoir sur le travail et sur la place que chacun occupe dans le travail à la capacité à se voir dans le métier d'enseignant par anticipation de la maîtrise à venir, nous touchons ici à ce que nous avons identifié précédemment dans notre travail sur l'évaluation. Si le degré de lucidité repose sur la technique évaluative, l'efficacité d'évaluateur se nourrit de la connaissance des difficultés du stagiaire à anticiper, à prendre de l'information sur le métier pour la transformer en savoir, à assumer une filiation symbolique le conduisant du syncrétisme à l'autonomie et ce dans des conditions de confiance permettant l'alternance entre identification et identification. Le passage du statut d'apprenant à celui d'impétrant prend, avec cette perspective, un relief particulier car à la différence de l'étude de Delbos et Jorion dans laquelle le jeune devient la métaphore du père lorsqu'il accède à la maîtrise, le stagiaire disparaît symboliquement avant de ressusciter comme pair de ces évaluateurs une fois la validation proclamée.

La perspective d'une influence de l'évaluation sur une partie de la construction identitaire se confirme donc. L'évaluation est alors à considérer comme un construit d'actions sociales (Clavier et Aubégnay, 2006). En effet notre étude montre que le formateur passe du statut d'expert qui conseille à celui de médiateur visant à émanciper le stagiaire entre savoir et pratique. L'apprenant devient un impétrant dans le sens où utiliser l'évaluation comme émancipation revient à dépasser les régulations académiques d'acquisition des savoirs afin de promouvoir sa propre régulation en termes d'expérience professionnelle.

Ce changement de posture chez le formateur devenu médiateur et chez l'apprenant devenu impétrant pose la question de la clarté de l'évaluation. Un médiateur est-il en situation de confronter le stagiaire à un référentiel risquant ainsi de perdre ce sentiment de confiance dont a besoin le stagiaire pour construire son identité professionnelle ? Il semble que la tentation d'abandonner ou de différer une démarche dialectique d'évaluation en utilisant un référentiel soit assez forte. Une partie de l'identité d'aspiration du stagiaire se trouve ainsi influencée implicitement.

Cependant, on ne peut négliger le recours au référent et la conscience de la filiation symbolique entre formateurs et stagiaires sous peine d'amener le « dire » en évaluation à être une épanadiplose du sujet... telle que le sujet soit... comme Fabrice à Waterloo.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Altet, M., Paquay, L. et Perrenoud, P. (éds) (2002). *Formateurs d'enseignants. Quelle professionnalisation ?* Bruxelles, De Boeck.
- Bovier, A. et Obin, J.-P. (éds) (1998). *La formation des enseignants sur le terrain*. Paris, Hachette Éducation.
- Jonnaert P. (2002). *Compétences et socioconstructivisme*. Bruxelles, De Boeck.
- Aubegny, J et Clavier, L. (éds) (2004). *L'évaluation en IUFM*. Paris, L'Harmattan, pp. 101-127.
- Aubegny, J. et Clavier, L. (éds) (2006). *L'évaluation entre permanence et changement*. Paris, L'Harmattan.
- Bertrand, N. (2004). Évaluation des CPE2 à l'IUFM des Pays de la Loire, in Aubegny, J. et Clavier, L. (éds) (2004). *L'évaluation en IUFM*. Paris, L'Harmattan, pp. 101-127.
- Chaix, M.-L. (2000). Anticiper l'identité pour construire les compétences : pratiques de tuteurs d'ingénieurs. In *Actes du Colloque AECSE «Pratiques dans l'enseignement supérieur»*, 2 au 4 octobre 2000, Toulouse (version numérique).
- Clavier, L. (2001). *Évaluer et former dans l'alternance*. Paris, L'Harmattan.
- Clavier, L. (2003). L'évaluation au service de l'émancipation : de l'efficacité à l'équité. In *Actes du 16^e colloque ADMEE «Évaluation entre efficacité et équité»*, du 4 au 6 septembre 2003. Version numérique «Les cahiers du service de pédagogie expérimentale», n^{os} 15 et 16, Université de Liège.
- Cohen-Scali, V. (1997). *Alternance et identité professionnelle*. Paris, PUF.
- Dubar, C. (1991). *La socialisation. Construction des identités sociales et professionnelles*. Paris, A. Colin.
- Jorro, A. (2005). Les compétences évaluatives des formateurs d'enseignants. In *Actes du Colloque international Recherche(s) et formation, «Former des enseignants-professionnels savoirs et compétences»*, du 14 au 16 mars 2005. IUFM des Pays de Loire (version numérique), Nantes.
- Lecointe, M. (1997). *Les enjeux de l'évaluation*. Paris, L'Harmattan.
- Perrenoud, P. (2001). Articulation théorie-pratique et formation de praticiens réflexifs en alternance. In Lhez, P., Millet, D. et Séguier, B. (éds).

Alternance et complexité en formation. Éducation – Santé – Travail social. Paris, Éditions Seli Arslan, pp. 10-27.

Perrenoud, P. (2001). *Développer la pratique réflexive dans le métier d'enseignant. Professionnalisation et raison pédagogique.* Paris, ESF.

Pineau, G. (éd.) (2005). L'alternance, une alternative éducative? *Éducation permanente*, n° 163.

Saussez, F., Ewen, N. et Girard, J. (2001). Au cœur de la pratique réflexive, la conceptualisation? *Recherche et Formation*, n° 36, 69-87.

Chapitre 12

Évaluer la formation initiale des enseignants

Stéphane Brau-Antony, Christine Jourdain

Depuis quelques années, la formation professionnelle des enseignants en France fait l'objet de nombreuses critiques. Un certain nombre d'opinions provenant de milieux divers apparaissent à l'heure actuelle pour mettre en évidence les insuffisances de la formation professionnelle dispensée en IUFM (Institut Universitaire de Formation des Maîtres). Trois types de sources peuvent être identifiés :

- des enquêtes journalistiques relayées par un certain nombre d'ouvrages à caractère polémique (Boutonnet, 2003) qui adoptent une attitude très critique vis-à-vis de la formation des enseignants. Les IUFM auraient ainsi relégué au second plan les finalités de l'école dédiées à la transmission du savoir au profit d'un discours axé sur les besoins de l'enfant. Le débat qui oppose les « Républicains » partisans d'une école sanctuaire privilégiant la mission d'instruction aux « Pédagogues » (Blais, Gauchet et Ottavi, 2002) adeptes d'une réforme de l'école en prise directe avec les évolutions de la société constitue une parfaite illustration des tensions qui traversent l'école française et qui affectent, en conséquence, les points de vue exprimés à propos de la formation des enseignants ;
- des rapports émanant du ministère de l'Éducation nationale soulignant l'impérieuse nécessité de redessiner les contours de la professionnalité enseignante (Obin, 2003) et de former plus efficacement les enseignants

de façon à ce qu'ils puissent assumer des missions qui se sont largement complexifiées. Le métier d'enseignant ne peut plus se réduire à l'exercice de sa responsabilité au sein de la classe, le travail enseignant subit aujourd'hui de profondes mutations et suppose la maîtrise de nouveaux champs de compétences liés au pilotage de projets et à la mise en œuvre de démarches pluridisciplinaires ;

- des enquêtes à caractère scientifique menées principalement par les sociologues sur la base d'entretiens de jeunes enseignants (Charles et Clément, 1997 ; Rayou et van Zanten, 2004) mettant en évidence la difficulté à exercer un métier pour lequel ces derniers se sentent insuffisamment préparés. La formation est perçue comme trop théorique, déconnectée des problèmes rencontrés lorsque les enseignants font la classe, elle est par ailleurs dispensée par des formateurs trop éloignés du « terrain ».

La formation initiale des enseignants français est donc mise sur la sellette. Elle fait globalement l'objet d'une évaluation négative. Nous pensons cependant que les jugements de valeur portés sur la formation des enseignants mais aussi sur l'institution qui en a la responsabilité sont souvent partiels voire partiiaux, c'est la raison pour laquelle ce texte a pour but de présenter un dispositif d'évaluation de la formation des enseignants dont l'objectif est de pointer, si possible de façon rigoureuse, la contribution de l'ensemble des actions de formation mis en place à l'IUFM Champagne-Ardenne à la construction d'un certain nombre de compétences professionnelles.

Le chapitre de cet ouvrage sera structuré de la manière suivante. Dans un premier temps nous exposerons les enjeux d'une évaluation de la formation des enseignants et notamment les intérêts qu'elle peut présenter pour l'ensemble des acteurs concernés (formateurs et formés). Dans un second temps nous évoquerons la méthodologie (enquête par questionnaire) et les référents conceptuels utilisés ayant permis d'élaborer le dispositif d'évaluation de la formation. La dernière partie de l'article sera consacrée aux principaux résultats de l'enquête.

1 ÉVALUER LA FORMATION DES ENSEIGNANTS : QUELS ENJEUX ?

La plupart du temps les évaluations des actions de formation fournissent une production foisonnante d'informations et de savoirs sur le déroulement de ces actions sans qu'il y ait intervention explicite des objectifs. Selon Guskey (2000), les pratiques d'évaluation les plus courantes se contentent d'utiliser des questionnaires de satisfaction des participants sans que soient véritablement mis en évidence ce qu'ont appris les enseignants. Alors que l'évaluation devrait rendre compte des changements à venir dans les pratiques professionnelles des enseignants et se rapporter au développement profes-

sionnel de ces derniers, elle n'accorde finalement que peu d'importance à des indicateurs de transformation de leur activité. S'il existe de redoutables problèmes méthodologiques pour mesurer précisément des effets de formation (Cauterman, Demailly, Suffys et Bliez-Sullerot, 1999) faut-il pour autant renoncer à évaluer la capacité qu'a l'enseignant à mobiliser, en situation de travail, les connaissances et compétences acquises en formation ?

Dans le cadre de l'évaluation d'une formation professionnelle il nous semble essentiel de poser comme principe qu'évaluer de la formation initiale des enseignants nécessite de repérer au préalable le processus de transformation des profils professionnels auxquels correspondent les actions de formation (Barbier, 1985). Apparaît donc ici un premier enjeu : permettre aux enseignants stagiaires de donner leur point de vue sur leur formation suppose de clarifier, en amont, les objectifs de formation et donc d'explicitier ce par rapport à quoi il leur sera possible d'évaluer. Il est donc impératif que les objectifs de formation soient définis de manière à ce que les évaluateurs puissent dire s'ils ont été atteints ou pas. En d'autres termes, il convient de répondre à ces trois questions : à quel métier forme-t-on ? Quelles sont les compétences professionnelles mises jeu dans l'exercice de ce métier ? Les dispositifs de formation contribuent-ils ou pas au développement de ces compétences professionnelles ?

Le deuxième enjeu porte sur la finalité d'une évaluation de la formation initiale. Plus qu'un contrôle ou une vérification de ce qui a été appris au cours de la formation, l'intérêt d'une telle évaluation porte sur la régulation du système de formation. L'évaluation se situe ici dans une logique de pilotage du système, l'évaluation par les enseignants stagiaires étant conçue comme une information en retour devant aider les acteurs de la formation à ajuster les différents dispositifs de formation. L'évaluation s'apparente dans ce cas à un travail de prise d'informations en vue de fournir des données utiles à la prise de décision (Stufflebeam, 1980). C'est une évaluation « dynamique et impliquante » (Bouvier, 1998) dont le but est d'aider l'établissement de formation et ses acteurs à évoluer et s'adapter afin d'améliorer la formation.

2 MÉTHODOLOGIE UTILISÉE ET RÉFÉRENCES CONCEPTUELLES

2.1 Le choix d'un questionnaire

L'enquête par questionnaire nous a paru être l'outil le plus pertinent pour recueillir les opinions de l'ensemble des enseignants stagiaires. Il existe bien évidemment d'autres méthodologies pour juger l'impact d'une formation sur l'activité professionnelle des enseignants. Si l'on souhaite évaluer des effets de formation, la méthodologie la plus valide consisterait à repérer ce qui est réellement réinvesti dans la pratique professionnelle en allant voir

dans les situations de travail réel ce que la formation a transformé chez les enseignants. Analyser ces processus de changement nécessiterait ainsi une observation prolongée des pratiques professionnelles des personnes. Cette démarche pour intéressante qu'elle soit nous apparaît trop lourde et peu valide compte tenu du contexte dans lequel s'inscrit cette évaluation.

Nous sommes par ailleurs conscients que les enquêtes par questionnaires présentent de nombreuses limites bien analysées par les travaux sur la psychosociologie des situations d'enquête (Ghiglione et Matalon, 1991), qui plus est quand il s'agit de juger la qualité d'une formation à visée professionnalisante. Nous avons cependant tenté de minimiser certains biais en nous démarquant d'une enquête d'opinion classique en posant comme préalable que l'appréciation portée par les stagiaires sur leur formation devait être systématiquement rapportée à des objectifs de formation s'articulant à des compétences professionnelles.

Le questionnaire est structuré en trois parties :

- une première partie concerne la connaissance du plan de formation : comment et par qui les stagiaires ont-ils été informés des dispositifs de formation ? Ont-ils eu connaissance des objectifs, contenus et modalités d'évaluation des différents dispositifs de formation ? Ont-ils eu des informations sur le processus de validation ?
- une seconde partie met en relation l'ensemble des dispositifs de formation avec la construction de compétences professionnelles réparties en 4 domaines (cf. annexe) ;
- une troisième partie appelée bilan général laisse la possibilité aux répondants de faire des commentaires et suggestions concernant leur année de formation (question ouverte).

2.2 Une entrée par les compétences professionnelles

Dans la seconde partie du questionnaire, dont nous présenterons les principaux résultats, les enseignants stagiaires avaient à répondre à la question suivante : « à travers quels dispositifs de formation et avec qui avez-vous appris à construire les compétences professionnelles suivantes ? ». Les dispositifs de formation concernent la formation dans les centres IUFM ainsi que les stages. La formation en centre est organisée autour de trois axes de formation principaux : une formation didactique disciplinaire, une formation en sciences humaines et sociales et le mémoire professionnel. Les formateurs en IUFM (tuteurs pour le suivi du stagiaire et la supervision des stages, directeurs de mémoire et formateurs dans les différents dispositifs de formation) et les conseillers pédagogiques et maîtres d'accueil en stage sont les principaux acteurs de la formation.

Le choix d'une entrée par les compétences professionnelles n'est pas neutre car il renvoie à la définition des objets d'évaluation. Donner l'occasion aux enseignants stagiaires d'évaluer leur propre formation nécessite de cerner avec précision le référent ou les critères de l'évaluation (Hadj, 1989). Ces critères d'évaluation ont été définis à partir de travaux ayant étudié les caractéristiques de la professionnalité enseignante (Bourdoncle, 1991 ; Perrenoud, 1999 ; Tardif et Lessard, 1999). Ces travaux assimilent l'enseignant à un professionnel qui dispose de savoirs et de savoir-faire hautement spécialisés.

Inciter les enseignants stagiaires à se prononcer sur la valeur de la formation c'est donc identifier ce qu'elle leur apporte du point de vue de la construction de compétences professionnelles. La compétence professionnelle est ici entendue comme la mise en synergie de ressources multiples et hétérogènes permettant de faire face à des situations complexes avec un maximum d'efficacité, elle est de l'ordre du savoir mobiliser (Perrenoud, 1999). Les compétences font par ailleurs l'objet d'apprentissages longs et coûteux qui passent par la médiation d'intervenants notamment dans le cas d'activités complexes sollicitant des savoirs composites. Les recherches sur le travail enseignant montrent que l'enseignement consiste en la mise en action de nombreux savoirs composant une sorte de réservoir dans lequel l'enseignant puise pour répondre à certaines demandes des situations de classe (Gauthier, Desbiens, Malo, Martineau et Simard, 1997).

Quatre domaines de compétences professionnelles ont ainsi été déterminés en prenant appui notamment sur les différents référentiels de compétences institutionnels (Référentiel des compétences professionnelles du professeur des écoles stagiaire en fin de formation initiale publié en 1994, Mission du professeur exerçant en collège, lycée d'enseignement général ou technologique ou en lycée professionnel paru en 1997). Le domaine 1 porte sur la conception et la planification de l'enseignement, le domaine 2 concerne la gestion de classe et la régulation des apprentissages, le domaine 3 se rapporte à l'auto-évaluation et l'analyse réflexive et le domaine 4 a trait à la responsabilité et l'éthique professionnelles. Chaque domaine de compétence a été décliné en compétences associées (cf. annexe).

L'apport des référentiels institutionnels nous est apparu indispensable car ils donnent à voir ce que l'on est en droit d'attendre d'un futur professionnel. Ces référentiels sont des instruments utiles pour penser les pratiques professionnelles, en délimiter les différentes facettes. Ils permettent aussi de guider la formation professionnelle et par conséquent de l'évaluer. Il est important de signaler que les référentiels ne rendent pas compte de la complexité des actes professionnels de l'enseignant, un référentiel de compétences est avant tout une construction théorique qui emploie un vocabulaire particulier, une sorte de sémantique de l'action dont l'objectif est de désigner ce que font les sujets. Les concepteurs des référentiels forma-

lisent les aspects jugés centraux pour exercer une activité particulière dans un temps donné et dans un contexte précis. Il faut également préciser que la définition des critères à partir desquels on définit, dans ces référentiels, le « bon enseignant » renvoie très souvent à une conception idéale-typique du métier. Paquay (2004) souligne à ce titre les risques de fixer des standards trop élevés ne correspondant pas au fonctionnement réel des enseignants en situation. Il s'avère donc nécessaire dans le cadre de la formation initiale d'ajuster le niveau d'exigence attendu. Les conditions d'exercice du métier (stage en responsabilité toute l'année pour les PLC et PLP, 9 semaines de stage pour les PE fractionnées en trois périodes) nécessitent, non pas de réviser les exigences à la baisse, mais de pointer ce que sont susceptibles de réaliser des enseignants novices. La difficulté est de taille car cela suppose de clarifier le seuil d'expertise visé, en tenant compte des situations réelles de travail et des possibilités des stagiaires. C'est la raison pour laquelle certains auteurs font parfois la distinction entre le référentiel métier et le référentiel de formation (Tarquinio, 2001).

Nous avons fait l'hypothèse que le métier d'enseignant possède une identité structurée autour d'un noyau stable de compétences que doit mobiliser chaque enseignant quel que soit le contexte d'exercice. Ces compétences sont fondées sur des valeurs et des représentations du métier partagées et participent à l'insertion des acteurs dans un groupe professionnel bien identifié. Partir de ce postulat ne veut pas dire que l'on ignore l'existence de cultures professionnelles spécifiques. L'identité professionnelle des professeurs des écoles, par exemple, est fortement marquée par la polyvalence mais aussi par la possibilité de faire appel à des intervenants extérieurs (Baillat, Espinoza et Vincent, 2001). Au sein des disciplines de l'enseignement secondaire se construisent des identités disciplinaires plurielles qui donnent parfois l'impression que les enseignants d'Education Physique (Roux-Pérez, 2004), les professeurs documentalistes (Frisch, 2003) ou les enseignants de lycée professionnel ne font pas tout à fait le même métier. De même on a souvent le sentiment qu'enseigner en milieu difficile est un métier à part où les enseignants sont parfois obligés de développer des « stratégies de survie » (Woods, 1986) pour faire face à un ensemble de situations auxquelles ils ne sont pas forcément préparés. Néanmoins, nous avons décidé d'estomper ces différences en énonçant des compétences génériques qu'il nous semble prioritaires de développer en formation initiale aussi bien dans le premier que dans le second degré, quelles que soient les disciplines enseignées.

Notons que la mise en mots des compétences n'a pas été chose aisée d'autant plus que nous ne les avons pas définies sous la forme de tâches à réaliser ou de comportements observables. La difficulté a été de traduire ces compétences de manière suffisamment explicite tout en essayant de nommer l'activité mise en jeu par le sujet pour construire les compétences professionnelles visées (présence de termes tels que concevoir, faire des choix,

adapter...). Enfin la formulation de la question «à travers quels dispositifs de formation et avec qui avez-vous appris à construire les compétences professionnelles suivantes?» est fondée sur l'idée que l'intervention des formateurs vise à enclencher un processus de construction de compétences. Dire que telle ou telle compétence est acquise en fin de formation initiale est important au regard de la titularisation des enseignants stagiaires, il est cependant nécessaire de signaler que chacun des domaines de compétences s'inscrit plus largement dans la perspective d'un développement professionnel à envisager dans le long terme (articulation formation initiale / formation tout au long de la vie).

3 RÉSULTATS DE L'ENQUÊTE

300 stagiaires professeurs des écoles (PE) et 314 stagiaires professeurs de lycée et collège (PLC) ou de lycée professionnel (PLP) ont répondu à cette enquête. La passation des questionnaires s'est déroulée au cours d'une journée de formation en mai ou juin 2004 (après la validation des stagiaires), d'où un taux de retour très élevé (89% pour les PE, 90% pour les PLC / PLP).

L'analyse des résultats se centre sur les effets des différents dispositifs de formation au niveau de la construction des compétences professionnelles. Il s'agira notamment d'identifier quels sont les dispositifs qui sont repérés par les stagiaires comme contribuant au développement des compétences de chacun des quatre domaines de compétences cités précédemment.

3.1 Les réponses des PE

Quel que soit le domaine de compétence, les stages en responsabilité et de pratique accompagnée sont les dispositifs de formation qui sont systématiquement mis en avant par les stagiaires comme contribuant au développement de l'ensemble des compétences professionnelles (entre 55 et 96% de citations pour le stage en responsabilité, entre 40 et 87% pour le stage de pratique accompagnée). L'importance des stages est d'autant plus forte que les compétences relèvent du domaine de la gestion de classe et de la régulation des apprentissages (entre 72 et 94% de citations) ou de l'auto-évaluation et de l'analyse réflexive (entre 66 et 96%).

On retrouve cette prégnance du «terrain» au niveau des acteurs de la formation puisque le maître d'accueil en stage est le premier acteur cité pour 13 compétences sur 22 (par 45 à 83% des stagiaires) et qu'il est l'acteur principal pour le domaine de la gestion de la classe et de la régulation des apprentissages avec en moyenne 68% de citations. Mais un autre acteur se dégage très nettement pour l'ensemble des compétences c'est le stagiaire lui-même qui met en avant son propre travail pour 7 compétences sur 22 (cité par 34 à 66% des stagiaires).

Si les stages sont des éléments incontournables de mise en œuvre des compétences, d'autres dispositifs jouent également un rôle important. Ainsi, dans le domaine de la conception et de la planification de l'enseignement (cf. figure 12.1), les formations didactiques disciplinaires et la dominante sont citées par plus de la moitié des stagiaires lorsqu'il s'agit *d'acquérir et d'actualiser des connaissances, de concevoir une séance, de faire des choix de contenus ou de concevoir des situations d'apprentissage*. C'est avant tout le travail individuel qui est mentionné par 61 % des stagiaires pour l'acquisition des compétences de ce domaine avant même le maître d'accueil en stage ou le formateur IUFM (respectivement 58 et 57 %).

Dans le domaine de l'auto-évaluation et de l'analyse réflexive, le mémoire professionnel est perçu comme étant un élément central de la formation (entre 49 et 74 % des réponses), notamment quand il s'agit de *questionner sa pratique* (74 % de citations). À noter qu'en dehors des stages, c'est le pourcentage le plus élevé tous dispositifs et toutes compétences confondus.

Deux acteurs principaux sont logiquement associés au développement de la réflexivité : le tuteur qui participe au suivi du stagiaire tout au long de son année de formation et le maître d'accueil en stage (entre 69 et 84 % de citations pour le premier, entre 53 et 82 % pour le second). Le tuteur est plus spécifiquement associé à *la formulation des besoins de formation* pour 69 % des stagiaires et au *questionnement de la pratique* pour 84 % des stagiaires. Ce dernier pourcentage est le plus élevé pour un acteur de la formation tout domaine de compétences confondu.

En ce qui concerne les compétences relevant de la responsabilité et de l'éthique professionnelles, les taux de réponses des stagiaires sont globalement moins élevés que pour les autres domaines. Deux dispositifs participent juste après les stages à la construction des compétences dans ce domaine : entrée dans la fonction publique et dimension éducative du métier (entre 20 et 58 % de citations, exception faite du travail en équipe qui relève davantage de la dominante pour 22 % des stagiaires). Pour plus de 50 % des stagiaires ce sont ces dispositifs qui contribuent à *inscrire les valeurs du service public dans leur pratique professionnelle*.

3.2 Les réponses des PLC / PLP

Comme pour les stagiaires du premier degré, les stagiaires du second degré mettent systématiquement en avant le stage en responsabilité pour la construction des compétences professionnelles et ce, quelle que soit la compétence considérée (entre 40 et 89 % de citations selon les compétences). En revanche contrairement aux PE, si le stage de pratique accompagnée arrive au second plan pour 9 compétences sur 20, d'autres dispositifs jouent un rôle plus important comme la formation didactique pour 10 compétences

sur 20, notamment dans le domaine de la conception et de la planification de l'enseignement (en moyenne 65 % de citations pour la formation didactique contre 58 % pour le stage de pratique accompagnée).

On retrouve au niveau des acteurs de la formation, l'importance de l'accompagnement sur le « terrain » puisque le conseiller pédagogique est systématiquement mis en avant par les stagiaires (par 26 et 81 % des stagiaires selon les compétences) sauf pour ce qui est d'*inscrire les valeurs du service public dans leur pratique professionnelle* où ce sont les formateurs IUFM qui sont mentionnés par 46 % des stagiaires.

Contrairement aux PE, la part du travail individuel arrive en second pour les domaines de la gestion de classe et de la responsabilité et de l'éthique professionnelle (respectivement entre 43 et 49 % et entre 18 et 36 % de citations), alors que les formateurs IUFM interviennent davantage dans la conception et la planification de l'enseignement (entre 43 et 68 %) ou au niveau de l'analyse réflexive grâce à la médiation du tuteur (en moyenne 52 % de citations contre 64 % pour le conseiller pédagogique).

Outre les stages et la formation didactique, d'autres dispositifs se démarquent pour les domaines relevant de la réflexivité et de l'éthique et de la responsabilité professionnelles (cf. figure 12.2). Ainsi, le mémoire professionnel est, comme pour les stagiaires du premier degré, un élément qui permet aux stagiaires de questionner et d'analyser leur pratique (entre 46 et 66 % des citations). En revanche, le conseiller pédagogique, référent du stagiaire sur son lieu de stage et tout au long de l'année, reste l'acteur principal devant le tuteur (respectivement 64 et 52 % des citations).

Comme pour les PE, le domaine de la responsabilité et de l'éthique professionnelles a donné lieu à moins de réponses que pour les autres domaines, néanmoins les dispositifs entrés dans la fonction publique et dimension éducative du métier sont clairement identifiés par 47 et 43 % des stagiaires comme leur permettant d'*inscrire les valeurs du service public dans leur pratique professionnelle* juste après le stage en responsabilité (54 % des citations).

4 DISCUSSION

Les résultats font apparaître des points communs mais aussi quelques divergences entre les PE et les PLC / PLP. Il n'est pas étonnant que les stages soient cités en premier chez les deux populations quels que soient les domaines de compétences considérés. C'est sur le lieu de stage que l'on apprend à « faire la classe ». Comme l'ont montré Tardif et Lessard (1999) la classe constitue la structure cellulaire du travail enseignant et concentre la plupart des activités réalisées par l'enseignant. Privilégier en tant que stagiaire le « terrain » est d'autant moins surprenant que la pré-professionnalisation à l'université reste

sporadique, c'est ce que montre notamment la première partie de l'enquête (26 % des PE et 33 % des PLC / PLP en ont bénéficié). Dans leur grande majorité les stagiaires découvrent le métier d'enseignant en deuxième année d'IUFM, la mise en stage constitue donc le dispositif de formation le plus saillant dans la mesure où il correspond à la situation de travail authentique à laquelle tout enseignant est confronté. Cette proximité avec les situations de travail ordinaire explique les raisons pour lesquelles les compétences relatives à la gestion de classe et à la régulation des apprentissages sont les plus associées à la mise en stage. On peut estimer que pour ces enseignants novices apprendre à faire son métier suppose de savoir exercer un contrôle sur la classe c'est-à-dire posséder un certain nombre de routines ou de tours de main afin d'instaurer un climat favorable à l'apprentissage. La construction de ce type de savoir-faire passe bien évidemment par l'expérience pratique. Ces préoccupations recourent les conclusions de travaux de recherche sur les jeunes enseignants (Rayou et van Zanten, 2004 ; Woods, 1986) : la construction d'un ordre scolaire (faire respecter la discipline), l'aptitude à s'adapter à un public hétérogène sont les conditions préalables pour que puissent se transmettre les connaissances.

L'accent mis sur le « terrain » va de pair avec le rôle exercé par le maître d'accueil de en stage ou le conseiller pédagogique. Ce rôle est d'autant plus marqué que les compétences relèvent de la gestion de classe. Les maîtres d'accueil en stage et les conseillers pédagogiques sont perçus comme des personnes ressources susceptibles de fournir une aide pour enseigner sous la forme de conseils validés par l'expérience et mobilisables dans l'ici et maintenant. Cette modalité de formation proche du compagnonnage semble plébiscitée par les stagiaires. Cependant cet accompagnement ne relève pas uniquement de ce domaine de compétences, maîtres d'accueil en stage et conseillers pédagogiques contribuent pour une large part à la formation professionnelle dans les autres domaines de compétences, preuve que les stagiaires leur donne une importance cruciale pour leur apprendre le métier.

Les stagiaires signalent également les bénéfices sur le plan professionnel que leur apportent les formateurs IUFM (tuteur et responsables de modules de formation) mais sur des domaines de compétences plus ciblés (conception et planification de l'enseignement et analyse réflexive). Tout se passe comme si les missions des formateurs IUFM et celles des conseillers pédagogiques et maîtres d'accueil en stage fonctionnaient sur un mode dichotomique en fonction de leurs compétences et statuts respectifs. Les formateurs IUFM seraient plutôt centrés sur la composante planificatrice de l'action (apports de connaissances sur les programmes scolaires et construction de contenus d'enseignement, d'où l'importance attribuée aux formations didactiques disciplinaires et à la dominante). Ces formateurs interviennent également au niveau du développement de la pratique réflexive en incitant

les stagiaires à mobiliser des savoirs théoriques. Ceci explique leur impact dans la démarche de questionnement qu'implique le mémoire professionnel avec la nécessité pour le stagiaire de convoquer des outils conceptuels pour désigner sa pratique et l'explicitier. Notons que cet objectif est très bien identifié par les deux populations. De leur côté, les conseillers pédagogiques et maîtres d'accueil en stage, comme nous l'avons vu précédemment, adopteraient des modalités de formation s'articulant plutôt autour d'une logique de rationalité technique (Schön, 1994) valorisant la transmission de compétences directement utilisables dans l'action.

Pour les PE comme pour les PLC / PLP les dispositifs de formation de type disciplinaire sont considérés comme essentiels et les amènent à changer fondamentalement leur rapport au savoir et par-là même leur rapport au métier (Bouissou et Brau-Antony, 2005) : passage d'une posture d'étudiant qui maîtrise un ou plusieurs champs disciplinaires conformément aux exigences du concours de recrutement à une posture de professionnel qui utilise des savoirs pour enseigner, comprendre sa pratique et ses effets.

On observe tout de même quelques différences entre les PE et les PLC du point de vue des apports des formations disciplinaires. Ces apports sont plus marqués chez les PLC / PLP et plus dilués chez les PE en fonction des disciplines. Chez ces derniers l'analyse détaillée discipline par discipline montre que les effets de formation perçus sont variables, certaines disciplines contribuant plus que d'autres à la construction des compétences professionnelles. Ce constat renvoie, nous semble-t-il, aux identités professionnelles des stagiaires. Pour les PLC / PLP l'appartenance à un groupe professionnel structuré autour d'une discipline « mère » permet à ses membres de se reconnaître et de faire reconnaître leur spécificité. L'identité professionnelle inclut ainsi l'identification à une discipline dans laquelle l'enseignant se reconnaît (Berger, 1989) et dans laquelle il souhaite être formé en priorité. Pour les PE, la polyvalence transforme radicalement leur rapport aux disciplines scolaires, leur position de généraliste les incite peut-être à minorer certaines formations disciplinaires au motif que certaines disciplines sont difficilement enseignables ou moins fondamentales (Garnier, 2003). Elle les rend également plus sensibles à la formation en sciences humaines et sociales.

Il convient enfin de noter que la part de travail individuel est loin d'être négligeable même si elle est plus accentuée chez les PE que chez les PLC / PLP. On peut ainsi faire l'hypothèse qu'une partie de la formation professionnelle est prise en charge de manière autonome par les stagiaires selon un processus d'auto-formation qui peut prendre des formes multiples : recherches sur internet, consultations de publications pédagogiques, sollicitation de leurs pairs ou de collègues expérimentés... Cette autoformation est également à relier à l'expérience qui se construit au fur et à mesure de l'exercice du métier.

Au terme de l'analyse des résultats de cette enquête dont nous n'avons livré que des résultats partiels, il apparaît nécessaire de revenir sur les deux enjeux que nous avons définis en présentant cette évaluation de la formation par les enseignants stagiaires. Cette enquête a effectivement permis de fournir de nombreuses informations sur la perception qu'ont les enseignants stagiaires de l'apport des différents dispositifs de formation à la construction de compétences professionnelles. Les résultats montrent quels sont les dispositifs les plus valorisés par les PE et les PLC / PLP. Il faut cependant se garder de généralisations hâtives à propos de résultats qui tendent à minorer ou majorer les effets de tel ou tel dispositif de formation car les durées de formation selon les dispositifs varient considérablement. Le temps de formation est de fait une variable déterminante qui agit sur la transformation de la professionnalité enseignante.

Par ailleurs, cette évaluation s'inscrit dans une logique de pilotage des formations. Il est cependant encore trop tôt pour évaluer les effets de cette enquête sur les orientations prises par l'IUFM en matière de politique de formation d'autant plus que des analyses plus approfondies seraient nécessaires notamment pour l'ensemble des disciplines du second degré.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Baillat, G., Espinoza, O. et Vincent, J. (2001). De la polyvalence formelle à la polyvalence réelle: enquête nationale sur les pratiques professionnelles des enseignants du premier degré. *Revue française de pédagogie*, 134, 6-11.
- Barbier, J.-M. (1985). *L'évaluation en formation*. Paris, PUF.
- Berger, G. (1989). Éléments de réflexion pour des dynamiques de changement dans la formation des enseignants: la notion d'identité professionnelle. *Recherche et formation*, 6, 17-22.
- Blais, M.C., Gauchet, M. et Ottavi, D. (2002). *Pour une philosophie politique de l'éducation*. Paris, Bayard.
- Bouissou, C. et Brau-Antony, S. (2005). Réflexivité et pratiques de formation. Regards critiques. *Carrefours de l'éducation*, 20, 113-122.
- Bourdoncle, R. (1991). La professionnalisation des enseignants: analyses sociologiques anglaises et américaines. *Revue française de pédagogie*, 94, 73-92.
- Boutonnet, R. (2003). *Journal d'une institutrice clandestine: IUFM, ou comment ne pas apprendre le métier d'enseignant*. Paris, Bayard.
- Bovier, A. (1998). Évaluation ou pilotage des organisations de formation? In G. Pelletier, *L'évaluation institutionnelle de l'éducation: défi, ouverture et impasse*. Montréal, AFIDES, pp. 137-150.

- Cauterman, M.M., Demailly, L., Suffys, S. et Bliez-Sullerot, N. (1999). *La formation continue des enseignants est-elle utile?* Paris, PUF.
- Charles, F. et Clément, J.-P. (1997). *Comment devient-on enseignant? L'IUFM et ses publics.* Strasbourg, Presses universitaires de Strasbourg.
- Circulaire n° 97-123 du 29 mai 1997. *Mission du professeur exerçant en collège, en lycée d'enseignement général et technologique ou en lycée professionnel.*
- Frisch, M. (2003). Didactique de la documentation : bâtir une éducation documentaire et informationnelle. *Perspectives documentaires en éducation*, 57, 79-87.
- Garnier, P. (2003). *Faire la classe à plusieurs. Maître et partenariats à l'école élémentaire.* Rennes, PUR.
- Gauthier, C., Desbiens, J.-F., Malo, A., Martineau, S. et Simard, D. (1997). *Pour une théorie de la pédagogie. Recherches contemporaines sur le savoir des enseignants.* Bruxelles, De Boeck.
- Ghiglione, R. et Matalon, B. (1991). *Les enquêtes sociologiques. Théories et pratique.* Paris, Armand Colin.
- Guskey, T.R (2000). *Evaluating Professional Development.* Thousand Oaks CA, Corving Press.
- Hadji, C. (1989). *L'évaluation, règles du jeu. Des intentions aux outils.* Paris, ESF.
- Note de service 94-271 du 16 novembre 1994. *Référentiel des compétences professionnelles du professeur des écoles stagiaire en fin de formation initiale.*
- Obin, J.-P. (2003). *Enseigner, un métier pour demain.* Paris, La Documentation française.
- Paquay, L. (2004). L'évaluation des enseignants et de leur enseignement : pratiques diverses, questions multiples ! In L. Paquay, *L'évaluation des enseignants. Tensions et enjeux.* Paris, L'Harmattan, pp. 13-57.
- Perrenoud, P. (1999). *Dix nouvelles compétences pour enseigner.* Paris, ESF.
- Rayou, P. et van Zanten, A. (2004). *Enquête sur les nouveaux enseignants. Changeront-ils l'école?* Paris, Bayard.
- Roux-Pérez, T. (2004). L'identité professionnelle des enseignants d'EPS : entre valeurs partagées et interprétations singulières. *Revue STAPS*, 63, 75-88.
- Stufflebeam, D.L. (1980). *L'évaluation en éducation et la prise de décision.* Ottawa : N.H.P.

- Schön D.A. (1994). *Le praticien réflexif. À la recherche du savoir caché dans l'agir professionnel*. Montréal, Éditions Logiques.
- Tardif, M. et Lessard, C. (1999). *Le travail enseignant au quotidien, Expérience, interactions humaines et dilemmes professionnels*. Bruxelles, De Boeck.
- Tarquinio, C. (2001). Le recours aux référents professionnels comme préalable à l'identification et l'évaluation des compétences. In G. Figari et M. Achouche, *L'activité évaluative réinterrogée. Regards scolaires et socioprofessionnels*. Bruxelles, De Boeck, pp.300-310.
- Woods, P. (1986). *Inside Schools. Ethnography in Educational Research*. Londres, Routledge and Kegan Paul.

Annexe

Référentiel de compétences professionnelles Professeur des Écoles stagiaires

A DOMAINE 1: CONCEPTION ET PLANIFICATION DE L'ENSEIGNEMENT

A.1 Acquérir et actualiser les savoirs disciplinaires nécessaires pour enseigner. **A.2** Concevoir une séance articulant objectifs d'apprentissage et moyens pédagogiques et didactiques (supports, travail de groupe...). **A.3** Concevoir une planification (à moyen et long terme) articulant objectifs d'apprentissage et moyens pédagogiques et didactiques. **A.4** Faire des choix de contenus adaptés aux élèves à partir des programmes de la discipline. **A.5** Concevoir des situations d'apprentissage adaptées aux possibilités des élèves. **A.6** Concevoir une évaluation permettant de rendre compte des acquis des élèves (contenus ou compétences évalués, épreuves, barèmes...).

B DOMAINE 2: GESTION DE CLASSE ET RÉGULATION DES APPRENTISSAGES

B.1 Adapter le contenu de la séance en fonction des réussites et des difficultés d'apprentissage des élèves. **B.2** Adapter le contenu de la séance en fonction des comportements des élèves. **B.3** Adapter l'organisation de la séance (gestion du temps, de l'espace, regroupement...) en fonction des réussites et des difficultés d'apprentissage des élèves. **B.4** Adapter l'organisation de la séance (gestion du temps, de l'espace, regroupement...) en fonction des comportements des élèves. **B.5** Instaurer dans la classe un climat de confiance (attitude, respect...) favorable aux apprentissages. **B.6** Adapter la communication pour favoriser les apprentissages (corps, voix, registre de langue...).

C DOMAINE 3: AUTO-ÉVALUATION ET ANALYSE RÉFLEXIVE

C.1 Accepter de questionner régulièrement sa pratique professionnelle. **C.2** Analyser sa pratique professionnelle en identifiant les écarts entre ce que l'enseignant a prévu et ce que les élèves ont réalisé. **C.3** Analyser sa pratique

professionnelle en émettant des hypothèses sur les causes possibles de ces écarts à partir de connaissances théoriques (didactiques, sciences humaines et sociales...). **C.4** Formuler des besoins pour sa formation professionnelle ultérieure.

D DOMAINE 4: RESPONSABILITÉ ET ÉTHIQUE PROFESSIONNELLES

D.1 Inscire les valeurs du service public dans sa pratique professionnelle (lutte contre les discriminations, égalité des chances, laïcité, prévention des conduites à risque...). **D.2** Participer à la vie de l'école et aux instances de décision et de concertation (conseils d'école, de cycle...). **D.3** Construire des relations avec des partenaires extérieurs (classes transplantées, échanges culturels, linguistiques, ateliers artistiques, partenariat avec les collectivités locales...). **D.4** Travailler en équipe (participation à l'élaboration ou à la mise en œuvre du projet d'école, de projets disciplinaires ou pluridisciplinaires...). **D.5** Assurer la sécurité des personnes (élèves, personnels...). **D.6** Participer au suivi et à l'intégration des élèves (conseil de cycle, réunion de parents, réseaux d'aide...).

Conclusion

Les questions posées dans cet ouvrage ne sont pas, il est vrai, entièrement nouvelles : la fonction de l'évaluation dans le système éducatif, la recherche d'outils permettant de prendre en charge les différentes dimensions impliquées dans ce projet, le souhait de traduire ces problématiques attachées à la formation des formateurs... toutes questions auxquelles sont confrontés tous ceux qui travaillent à de nouvelles approches de la régulation de l'activité de formation. On le sait, ce projet reste ambitieux dans la mesure où ces questions, complexes, appellent des réponses prudentes et nuancées.

Les travaux présentés ici montrent cependant le développement de la réflexion produite par le surgissement de nouveaux objets, comme le portfolio, et surtout par l'actualisation de questions devenues classiques, telles que l'évaluation des compétences ou celle des formations.

L'approche par compétences qui se développe aujourd'hui avec force dans plusieurs pays ne se dément pas comme le confirmait le colloque ADMEE d'octobre 2005 : ainsi en France par exemple, la mise en place du « socle commun des connaissances et compétences » et l'introduction de l'approche par compétences dans la formation des enseignants témoignent d'une actualité brûlante. De ce point de vue, les contributions de cet ouvrage ont bien un caractère heuristique dans la mesure où, si elles reflètent bien un état des connaissances à un moment donné, elles constituent aussi une photographie des questions auxquelles se trouvent soumises l'école et la formation des enseignants. À cet égard, comme plusieurs contributions l'ont illustré, la dimension internationale est aujourd'hui un levier essentiel pour quiconque veut percevoir l'origine des mouvements en cours : lorsque le ministre Xavier Darcos prônait sa réforme de l'école primaire, c'est bien à l'aide de la comparaison proposée par PISA qu'il argumente les choix fondamentaux présidant à la réforme. L'approche par compétences a cependant

montré, aussi bien en ce qui concerne le champ de la scolarité obligatoire que celui de la formation des enseignants, que les obstacles à surmonter n'étaient pas aisés à surmonter et faisaient souvent écho aux réflexions des spécialistes et des chercheurs. La question de l'évaluation des compétences ne va pas en effet de soi : les contributions ont montré l'intérêt de se poser la question du pourquoi de l'évaluation mais, aussi la question éthique qui surgit très rapidement dès lors que la démarche d'évaluation se met en place. Cette question éthique interpelle les acteurs divers de l'évaluation quand ils se trouvent confrontés aux contradictions inhérentes à l'action.

Au total, les questions soulevées ici ne constituent pas seulement, on l'aura deviné, le chapitre final d'un ouvrage qui s'achève mais bien plutôt une perspective à partir de laquelle d'autres recherches peuvent être envisagées.

Liste des figures et tableaux

FIGURES

Figure 2.1	Matrice de l'agir du professionnel	32
Figure 2.2	Trois focales pour évaluer l'activité enseignante	39
Figure 3.1	Situation des élèves français selon le retard scolaire (échelle internationale de PISA 2003, compréhension de l'écrit)	51
Figure 6.1	Extrait d'un exercice de l'évaluation en CE2	89
Figure 6.2	Les scores de Morgane et d'Amandine (CE2 en mathématiques)	90
Figure 8.1	La compétence	113
Figure 8.3	Le triangle d'adéquation	114
Figure 8.2	Compétence et piliers de la compétence	114
Figure 8.4	Le référentiel de compétences comme cadre de référence pour la formation continue et la formation initiale	115
Figure 8.5	Analyse des niveaux d'écarts dans l'action pédagogique en santé	116
Figure 9.1	Une planification effectuée pendant un stage	151

TABLEAUX

Tableau 3.1	PISA 2003, répartition des élèves de 15 ans selon le niveau scolaire	50
Tableau 5.1	Le modèle DISA : un cadre conceptuel pour les outils d'évaluation des apprentissages	75
Tableau 5.2	Exemple d'évaluation par portfolio dans le cadre d'entretiens tripartites à l'école primaire (basé sur Allal et al., 1998)	77

Tableau 5.3 Exemple d'évaluation par portfolio dans le cadre d'entretiens de coévaluation en formation des enseignants (basé sur Mottier Lopez et Tessaro, 2007)	79
Tableau 6.1 Exemple de tableau de compétence : Français CE2 (extrait)	85
Tableau 6.2 Résultats de six élèves d'une classe de CE2 à l'évaluation de français en CE2 (extrait)	87
Tableau 6.3 Groupe de besoins pour trois élèves d'une classe de CE2 à l'évaluation de français en CE2 (extrait)	88
Tableau 6.4 Tableau commun des compétences	96
Tableau 8.1 Tableau de spécification des paliers de compétences à mobiliser et à certifier dans les cours	120
Tableau 8.2 Tableau de spécification des paliers de compétence à mobiliser et certifier dans l'enseignement pratique	124
Tableau 8.3 Exemple d'un modèle de bulletin	129
Tableau 9.1 Contenus du portfolio et échanges de co-évaluation	150

Table des matières

Introduction	7
<i>Gilles Baillat, Jean-Marie De Ketele, Léopold Paquay, Claude Thélot</i>	
1 À quoi l'évaluation peut-elle et doit-elle servir ?	7
2 Les outils au service de l'apprentissage	10
3 Les outils d'évaluation au service de la formation des enseignants et des formateurs	12
 Partie 1 LES DÉMARCHES D'ÉVALUATION : POUR QUOI? POUR QUI?	 15
 Chapitre 1 L'évaluation dans le système éducatif : pour quoi faire ? À quelles conditions l'évaluation est-elle utile ?	 17
<i>Claude Thélot</i>	
1 Conditions pour que l'évaluation informe réellement la société sur l'état de l'École	18
1.1 Être crédible	19
1.2 Diffuser et rendre compréhensible	21
2 Conditions pour que l'évaluation soit un levier d'amélioration interne de l'École	22
2.1 Quelle utilisation des évaluations des pratiques, des évaluations des enseignants et des évaluations des établissements ?	24
2.2 Conditions pour que les évaluations de pratiques, celles d'enseignants, celles d'établissements servent vraiment	25
3 Une remarque pour conclure	27
 Chapitre 2 De l'analyse à l'évaluation d'une pratique professionnelle	 29
<i>Anna Jorro, Jean-Jacques Maurice</i>	
1 Distinguer deux processus majeurs	29
2 Quels « grains d'analyse » de l'activité professionnelle ?	30

2.1 Des analyseurs pour comprendre les détails de l'activité	31
<i>Les gestes langagiers</i>	32
<i>Les gestes de mise en scène du savoir</i>	32
<i>Les gestes d'ajustement de l'action</i>	32
<i>Les gestes éthiques</i>	33
2.2 Les savoir-faire développés dans la conception des tâches	33
3 Les enjeux de l'évaluation des pratiques professionnelles	36
3.1 Exigences méthodologiques	36
3.2 Exigences éthiques	38
4 Une recherche en cours	38
5 Conclusion	40
Références bibliographiques	41
 Chapitre 3 Que fait-on des évaluations internationales dans le système éducatif français?	 43
<i>Jean-Claude Emin</i>	
1 Qu'a-t-on tiré de PISA pour le système éducatif français?	44
1.1 Un questionnement des objectifs de l'enseignement français	44
1.2 Une contribution à la définition d'un socle commun de compétences et de connaissances pour la fin de la scolarité obligatoire	46
1.3 Un questionnement de pratiques d'enseignement	47
1.4 Des interrogations renforcées sur la pratique du redoublement	49
2 Mieux utiliser PISA ?	52
2.1 Rendre PISA plus pertinent au regard des décisions qui incombent aux responsables nationaux des systèmes éducatifs	53
2.2 Participer plus activement à la conception de PISA, et plus généralement des enquêtes internationales	55
2.3 Insister pour enrichir PISA et le rendre plus pertinent	55
3 Combiner évaluations internationales et évaluations nationales	55
 Chapitre 4 Quelques aspects éthiques de l'évaluation	 57
<i>Bernard Rey</i>	
1 Les pièges de la technicité	58
2 Le rapport à autrui dans l'évaluation	61
3 Les pièges du contrat	64
Conclusion	66

Partie 2	LES OUTILS D'ÉVALUATION AU SERVICE DE L'APPRENTISSAGE	69
Chapitre 5	Conceptualiser les outils d'évaluation des apprentissages	71
	<i>Linda Allal</i>	
1	Outils, un concept polysémique	72
2	Un modèle de conceptualisation des outils d'évaluation	74
3	Exemples illustrant le modèle DISA	75
3.1	Exemple d'une évaluation par portfolio à l'école primaire	76
3.2	Exemple d'une évaluation par portfolio dans la formation des enseignants	78
4	Éléments de conclusion	80
	Références bibliographiques	81
Chapitre 6	Les évaluations diagnostiques en France : le diagnostic, une notion et une exploitation en classe qui évoluent depuis 1989	83
	<i>Gérard Brézillon, Françoise Champault</i>	
1	Les évaluations de masse de rentrée	84
1.1	Les évaluations diagnostiques de rentrée par protocole	84
1.2	Les tableaux de compétence	84
1.3	La constitution de groupes de besoin indépendamment des scores moyens globaux	86
1.3.1	<i>Groupe de niveau / groupe de besoin</i>	86
1.3.2	<i>Le diagnostic est affiné (rôle des codages intermédiaires)</i>	89
1.4	Comment lier un résultat global et le profil d'un élève ?	90
1.4.1	<i>Morgane et Amandine (évaluation CE2 en mathématiques)</i>	90
1.4.2	<i>L'analyse des difficultés</i>	91
2	La banque d'outils d'évaluation diagnostique (www.banqoutils.education.gouv.fr)	92
2.1	Des protocoles de rentrée aux outils de la banque	92
2.2	Le diagnostic dans le contexte de la banque d'outils	93
2.2.1	<i>Des différences entre évaluation par protocole et évaluation banque d'outils</i>	93
2.2.2	<i>Des outils diagnostiques : principales caractéristiques</i>	93

2.2.3 Les compétences retenues pour la conception d'outils Banque	94
2.2.4 Tableau commun des compétences : un tableau d'indexation des outils	95
2.3 Construction des outils	95
2.4 L'utilisation des outils de la banque dans les classes	97
 Chapitre 7 Les outils d'évaluation ouverts, ou la nécessité de clés de fermeture	 99
<i>François-Marie Gerard</i>	
1 Trois outils ouverts	100
1.1 Le portfolio	100
1.2 L'évaluation des compétences à travers des situations complexes	101
1.3 L'évaluation de projets, de programmes ou de systèmes	101
2 Des caractéristiques communes, pour des clés de fermeture	102
2.1 Les outils d'évaluation ouverts privilégient les trois fonctions de l'évaluation	102
2.2 Les outils d'évaluation ouverts deviennent des outils d'action	103
2.3 Les outils d'évaluation ouverts intègrent le sujet dans la démarche	104
2.4 Les outils d'évaluation ouverts se fondent sur des critères et des indicateurs	104
2.5 Les outils d'évaluation ouverts portent sur des dimensions multidimensionnelles	105
2.6 Les outils d'évaluation ouverts favorisent la communication	105
Références bibliographiques	106
 Chapitre 8 Formation, compétences et constructivisme : le référentiel d'évaluation, outil de cohérence dans les programmes de formation en santé	 111
<i>Florence Parent</i>	
1 Introduction	111
2 Objectif	112
3 Cadre conceptuel de référence	112
4 L'évaluation dans la formation initiale des professionnels de santé : exemple d'un référentiel d'évaluation pour l'enseignement infirmier	117
4.1 Caractéristiques principales du programme	117
4.1.1 Les outils du référentiel d'évaluation	118

4.1.2 Particularités de l'approche pour l'enseignement théorique	122
4.1.3 Modéliser l'intégration	127
4.1.4 Synthèse et discussion	131
4.2 Synthèse	135
Références bibliographiques	136
 Partie 3 LES OUTILS D'ÉVALUATION AU SERVICE DE LA FORMATION DES ENSEIGNANTS ET DES FORMATEURS	 141
 Chapitre 9 Portfolios et entretiens de co-évaluation : des leviers de la professionnalisation des jeunes enseignants	 143
<i>Lucie Mottier Lopez, Sabine Vanhulle</i>	
1 Le portfolio en formation à l'enseignement : quelques éléments de consensus	144
2 Un ouvrage évolutif pour un processus de développement	145
2.1 Les deux facettes de l'instrument : instrumental et sémiotique	145
2.2 Portfolio et didactique des savoirs professionnels	146
2.3 Des écrits de formation balisés par des critères	146
3 Le portfolio, un outil de médiation des évaluations conjointes	147
3.1 Évaluations conjointes	147
3.2 Médiation du portfolio	149
4 L'exemple d'un entretien de co-évaluation, celui de Véra	150
4.1 Contenus du portfolio	151
4.2 Échanges de co-évaluation	152
5 Éléments de conclusion	155
Références bibliographiques	156
 Chapitre 10 Pertinence et validité de tâches complexes dans l'évaluation de futurs enseignants	 159
<i>Pierre-François Coen, Aline Galland, Isabelle Monnard, Nicolas Perrin, Yviane Rouiller</i>	
Introduction : curriculum en termes de compétences et évaluation	159
1 Plusieurs manières d'envisager la pertinence des tâches complexes	160
1.1 La formation professionnelle comme vecteur de changement	161
1.2 Se référer à l'activité des professionnels	162
1.3 La didactisation des situations	162

2	Le regard des enseignants comme moyen pour vérifier la pertinence des tâches complexes	163
2.1	Tâches complexes proposées à l'examen des enseignants expérimentés	164
2.2	Analyse de la pertinence aux yeux des enseignants	165
2.2.1	<i>Pertinence en lien avec l'exercice du métier</i>	165
2.2.2	<i>Pertinence en lien avec les pratiques réelles</i>	166
2.2.3	<i>La pertinence en lien avec le moment choisi pour cette tâche dans la formation</i>	168
3	Questionnement des notions de pertinence... et de validité	170
3.1	La notion de famille de situations	170
3.2	La notion de ressource	171
3.3	Pertinent... par rapport à quel référentiel?	172
3.4	Vers la définition de situations emblématiques ou de concepts pragmatiques	173
	Références bibliographiques	174

Chapitre 11 **Évaluation des compétences et changements identitaires** 177

Christine Baeriswyl, Loïc Clavier, Marlyse Pillonel

1	Analyse du vécu des acteurs suite à des changements dans les dispositifs et l'évaluation de la formation pratique en Haute École spécialisée	178
1.1	Du côté des professeurs	178
1.2	Du côté des praticiens	179
1.3	Synthèse – discussion	180
2	La validation en IUFM : de l'influence des pratiques d'évaluation sur la construction de l'identité professionnelle de l'enseignant stagiaire	182
2.1	L'évaluation en IUFM	182
2.2	Identité professionnelle, quelques repères conceptuels	185
	Références bibliographiques	188

Chapitre 12 **Évaluer la formation initiale des enseignants** 191

Stéphane Brau-Antony, Christine Jourdain

1	Évaluer la formation des enseignants : quels enjeux ?	192
2	Méthodologie utilisée et références conceptuelles	193
2.1	Le choix d'un questionnaire	193
2.2	Une entrée par les compétences professionnelles	194

3	Résultats de l'enquête	197
3.1	Les réponses des PE	197
3.2	Les réponses des PLC / PLP	198
4	Discussion	199
	Références bibliographiques	202
Annexe	Référentiel de compétences professionnelles Professeur des Écoles stagiaires	205
A	Domaine 1 : Conception et planification de l'enseignement	205
B	Domaine 2 : Gestion de classe et régulation des apprentissages	205
C	Domaine 3 : Auto-évaluation et analyse réflexive	205
D	Domaine 4 : Responsabilité et éthique professionnelles	206
	Conclusion	207
	Liste des figures et tableaux	209