

Mesure et évaluation en éducation



Dix années de travaux de recherche en évaluation (1986-1996) Bilan tiré des travaux du X^e colloque ADMEE-Europe (septembre 1996)

G. Figari et M. Achouche

Volume 19, numéro 3, 1997

Variations culturelles sur le thème ADMEE

URI : <https://id.erudit.org/iderudit/1091394ar>

DOI : <https://doi.org/10.7202/1091394ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer ce document

Figari, G. & Achouche, M. (1997). Dix années de travaux de recherche en évaluation (1986-1996) : bilan tiré des travaux du X^e colloque ADMEE-Europe (septembre 1996). *Mesure et évaluation en éducation*, 19(3), 5-39.
<https://doi.org/10.7202/1091394ar>

Tous droits réservés © ADMEE-Canada - Université Laval, 1997

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

<https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche.

<https://www.erudit.org/fr/>

Dix années de travaux de recherche en évaluation (1986-1996)

**Bilan tiré des travaux du X^e colloque ADMEE-Europe
(septembre 1996)**

G. Figari et M. Achouche
Université P. Mendès France, Grenoble

L'occasion était trop belle de profiter de ce dixième anniversaire pour tenter de mesurer le chemin parcouru dans le cadre d'un champ de recherche qui constitue la préoccupation essentielle de l'ADMEE, celui de l'évaluation.

Autour du thème « dix années de recherche en évaluation : quelles lignes de force? Quelles perspectives? », il s'est agi de réunir de manière la plus large possible les chercheurs travaillant sur l'évaluation. La rencontre a permis également de solliciter des professionnels et des praticiens l'expression qu'ils entretiennent avec la recherche dans ce domaine et d'inviter les organismes nationaux et internationaux dont la fonction est d'évaluer les systèmes éducatifs à présenter leurs ressources, leurs méthodes et leurs résultats.

Malgré le nombre important de communiquants, il aurait été difficile de prétendre que toute la recherche effectuée ces dix dernières années était présente. Cette vision prenant en compte « les travaux européens et nord-américains » de ces dix dernières années a été présentée, dès l'abord, par J. Cardinet et D. Laveault qui ont souligné que, « du point de vue de la pratique », cette période a vu le « le triomphe de la culture de l'évaluation » au service d'une « nouvelle philosophie politique » et que, « du point de vue de la recherche », c'est surtout « le triomphe de l'approche cognitiviste » (définie au sens large) qui est à remarquer. Des progrès dans un certain

Ce texte est publié sur invitation à ses auteurs par la revue *Mesure et évaluation en éducation*. La richesse et la diversité des communications présentées au 10^e colloque de l'ADMEE-Europe tenu à Grenoble du 18 au 20 septembre 1996 sont telles qu'il est apparu utile d'en faire part à nos lectrices et à nos lecteurs.

nombre de « compréhensions » ont été repérés : « une meilleure compréhension que les scientifiques ont du savoir, de l'évaluateur et de l'élève », « une meilleure compréhension, enfin, du pôle de la société ».

Les textes proposés par la suite s'inscrivent apparemment bien dans ces grandes lignes. En tout cas, grâce aux participants ainsi rassemblés, un bilan des productions les plus récentes a pu être réalisé à travers les conférences, les tables rondes, les communications et les discussions qui ont accompagné ces différentes activités. Il restait, la prise de distance aidant, à dresser le bilan du bilan: c'est ce qui est tenté ici.

En attendant la publication des travaux du colloque, il n'était pas intéressant d'en rendre compte autour d'un inventaire thématique : ce qui importait surtout consistait à réfléchir sur les problématiques autour desquelles les recherches sur l'évaluation semblaient se développer. C'est pourquoi une lecture, approfondie certes, mais demeurant une « lecture » au sens interprétatif du terme, est proposée dans le cadre de cet article, autour de ce qui a paru comme la « dominante d'ambiance » semblant caractériser la façon de parler de l'évaluation aujourd'hui. On pourrait penser en effet, en effectuant cette lecture, que, après les périodes « fondatrices » et militantes qui ont invité à « évaluer autrement », après celles où l'on s'est attaché à définir des modèles et à reconstruire des méthodes, est venue l'ère de la maturité où l'on peut se permettre le doute scientifique. Dans cet esprit, il a été pris le parti de « poser des questions » à l'évaluation, et de se demander à travers les travaux conduits, quels en sont *les déterminants* (§1), quelles en sont *la nature* (§2) et *la visée* (§3), quelle implication l'évaluation réalise dans *le contexte* (§4) et enfin comment elle pose *les problèmes de méthode* (§5). Les auteurs (transformés pour un temps en « méta-analystes ») ont voulu ainsi construire une signification que l'accumulation de l'information ne pouvait pas laisser apparaître. Cependant le souci de ne pas trahir le contenu même des communications a considérablement limité cette ambition : l'inventaire de ce qui a été écrit reste visible dans la synthèse présentée.

1. Questions portant sur les déterminants de l'évaluation

Qu'est-ce qui influence l'évaluation? Les éléments susceptibles de peser sur l'évaluation sont nombreux et pour certains fort connus dans le monde de l'éducation (circonstances de l'évaluation, origine socio-économique des élèves, connaissance de leurs performances antérieures, relation maître-élèves, prégnance de la décision, etc.). Les recherches sur

l'évaluation débordent cependant le strict cadre des performances scolaires et s'intéressent de plus en plus à l'explication des comportements sous-jacents. Ainsi, l'adhésion à des normes sociales oriente l'explication des comportements, ceux d'autrui et les siens (1.1). Les facteurs socioculturels, d'autre part, sont à l'origine de perceptions attributionnelles qui participent aux attitudes de résistance, causes de l'échec scolaire (1.2). Le contrôle et le jugement scolaires sont aussi investis par l'imaginaire, formant de la sorte le lieu d'un enjeu dont il importe d'élucider la fonction dans la détermination de l'activité évaluative (1.3).

1.1 Les déterminants psycho-sociaux : importance de l'internalité

Les théories de l'attribution, qui étudient les processus normatifs constituant les mécanismes de détermination du jugement, mettent l'accent sur la « norme d'internalité » dont l'intervention conduit, en matière d'explication causale des comportements, à une surestimation des facteurs dispositionnels au détriment des facteurs situationnels. Leur poids explique l'intérêt porté à leurs effets sur l'activité évaluative en milieu scolaire, aux mécanismes de leur acquisition ainsi qu'aux éventuels moyens d'en corriger les impacts.

1.1.1 La mise en valeur des effets de l'internalité sur l'évaluation et ses mécanismes d'acquisition

La mise en évidence des effets de l'internalité sur les pratiques évaluatives en milieu scolaire rencontre sans aucun doute une résonance dans le champ des recherches sur l'évaluation comme sur celui plus général des sciences de l'éducation.

- *La valorisation de la norme d'internalité en milieu scolaire*, étudiée par N. DUBOIS (« Internalité et évaluation scolaire »), est en effet un phénomène massif et généralisé, dont les élèves sont d'ailleurs tout à fait conscients et qu'ils intègrent dans leurs comportements scolaires. « Constitutive de la pratique évaluative des enseignants », l'internalité marque « l'intervention subreptice d'un critère implicite ».

- *L'intervention de la norme d'internalité dans la construction du jugement scolaire*, conjointement avec les caractéristiques socio-démographiques des élèves et le contexte de la classe, constitue le sujet de la recherche, menée en situation « écologique » et non pas expérimentale, par P. BRESSOUX et P. PANSU (« Jugement scolaire et valeurs normatives d'internalité »). Les résultats confirment que, « toutes choses égales par ailleurs, les enseignants privilégient dans leur évaluation les élèves qui choisissent préférentiellement les explications internes ».

- La réalité et les mécanismes de son influence établis, il reste à s'interroger sur *les conditions d'acquisition de la norme d'internalité* et plus particulièrement sur les liens entre cette acquisition et les styles (ou types) pédagogiques pratiqués (« coopératif », usant d'un mode de fonctionnement démocratique, ou bien « standard », basé sur des modes de conduite plus autoritaires). Pour P. Y LE MANIO et A. SOMAT (« Clairvoyance, climats pédagogiques et utilisation de registres évaluatifs/descriptifs de la connaissance personnalologique »), la relation entre internalité et types pédagogiques est conditionnée par « la clairvoyance normative », qui, croisée avec celle d'internalité, permet de distinguer les « véritables internes » (« internes non clairvoyants ») des « internes stratégiques » (« internes clairvoyants »). À partir de là, les liens pédagogie/internalité sont ainsi précisés : « [...] les pédagogies standard, par leur mode évaluatif, produisent des individus clairvoyants alors que les pédagogies coopératives, par une distribution adaptée des renforcements, provoquent moins de clairvoyance ».

Ces travaux contribuent à la mise en question des pratiques et à la relativisation des résultats de l'évaluation. Ce faisant, ils ouvrent une voie de recherche prometteuse, non seulement pour l'évaluation, mais aussi, comme le note J. PY (« Les apports de la psychologie sociale à l'évaluation : une vue d'ensemble ») pour ce qui concerne la formation et les pratiques pédagogiques dans la mesure où les dispositifs socio-éducatifs sont l'un des principaux vecteurs de l'acquisition et de la valorisation de la norme d'internalité.

1.1.2 La prise en compte des effets de l'internalité

Le poids de ces déterminations implicites incite des chercheurs à en explorer les possibilités et à étudier les moyens d'en corriger ou du moins d'en limiter les effets.

- À cet égard, l'étude expérimentale des caractéristiques d'une évaluation effectuée en l'absence de critères définis démontre, selon A. CHANNOUF et C. MANGARD (« L'influence du traitement non conscient sur l'évaluation »), l'influence du contexte « notamment lorsque l'individu n'a pas conscience de telles influences », le contexte émotionnel étant là prépondérant.

- Dans ces conditions, pour J. L. BEAUVOIS et L. CAMBON (« Orientation "on-line" vs "memory-based" et accès aux déterminants du jugement : quelques implications pour l'évaluation des personnes »), l'accès aux déterminants du jugement reste limité aux évaluations pratiquées sur une base mémorielle, qui permettent effectivement la mobilisation des théories

causales. Une telle conclusion laisse penser que « [...] des formations proposant l'apprentissage de théories causales adéquates au jugement de compétence (conjointement à l'éradication des théories parasites) permettraient des jugements d'évaluation plus fiables dès lors qu'ils seraient pratiqués à partir d'une base mémorielle (jugement "memory-based") ».

1.1.3 L' acceptabilité sociale de l'internalité

Dès lors, la question posée par l'existence de ces déterminants n'est pas tant celle de la fiabilité de l'évaluation mais bien plutôt celle de « l'acceptabilité sociale de tels jugements » dans la mesure où, pour J.L. BEAUVOIS (« La désirabilité et l'utilité sociale dans l'activité évaluative »), l'utilité sociale guide davantage l'activité évaluative que celle de la désirabilité et l'internalité est valorisée par le milieu scolaire parce qu'elle est socialement utile : « [...] la norme d'internalité définit des événements qui sont d'abord utiles parce qu'ils sont en adéquation [...] avec une option fondamentale des démocraties libérales ».

Ces conclusions rejoignent celles de N. DUBOIS pour qui la norme d'internalité conditionne sans doute l'exercice de l'activité évaluative. Par son adhésion à la norme d'internalité, l'enseignant « participe au consensus normatif qui, dans notre société, veut que les explications internes soient socialement valorisées ». La question porterait plutôt sur le « degré d'acceptabilité sociale » de l'adhésion à la norme d'internalité.

Les travaux sur la norme d'internalité soulignent le caractère profondément social de l'acte d'évaluation implicitement influencé par l'environnement social. Les questions habituellement posées à l'évaluation, par les recherches docimologiques par exemple, s'en trouvent du même coup déplacées. Dans cette perspective, en effet, la question devient celle du consensus normatif qui oriente la formation du jugement, c'est-à-dire des conditions d'exercice (« le degré d'acceptabilité ») de l'activité évaluative.

1.2 Les déterminants socioculturels

Les activités évaluatives en milieu éducatif sont tributaires des valeurs sociales dominantes, comme cela a été montré à propos de la norme d'internalité. Mais l'école elle-même peut constituer le lieu d'enjeux identitaires et devenir l'objet de jugements inconscients qui influencent son fonctionnement. D'autres processus en effet interviennent, à un niveau certes différent, dans la détermination du devenir de certaines populations scolaires. Ainsi des facteurs socioculturels et en particulier du phénomène d'acculturation dont A. GOUGNONBE (« acculturation, attribution de causalité et échec scolaire en Afrique noire ») étudie le rôle « d'obstacle

exogène à la réussite scolaire en Afrique noire ». L'exploration des représentations spécifiques à une telle situation et l'analyse des « perceptions attributionnelles issues du culturel » font ressortir les distances de tout ordre que l'élève doit franchir pour accéder à la réussite scolaire. Dans ce contexte, les pratiques scolaires en matière d'évaluation subissent aussi « l'influence inconsciente de ces facteurs qui se trouvent au carrefour du psychologique et du sociologique ».

1.3. Les déterminants relevant de l'imaginaire

Un autre déterminant du contrôle scolaire a été avancé avec l'évocation de la notion de « signification imaginaire sociale » chez l'enseignant-évaluateur. Le jugement porté sur l'élève serait marqué, de manière peut-être plus pernicieuse que dans le cas des processus normatifs, par l'imaginaire de l'enseignant. D. TSAKIRIS (« L'enjeu imaginaire dans le contrôle scolaire : esquisse d'une approche ») s'attache à élucider les mécanismes qui, en situation de contrôle et en position de jugement, aboutissent à la transformation du sujet-élève en un « cas qui tout à la fois constitue un objet descriptible et un objet susceptible d'être assujéti à la norme ». Ces mécanismes apparaissent comme le produit d'une « élaboration imaginaire » au sens de Castoriadis, le résultat d'une « articulation particulière du psychique et du social » qui masque les rapports intersubjectifs et les interactions qu'implique de fait la situation de contrôle scolaire.

L'évaluation peut donc apparaître comme « une activité sous influences » qu'aucun effort d'objectivité ne saurait faire échapper aux biais introduits par les divers déterminants qui y participent. En quelque sorte, tout se passe comme si, après avoir montré avec la docimologie que les notes étaient injustes, on cherchait aujourd'hui à comprendre la part de responsabilité des acteurs dans cette injustice. Cependant, le caractère socialement utile de certains de ces processus mais aussi leur intervention dans toutes les interactions et comportements quotidiens invitent à plus de circonspection. Si la recherche d'une objectivité totale, en ce domaine comme en d'autres, est une entreprise jugée toujours aussi illusoire que vaine, les travaux qui ont mis en valeur les déterminants de l'évaluation complètent la connaissance des mécanismes constitutifs de l'acte d'évaluer, plaidant ainsi pour plus de lucidité et plus d'exigence.

2. Questions portant sur la nature de l'évaluation

Qu'est-ce que l'évaluation? Si sa nature ne donne pas lieu à des travaux qui en proposeraient d'autres définitions, l'évaluation est cependant considérée dans trois de ses dimensions : l'évaluation comme éthique (2.1), comme modélisation (2.2), comme pratique sociale (2.3) et interrogée sur ses rapports avec la recherche (2.4).

2.1 Une éthique ?

C'est un véritable renversement de la logique de l'évaluation que propose M. LECOINTE (« Le paradigme de la valeur dans l'évaluation »), invitant à accorder le primat à l'axe éthique, « véritable opérateur de l'évaluation » sur l'axe technique qui le masque. L'affirmation d'un « principe de valorabilité, choix axiologique, parti pris éthique, pédagogique et formateur » institue une triple articulation de l'évaluation dans laquelle l'axe de la référence représentant « la constante évaluation », englobe ceux du référent et du référé. Ce retournement de logique, qui concerne aussi bien la problématique et les finalités que la démarche ou les acteurs, ne signifie pas négation mais « changement de sens ou de polarité dans un bipôle ». Ainsi, la reconnaissance de l'existence d'un paradigme de la valeur constitutif de la nature de l'évaluation conduit M. LECOINTE à proposer une démarche axiologique aboutissant à ce qu'il nomme, en démarquant HABERMAS, « un agir évaluatif ».

2.2 Une modélisation ?

Même si elle n'est pas construite sous nos yeux, la modélisation est évoquée dans plusieurs travaux et soumet l'évaluation à des choix épistémologiques. Ainsi, la réflexion sur la nature de l'évaluation se fonde pour J. F. COUET (« Esquisse d'un nouveau paradigme pour l'évaluation. Vers une modélisation des concepts de temporalité et de projet ») sur une conception de la démarche scientifique regardée comme étant essentiellement une activité de modélisation. Or, « la démarche évaluative tenant plus à la culture (ici culture scientifique) qu'à un dispositif, soulève avant tout un problème épistémologique ». La critique du « paradigme dominant de la scientificité », (rejeté parce que « objectivant et spatialisant », il se « réfère plus à une logique de contrôle »), conduit alors à la recherche d'un « nouveau paradigme pour l'évaluation » qui donnerait lieu à une autre modélisation des concepts de temporalité et de projet et ouvrirait le passage « d'un modèle statutaire à un modèle identitaire pouvant rendre raison de la multiréférentialité et de la temporalité de l'être ».

2.3 Une activité sociale?

Explorant des pistes pour la recherche en évaluation, J.M. BARBIER (« Quelques questions pour la recherche en évaluation ») traite cinq questions qui, à ses yeux, mériteraient d'être théoriquement approfondies : « les contreparties évaluatives », « les rapports entre évaluations réciproques et transactions de reconnaissance », « la polyfonctionnalité des actes d'évaluation », « l'évaluation comme lieu de médiations » et enfin, « les paradigmes de l'évaluation ». C'est à travers ce dernier point qu'il s'intéresse au statut de l'acte d'évaluation, trop vite assimilé selon lui à l'activité scientifique par un discours dominant qui aurait pour fonction de « laisser ses destinataires "dans" la pratique promue ». Cherchant alors à spécifier cet acte, il le définit « comme un acte social au même titre que les actes qu'il se donne pour objet », ce qui renvoie à la question fondamentale de la valeur.

2.4 Les rapports entre recherche et évaluation

Nombre de travaux alimentent le débat sur le problème des rapports entre recherche et évaluation. Faisant le point sur la question, C. HADJI souligne la difficulté de penser l'articulation recherche/évaluation, tant les ambiguïtés qui pèsent sur cette relation sont nombreuses : « On peut envisager des recherches sur l'évaluation, comme pratique, et des recherches avec l'évaluation (l'utilisant comme moyen). Si « l'évaluation, qui est d'ordre praxéologique, n'est pas, et ne sera jamais, scientifique », il convient cependant « [...] de distinguer plusieurs types de recherche, et la recherche évaluative est tout à fait capable de produire des connaissances, par exemple, sur les effets établissement, maître, etc... Il faut donc se méfier des typologies enfermant chacune des activités distinguées dans une conception restrictive [...]. Alors que le chercheur n'est pas à proprement parler un expert, l'évaluateur peut éclairer la prise de décision, ce qui expose au risque de la réduction simplificatrice. Tout cela conduit à redéfinir les termes de connaissance, de science, et à repenser la scientificité. L'évaluation peut faire l'objet d'un travail scientifique ; et certaines évaluations de politiques appartiennent bien au champ de la recherche ».

Les investigations portant sur la nature de l'évaluation vont dans le sens de celles ayant trait à ses déterminants : elles confirment en quelque sorte la mise en doute de la clarté du processus d'évaluation. S'y ajoutent plus spécifiquement une contestation de l'unidimensionnalité de ce processus et la tentative, que suggère l'expression « d'agir évaluatif » avec la référence à Habermas, d'ancrer l'évaluation dans une théorie générale de l'action.

3. Questions portant sur la visée de l'évaluation

À quoi sert l'évaluation? Si l'on examine les travaux présentés en cherchant les réponses à cette question, on retrouve deux grandes visées qui ne sont pas nouvelles et qui recouvrent, sans toujours les nommer, l'évaluation formative, à travers l'aide aux apprentissages et à l'élaboration de l'enseignement (3.1) et l'évaluation sommative, à travers le dévoilement de sa fonction d'interface entre école et société (3.2), sous des formes actualisées et des problématiques plus précises. Une place à part doit être réservée à une préoccupation émergente : l'évaluation comparative (3.3).

3.1 L'évaluation formative revisitée

Apparaît ainsi le rôle joué par l'évaluation dans l'aide aux apprentissages (3.1.1) et dans l'élaboration ou l'amélioration de l'enseignement (3.1.2), rappelant les visées de l'évaluation formative au sens large, mises en avant par les travaux de Bloom et Scriven, puis plus tard, par ceux de Allal, Cardinet, De Ketele et Scallon (parmi d'autres). Cette distinction s'avère difficile à opérer dans les travaux évoqués, l'évaluation des apprentissages poursuivant souvent simultanément les deux visées : pour la clarté de l'exposé et au vu du grand nombre de contributions, il est apparu néanmoins utile de la conserver, suivant la dominante qui semble se dégager des textes présentés.

3.1.1 Aide aux apprentissages

Deux tendances se dégagent là aussi, selon que l'accent est mis, spécifiquement sinon explicitement, sur l'autoévaluation ou sur des problématiques plus générales d'évaluation et de facilitation des apprentissages.

AUTOÉVALUATION

- *L'autoévaluation vue sous l'angle de la métacognition* : l'autoévaluation, dans la mesure où elle marque l'importance de la connaissance par l'apprenant de ses stratégies en relation avec ses performances, entraîne un intérêt accru pour la métacognition. B. NOEL (« L'autoévaluation comme composante de la métacognition, essai d'opérationnalisation ») étudie le fonctionnement d'une opération métacognitive (« le mode métacognitif ») et souligne l'influence, à côté de la compréhension de ses propres stratégies, du contexte d'apprentissage pour l'acquisition de l'autonomie. P.F. COEN (« AutoEval, un didacticiel pour instrumenter la métacognition ») présente le modèle qui a présidé à la conception, effectuée à partir d'une analyse des activités mises en oeuvre dans la réalisation d'une tâche, d'un instrument

informatisé d'autoévaluation, en explique le fonctionnement et présente les résultats des expérimentations auxquelles il a donné lieu.

- *L'autoévaluation vue sous l'angle des changements qu'elle provoque* : F. CAMPANALE (« L'autoévaluation, facteur de transformation des conceptions et des pratiques pédagogiques ») élabore un modèle des processus (« distanciation » et « décentration ») en jeu dans l'autoévaluation et en étudie les effets sur un groupe d'enseignants en formation pour conclure à des transformations affectant « les conceptions de l'évaluation, les conceptions pédagogiques et, au-delà, un changement de paradigme ».

- *L'autoévaluation comme pratique du « soupçon »* : M. ZEMMOUR (« le DÉFI, proposé à des élèves de CM2 en processus rédactionnel, comme critère d'autoévaluation ») propose d'intégrer le mot défi (au sens de défier et se défier d'un mot ou d'un concept, en découvrir la polysémie et dévoiler de cette manière la complexité du réel) comme critère de l'autoévaluation. La « pratique du "soupçon" qu'entraîne l'utilisation du critère défi » est en effet « autoévaluative, dans la mesure où elle renvoie le sujet à un système de valeurs qui oriente *sa méthode, son exode*. » Se référant au paradigme de la complexité, le « critère défi » ancre son action dans les processus de l'évaluation-régulation

- *L'autoévaluation par la mesure du doute et de la certitude* : l'intérêt pour « les techniques de mesure du doute et de la certitude se justifie, selon D. LECLERCQ et J.L. GILLES, par le rôle que « la connaissance partielle joue [...] dans l'apprentissage ». Si on considère que « l'ignorance connue n'est pas dangereuse » et que le doute « est le moteur même de la connaissance », mettre l'apprenant en face des conséquences de ses choix apparaît comme le moyen le plus sûr de favoriser l'autoévaluation. La recherche portant sur des techniques éprouvées de mesure du doute et de la certitude s'inscrit ainsi dans le cadre d'une « éduométrie » soucieuse de « [...] disposer d'instruments adaptés à l'objet même de l'évaluation des apprentissages : les modifications intraindividuelles ».

PROBLÉMATIQUES DE FACILITATION DES APPRENTISSAGES : ÉVALUATION ET RÉGULATION

Dans le prolongement des travaux de Allal, de Bonniol et de Genthon, le modèle systémique continue à inspirer des approches de l'évaluation formative comme moyen de régulation des processus d'apprentissage. La référence à Galpérine et Vigotsky constitue un autre point commun aux thèmes développés autour de la régulation : « l'ouverture du sens », l'évaluation continue, le rôle des critères, « l'évaluation-suivi ».

- *L'ouverture du sens comme modalité de l'évaluation-régulation* : l'intérêt porté aux processus et non plus aux résultats ou aux procédures renvoie à la question du sens. Encore convient-il de distinguer entre des modalités d'évaluation attachées au sens « formalisé » et visant l'objet de savoir et des approches occupées à « décrire les processus de régulation des élèves dans leur parcours d'élaboration du sens ». Se situant dans les perspectives ouvertes par J.J. Bonniol et M. Genthon, A. JORRO (« L'évaluation comme ouverture du sens »), opère d'emblée une distinction entre évaluation formative et évaluation-régulation qui s'intéresse à la signification, au sens, mais du point de vue « des projets de connaissance des élèves et leur confrontation avec les savoirs en jeu ». Dans ce cadre, l'évaluation ouverture du sens favorise l'autoquestionnement, elle « joue alors comme un processus d'apostrophe du sujet sur lui-même, le mobilisant dans sa relation à l'écrit, au texte, à la situation ».

- *Du contrôle continu à l'évaluation permanente* : l'appui sur les théories sociocognitives de Galpérine et en particulier la référence à la régulation du processus de construction de l'action mentale du sujet, autorise C. PIARD (« Trente ans d'évaluation "sur le terrain" : de la gymnastique scolaire à la gymnastique universitaire ») à effectuer un déplacement de l'instant d'évaluation, de celui fixé par l'institution à celui décidé par l'étudiant. Ce déplacement fait que « [...] l'évaluation devient bien la phase ultime de la constitution de l'action mentale et s'intègre dans le processus pédagogique qu'elle conforte et valide à la fois ».

- *Les critères, instrument privilégié de la démarche évaluative* : l'utilisation de critères définis « comme des instruments pour la construction personnelle du sujet » et la référence à Vigotsky et Galpérine assurent la cohérence des évaluations conduites par le Groupe Évaluation et Apprentissage (GREVA). Qu'il s'agisse d'une évaluation de tâches scolaires, de formation des enseignants à l'évaluation ou d'évaluation d'une action de formation, les critères sont utilisés comme des « points de repère » et mis « à la disposition de celui qui fait, pour qu'il oriente, réalise et régule son action ».

- *L'évaluation et le suivi pédagogique* : l'évaluation comporte une dimension dont l'intérêt n'est peut-être pas toujours bien mesuré : la communication. L'instauration dans certains pays d'un suivi pédagogique révèle les enjeux attachés à la prise en compte de cette dimension. Pour M. DERYCKE (« Le suivi pédagogique : la communication prise au piège »), les fonctions de communication et de liaison dévolues en France, par exemple, au livret scolaire, impliquent « une dialectique évaluative

originale » dès lors que l'élève est placé au centre d'un dispositif réunissant non plus un seul enseignant mais une équipe en relation avec les parents. Cependant, loin de toujours correspondre aux exigences « formatives et formatrices de l'évaluation », les pratiques réduisent souvent le suivi « [...] à la transmission d'un bilan des acquisitions [...] ». L'analyse, tant en synchronie qu'en diachronie, des « contraintes énonciatives de la régulation », fait pourtant apparaître la nécessité « [...] de faire entrer dans le calcul du suivi, et particulièrement de la régulation, le seul élément constant, l'élève lui-même, pôle permanent d'interaction avec la série d'enseignants successifs qu'il rencontrera ».

- *La régulation et la communication* : la recherche menée par M. ALTET (« Évaluation formative et communication en classe : rôle des prises d'information et des modes de régulation ») aboutit également à souligner le rôle prépondérant de la communication dans la conduite de l'évaluation formative. Basée sur une observation des pratiques, la recherche confirme l'hypothèse selon laquelle « [...] la mise en pratique de l'évaluation formative passe par un changement de représentations, d'attitudes en ce qui concerne l'ensemble du processus intégré enseignement-apprentissage-évaluation et par une prise de conscience de la nécessité d'un changement de l'ensemble de la communication et de la régulation pédagogique et pas seulement de ce qui est identifié comme champ de l'évaluation ». Elle met en évidence « [...] la nécessaire mise en cohérence entre les modes d'évaluation, les stratégies pédagogiques et les modes de communication » et montre que « l'évaluation formative se situe avant tout dans un rapport social et pédagogique de communication ».

- *Les nouveaux courants en évaluation formative : les résultats d'une innovation* dans l'enseignement primaire en Suisse romande. J. WEISS et M. WIRTHNER (« Des intentions aux réalisations. Dix ans de recherches et d'innovations en Suisse dans le domaine de l'évaluation ») étudient les changements de pratiques intervenus à la faveur d'innovations touchant aux « fonctions et formes » de l'évaluation et dans le prolongement de travaux de chercheurs comme Allal, Cardinet et Perrenoud. S'intéressant plus particulièrement à l'évaluation formative, ils analysent les apports de l'utilisation du portfolio comme outil d'apprentissage, la fonction régulatrice de l'évaluation intégrée à la didactique et le rôle des parents dans les dispositifs d'évaluation. Des difficultés sont signalées et notamment la « contradiction entre une pratique sélective et un projet de promotion de l'évaluation formative qui doit favoriser les apprentissages ». La solution, « véritable révolution copernicienne de l'évaluation », consisterait à confier la responsabilité de son devenir scolaire à l'enfant et à sa famille : « Pour bien

remplir sa mission formative, l'école devrait d'une part renoncer à orienter et sélectionner, d'autre part repousser au plus tard les échéances certificatives. »

3.1.2 L'aide à la conception ou à la régulation de l'enseignement

Bien qu'il s'agisse toujours d'évaluation concernant les apprentissages, cette dernière est évoquée, dans les travaux suivants, à travers des préoccupations didactiques du type « comment analyser les compétences? » ou « comment évaluer les besoins, le suivi ou les résultats de la formation? ».

L'ÉVALUATION DES COMPÉTENCES SCOLAIRES

La question sous-jacente aux travaux qui figurent ici pourrait être la suivante : comment, au-delà des phénomènes observables, rendre compte des stratégies et modes de conduite de celui qui effectue une tâche? Question récurrente en matière d'évaluation des compétences et à laquelle n'échappe pas l'analyse des tâches scolaires, que ces dernières soient abordées dans leur dimension cognitive ou par leurs problématiques spécifiques du type « prédiction des difficultés », ou de type « évaluation de la compréhension en lecture » ainsi que « des processus d'écriture », de « communication orale » ou encore des acquis de « l'éducation physique ».

- *L'évaluation de la « compétence cognitive »* : C. NIDEGGER et E. SAADA (« Observation des performances et estimation des compétences des élèves à l'entrée de la scolarité obligatoire ») tentent de saisir une compétence scolaire en quelque sorte globale, définie par rapport à la capacité de résoudre des problèmes ou d'atteindre les objectifs du programme. L'analyse doit donc se faire en tenant compte de la distinction entre un niveau de « compétences directement actualisables » et accessibles par l'observation des performances et un niveau de « compétences cachées » dont l'inférence se fait « à partir d'un faisceau d'indices fournis par des performances reliées entre elles ». L'évaluation des « stratégies et (des) représentations actualisées » peut alors ouvrir des « [...] pistes didactiques [...] pour différencier et gérer les différents obstacles rencontrés par les élèves ».

- *L'évaluation prédictive des difficultés d'apprentissage* : Une visée semblable, diagnostique et prédictive, se retrouve dans le projet DIVA (Didactique Individualisée avec Évaluation Analogique » présenté par E. NARDI (« L'évaluation analogique : projet DIVA »). L'ambition est de réaliser un diagnostic précis des difficultés qui attendent l'élève, pour concevoir un parcours individualisé d'apprentissage et ce, avant l'apparition des difficultés en question. À partir du profil de l'apprenant obtenu grâce à

l'évaluation analogique, « le système [...] permet de trier automatiquement pour chaque segment la variante (haute, moyenne, basse) qui se prête le mieux aux caractéristiques personnelles de l'apprenant ». Outil de cette « didactique individualisée » (issue de « la stratégie du mastery learning » conçue aux Etats-Unis pour soutenir la qualité de l'enseignement), l'évaluation analogique « exploite les propriétés logiques de l'analogie, du fait qu'elle se base sur des questions qui demandent des prestations *analogues* à celles sur lesquelles se basera l'intervention de l'apprentissage ». Le projet DIVA, actuellement expérimenté en Italie, s'applique « à n'importe quel contenu et à n'importe quel type d'école ».

- *L'évaluation de compétences liées à la lecture ou à l'écriture* : évaluer la compréhension d'un texte suppose que soient précisément désignés les « produits », les « traces » à recueillir et analysés selon une perspective théorique bien définie. La notion de « représentation interprétative verbalisée » permet à J.P. SIMON et F.GROSSMAN (« Évaluer la lecture au collège et en formation d'adultes : principes d'une démarche formative intégrée ») de caractériser les « produits de lecture », recueillis grâce aux « verbalisations écrites », qui servent de base à l'analyse des processus cognitifs. Désignée comme « relevant de l'interaction sociale », l'activité de lecture intègre l'évaluation comme « une composante à part entière » et nécessite donc « la mise en commun des « lectures singulières » auparavant objectivées à partir des traces relevées.

C'est cette même notion de verbalisation écrite, « trace de l'activité interprétative du scripteur » qui sert de point de départ à la recherche de F. BOCH (« L'écrit comme construction de savoirs à l'université. Évaluation de la cohérence dans la tâche de réécriture »). L'analyse des pratiques scripturales, étape avant l'étude de la problématique des représentations, est ici détachée de toute préoccupation immédiatement didactique ou formative et aboutit à la description des catégories de verbalisations écrites produites.

La préoccupation formative est fortement marquée dans l'évaluation effectuée par T. RIBAS i SEIX (« L'utilisation de grilles de révision du texte dans une séquence didactique d'enseignement de la composition écrite »), qui lui assigne la finalité « d'aide aux processus d'enseignement et d'apprentissage » où la communication joue un rôle décisif. Composante de la séquence didactique (qui comprend en outre « la préparation et la réalisation »), l'évaluation y assure une fonction à la fois « diagnostique et proactive », régulatrice et sommative. Exploitant la transcription des échanges réalisés dans les groupes d'élèves, la grille de révision mais aussi l'ensemble des éléments utilisés lors de la séance, l'analyse se termine sur une

série de remarques prônant en particulier une diversification « des instruments d'évaluation et des moments de leur utilisation », en fonction des spécificités et des divergences des processus d'écriture observés.

- *L'évaluation de compétences liées à la communication orale* : l'évaluation de compétences attachées à la communication orale se heurte à des difficultés spécifiques. Quant à la définition de l'objet et aux modalités de son acquisition d'abord : que signifie en effet apprendre à parler et à communiquer? Quant à son évaluation ensuite, qui soulève le problème de la mesure d'une capacité liée à des facteurs tant affectifs que socioculturels. M.C. PRESSE (« Évaluer la communication en petite section de maternelle ») étudie les pratiques d'évaluation de l'oral à travers notamment les grilles utilisées par les enseignantes de maternelle et l'observation de « séances de langage collectif ». Elle note que l'usage de grilles, portant référence à deux normes, objective et subjective, n'empêche pas l'intervention de biais subjectifs qui font se demander si la fonction de l'évaluation n'est pas dans ce cas de « confirmer la représentation que l'enseignant a de l'élève » et de conforter les théories sous-jacentes de l'apprentissage qui président à l'élaboration de ces grilles.

Malgré l'importance de son rôle dans un contexte scolaire marqué par la diversité socioculturelle et linguistique, la compréhension de l'oral reste une compétence rarement évaluée. La recherche, de type « bilan ou pronostique », menée en Suisse romande par A. SOUSSI (« Et s'ils ne comprenaient pas ce qu'on leur disait? Une évaluation de l'inévalué ») auprès d'un échantillon de 2 000 élèves de 11-12 ans, constate d'abord « comme l'on pouvait s'y attendre, un écart relativement important entre le rendement à la situation écrite et celui à la situation orale ». Une analyse plus poussée des résultats semble faire ressortir le fait que les élèves non francophones souffrent d'un handicap aggravé par l'absence d'évaluation (« critères difficiles à préciser ») et de travail spécifique sur la compréhension de l'oral.

On sait que la question des divergences inter et intra évaluateurs occupe une place de choix dans la littérature consacrée à la docimologie. C'est de ce point de vue, celui de la fidélité inter-évaluateurs, que J. BERNIER O'DONNELL choisit d'aborder le problème, plus rarement étudié, de l'évaluation de la production orale en langue étrangère, en anglais en l'occurrence. Pour ce faire, il soumet les productions d'un échantillon de 12 étudiants à divers groupes d'évaluateurs notant soit en « direct », soit à partir d'un enregistrement vidéo : des enseignants internes, des « non-natifs » naïfs (non initiés à l'évaluation), des enseignants externes, des natifs naïfs et des non-natifs naïfs jugeant en différé. À quoi s'ajoute l'autoévaluation à laquelle

devait procéder chaque étudiant de l'échantillon. Les nombreux écarts constatés tiendraient pour l'essentiel à des difficultés « inhérentes à l'évaluation de l'oral : nature éphémère de la parole (absence de traces) agissant pour le locuteur comme pour le récepteur, problèmes liés à l'appréciation de l'accent, etc. Ces résultats ne signifient cependant pas que toute évaluation soit, en ce domaine, condamnée, car « il semblerait que la fidélité "inter-évaluateur" peut être améliorée ».

- *L'auto-estimation des compétences et l'évaluation des degrés de certitude* : une évaluation qui intègre l'estimation des degrés de certitude accompagnant la formulation d'une réponse permettrait « de contrer efficacement la divination au pur hasard par les QCM » et autoriserait « un diagnostic plus subtil et par là différents niveaux de remédiation ». Pour J.L.GILLES (« Utilisation des degrés de certitude et normes de réalisme en situation d'examen et d'auto-estimation »), « Les enjeux sociaux liés à l'auto-évaluation du degré de certitude sont réels (car) cette capacité constitue une compétence métacognitive d'ordre stratégique qui soulève le problème de "l'après diagnostic", de la prise de décision en situation d'incertitude ». Après en avoir posé les conditions méthodologiques, l'auteur étudie, à partir de données recueillies en situation d'examen, les résultats des procédures utilisées à la Faculté de Psychologie et des Sciences de l'Éducation de l'université de Liège (F.A.P.S.E) en concluant : « il semble que des changements s'opèrent avec le niveau de familiarisation des étudiants à la procédure d'auto-évaluation des compétences utilisée à la F.A.P.S.E ».

- *L'évaluation des compétences en EPS* : une approche quantitative et qualitative des compétences acquises par les élèves peut être un moyen de vérifier l'adéquation des normes servant de référence à l'enseignement et à l'évaluation des apprentissages. C'est la démarche, située dans la « perspective du développement d'un "empirisme objectif" en pédagogie », adoptée par D. MARCHAL et J. F. GREHAIGNE (« L'évaluation des apprentissages en Education Physique : approches quantitatives et qualitatives des compétences acquises par les élèves »). L'étude statistique et comparative de deux populations, athlètes de haut niveau d'une part et élèves non spécialistes de l'autre, entend montrer « les limites des normes de références issues de l'analyse technique ou de la pratique de haut niveau dans le cadre des activités physiques proposées aux élèves ».

- *Évaluation de la qualité des apprentissages* : l'évaluation de la qualité des apprentissages se heurte à l'écueil du nombre et de la diversité des variables qui interviennent dans ce processus. M.L.CABANI et M.R. TERRADELLAS PIFERRER exposent les principes et la méthodologie

utilisée pour apprécier les apports de la taxonomie « SOLO », un des instruments susceptibles de permettre en ce domaine, « une mesure objective de ce qu'apprennent les élèves et de comment ils l'apprennent » en même temps qu'une « analyse des facteurs qui ont une influence sur la qualité de l'apprentissage et offrent des critères d'évaluation et des indicateurs de qualité [...] ».

On constate par la richesse des recherches s'intéressant au rôle de l'évaluation dans l'aide aux apprentissages et dans l'amélioration de l'enseignement, que l'évaluation « formative » reste une préoccupation importante avec l'apparition de nouvelles questions, la désignation de nouveaux objets à évaluer comme la compétence, la communication, les processus de régulation. Il nous paraît important de rappeler ici la question que J.M. DE KETELE (« L'évaluation conjuguée en paradigmes ») pose au terme de sa synthèse des travaux effectués dans le domaine de l'évaluation : « On peut se demander si un système de formation n'est pas caractérisé par un ensemble de paradigmes (et donc de référentiels) plus ou moins indépendants et si un des problèmes majeurs pour les travaux à venir n'est pas de les articuler théoriquement et opérationnellement ». Et l'auteur ajoute : « [...] le paradigme de la régulation nous semble une voie intéressante pour répondre à cette question ».

L'évaluation formative n'apparaît peut être plus alors comme une fonction claire, rassurante et nécessaire, mais elle est « revisitée » : sa dimension cognitive en sort certainement renforcée, le doute, concernant ses modalités et ses jugements, accentué.

3.2 L'évaluation sommative, enjeu de communication sociale

Les enjeux sociaux actuels constitués par les exigences de transparence des organisations aussi bien que par les difficultés de l'orientation et de l'insertion confèrent à l'évaluation sommative une nouvelle légitimité. Certes, le vocabulaire est plus neutre, plus technique, plus pragmatique : on parlera d'évaluation « d'effets » (3.2.1), de « validation des acquis » (3.2.2), de « certification » (3.2.3), ainsi que, toujours, « d'analyse des besoins » (3.2.4), les débats que cette notion avait suscités au cours des années 70 n'ayant apparemment plus cours.

3.2.1 L'évaluation des résultats et des effets

- La régulation de l'enseignement utilise aussi l'évaluation de la formation elle-même. F. GOMEZ analyse la démarche et les résultats de « L'évaluation d'une formation commune dans un IUFM ». Conçue comme

une évaluation globale intégrant des interventions plus ponctuelles, l'opération s'inspire de « l'évaluation répondante » de STAKE en ce qui concerne la position de l'évaluateur (« un facilitateur » et un « régulateur d'échanges ») et les domaines à considérer (« les antécédents », « les transactions » et « les résultats »). Au-delà des résultats, cette évaluation aura permis de constater que « si les effets sont toujours décalés par rapport aux attentes, ils sont loin de provoquer la déception, « l'inattendu » de la formation étant jugé positivement par chacun des groupes d'acteurs ».

3.2.2 La validation des acquis

Située à l'interface École/Société, la validation des acquis interroge les démarches d'évaluation en les confrontant à des pratiques sociales et professionnelles qui relèvent de logiques autres que celles du système éducatif. En organisant la validation d'acquis de formation et en lui conférant « le sens fort de la certification », les textes officiels contribuent selon J.AUBRET (« validation des acquis; le cas de l'université ») à l'émergence de questions importantes pour la recherche en évaluation. Ces questions concernent d'abord « la définition de l'objet et des observables à valider » dans la mesure où l'expression « validation des acquis » recouvre des représentations et des pratiques différentes. Ainsi, l'entreprise parle de compétences et d'expériences professionnelles, notions qui évoquent des dimensions personnelles certes essentielles mais difficiles à prendre en compte : « L'expérience professionnelle nous renvoie vers le sujet et son activité cognitive organisatrice. Il n'y a donc pas de cadre pour l'analyser et l'évaluer. » L'autre problème réside dans l'utilisation de la mesure comme instrument de « différenciation interindividuelle » alors que la validation des acquis sollicite « la partie non visible » de l'évaluation, c'est-à-dire « le système de valeurs qui structurent l'acte d'évaluer ». À travers le statut respectif du candidat et de l'évaluateur, la validation des acquis met en jeu le caractère social de l'évaluation et laisse entrevoir de nouveaux enjeux pour la recherche : « L'approche par la validation des acquis ouvre des perspectives nouvelles sur un objet d'étude considéré le plus souvent sous l'angle de la mesure et de la notation. »

3.2.3 La certification

- *La certification en cours de formation, un autre mode d'attestation des acquis* : l'institution du contrôle en cours de formation représente plus qu'un simple aménagement technique des modalités de l'évaluation certificative. Elle signifie une innovation majeure, un changement de culture qui met à jour les enjeux liés à l'acte d'évaluation. La réflexion de M.C. DAUVISIS (« Dix ans de certification en cours de formation : des pratiques pédagogiques aux enjeux identitaires ») sur les pratiques portées par l'instauration du contrôle

continu en cours de formation dans l'enseignement agricole fait ressortir que « la part de certification confiée au formateur dans cet espace de la formation induit un contrôle social dans le cadre d'une négociation, de comptes à rendre et de résultats à faire valider par le regard externe d'un jury ». Ces « enjeux identitaires » recouvrent les questions proprement pédagogiques et définissent l'évaluation comme « une pratique sociale fortement contextualisée ». Dans ce sens, elle appelle une « approche de la complexité qui mobilise les modèles issus des diverses sciences sociales ».

- *Le contrôle en cours de formation et l'épreuve terminale* : le contrôle en cours de formation a également pour fonction de moduler les effets des examens-sanctions en diversifiant les situations d'évaluation et en facilitant leur adaptation aux contextes professionnels. Tels étaient en tout cas les objectifs de la réforme de l'enseignement agricole, où coexistent désormais contrôle certificatif en cours de formation et évaluations terminales. L'enquête conduite par N. DROYER et G. FROSSARD (« Contrôle en cours de formation, épreuve terminale : quelle(s) différence(s)? ») cherche à relever et à expliquer les éventuels écarts entre ces deux types d'évaluation. Leurs résultats révèlent que « les notes de contrôle en cours de formation sont presque toujours supérieures [...] aux notes obtenues en épreuve terminale » mais ne dénotent cependant aucun dysfonctionnement, les écarts relevant plutôt de la différence d'objets de savoirs évalués par les deux épreuves et des modalités plus innovantes adoptées par les contrôles certificatifs.

- *L'évaluation certificative, un fait didactique* : l'évaluation certificative ne relève pas que d'une problématique docimologique. Elle investit le champ didactique et devient à ce titre un objet de recherche légitime pour le didacticien. B. DAVID (« Observatoire des pratiques d'évaluation certificative en Éducation Physique et sportive ») définit les principes et le cadre conceptuel d'une recherche qui « voudrait contribuer à jeter les bases d'une théorie de l'évaluation certificative en EPS, permettant de lire à terme les pratiques évaluatives dans leur diversité empirique ». Partant d'une conception du phénomène didactique qui intègre, outre la relation enseignant/enseigné, le savoir en jeu, il décrit les principes d'une démarche de recherche s'inscrivant « dans l'ambiance des recherches modernes en sciences humaines, croisant des analyses qualitatives et quantitatives ».

3.2.4 *L'analyse des besoins*

- *L'analyse des besoins de formation* : celle-ci reste une démarche souvent pratiquée dans le cadre de la formation professionnelle et de la formation de formateurs, notamment pour définir un contenu d'enseignement. Distincte de celles d'attente, envie, demande, motivation, etc., la

notion de besoin de formation doit, selon Y. PIN (« Analyse des besoins de formation des enseignants ») s'entendre comme « [...] la résultante d'un processus associant les différents acteurs concernés, et traduisant un accord entre eux sur des manques à combler par le moyen de la formation ». Adoptant cette perspective participative, l'étude s'effectue en deux étapes, des entretiens de groupes succédant au dépouillement des résultats de l'enquête par questionnaires. Les résultats amènent à constater que « les nouvelles problématiques obligent à repenser la formation » car l'analyse des besoins amène à considérer que « l'objectif est la construction d'une nouvelle identité professionnelle ».

3.3 La visée comparative

La comparaison (de modèles d'apprentissage ou d'enseignement, de niveaux de performance, etc.) tend d'autant plus à se généraliser que nombre de pays se posent des questions quant à l'efficacité de leur système éducatif et à l'opportunité de sa réforme. Cette visée comparative soulève, dans les textes évoqués ici, des problèmes dont la complexité laisse parfois planer un doute quant à la faisabilité de telles évaluations.

- *Une évaluation exploratoire dans des systèmes éducatifs différents* : exploratoire, l'évaluation menée dans quatre pays francophones et étudiée ici par G. LEGROS et J.P. LALANDE (« Savoir écrire en 3^e secondaire : une évaluation comparée dans des pays francophones ») n'avait pas pour but la mesure de l'efficacité proprement dite des systèmes considérés. Elle se proposait plutôt de restituer le contexte et de « dresser un portrait comparatif des pratiques déclarées » en matière « d'enseignement/apprentissage de l'expression écrite » pour « étudier [...] les corrélations statistiques entre ces diverses données et les performances des élèves ». Même si les instruments d'évaluation (questionnaires aux enseignants, jeu d'épreuves formelles et rédaction pour la mesure des performances des élèves) et la correction des épreuves ont été conçus en tenant compte de toutes les contraintes (formation des enseignants, mesure de l'homogénéité des jugements, ajustement des scores, etc.) et ont rendu possible la comparaison recherchée, il reste que les résultats doivent être interprétés avec prudence, car « plusieurs des aspects de cette démarche sont nouveaux et ont été développés pour les besoins de cette recherche ».

- *L'évaluation comparée de performances en langue seconde* : cette évaluation se heurte à une double difficulté. On ne peut d'une part pas comparer des résultats obtenus au moyen d'épreuves et dans des circonstances d'évaluation différentes et on ne peut pas non plus utiliser les

mêmes épreuves sans tenir compte des contextes d'enseignement. J. MONNARD, C. TIECHE et J.L. GURTNER (« Faire face aux exigences de comparabilité dans l'évaluation d'une langue seconde ») exposent les principes ayant guidé la conception et l'élaboration d'un instrument qui « cherche à annuler l'effet de (ces) différences afin de permettre les comparaisons » et qui, au-delà de certaines « faiblesses » résiduelles, pourrait s'avérer utile pour la régulation de l'enseignement et de l'apprentissage.

Des travaux présentés ressort une nouvelle façon de s'intéresser à l'évaluation sommative : plutôt que sa pertinence ou sa signification internes, c'est sa fonction externe de communication sociale qui est principalement traitée. En effet, les résultats, les effets, les acquis ne sont évoqués qu'à travers leur « certification », privilégiant ainsi, plutôt que sa signification « scolaire », son interprétation et son utilisation sociale.

4. Questions portant sur l'implication dans un contexte

Les réponses aux questions précédentes sur les visées de l'évaluation sommative viennent de mettre en évidence l'intérêt croissant pour la signification et l'utilisation sociale de l'évaluation. Cette tendance s'amplifie encore lorsque sont étudiés les rapports avec le contexte. En effet, deux types de travaux au moins témoignent de l'importance croissante de cette implication : ceux qui ont pour objet l'élucidation de comportements professionnels et sociaux à travers l'évaluation de la « compétence » (4.1) et ceux qui relèvent de plus en plus des « politiques » éducatives et culturelles (4.2). Un troisième domaine pose le même type de questions sur les rapports de l'évaluation avec le contexte : celui de « l'éducation-développement » (4.3).

4.1 L'élucidation de comportements professionnels et sociaux

Il est en effet indéniable que l'évaluation n'est plus spécifique au monde scolaire : comme le loup qui est sorti du bois, elle se déploie maintenant au grand jour dans les organisations professionnelles et sociales. Les préoccupations des travaux qui en relèvent vont même jusqu'à émettre des doutes sur le rôle de la formation dans la préparation aux tâches professionnelles. Dans une société où l'école est sans cesse mise en cause pour ses insuffisances en matière de qualification, on aurait tort de ne pas prêter attention à ce qui pourrait bien devenir une remise en question de l'évaluation scolaire... C'est autour de la notion de « compétence » que ce problème est abordé dans les travaux présentés ici. On évoquera les rapports entre

psychologie cognitive et évaluation des compétences (4.1.1), l'analyse des compétences dans l'organisation du travail (4.1.2), comme préalable à la constitution de référentiels (4.1.3), et comme évaluation (impossible ?) de la formation (4.1.4).

4.1.1 La psychologie cognitive et l'évaluation des compétences

- *L'analyse des compétences*, « une descente vers le cognitif » : dans ce domaine, l'évaluation ne peut procéder que par l'analyse. Les conceptions classiques inspirées par les modèles scolaires sont là inopérantes car le développement des compétences (scolaires ou professionnelles, celles de l'ouvrier peu qualifié ou celles de l'ingénieur expert) s'effectue à long terme et implique des processus cognitifs complexes aboutissant à la constitution « d'un répertoire de formes d'organisation de l'activité dans des registres variés : les gestes, les affects et les émotions, le langage, les relations avec autrui, les savoirs et savoir-faire scientifiques et techniques ». L'impossibilité d'explicitier la plus grande partie des connaissances composant la compétence ajoute au fait que son évaluation « [...] appelle non seulement l'analyse du résultat de l'activité mais aussi celle de l'organisation de l'activité ». À travers la définition du concept de « schème » et de ses éléments constitutifs, G. VERGNAUD (« Psychologie cognitive et évaluation ») montre la nécessité du recours à l'inférence pour la description des opérations cognitives impliquées dans la mise en oeuvre d'une compétence et caractérise comme suit la méthode d'analyse, qualifiée de « descente vers le cognitif » : pas d'analyse des compétences sans analyse de l'activité, pas d'analyse de l'activité sans analyse des conceptualisations sous-jacentes [...] ». C'est l'absence de prise en compte de cette dimension qui rend, par exemple, peu « économiques » et peu « opératoires » les référentiels de métier souvent construits sur une longue énumération « de gestes et de compétences élémentaires ».

4.1.2 L'analyse des compétences dans l'organisation du travail

- *L'évaluation de la formation et de la compétence des opérateurs* : le dépassement du mode d'organisation tayloriste du travail a conduit à la considération de la dimension sociale de la compétence, désormais « collective et contextuelle au lieu d'être individualisée et générale ». Le mode d'organisation du travail participe, au même titre que la formation et l'expérience professionnelle, à la construction de la compétence. T. BOLLON (« Évaluation de la formation et de la compétence des opérateurs : une approche cognitive ») utilise une approche cognitive pour « évaluer le poids respectif de chacune de ces trois variables [...] et définir en termes d'acquisition de connaissances l'impact de l'organisation du travail, des actions de formation et de l'expérience professionnelle ». Les résultats

« assez inattendus » autorisent à conclure « [...] qu'une action de formation ne doit être définie qu'en fonction de l'organisation du travail » dans la mesure où elle « [...] ne demeure efficace que si l'organisation du travail laisse à l'opérateur une marge d'autonomie suffisante pour utiliser ses connaissances ».

C'est à une conclusion semblable qu'aboutit le travail de M. DUBOIS et T. BOLLON (« Pour une meilleure prise en compte des modes d'organisation dans un programme de formation »). S'interrogeant sur « l'efficacité de l'outil formation », leur enquête fait ressortir l'importance « des modes d'organisation implicites du travail dans la construction du savoir ouvrier ». Du point de vue de la formation, ceci implique pour le formateur la nécessité « d'intégrer dans son analyse préalable du poste ou de la fonction la variable organisationnelle » et pour le dirigeant d'entreprise de « gérer la politique de formation de l'entreprise en fonction du mode d'organisation du travail et du fonctionnement réel des équipes [...] ».

4.1.3 L'analyse des compétences comme préalable à l'élaboration de référentiels

La notion de « référentiel » s'est répandue, depuis le milieu des années 80, dans le monde de la formation en France, pour désigner toutes sortes de documents d'origine et de légitimité diverses proposant, en guise de cadre de références, des listes de capacités, de compétences, d'objectifs de formation relatifs à des fonctions professionnelles déterminées. Rien d'étonnant si les travaux sur l'évaluation s'en inspirent, en constituent le reflet ou remettent en cause la notion (cf. Figari, 1994) tout en soulignant les liens qui s'établissent entre les catégories de ces « référentiels » et les critères de l'évaluation qui les concerne.

- *L'analyse de compétences, « référentiel métier » et « référentiel formation »* : l'analyse des compétences est de plus en plus considérée comme un préalable à la construction de référentiels métiers et de référentiels de formation. Travaillant dans cette optique, TARQUINIO CYRIL (« Les emplois familiaux en Europe : de la constitution d'un référentiel métier à l'élaboration d'un référentiel de formation ») procède à une étude des compétences professionnelles guidée par l'objectif d'établir les « critères permettent d'envisager à la fois leur caractère transversal et [...] transférable ». Cette précision se fonde sur l'analyse des démarches cognitives qui « montre souvent que les métiers se caractérisent uniquement par quelques démarches intellectuelles dominantes ». Les capacités du référentiel métier sont ensuite traduites, avec l'aide de professionnels, en capacités de formation pour « contribuer à l'ajustement des contenus

pédagogiques ». L'élaboration de référentiels se heurte à des exigences difficilement compatibles et qu'il faut pourtant concilier : l'exigence de précision d'une part, la nécessité de disposer d'outils souples et adaptables d'autre part.

- *L'élaboration d'un « référentiel d'employabilité » et l'évaluation/validation de l'employabilité* : la notion de référentiel d'employabilité introduit la « logique des personnes » en complément de la « logique du métier » présente dans les référentiels du même nom. Pour P. RABOT et F. CAMPANALE (« Référentiel d'employabilité : agent d'étage ») le référentiel d'employabilité, situé au croisement de deux logiques exerce, « dans le cadre d'une démarche d'accompagnement de la formation [...] les fonctions d'orientation, d'évaluation et de validation des acquis ».

L'utilisation d'un référentiel d'employabilité comme guide de la formation demande que soient définies en parallèle une démarche et des modalités d'évaluation cohérentes avec le contenu de la notion. Ainsi, pour J. CAMPANALE et F. RABOT, l'évaluation de l'employabilité ne doit pas se résumer à un contrôle mais doit aussi faire intervenir la fonction de régulation qui « [...] en impliquant le stagiaire, favorise le développement de sa capacité d'auto-évaluation prise dans le sens de la régulation de ses pratiques et de ses démarches d'action [...] ». La validation quant à elle se fait par une commission sur la base du profil d'employabilité établi à l'issue de la formation.

4.1.4 L'analyse des compétences comme évaluation de la formation?

- *L'évaluation de compétences professionnelles à l'issue d'une formation* : la formation dispensée dans les IUFM vise l'acquisition de compétences professionnelles. Celles-ci sont-elles pour autant mesurables à l'issue de la formation? L'expérience de directeur de mémoire professionnel, l'observation et le suivi durant cinq ans de stagiaires se destinant à l'enseignement en lycées techniques incitent L. DAUZET (« Recherche sur l'évaluation ») à répondre par la négative. Même si le dispositif de formation constitue un « véritable parcours en alternance », les « compétences réellement acquises pendant la formation ne se révéleront que lors de l'exercice professionnel, après plusieurs années ». L'évaluation ne peut ainsi porter que sur « la qualité de la formation » et ne peut, à travers les productions des stagiaires, apprécier qu'un « potentiel ».

- *L'évaluation/validation de compétences pédagogiques* : l'association des termes « évaluation » et « validation » désigne clairement le but poursuivi par ces procédures : « [...] le transfert immédiat sur le terrain ». La démarche

présentée par J.M. GAUDIN et P. BLANCHET (« l'évaluation des compétences pédagogiques dans la formation de formateurs Pedagogi-A : évolution d'une réflexion et d'une pratique ») s'inscrit dans le cadre d'une formation de formateurs qui entend « garantir aux employeurs des compétences professionnelles attestées permettant de mettre directement le formateur en situation de gestion d'un groupe d'apprenants ». Définissant la pédagogie comme étant « avant tout de l'action concrète », la réflexion conduit à privilégier les modalités d'évaluation du savoir-faire, « garant du savoir tandis que l'inverse n'est pas toujours certifié ».

Nécessaire à l'élaboration de la formation comme à l'appréhension de ses résultats et de ses effets, l'analyse des compétences, telle qu'elle ressort des textes présentés ici, apparaît comme un outil privilégié de l'évaluation. Conçue dans une optique cognitiviste, attentive aux exigences du métier et aux caractéristiques des individus, « aux résultats de l'activité comme à l'organisation de l'activité » (G. Vergnaud), l'évaluation des compétences montre l'exemple d'une démarche qui ne sépare pas l'examen du produit de celui des procédures et des processus qui ont permis d'y parvenir.

4.2. Questions portant sur l'évaluation des politiques publiques, éducatives et culturelles

L'évolution des systèmes éducatifs sous le poids des changements qui affectent nos sociétés tant sur le plan économique et social que technologique ne pouvait pas éviter un intérêt grandissant pour l'évaluation des politiques publiques. Ceci est particulièrement sensible dans les pays de l'OCDE, au palier national et international. Les travaux présentés font porter l'évaluation sur les « actions publiques » (4.2.1), analysant les politiques dans leur processus de prise de décision et sur les « systèmes éducatifs » (4.2.2), s'intéressant alors aux effets et à l'efficacité.

4.2.1. L'évaluation des actions publiques

L'évaluation des « politiques publiques » a donné lieu à des questions « ouvertes », proposant des pistes de réflexion et des investigations prometteuses sur l'évaluation vue elle-même comme pratique sociale, sur son rôle dans la prise de décision, sur ses nécessaires adaptations aux nouvelles formes de politiques de formation et de politiques culturelles.

- *L'évaluation de l'action publique, une pratique sociale et un acte stratégique* : La diversité des objectifs mais aussi des formes que revêt l'évaluation des politiques publiques contredit les définitions habituelles de l'évaluation, y compris celles privilégiant « l'aide à la prise de décision »,

cette dernière expression suggérant un mécanisme selon lequel la décision interviendrait après l'évaluation. L'examen des pratiques réelles semble montrer au contraire que l'évaluation est avant tout « un acte stratégique au coeur d'un "champ d'action concret" ». C'est de cette position théorique que L. DEMAILLY (« L'approche stratégique de l'évaluation ») appréhende la question de l'efficacité et des effets de l'évaluation. Elle attribue l'impact souvent faible des évaluations aux « contradictions stratégiques » dans lesquelles se trouve pris l'évaluateur mais surtout aux « faiblesses méthodologiques » et théoriques supportant une « conception linéaire du social » ou supposant une « équivalence forte entre social et opinion publique ». Classés selon le type de production (« efficacité scientifique », « efficacité idéologique », « efficacité stratégique », « efficacité culturelle ou formative » et « efficacité de rationalisation »), les « outputs » sont étudiés sous l'angle de leurs produits et de leurs effets pour déterminer les « conditions méthodologiques d'une évaluation qui vise rigueur et rationalisation ».

- *Le rôle de l'évaluation dans la prise de décision politique* : dans le prolongement de l'idée précédente, l'évaluation s'intègre dans le processus décisionnel, participant ainsi à la gestation de la décision. À travers le questionnement des acteurs, nécessaire à l'élaboration d'un référentiel spécifique, V.K. IFRAH (« L'aide à la décision politique : l'évaluation comme induction de changements dans les systèmes décisionnels locaux ») analyse la façon dont l'intervention d'un expert conseiller, par ailleurs inscrit dans le système scientifique, interfère dans le système décisionnel local, inséré, lui, dans un autre environnement. La confrontation dialectique des deux systèmes favorise, tout au long du questionnement, l'activité réflexive des acteurs conduisant ainsi à la prise de décision proprement dite : « [...] l'enquête, une des composantes du processus d'évaluation, devient une dimension inductrice de changements dans le sens où elle prépare le terrain aux décisions possibles. »

- *L'évaluation des politiques publiques de formation* : reconnaissant que l'évaluation des politiques publiques a été marquée dès ses débuts en Amérique du Nord par l'évaluation de la formation dont elle constituait en quelque sorte une extension, Y. BION (« Qu'apporte l'évaluation des politiques publiques à l'évaluation de la formation? ») entreprend de tirer les enseignements, pour l'évaluation de la formation, de l'expérience de la France en ce domaine. Il remarque que la tentative d'y ancrer une évaluation fondée sur le paradigme positiviste et attachée à la mesure des effets s'est heurtée à de multiples difficultés, tant internes (tenant à la nature de l'objet et à la visée adoptée) qu'externes (dues au contexte institutionnel, aux

résistances et aux réticences face à la fonction de contrôle que ce type d'évaluation peut exercer). Une autre approche apparaît alors possible, basée sur une conception de l'évaluation comme « un processus de régulation du consensus » au sein d'une réalité posée comme « un construit social multiple ». Pour Y. BION, l'évaluation de la formation, qui connaît une évolution semblable, s'inscrit elle aussi dans des réseaux de pouvoir dont la légitimité devrait être considérée au lieu « que l'institution ne donne le pouvoir légitime et formel d'évaluer qu'à un certain nombre de personnes ». De même, l'évaluation de la formation peut utilement tirer parti de distinctions opérées dans l'évaluation de programme, telle celle entre résultats (« outputs ») et effets (« outcomes ») pour apprécier l'impact d'une action.

- *L'évaluation d'une politique culturelle régionale* : la politique culturelle constitue désormais, la décentralisation aidant, un champ d'intervention. L'évaluation dont rend compte E. AUCLAIR (« Évaluation de la politique culturelle de la Région Rhône-Alpes ») a pour objectif la réalisation d'un diagnostic tenant compte « des attentes et des perceptions des partenaires » et la mise en place « d'outils d'analyse et de suivi des actions » en relation avec les divers acteurs. La « démarche qualitative » suivie obéit aux « logiques d'analyse mises en oeuvre en parallèle » et qui consistent dans « l'étude de la cohérence interne de la politique culturelle, l'étude de sa cohérence externe, l'étude des moyens affectés à cette politique, l'étude des résultats ».

4.2.2 L'évaluation des systèmes éducatifs

On dispose aujourd'hui de bilans (OCDE, ministères nationaux de l'Éducation) mais aussi de recherches en cours sur l'efficacité et la rentabilité des systèmes éducatifs donnant accès à une meilleure connaissance de cet objet, tant dans les cadres nationaux qu'internationaux.

L'ÉVALUATION DES OUTPUTS : L'ÉVALUATION DANS UN CADRE NATIONAL

- *L'évaluation, une aide à l'action publique et un outil de changement au sein du système éducatif* : l'objectif de l'évaluation menée par la Direction de l'Évaluation et de la Prospective (« Dix ans d'évaluation au ministère de l'Éducation nationale : l'expérience de la Direction de l'Évaluation et de la Prospective ») est de constituer un instrument pour l'action publique et un facteur de changement au sein du système éducatif français. L'aide à l'action publique appelle l'évaluation de l'efficacité du système à travers la mesure de ses résultats. Cette mesure de l'efficacité s'effectue à travers quatre « approches de complexité croissante » : mesurer l'atteinte des objectifs en évaluant les compétences et les connaissances des élèves, mesurer l'impact

des politiques éducatives sur les acquis des élèves, mesurer l'impact des pratiques des acteurs sur les acquis des élèves ou le fonctionnement du système, apprécier l'influence des organisations sur la réussite des élèves. Les évaluations annuelles de compétences effectuées à l'entrée de la 3^e année d'enseignement primaire (CE2), de la 1^{re} année d'enseignement secondaire (6^e) et de la 5^e année d'enseignement secondaire (seconde) s'accompagnent de la mise à disposition d'instruments de pilotage et d'outils pour une évaluation conçue comme aide à l'apprentissage. La réussite d'un programme aussi ambitieux dans ses objectifs que complexe dans les modalités de sa réalisation est due en grande partie au dépassement des clivages entre une approche « scientifique » et une approche « sociologiste », la DEP utilisant « à la fois des méthodes quantitatives pour l'échantillonnage des populations évaluées et le traitement des données, et des méthodes qualitatives pour la construction des protocoles d'enquête ».

- *Évaluer l'efficiencia de l'enseignement universitaire* : s'intéresser à l'efficiencia du système éducatif reste toujours socialement difficile. Or, pour M. DURU-BELLET et A. MINGAT (« L'évaluation de l'enseignement universitaire dans une perspective analytique »), « [...] établir une relation entre les résultats obtenus et l'allocation de ressources, qu'elles soient financières, humaines ou organisationnelles [...] est pourtant [...] ce qui définit de façon centrale l'évaluation ». C'est en partant de cette problématique, appliquée à l'enseignement universitaire, qu'est exposée la démarche nécessaire à la conduite d'un programme de recherche axé « prioritairement » sur la production de connaissances mais dont les résultats peuvent également « [...] permettre d'évaluer les marges pour une meilleure utilisation des moyens disponibles en vue d'une amélioration de la réussite des étudiants ».

- *L'évaluation de l'impact d'une réforme curriculaire* : les modes mais aussi les stratégies d'apprentissage ne sont sans doute pas indifférents au type de curriculum qui dicte leur progression. Une réforme curriculaire agit à tous les niveaux de la sphère éducative. L'équipe de l'Institut d'Éducation et de Psychologie de l'Université du Minho-Portugal (« L'impact de la réforme curriculaire sur la pensée et sur l'action des enseignants ») choisit d'étudier cet impact en se centrant sur la pensée et l'action des enseignants, « acteurs privilégiés » du développement curriculaire. Interrogeant leurs pratiques et leurs attitudes à l'aide d'une méthodologie quantitative (questionnaires) complétée par une approche qualitative (entretiens semi-directifs), l'étude conclut que « cette réforme curriculaire n'a pas produit les effets escomptés ».

- *L'employabilité, critère d'évaluation de l'efficacité de la formation?* : cette question renvoie à celle de la légitimité d'une évaluation « externe » qui consisterait à évaluer un système par un autre. Si P. BAILLY (« L'employabilité est-elle une évaluation de l'efficacité du système de formation? ») ne conteste pas, dans le cadre d'une École pratiquant la scolarisation de masse, la légitimité de l'évaluation du système de formation par un critère comme l'employabilité, il fixe les limites d'une approche « unidimensionnelle » qui ignorerait les aspects proprement éducatifs, culturels et sociaux de l'éducation. Analysant l'évolution des rapports entre système éducatif et système productif (« deux espaces sociaux qui se situent sur deux plans différents ») pour en faire ressortir la configuration sociale actuelle, il défend l'idée selon laquelle, « s'il est légitime de poser des questions sur l'efficacité économique et sociale du système de formation, cela ne signifie pas qu'il faille lui imposer des logiques du monde économique [...] ».

- *La rentabilité, critère d'évaluation de l'efficacité du système éducatif?* : l'éducation possède une dimension économique qui fait l'objet d'évaluations menées à partir de la notion de rentabilité. Ce critère peut-il permettre de mesurer l'efficacité du système éducatif? A. AYAICHIA (« La notion de rentabilité en éducation ») analyse les démarches et les méthodes utilisées dans ce domaine pour illustrer les difficultés d'une telle approche. Le critère de rentabilité apparaît ainsi « lié à une vision restrictive de l'éducation » car fondé sur « une rationalité individuelle » qui ignore les spécificités de l'éducation. Surtout, son caractère « unidimensionnel » ne considère qu'un aspect d'une question autrement plus complexe. Cette complexité et cette multidimensionnalité appellent plutôt une évaluation qui « requiert des analyses interdisciplinaires dans lesquelles l'économie prend sa place ».

L'ÉVALUATION DANS UN CADRE INTERNATIONAL

- *Les problèmes d'indicateurs : l'OCDE.* L'analyse dans une optique historique des activités d'évaluation de l'OCDE dans le domaine de l'éducation montre l'évolution des stratégies et des méthodes utilisées en fonction d'une demande de plus en plus forte de « reddition des comptes ». P. LADERRIERE (« Les activités d'évaluation à l'Organisation de Coopération et de Développement Économiques ») passe en revue les différents modes d'intervention de l'OCDE qui, rappelle-t-il, « tout en reconnaissant et en prenant nécessairement en compte la multidimensionnalité des objectifs de l'éducation, [...] centre essentiellement son action sur la dimension socio-économique ». S'agissant de la politique et des

stratégies de l'évaluation, les demandes ont amené l'OCDE à « analyser comment mieux évaluer les résultats du fonctionnement des systèmes d'éducation ». Les travaux effectués dans ce domaine mettent l'accent sur l'importance de « l'autoévaluation comme moment crucial devant précéder une évaluation externe et interagir avec elle ». La production d'indicateurs du système d'enseignement répond à des préoccupations centrées sur « la qualité, l'efficacité, voire l'efficience des systèmes d'éducation ». Cette production, « au carrefour du quantitatif et du qualitatif » appelle de grandes précautions quant à son utilisation.

- *Problèmes de la comparabilité* : divers facteurs contribuent à limiter la faisabilité des enquêtes internationales, pourtant indispensables au pilotage des systèmes éducatifs. BERTRAND et C. JEANRIE (« Comment faire échec aux contraintes limitant la faisabilité des enquêtes internationales: le cas des biais d'adaptation culturelle ») s'attaquent plus précisément au « problème de traduction ou mieux d'adaptation culturelle » des instruments de recueil de données en proposant des pistes pour obtenir « des versions vraisemblablement équivalentes ».

L'ÉVALUATION DU FONCTIONNEMENT

- *Le projet individuel et la différenciation sociale* : l'introduction d'un nouveau dispositif d'orientation scolaire, basé sur les notions d'éducation des choix et de projet individuel, avait pour objectif une meilleure prise en compte des aptitudes et des aspirations des élèves, ce qui supposait de la part des établissements le développement d'actions centrées sur le monde professionnel. Or, l'enquête effectuée par S. DEVINEAU (« L'éducation des choix. Le projet de l'élève : illusions individuelles et réalités de classe ») auprès d'une cohorte de 6 570 élèves, de la fin de l'élémentaire à la fin de la 5^e, complétée par un questionnement des enseignants, confirme l'hypothèse selon laquelle : « le rôle joué par le projet individuel de l'élève tend vers la transmission de l'acceptation, et l'adaptation à l'état de l'emploi, cette mission de l'école s'effectuant inégalement selon l'origine sociale des élèves ». L'analyse du fonctionnement et des effets du « mode de communication scolaire » révèle en effet qu'un tel processus de « construction de la conformité » paraît inéluctable car « sans le souci des contraintes qui font les goûts et aspirations des élèves d'origine sociale différente, en apparence neutre et universelle, l'éducation des choix oriente le projet de l'élève [...] par l'adaptation plus pessimiste que réellement pragmatique à la fatalité d'un destin social ».

- *L'école et l'entreprise* : les relations école-entreprise forment, en cette période de crise économique et de difficultés d'insertion, une des questions

les plus débattues au sein du monde éducatif. C'est en y travaillant, en lien avec le projet de l'élève, que P. LAMOUR (« Benchmarking et système universitaire finlandais ») a été amené à s'intéresser au concept de « Benchmarking » qui définit les principes permettant aux entreprises de se situer « dans une perspective de progression en excellence en vue de certification ». Ces principes (« apprendre des autres », « s'inspirer des meilleures pratiques », « évaluer en permanence ses performances », etc.) seraient susceptibles d'un transfert qui, au-delà du fonctionnement de l'école, agirait sur les méthodes d'apprentissage.

- *Le monitorat, un encadrement personnalisé* : le « métier d'étudiant » nécessite aussi un encadrement et un accompagnement. À l'issue d'une étude sur l'impact d'un programme de monitorat créé et encadré par les étudiants eux-mêmes, M. BRUNEAU (« Le monitorat : un accompagnement nécessaire à l'apprentissage et à l'enseignement »), souligne qu'il constitue « un facteur de réussite académique et de rétention aux études « en même temps qu'il contribue à l'évolution des discours et des pratiques pédagogiques ».

- *Éducation et socialisation* : La notion de vie scolaire désigne une fonction de socialisation à l'école, désormais conçue aussi comme un lieu de vie. La confrontation des textes prescriptifs et descriptifs suivie d'une étude de cas dans un collège et un lycée constitue une première approche de ce qui représente une extension du rôle de l'école et un réaménagement de ses tâches. V. BARTHELEMY (« Vie scolaire et socialisation : étude de la manière dont est prise en charge par l'établissement la socialisation de l'élève ») conclut son étude par une interrogation relative aux choix qu'entraîne cette évolution des missions de l'école : « Rôle purement éducatif ou rôle social? Peut-elle intégrer toutes les nouvelles données sociales? Ou ne serait-il pas préférable d'éviter de trop faire rentrer la société dans l'enceinte scolaire? »

4.2.3 L'évaluation dans le cadre de l'éducation-développement

- *Le contexte d'action éducative, un cadre pour l'éducation-développement* : la notion de système éducatif apparaît trop restrictive pour rendre compte des formes que revêt l'éducation dans les sociétés traditionnelles africaines. A. BABA-MOUSSA (« Analyse organisationnelle et référentialisation. Un référent théorique et une méthode pour évaluer les dispositifs d'éducation informelle au Bénin ») utilise les termes « d'action éducative » et « d'éducation-développement » pour tenter de cerner toutes les dimensions et toutes les interactions des associations de jeunesse avec le reste du « système éducatif » au Bénin. S'appuyant sur l'analyse organisationnelle et la référentialisation, il propose ici une démarche pour la construction

« d'un modèle d'analyse [...] permettant d'évaluer des dispositifs d'action informelle », révélant un autre modèle d'éducation que celui importé du système européen.

Les travaux sur l'évaluation des actions publiques font principalement ressortir une *problématisation* de la notion de contexte, désormais considéré comme un élément majeur pour la compréhension des processus mis en jeu dans la conduite d'une activité éducative, culturelle ou politique. L'interrogation sur le rôle et la fonction du contexte incite à une réflexion sur les conceptions des catégories du sujet et du social ainsi que sur les modalités de leur articulation. Cette problématique, aujourd'hui centrale dans les sciences sociales, se répercute en premier lieu sur le choix des méthodes de l'évaluation et permet par exemple une approche différente du débat quantitatif/qualitatif.

5. Questions portant sur les méthodes de l'évaluation

Il est toujours difficile d'isoler les méthodes comme caractéristiques dominantes des recherches : tous les travaux présentés utilisent, décrivent et justifient les méthodes.. Il ne faudra donc pas chercher ici un panorama des méthodes utilisées dans l'ensemble d'entre eux. Mais, s'agissant d'évaluation, les réflexions prolongeant des méthodes connues (la généralisabilité) ou proposant l'élaboration de méthodes nouvelles méritent d'être relevées comme telles.

- *Les apports du modèle de la généralisabilité à l'évaluation* : dans le prolongement des travaux de Cardinet, Allal et Tourneur, D. BAIN (« La généralisabilité, un modèle bien pratique pour élaborer ou évaluer vos évaluations ») rappelle que le principal apport de la généralisabilité réside dans son incitation à une plus grande rigueur dans le statut des données manipulées par le chercheur ou le praticien, corrigeant ainsi les faiblesses du modèle classique d'approche de la mesure. Il conclut ainsi la présentation des exemples illustrant les possibilités du modèle : « Le grand mérite de la généralisabilité est d'obliger le chercheur à se poser toutes sortes de questions, souvent fondamentales, sur les données qu'il traite et les conditions dans lesquelles elles ont été recueillies et échantillonnées ».

- *Les apports de nouvelles techniques de modélisation à l'évaluation en éducation* : Les modèles multiniveau, élaborés dans les années 80, constituent un progrès sensible dans l'approche des situations éducatives et pour la

recherche en sciences sociales en général. P. BRESSOUX (« Les modèles multiniveau pour analyser les effets de la formation des maîtres ») s'intéresse plus spécifiquement à l'application du modèle à la question de « l'effet-maître » et montre comment il permet de résoudre les problèmes rencontrés dans ce domaine, tels que « les biais d'agrégation », « la non-indépendance des résidus » ou encore « l'hétérogénéité des relations ». L'exemple des effets de la formation initiale des instituteurs sur les acquis de leurs élèves illustre la façon dont les modèles multiniveau peuvent « rendre compte de la dimension sociale de l'apprentissage et mettre en adéquation les traitements empiriques avec les théories des sciences sociales ».

- *Le traitement des données textuelles* : les données textuelles constituent un matériau privilégié pour l'évaluation d'un dispositif éducatif, dans la mesure où elles fournissent l'essentiel de l'information nécessaire à la construction du référentiel. L'importance de ce moment-clé de la démarche évaluative repose la question de la place respective des méthodes quantitatives et des approches qualitatives. À travers l'exemple d'une « analyse descendante hiérarchique utilisant le logiciel *Alceste* » appliquée à un dispositif de formation de chercheurs, P. PENNEL (« Les apports de l'analyse statistique des données textuelles. L'analyse descendante hiérarchique dans l'évaluation des dispositifs éducatifs ») montre les apports mais aussi les limites, dans ce domaine, d'une approche uniquement quantitative. L'analyse du corpus permet certes de vérifier la pertinence de l'outil; elle illustre aussi le fait que « cette approche générale du contenu d'un corpus textuel ne permet pas à l'analyste de repérer l'ensemble des informations spécifiques susceptibles d'éclairer une problématique donnée », rendant « souhaitable l'utilisation de méthodes complémentaires d'analyse de contenu ».

- *Une autre méthode de recherche en sciences de l'éducation : l'entretien complexe de recherche* : L'entretien semi-directif est une méthode couramment utilisée dans les recherches en sciences de l'éducation. Son étayage théorique est cependant rarement questionné, sa généralisation en garantissant en quelque sorte la validité. Amenée à effectuer pareil questionnement, M. KHAMKHADJI ZEMMOUR (« Proposition d'une méthode différente pour la recherche en sciences de l'éducation : l'entretien complexe de recherche ») met à jour un modèle linéaire de la communication conçue comme un simple outil de transmission de l'information d'un émetteur à un récepteur, modèle jugé incompatible avec les avancées actuelles des théories de la communication. Intégrant les concepts de « dialogue dialogique », de « co-construction » du sens par les sujets de l'interlocution, l'entretien complexe de recherche prend acte de

l'engagement du chercheur dans « la construction du sens et de la référence » et de la complexité de la situation ainsi élaborée. Non directement affiché comme une problématique d'évaluation, le sujet peut sans aucun doute s'y appliquer.

Les travaux consacrés à la question des méthodes de l'évaluation se signalent d'abord par le rôle accordé à la modélisation, entendue comme « une construction élaborée à partir d'hypothèses de recherche », qui vise à « [...] expliquer le mieux possible les phénomènes étudiés et à être le plus simple possible » (P. Bressoux). La sortie de cette contradiction par l'acceptation d'une certaine complexité propre à assurer une meilleure représentation de la réalité se traduit, dans ces travaux, par un dépassement de l'opposition traditionnelle entre qualitatif et quantitatif. S'appuyant sur les théories des sciences sociales, la modélisation utilise les approches quantitatives et qualitatives pour appréhender une réalité sociale comprise comme une construction interactive.

En conclusion : prospective et contrepoint

PROSPECTIVE

M. VIGEZZI (« Les rôles et les fonctions de la formation : prospectives ») étudie les « six formes et évaluations de la formation qui peuvent être imaginées et proposées comme scénarios globaux » à partir d'une analyse prospective des domaines économique et social d'une part, du travail et de l'emploi d'autre part. Soulignant en conclusion le risque inhérent à toute entreprise de prospection de l'incertain, il note que celle-ci est aussi « synonyme de prendre un rendez-vous » : « Mais cet incertain statistique n'est-il pas lui aussi élément d'évaluation puisqu'il permet de repérer les capacités d'invention d'un corps social? » Il n'est évidemment pas question de s'essayer à un pareil exercice de prospective au terme d'une étude conçue, rappelons-le, comme un compte rendu et une lecture des problématiques soulevées lors de ce X^e colloque de l'ADMEE. On peut cependant, au prix d'une perte des nuances, tenter de souligner des évolutions, caractériser des lignes de force qui apparaissent, à travers les travaux présentés, comme particulièrement significatives des tendances actuelles de la recherche en évaluation.

CONTREPOINT

Il est frappant de constater la part relativement réduite des travaux précisément et spécifiquement consacrés aux questions portant sur les

méthodes de l'évaluation. Non que ces questions soient pour autant absentes des recherches présentées. Mais intégrées aux problématiques soulevées, elles n'en constituent pas la préoccupation dominante. Autrement dit, les questions de méthode se trouvent subordonnées aux questions d'ordre épistémologique dont la prégnance paraît frappante. J. ARDOINO, dans « Contrepoint », met l'accent sur ce qui selon lui caractérise l'évolution en cours dans le domaine de l'évaluation : l'attention portée aux processus et non plus aux seuls produits, ce qui signifie une prise en compte de la complexité au détriment de la simplicité et du compliqué : « (...) on oscillait tout au plus, entre le simple et le compliqué [...] L'évaluation, dont nous sommes en train de parler, se propose bien, elle, une intelligence de la complexité, au sens propre de ce dernier terme ». L'accent mis sur la dimension éthique de l'acte évaluatif, l'importance accordée aux logiques d'organisation et de fonctionnement des sujets considérés dans leur contexte social, indiquent assez la recherche de sens à laquelle tendent aujourd'hui les travaux sur l'évaluation. Ce faisant, cette dernière intègre les avancées théoriques réalisées dans le champ des sciences sociales quant à la conception du langage et de la communication, au rôle des interactions sociales, à l'articulation de l'individuel et du social... Échappant aux tentations normatives et réductrices du contrôle et du positivisme, la recherche en évaluation intègre la complexité des phénomènes étudiés comme un point de vue nécessaire à leur compréhension.