

Méthodologie d'évaluation en éducation et formation

ou l'enquête
évaluative

Pédagogies en développement

Gérard Figari
Dominique Remaud
Avec la collaboration de
Claire Tourmen

Préface de Patrick Mayen



de boeck

Méthodologie d'évaluation en éducation et formation

Collection dirigée par
Jean-Marie De Ketele

Comité scientifique international
Daniel Chevrolet (Rennes II); Claude Tapia (Tours); Ben Omar Boubker (Rabat);
André Girard (Québec); Claudine Tahiri (Côte d'Ivoire)

Méthodologie d'évaluation en éducation et formation

ou l'enquête
évaluative

Préface de Patrick Mayen

Gérard Figari
Dominique Remaud
Avec la collaboration de
Claire Tourmen

Pour toute information sur notre fonds et les nouveautés dans votre domaine de spécialisation, consultez notre site web: www.deboeck.com

© De Boeck Supérieur s.a., 2014
Rue des Minimes 39, B – 1000 Bruxelles

1^{re} édition

Tous droits réservés pour tous pays.

Il est interdit, sauf accord préalable et écrit de l'éditeur, de reproduire (notamment par photocopie) partiellement ou totalement le présent ouvrage, de le stocker dans une banque de données ou de le communiquer au public, sous quelque forme et de quelque manière que ce soit.

Imprimé en Belgique

Dépôt légal:

Bibliothèque nationale, Paris: février 2014

Bibliothèque royale de Belgique, Bruxelles: 2014/0074/082

ISSN 0777-5245

ISBN 978-2-8041-8477-3

Sommaire

Sommaire	5
Remerciements	7
Préface	9
<i>Patrick MAYEN</i>	11
Introduction	13
Partie 1. Les fondements de l'activité évaluative	19
Introduction	21
1 L'évaluation en débats	25
2 Des cultures de l'évaluation ?	33
3 Des définitions de l'évaluation	39
4 Évaluation et théorie	49
5 L'évaluation, une activité langagière	53
Conclusion : une évaluation influencée par ses fondements théoriques	57
Partie 2. Méthodologie d'enquête évaluative par la référentialisation	59
Introduction	61
6 La place des référentiels dans l'évaluation	65
7 Méthodologie d'évaluation par la référentialisation	71
Conclusion : le sens et l'intérêt d'une « enquête évaluative »	105

Partie 3. Méthodologie appliquée à différents champs de l'éducation et de la formation	107
Introduction	109
8 Première application : Évaluation de politiques éducatives	113
9 Deuxième application : Évaluation de projets éducatifs	127
10 Troisième application : Évaluation des dispositifs et programmes	139
11 Quatrième application : Évaluation des outils pédagogiques	151
12 Cinquième application : Évaluation de compétences	163
Conclusion : référentialisation et exemplification	177
Conclusion	179
Bibliographie	185
Table de figures	195
Table des matières	197

Remerciements

Cet ouvrage a bénéficié de concours précieux ayant permis de confronter des concepts, des méthodes et des expériences qui en constituent le corps essentiel.

Que soient remerciés, ici, ceux qui ont bien voulu y participer :

- Claire Tourmen, qui a accepté de rédiger le chapitre sur l'évaluation des compétences ;
- Gildas Gautier, qui a contribué activement au lancement du projet, en y apportant sa large expérience et sa réflexion ;
- Michel Jolland, qui a su exercer, de bonne grâce et avec pertinence, un jugement critique éclairé, au moment des choix organisateurs du texte ;
- Patrick Mayen, qui a suivi, à l'occasion de discussions sans concessions, l'évolution de la conceptualisation ayant marqué la pensée conductrice de l'ouvrage et a rédigé une préface qui en enrichit le sens même.

Préface

La méthode de la référentialisation :
principes et instruments pour une conduite
démocratique de l'enquête évaluative

« Quand nous disons que la pensée et les croyances devraient être expérimentales, et non absolutistes, c'est une certaine logique de la méthode que nous avons à l'esprit, et non prioritairement l'exécution d'une expérimentation du type de celles qui sont menées en laboratoire. Une telle logique implique les facteurs suivants : premièrement, que les concepts, les principes généraux, les théories et les développements dialectiques qui sont indispensables à toute connaissance systématique soient formés et mis à l'épreuve en tant qu'outils d'enquête. Deuxièmement, que les mesures politiques et les mesures d'action sociale soient traitées comme des hypothèses de travail, non comme des programmes auxquels il faudrait adhérer et qu'il faudrait exécuter de façon rigide. Elles seraient expérimentales au sens où l'on admet qu'elles soient sujettes à une observation constante et bien outillée des conséquences qu'elles produisent quand on agit sur elles, et qu'elles soient sujettes à une révision prompte et flexible à la lumière des conséquences observées. Si ces deux conditions étaient remplies, les sciences sociales pourraient former un dispositif capable de conduire les investigations, d'enregistrer et interpréter (organiser) leurs résultats. Dans ce cas, on ne considérerait plus ce dispositif comme la connaissance elle-même, mais comme l'ensemble des moyens intellectuels destinés à découvrir les phénomènes ayant une portée sociale et à comprendre leur signification » (Dewey, 2010, pp. 304-305).

Le livre qui est proposé ici présente ce qu'on pourrait appeler la méthode de la référentialisation, créée et développée par Gérard Figari. Plus précisément : créée, expérimentée et développée par Gérard Figari et par ses collègues. La méthode de la référentialisation constitue un cas de ce que John Dewey appelle de ses vœux dans les propos cités en introduction de cette préface. Elle répond aux deux conditions qu'il énonce : les principes, les concepts qui la fondent ont été formés et mis à l'épreuve en tant qu'outils d'enquête, à la fois dans un grand ensemble de chantiers effectivement conduits et présentés ici, mais ils ont été mis à l'épreuve par les travaux de recherche et l'analyse rigoureuse auxquels ils ont été soumis. C'est une méthode car elle est tournée vers l'action, mais c'est une méthode qui dépasse le niveau des règles d'action pour se tenir au niveau

conceptuel. C'est ce qui en fait la force et l'originalité. La méthode de la référentialisation est à la fois expérimentée au sens où elle a été utilisée et mise à l'épreuve de l'intervention sociale dans le domaine de l'éducation, et examinée et mise à l'épreuve de l'analyse et de la réflexion conceptuelle et méthodologique. C'est aussi une méthode ouverte et non un protocole standardisé ou une procédure. Ce qui signifie qu'elle donne les moyens pour être réinventée et développée selon les caractéristiques des situations dans lesquelles elle peut être utilisée. Ceux qui chercheraient un protocole d'action standardisé pourraient découvrir assez vite qu'ils se sont trompés.

La méthode de la référentialisation répond aussi à la deuxième condition formulée par John Dewey : elle considère précisément que les mesures politiques et les mesures d'action sociale, comme il l'écrit, doivent être traitées comme des hypothèses. Ce qui est une condition nécessaire pour qu'elles soient effectivement soumises à l'évaluation, c'est-à-dire à l'observation des conséquences qu'elles produisent quand on agit sur elles.

La méthode de la référentialisation est une proposition pour traiter les dispositifs, les programmes et les politiques de formation comme des hypothèses, et pour les soumettre à l'évaluation. Ce qui est une proposition ambitieuse. Le livre présente la grande qualité de mettre à nu les principes, les concepts, les phases de l'action évaluative dans le cadre de la méthode de la référentialisation, et de les soumettre à l'épreuve de situations d'évaluation effectivement conduites par les auteurs. Comme l'écrit Gérard Vergnaud (2008), un concept est aussi l'ensemble des situations pratiques qu'il permet de traiter et qui contribuent à valider sa pertinence.

La méthode de la référentialisation est un ensemble organisé de moyens pour conduire des enquêtes évaluatives, autrement dit, pour reprendre encore John Dewey, comme des possibilités de transformation et de reconstruction du matériel qu'elles traitent. Ce qui fait la force de la référentialisation, c'est de construire et reconstruire l'objet, les fins et les référents de l'évaluation, autrement dit de construire la méthode pour chaque situation nouvelle qui se présente. C'est une manière de penser et d'aborder les problèmes qui s'oppose à plusieurs autres approches : celle de *l'expertise* toute-puissante de l'évaluateur qui aurait la méthode universelle applicable à toute situation, celle de la *spontanéité* qui, ignorant tout des phénomènes à propos desquels l'évaluation est conduite, produit des jugements de bon sens, celle qui, enfin, part de l'idée que l'objet, les fins et les référents sont déjà là, donnés, et qu'il ne reste qu'à *en faire la cueillette avant de mesurer et de juger*. La référentialisation pourrait aussi être définie comme la méthode de construction et de reconstruction des significations. Celles-ci n'étant jamais données une fois pour toutes, les critères et les indicateurs sont toujours à définir et à redéfinir en fonction des finalités, des conditions et des actions telles qu'elles se manifestent effectivement dans la dynamique des processus.

La référentialisation est une pratique conceptualisée, et, dans le livre proposé ici, les auteurs font le pari de nous faire partager les moyens de nous l'approprier. Dans ce sens, ils sont cohérents avec un autre aspect de la méthode elle-même : son caractère démocratique. On peut retrouver dans le texte de nombreux passages où l'on comprend que la référentialisation consiste fondamentalement à aider les acteurs concernés par une mesure politique, une action, un dispositif, à être partie prenante de l'enquête évaluative et à tirer, pour eux-mêmes et leur action à venir, les conséquences de ce qu'ils auront compris par la participation à l'enquête. Il y a aussi là quelque chose de démocratique dans la référentialisation.

On aura compris que le concept et la méthode de référentialisation méritent beaucoup d'attention et constituent une ressource précieuse pour l'action et la réflexion dont j'ai pu éprouver la pertinence et la richesse dans plusieurs chantiers que j'ai partagés, notamment avec Gérard Figari, sur des questions de Validation des Acquis de l'Expérience, de conception de référentiels et de situations d'évaluation de compétences professionnelles.

Patrick MAYEN

Introduction

1 ORIGINES CROISÉES D'UNE FAÇON DE PENSER LA MÉTHODOLOGIE DE L'ÉVALUATION EN ÉDUCATION

Cet ouvrage est l'aboutissement du cheminement lent et ardu d'une façon de penser la méthodologie de l'évaluation en éducation. Ce cheminement s'est déroulé à travers une longue expérience de pratiques évaluatives dans l'enseignement secondaire, l'enseignement supérieur et la formation professionnelle ainsi que selon une non moins longue succession de recherches sur le sujet. Une approche de l'évaluation par la « référentialisation » (notion définie p. 72) a fait l'objet d'une première présentation théorique (Figari, 1994) mais s'est considérablement modifiée, enrichie et structurée au point de rendre nécessaire le passage de l'idée initiale à sa traduction méthodologique.

Nous arrivons, en effet, au point d'orgue d'une double histoire : d'abord, celle d'une rencontre à la fois scientifique et pédagogique entre un enseignant-chercheur et trois de ses anciens étudiants (dans le cadre d'un master et d'un doctorat de l'Université de Grenoble) qui ont décidé de confronter leur expérience du même sujet constitué par l'évaluation en éducation et formation ; ensuite, celle de l'évolution d'une pensée sur l'évaluation.

L'enseignant-chercheur, premier auteur de ce livre, a étudié, pendant de longues années, le rôle des référentiels dans l'évaluation ; pendant de longues années, il a proposé dans ses cours et ses interventions extérieures, un travail de conception critique et partagée des référentiels d'évaluation dans diverses situations scolaires, universitaires ou relevant de la formation professionnelle. Et comme des tableaux ou des supports de conférences qui en sont les traces ont commencé à être reproduits, ici et là, par des participants à ces activités, il était temps de livrer ses notes, expériences et réflexions en même temps que de résumer ses travaux sur ces sujets. L'occasion se présente donc, ici, de réaliser cette mise à plat.

Les anciens étudiants, étant devenus consultants¹ ou maîtres de conférences², ont utilisé, dans leur métier, des procédures du type de celles qui vont être développées ci-dessous et avaient le désir de faire part de leur expérience des pratiques d'évaluation dans les milieux de l'éducation et de la formation. Leur vision plus professionnelle venait interpellier les modèles théoriques du chercheur et l'idée est née qu'une confrontation des deux types d'expériences pouvait donner lieu à la publication d'une *méthodologie de l'évaluation en éducation* équilibrée et, somme toute, plus utile.

Les quatre interlocuteurs ont partagé une double orientation : d'abord, celle d'une *pensée sur l'évaluation*, construite au fil des années d'expérience, consistant à considérer qu'il s'agit d'un phénomène partagé par des acteurs, d'une « pratique dialogique » (Mottier Lopez, 2013b) accompagnée d'une méthodologie interactive et qui nécessite une remise en cause permanente, aidée en cela par les apports de la recherche³. Cette pensée est principalement axée sur le rôle des référentiels (explicites ou implicites) dans les pratiques d'évaluation en éducation. La deuxième orientation consiste en un intérêt commun pour une *pédagogie de l'évaluation*. Comment partager une vision de cette dernière telle qu'elle vient d'être esquissée et telle qu'elle va être développée dans le détail, avec des étudiants, des enseignants, des éducateurs, des décideurs, des experts, des consultants, des inspecteurs, etc. ? Comment présenter et expliquer les possibilités de recentrage méthodologique dans ce domaine ? Et, enfin, avec la même armature, comment aborder, accompagner et analyser des activités d'évaluation ?

Le croisement des expériences de formation, d'expertise et de recherche des quatre protagonistes sur ce thème a débouché sur l'idée et la réalisation du présent ouvrage, avec le second auteur.

2 MÉTHODOLOGIE DE L'ÉVALUATION : PAS SEULEMENT UNE TAUTOLOGIE

Avant d'entreprendre la lecture d'un livre sur la « méthodologie de l'évaluation », on peut déjà s'interroger sur le fait de savoir si l'évaluation n'est pas, d'abord et avant tout, elle-même, une méthodologie. Certes, cette

1. Il s'agit de Dominique Remaud (second auteur) et de Gildas Gautier (ayant participé au lancement de ce projet).
2. Il s'agit de Claire Tourmen ayant collaboré à l'ouvrage par l'écriture du chapitre sur l'évaluation des compétences.
3. Les quatre protagonistes ont participé au GREIF, groupe de recherche sur l'évaluation et l'ingénierie de la formation au Laboratoire des Sciences de l'Éducation de Grenoble Alpes, au début des années 2000. Cf. notamment des publications sur la recherche en évaluation : Figari & Mottier Lopez (2006) ; Mottier-Lopez & Figari (2012) ; Tourmen (2009).

dernière est partout dès que l'on parle de ce sujet et c'est peut-être bien ce qui pose problème : la chose paraît tellement évidente qu'on pourrait être tenté de faire l'économie d'un questionnement, d'une remise en cause, d'une recherche quelconque de fondements, pour se demander si les outils ne sont pas trompeurs, si d'autres phénomènes que la simple mesure n'interviennent pas dans le processus de l'évaluation lui-même. Avec Machado (2007) on se demande, néanmoins, si un « cadre d'interprétation cohérent et fondé sans lequel l'évaluation devient subjective et arbitraire » n'est pas nécessaire à la pratique de n'importe quelle méthode. Même s'il faut bien reconnaître que les « cadres d'interprétation » sont potentiellement nombreux à servir de référence à une méthodologie de l'évaluation et que la question de savoir comment en choisir un reste entière.

On voit déjà s'esquisser la thématique de notre propos : la méthodologie de l'évaluation n'est pas une tautologie ; l'évaluation est un processus qui ne se réduit pas à une méthode et il existe, assurément, au sein même de son déroulement, un espace pour la réflexion et la conceptualisation. On retrouvera, au long de l'ouvrage, cette volonté de commencer par la réflexion et la conceptualisation avant de construire et d'utiliser une méthode.

3 UN LIVRE DE « MÉTHODOLOGIE GÉNÉRALE SUR L'ÉVALUATION EN ÉDUCATION »

Il existe déjà une littérature variée décrivant, de manière distincte, des méthodes d'évaluation des apprentissages, des établissements, des politiques éducatives. On y fera référence, au fur et mesure de l'exposé. Mais que peut-on répondre à un professionnel de l'éducation et de la formation qui demande un ouvrage introductif à la mise en route d'un dispositif d'évaluation, qui puisse proposer comment commencer et comment finir, qui puisse ouvrir à la réflexion et au doute un ensemble de pratiques ancrées dans le fonctionnement social et qui ne rencontrent pas toujours d'adhésion ou de compréhension ? Il manquera, toujours, dans le paysage éditorial de l'évaluation un ouvrage collectif, consensuel, organisant autour des différents domaines éducatifs, la présentation d'une méthodologie générale à l'usage des étudiants, chercheurs, enseignants, formateurs et autres professionnels ayant en charge des activités de ce type. Ce livre ne prétend pas en tenir lieu et, sur ce thème, en appelle d'autres, plus complets. Mais il constitue, pour l'instant, une proposition de « méthodologie générale » de l'évaluation en éducation à l'intention des praticiens et des chercheurs intéressés par une approche globale de cette activité.

Précisons que cette méthodologie a été conçue et expérimentée dans le champ des *dispositifs éducatifs* et que, si elle peut s'appliquer à des objets très divers comme les *politiques*, les *projets*, les *dispositifs*

ou programmes, les outils pédagogiques ou les compétences, elle ne concerne les apprentissages qu'indirectement, c'est-à-dire par l'intermédiaire des objets précédents qui les influencent et les organisent. On ne s'occupera pas de l'évaluation des séquences didactiques au sein de la classe, traitant d'une discipline particulière, mais on s'intéressera au projet pédagogique qui favorisera un apprentissage, au programme qui orientera des stratégies d'apprentissage, aux outils pédagogiques qui serviront de supports et d'aides aux apprentissages ainsi qu'aux effets de ces apprentissages en termes d'acquisition de compétences.

4 UNE VOLONTÉ DE « FONDER » SUR DES CONCEPTS TOUTE MÉTHODOLOGIE DE L'ÉVALUATION

Réfléchir avant d'instrumenter : cette orientation existe déjà dans les travaux d'Allal (2008). Rien de nouveau ni d'original à cela, certes. Mais nous nous permettrons deux remarques à ce sujet : la première est qu'il peut exister de grands écarts entre le dire et le faire et la seconde est que la simple réflexion ne nous semble pas suffire en l'occurrence. Seule la deuxième remarque mérite un certain développement. Dans la continuité avec le concept de « praticien réflexif » lancé par Schön (1993), on parlera, ici, de *l'évaluateur réflexif*, dans les deux sens que revêt cette expression : le premier sens donné par l'auteur « comment les professionnels pensent dans l'action » et le second sens « comment les professionnels réfléchissent sur leur action ». L'évaluation réflexive sera présente tout au long de l'argumentation en faveur du rôle de la conceptualisation, enclenchée dès la première partie dans le cadre d'une revue des concepts sous-jacents aux débats et aux préoccupations des praticiens de l'évaluation et intégrée, par la suite, dans la présentation des différentes phases méthodologiques.

5 UNE APPROCHE INFLUENCÉE PAR LA THÉORIE DE L'ENQUÊTE DE DEWEY

En concevant l'évaluation comme une pratique de pensée s'inscrivant dans une recherche de sens des situations particulières, nous nous rapprochons de la définition que donne Dewey de *l'enquête* comme étant la transformation d'une situation indéterminée en une situation clarifiée grâce à une observation rigoureuse des faits et dans une interaction permanente avec le contexte. Ce que Gros (2011, § 13-18) décrit comme le *schème de l'enquête* chez cet auteur constitue un modèle particulièrement adaptable à l'approche de la méthodologie de l'évaluation que nous proposons dans les chapitres qui suivent. Selon l'auteur, les étapes structurelles du processus d'enquête se succèdent ainsi : 1) l'antécédent de l'enquête (une situation indéterminée) ; 2) l'institution d'un problème ; 3) la détermination de

la solution du problème (= l'observation des faits en corrélation avec la suggestion des idées) ; 4) le raisonnement (= l'examen des hypothèses) ; 5) l'expérimentation (test des hypothèses par rapport aux faits observés). On va retrouver dans ce qui va être appelé « référentialisation » de l'évaluation une similitude de démarche intellectuelle avec le schème de l'enquête, selon Dewey.

6 D'AUTRES INFLUENCES À ASSUMER

Il est notoire que les approches de l'évaluation en éducation ont considérablement évolué depuis quelques années. De nouveaux courants de pensée ont envahi ce domaine : ils sont largement représentés dans la publication internationale de langue française de l'ADMEE (Europe et Canada), la revue *Mesure et évaluation en éducation*, qui accueille les travaux des chercheurs spécialisés sur l'évaluation et à laquelle nous sommes redevables de leur actualisation. Ne pouvant les citer tous, nous renvoyons à la lecture du numéro spécial qui a marqué les trente ans de la revue⁴ ; les travaux sur les *apprentissages situés* (Allal & Mottier Lopez, 2007) ont provoqué des retombées sur l'évaluation (Allal, 2007) ; le repérage des phénomènes de la *métacognition* (ou prise en compte des processus cognitifs) a redonné à l'apprenant un rôle essentiel dans le processus d'évaluation (Noël, 1997) ; l'introduction du concept déjà évoqué de réflexivité (Schön, 1993) appliqué à la lecture des pratiques fait penser, tout naturellement aussi, à l'*évaluateur réflexif*. La didactique professionnelle avec ses analyses nouvelles de la compétence et de l'expérience bouleverse toutes les démarches d'évaluation tributaires des anciennes cultures docimologiques de mesure et de contrôle des performances. « Elle cherche à articuler de façon très forte deux dimensions qui ne vont pas forcément ensemble : la dimension théorique et la dimension opératoire... » À ce propos, elle « présente un cadre théorique de la conceptualisation dans l'action » (Pastré, Mayen, & Vergnaud, 2006). Dans le même champ d'études, les travaux de Mayen sur l'expérience (Mayen, 2009b), en rencontrant les conceptions de Dewey (1967), ont apporté leur contribution à la notion d'« expérience d'évaluation » qui sera invoquée, au moment de synthétiser l'ensemble des opérations expérimentées lors du déroulement des différentes phases de la méthodologie exposée ci-après.

7 UNE EXPOSITION EN TROIS MOMENTS

Cet ouvrage comprend trois moments, présentés dans trois parties ainsi organisées :

4. Numéro thématique de la revue *Mesure et évaluation en éducation*, paru en 2008 (vol. 31, n° 3).

Une *première partie*, intitulée « Les fondements de l'activité évaluative », questionne les différentes modalités de réflexion, de débats, de définitions, de culturalité spécifique relevant de l'évaluation en éducation, pour progresser vers un raisonnement théorique général, en suivant un slogan majeur : « des concepts avant des outils ».

Une *deuxième partie*, intitulée « Méthodologie d'enquête évaluative par la référentialisation », développe, après avoir analysé la place occupée par les « référentiels » dans le monde de l'enseignement et de la formation, un protocole méthodologique d'évaluation jouant sur trois dimensions : la conceptualisation, la modélisation, l'instrumentation.

Une *troisième partie*, intitulée « Méthodologie appliquée à différents champs de l'éducation et de la formation », décrit et analyse cinq cas d'évaluation ayant suivi, approximativement, le protocole précédent. Il s'agit d'exemples d'évaluation d'une politique éducative, d'un projet éducatif local, d'un curriculum d'enseignement professionnel, d'un ensemble d'outils pédagogiques et, enfin, de compétences.

8 TROIS CHEMINS DE LECTURE

Il nous paraît possible de lire cet ouvrage en suivant au moins trois chemins différents :

- la première possibilité consiste à lire les trois parties dans l'ordre logique (1, 2, 3) en s'appuyant sur un raisonnement allant du plus conceptuel au plus concret ;
- la deuxième possibilité consiste à lire selon une stratégie privilégiant la familiarisation avec ce qui constitue l'apport central : la méthodologie de l'enquête évaluative par la référentialisation, en suivant, donc, pour entrer immédiatement dans le vif, l'ordre 2, 1, 3 ;
- la troisième possibilité revient à lire à l'envers, dans l'ordre 3, 2, 1, pour commencer par des pratiques, retrouver la méthodologie qui était censée les organiser et prendre alors du recul pour aborder des conceptualisations.

Partie

1

LES FONDEMENTS
DE L'ACTIVITÉ ÉVALUATIVE

Introduction

1 UNE GRANDE DIVERSITÉ DE PRATIQUES D'ÉVALUATION

L'évaluation, dans le contexte éducatif, donne lieu à toutes sortes de conceptions, de pratiques, d'attitudes, de méthodes, portées par des institutions ou des individus, chacun proférant sa définition et légitimant ses pratiques et ses techniques. Dans cette jungle, où tous les chemins sont possibles même s'ils ne se croisent pas toujours, on entend donc parler tous les langages, utiliser une diversité infinie d'outillages, se réclamer de courants ou d'idéologies très différents selon que l'évaluateur se situe dans tel ou tel univers :

- celui de l'institution politique qui cherche à réguler le fonctionnement d'un système en se référant à tel ou tel courant national ou international ;
- celui du chercheur attaché à expliquer les résultats des dispositifs et des méthodes d'éducation et de formation en s'inspirant de méthodologies se réclamant de disciplines diverses, sociologiques et/ou psychologiques, notamment ;
- celui du décideur qui cherche à atteindre ses objectifs en s'appuyant sur les informations synoptiques qu'il a réunies dans son tableau de bord ;
- celui du pédagogue soucieux d'apprécier les progrès des apprentissages de ses élèves, en cherchant des outils dans des écrits spécialisés ;
- celui de l'expert ou du consultant devant répondre à des demandes d'audits institutionnels ;
- celui de l'apprenant impliqué dans la compréhension de ses difficultés d'apprentissage et dans la mise en valeur de ses compétences.

Par ailleurs, à l'intérieur même de chaque univers, l'interprétation et la mise en œuvre d'une évaluation sur un même objet peuvent varier d'un praticien à l'autre : rythme de passation différent, explicitation aléatoire des critères et des règles du jeu, niveau de participation inégal des acteurs et/ou des apprenants au processus même d'évaluation.

2 DES PRATIQUES PAS TOUJOURS RÉFLÉCHIES

Dans le monde scolaire, en général, mais aussi dans le monde professionnel, les pratiques d'évaluation pullulent d'autant plus qu'elles sont vécues comme allant de soi, attendues et obligatoires, séculaires et légitimes : en un mot, une *doxa*¹ tellement établie que les praticiens n'ont parfois même pas l'idée de la remettre en cause ni, au moins, de lui appliquer la moindre mise en doute. Les nombreux débats qui la mettent en scène remplacent trop souvent une démarche réfléchie sur l'évaluation : il n'est que de citer quelques publications au titre évocateur qui peuvent contribuer à jeter le doute sur la fonction intrinsèque ou l'utilité de l'évaluation : « Évaluer tue » (Pfauwadel, 2010) ou « La tyrannie de l'évaluation » (Del Rey, 2013). Ce phénomène est tellement répandu qu'il atteint même les sphères de la recherche, qui hésite parfois à traiter la relation entre évaluation et science autrement qu'en les distinguant de manière prudente.

Quant à la prédominance du recours systématique et immédiat aux outils, elle révèle un autre aspect de la *doxa* qui relaie cette confusion entre évaluation et instrumentation, c'est-à-dire une autre façon d'évaluer sans réfléchir, ou, du moins, sans invoquer de théorie suivie.

3 UN BESOIN DE VISIONS CLAIRES ET GLOBALES

Pourquoi, aujourd'hui, l'évaluation en éducation, en particulier, se pratique-t-elle de manière si éclatée ? Comment l'acteur éducatif, débutant ou réflexif², peut-il se situer et se frayer un chemin dans ce labyrinthe ? Comment peut-il affronter cette crise de sens (Figari, 2009) affectant un champ d'action aussi utilisé et aussi important dans le pilotage du système éducatif en général ? On ne pourra pas développer ici toutes les réponses apportées sur ce sujet : il s'agit d'une question relevant à la fois de l'histoire des disciplines des sciences humaines et sociales et de ce qu'il faudra bien appeler, un jour, l'épistémologie de l'évaluation. Le lecteur intéressé par le débat en cours à ce propos pourra consulter l'ouvrage de « Raisons éducatives », *Modélisations de l'évaluation en éducation* (Mottier Lopez et Figari, 2012).

Il n'en reste pas moins que toute tentative de repenser la pratique de l'évaluation nous paraît pertinente. Avant d'aborder la dimension méthodologique du processus qui fera l'objet de la deuxième et de la troisième partie de cet ouvrage, nous avons besoin de voir clair dans cette multitude de pratiques et de trouver un fil directeur capable de relier les discours des

1. La *doxa* est une opinion confuse sur des aspects de la réalité, construite sur des représentations collectives socialement reconnues mais ne relevant pas généralement de la science.

2. Au sens de Schön (1993)

différents partenaires de l'évaluation en éducation. Enfin, en même temps que « voir clair », nous avons aussi besoin de « voir large », c'est-à-dire de nous doter de visions globales des situations d'évaluation :

- un apprentissage ? Avec une prise en compte de ses dimensions cognitives et sociales ;
- un dispositif ? Avec une inscription dans une perspective situationnelle, systémique et politique ;
- une compétence ? Avec un détour par l'analyse du travail et plus généralement, la didactique professionnelle, pour comprendre les processus de construction des compétences (Pastré, Mayen et Vergnaud, 2006).

4 LA NÉCESSITÉ D'APPUYER LES MÉTHODES D'ÉVALUATION SUR DES CONCEPTUALISATIONS GÉNÉRATRICES DE SENS

C'est dans ce but que le présent chapitre s'organise : avant de se précipiter dans les choix méthodologiques et instrumentaux, il s'agira pour concevoir une évaluation, quelle qu'elle soit, de poser, d'abord, la question de son sens.

Et, parce que l'évaluation est d'abord un concept et que même les praticiens qui agissent sans recours explicite à une théorie utilisent cependant des concepts en actes et des représentations sociales, on visitera un certain nombre de ces concepts organisateurs de ce qui tient lieu d'une réflexion théorique. Ainsi, ce chapitre traitera essentiellement de la manifestation du rôle de la conceptualisation dans l'organisation d'une « réalité », celle des activités d'évaluation, à travers le filtre de plusieurs thèmes dont il est souvent question sur le terrain.

Dans ce but, on partira des débats bien connus qui agitent le petit monde de l'évaluation (chap. 1) pour décrypter les conceptions sous-jacentes qui s'y affrontent. On évoquera ensuite les discours sur les cultures de l'évaluation qu'affectionnent les institutions nationales et internationales pour repérer les influences qu'elles peuvent exercer sur les professionnels (chap. 2). On se demandera également quel peut être le statut des nombreuses définitions de l'évaluation inaugurant presque tous les documents qui en relèvent (chap. 3). On examinera, enfin, avec intérêt, le rapport que l'évaluation entretient avec la théorie afin d'essayer de formuler les fondements des méthodologies qui seront proposées et appliquées dans les parties 2 et 3 de l'ouvrage (chap. 4).

Chapitre 1

L'évaluation en débats

Le premier niveau de réflexion qui s'offre aux évaluateurs est constitué de « débats », de controverses. Ces controverses tentent toujours les pédagogues : même s'ils risquent de s'y enfermer, ces oppositions binaires peuvent leur servir de prises de conscience, de remises en question et, tout compte fait, de réflexions que nous appellerons « préthéoriques ». Parmi elles, citons « mesure *versus* évaluation », « contrôle *versus* évaluation », « quantitatif *versus* qualitatif », « objectif *versus* subjectif », « externe *versus* interne ». Examinons-les et voyons en quoi elles peuvent servir de « pensée provisoire » sur l'évaluation.

1 MESURE/ÉVALUATION

La mesure est un ensemble de procédures d'observation, de recueil et de traitement de données à partir desquelles l'évaluation va pouvoir se construire ; sans s'y opposer, elle n'est pas à confondre avec l'évaluation. Dans le système scolaire traditionnel, elle a longtemps été assimilée à la notation que la docimologie des années 1970 (De Landsheere, 1972 ; Bonboir, 1972) a longuement analysée et critiquée. Il a été prouvé, à l'époque, que la note était injuste, dangereuse et faussée, ne serait-ce que parce qu'elle ne pouvait servir qu'à un classement des élèves autour d'une courbe de Gauss (présentation graphique d'une répartition « normale » des sujets mettant en valeur le fait que, statistiquement, les élèves très faibles et très forts sont toujours peu nombreux alors que la majorité d'entre eux se

retrouvent autour de la moyenne), courbe qui ne rendait pas compte des marges de progrès des plus faibles par rapport aux plus forts.

La notation n'est, en effet, que le résultat d'une mesure ayant l'apparence de l'objectivité et de la rationalité : on sait, depuis l'époque de la docimologie critique, que les notes ne sont pas fiables (exemple connu, celui de la Commission française de l'enquête Carnegie, 1936, qui avait noté qu'il faudrait 127 corrections pour arriver à cerner la valeur d'une copie de philosophie, et 13 pour une copie de maths, par exemple). Ces travaux ont d'ailleurs été repris et confirmés par Merle (1996) et Suchaut (2008). « Les chercheurs identifient, en général, trois sources d'erreur », selon Demeuse *et al.* (2004) : « les enseignants¹, le système scolaire et les élèves ». En effet, on a noté un certain nombre d'influences pouvant constituer des sources d'erreur dans l'évaluation des élèves, intervenant sur ces trois facteurs :

- l'influence des intérêts et des valeurs propres de *l'enseignant* ainsi que des arrangements et des procédures d'ajustement des résultats qu'il est amené à pratiquer ;
- l'influence du *contexte scolaire* sur les procédures d'évaluation en général (classe, établissement) ;
- l'influence de l'origine sociale et du comportement de *l'élève*.

Face à ces critiques, les tenants d'une évaluation sans notes, organisée autour de critères, dans le cadre d'une régulation interactive, corrective, métacognitive et valorisante, ont beau jeu de proposer des systèmes *par compétences* (cf. expérimentations de l'évaluation par compétences : Eduscol, 2011). Mais ne s'agit-il pas, néanmoins, d'une autre forme de mesure ? Faire le bilan de l'acquisition progressive d'une compétence à l'aide d'une *grille de compétences* ayant recours à une symbolique par lettres qu'il s'agit d'accumuler en vue de constituer un *capital* d'éléments de compétences, ne revient-il pas également à opérer une *mesure* sur des variables qualitatives ?

On voit bien, ici, que le débat sur la confusion entre mesure et évaluation n'est pas clos. Même s'il n'occupe plus la première place, il reste l'enjeu des dispositifs d'évaluation lancés encore trop rapidement par les institutions à la recherche de légitimité. En tout cas, l'idée que l'évaluation ne peut pas se passer de mesure s'impose *a fortiori*, à partir du moment où l'on postule que le jugement évaluatif n'est acceptable par tous les partenaires du système scolaire et de formation que s'il est étayé sur des données observables, quantifiables et comparables, issues d'une opération de mesure. La question qui demeure alors consiste à se demander quoi et comment mesurer l'objet à évaluer.

1. Cf. « l'effet Posthumus » commenté par Crahay (1996).

2 CONTRÔLE/ÉVALUATION

Une deuxième opposition, longtemps féconde, a nourri les débats sur les postures d'évaluation : celle entre *contrôle* et *évaluation*, opposition qui fonctionne encore aujourd'hui. Ardoino et Berger (1989) en sont les commentateurs les plus connus. Le contrôle y est défini comme « un dispositif... constitué par un ensemble de procédures ayant pour objet d'établir la conformité... entre une norme...un modèle et les phénomènes ou les objets auxquels on les compare » (p. 12), alors que l'évaluation constitue « une appréciation ou une estimation ».

La mise en valeur de cette opposition a permis de sensibiliser les opérateurs à l'illusion que pouvait représenter le fait de nommer évaluation une procédure qui ne consistait, en réalité, qu'en une série de vérifications de conformité à des normes préétablies. Le problème soulevé est que l'opposition contrôle-évaluation sous-entend que cette dernière exclurait tout contrôle (métrique, technique, qualitatif, etc.) et que, pour leur part, les résultats d'un contrôle ne pourraient pas être utilisés comme données représentatives d'un objet ou d'un phénomène à évaluer. L'expérience des pratiques professionnelles d'évaluation montre pourtant que ces deux dimensions ne sont pas alternatives mais cohabitent souvent : chronologiquement, alternativement ou même conjointement.

Une autre distinction, plus récemment suggérée par Torrès (2010, p. 42), oppose, cette fois, « l'évaluation contrôle » à « l'évaluation critique ». Ici, la notion de contrôle renvoie, à peu près, à la définition précédente pendant que s'installe, en contrepoint, la notion d'« évaluation critique », donnant de l'importance à l'analyse, caractérisée comme « participative » et ayant recours à des « référentiels adaptés à la réalité » donnée à évaluer.

On constate, déjà, que ces oppositions ne constituent pas seulement des objets de débats de spécialistes : elles illustrent des « postures » différentes, parfois même opposées, qui se cachent derrière l'emploi du terme « évaluation » et qui peuvent expliquer que les protocoles qui en sont inspirés peuvent déboucher sur des résultats à significations différentes. Par exemple, le « contrôle », symbole du pouvoir et qui constitue un véritable « marqueur » d'une évaluation qui exclut du jeu le sujet concerné par l'évaluation, peut, par ailleurs, accompagner des processus de qualité.

3 QUALITATIF/QUANTITATIF

Ce couple de concepts se retrouve souvent dans les discours sur l'évaluation, soit pour exprimer une alternative (par exemple : va-t-on procéder à une évaluation quantitative ou qualitative pour examiner les effets d'une expérience sur la progression de la compétence d'un élève ?), soit pour opposer deux catégories de résultats attendus (résultats quantitatifs

mesurables et comparables *versus* résultats qualitatifs de type *étude de cas*), soit, parfois pour qualifier de scientifiques les premiers et de non scientifiques les seconds.

Pourtant, les méthodologies d'évaluation ont toujours la possibilité d'opter, selon le sujet et l'enjeu, pour une *approche quantitative* (du type des évaluations quantitatives internationales des systèmes éducatifs : cf. Mons, 2004) ou pour une *évaluation qualitative* de l'acquisition de compétences professionnelles chez un candidat à une validation d'acquis professionnels (cf. Mayen et Savoyant, 2009).

En fait, la relation entre évaluation quantitative et évaluation qualitative constitue un objet de réflexion particulièrement pertinent parce que l'une et l'autre sont moins opposées que complémentaires. En effet, les deux postures ne sont pas exclusives l'une de l'autre ; certaines études comportant un traitement quantitatif de données complètent ou cherchent à compléter ou à corroborer leurs résultats en procédant à des enquêtes qualitatives utilisant l'entretien non directif ou l'observation participante. Par ailleurs, les procédures quantitatives ont souvent besoin de s'appuyer sur des modélisations relevant de constructions qualitatives. Crahay (2006) situe, d'emblée, cette opposition au niveau scientifique et épistémologique ; il préconise de fonder les recherches en éducation sur une pluralité de méthodes complémentaires, à la fois quantitatives et qualitatives. On ne peut que souscrire à cette position et rappeler que ce qui vaut pour la recherche en sciences sociales vaut aussi pour l'évaluation qui s'inspire de ses méthodes.

4 EXTERNE/INTERNE

S'interroger sur l'alternative entre une évaluation externe ou interne ne constitue pas seulement une option méthodologique : on entre, par là, dans des choix théoriques et scientifiques majeurs en matière de conception même de l'évaluation.

Jean Cardinet (1990) formule très clairement l'importance de cet enjeu : « L'évaluation traite de problématiques très différentes selon qu'elle vise une fonction sommative, formative ou prédictive ou qu'elle traite du système, de l'école ou de l'élève » (Cardinet, 1990, p. 139), mais aussi selon que le point de vue initial sur l'objet à évaluer sera « externe, interne ou négocié ». Cette remarque vaut aussi pour l'évaluation des autres situations que la situation scolaire (audit interne/audit externe/audit participatif).

- *L'évaluateur externe* se place, par définition, en dehors de la situation d'apprentissage ou de formation et n'y joue, par conséquent, aucun rôle : il s'agit de « contrôler tous les facteurs qui interviennent dans l'apprentissage » et/ou dans la formation et de mesurer l'effet des « variables modifiables » comme, par exemple, le nombre

d'apprenants, la méthode pédagogique, le type de recrutement... L'évaluateur organise alors un dispositif d'observation et de questionnement sur ces variables et il traduit les résultats ainsi obtenus en comparaisons, en évolutions, en classements de toutes sortes. Il se réfère ainsi aux modèles des sciences de la nature et adopte une conception explicative de l'évaluation.

Cardinet (1990) a discuté cette conception lorsqu'elle se présente comme exclusive. Il remarque ainsi que, dans les situations éducatives, « si l'on veut isoler les faits scientifiques objectifs, les relations causales réelles sont difficiles à dégager » (Cardinet, 1990, p. 143). L'auteur dénonce, par la suite, les illusions que nourrit l'évaluation externe : illusions de « simplicité » (les phénomènes, eux, sont toujours complexes), de « généralité » (« les recherches expérimentales mettent en doute, elles-mêmes, la validité de leurs conclusions »), d'« objectivité » (« il suffit de changer la forme du raisonnement pour transformer la réussite en échec »).

- *L'évaluateur « interne »* va construire son point de vue autour de l'acteur (individuel ou institutionnel), avec une prise en compte des objectifs de ce dernier ; il va donc choisir comme objets à évaluer les « représentations » des acteurs concernés et les « analyser de façon clinique ». Mettre l'apprenant « en situation d'évaluation autonome », inciter la méso-structure éducative ou de formation à une auto-évaluation et à l'adaptation aux particularités locales, tels seront les principes organisateurs de l'évaluation interne. Celle-ci se réfère, quant à elle, aux modèles des sciences humaines : ainsi, l'explication cherchera moins les causes que les raisons.

Cardinet relève ensuite les risques de l'évaluation interne : pour les élèves (manque de comparaisons avec des normes sociales extérieures à la situation éducative) ; pour les établissements scolaires ou de formation (qui ne sont pas toujours préparés à apporter des correctifs collectifs) ; pour les systèmes (avec une autogestion entraînant l'éventuelle « parcellarisation des initiatives »). L'auteur montre bien que l'alternative entre les deux modèles externe et interne ne peut être tranchée que par la recherche de leur compromis, qu'il nomme l'évaluation négociée. Celle-ci prend en compte les représentations des différentes catégories d'acteurs concernées et se trouve amenée à « confronter tous les points de vue possibles ». « La réalité » soumise à évaluation sera alors constituée par « la somme des faits objectifs de la première approche (externe) et des représentations subjectives de la seconde (interne) ».

5 OBJECTIF/SUBJECTIF

Il s'agit du débat le plus ancien, le plus courant, le plus utilisable, aussi bien dans un colloque de spécialistes que dans une conversation de salon :

l'évaluation doit-elle ou peut-elle être objective ? Autrement dit, le jugement évaluatif doit-il être impartial ou peut-il l'être ? On peut discuter à l'infini sur la question ainsi posée et on comprend bien qu'aucune réponse définitive ne pourra être trouvée même s'il existe des voix pour affirmer toutes les positions possibles. Nous n'allons donc pas nous risquer, ici, à prendre un parti (de plus), à ce sujet, mais il nous semble plus utile de profiter du débat pour introduire (dans l'esprit du présent chapitre) une nécessaire réflexion théorique préalable. Examinons donc, d'abord, les concepts en cause.

5.1 Objectif

Est objectif ce qui se rapporte à l'objet de la connaissance et qui, par conséquent, ne dépend pas du sujet.

L'objectivité pourrait donc s'obtenir en respectant les règles de validité de la connaissance d'un objet (approche scientifique) et en observant l'objet à évaluer sans entretenir aucun lien avec lui.

5.2 Subjectif

Est subjectif ce qui se rapporte au sujet connaissant et qui, par conséquent, dépend du sujet (ses choix, ses préjugés, etc.).

La subjectivité, en évaluation, est une posture ambiguë, car elle peut qualifier à la fois :

- un comportement positif, parfois revendiqué par les évaluateurs (Gérard, 2002), consistant à se montrer compréhensif et même empathique face au sujet apprenant. L'évaluation serait à reconnaître comme un jugement, une appréciation, par essence subjective ;
- un comportement critiqué dans la mesure où 1) elle consisterait à porter des jugements influencés par des facteurs personnels (conceptions, état d'esprit...) et insuffisamment étayés par des informations indiscutables ; 2) elle conférerait à l'évaluateur un pouvoir discrétionnaire.

L'intersubjectivité (concept husserlien commenté par Mutelesi, 1998) vient atténuer cette apparente opposition : elle consisterait en une certaine forme d'objectivité se rapprochant de la somme des subjectivités des acteurs. Utilisé à la faveur des travaux d'inspiration constructiviste, ce concept issu du courant phénoménologique définit la relation réciproque des consciences, les unes avec les autres, comme étant à l'origine de la constitution et du partage d'un monde commun. Cette idée sera reprise, plus loin, lorsqu'il s'agira de fonder théoriquement les approches interactives de l'évaluation (par exemple, à propos des propositions de Rodrigues (p. 42).

6 DES OPPOSITIONS RÉDUCTRICES OU DES AMORCES DE THÉORISATION ?

On voit combien de tels débats structurés par des oppositions binaires risquant de donner l'illusion de discussions fondamentales alors qu'ils ne feraient que confisquer la recherche d'explications alternatives ou intermédiaires, sont à considérer avec précaution. Ces réductions quelque peu manichéennes relèveraient-elles d'une novlangue à la manière de la novlangue d'Orwell (1950) dans laquelle l'évaluation serait volontairement réduite à quelques concepts englobants ou peuvent-elles révéler des choix théoriques implicites ?

Faisons l'hypothèse qu'elles révèlent, parfois, des choix théoriques ou idéologiques. Ceux-ci sont tantôt explicites et assumés, tantôt implicites, voire non conscients. Mais, dans tous les cas, ils peuvent conditionner sensiblement les pratiques d'évaluation qui s'ensuivent (comme on l'a remarqué : mesurer et contrôler *versus* évaluer ? observer le sujet de l'extérieur ou tenter de l'associer au diagnostic ? choisir des outils qualitatifs ou quantitatifs ? chercher une forme d'objectivité ?). Le fait de repérer, de nommer et de définir les concepts qui sous-tendent ces catégories simplifiées de problèmes, permet à l'évaluateur de s'inscrire dans une démarche de construction plus éclairée des méthodes qu'il va utiliser. Enfin, l'explicitation de ces concepts constitue une manière plus abordable de chercher à définir les fondements théoriques des pratiques d'évaluation.

Chapitre 2

Des cultures de l'évaluation ?

1 CULTURE DE L'ÉVALUATION : UN THÈME FRÉQUENT

L'usage de la notion de *culture de l'évaluation* est très fréquent, soit au sein du système scolaire (et plus particulièrement dans les institutions comme les ministères de l'Éducation), soit dans les entreprises et les administrations (Conseil scientifique de l'évaluation, 1996), dans le cadre d'un modèle managérial « moderne ». Ici encore on a affaire à une expression aux multiples sens : un ouvrage de 1952 dénombrait déjà deux cents définitions du mot *culture*. Quant à la *culture de l'évaluation*, elle n'est définie que rarement : Mottier Lopez (2013a) le confirme en élaborant une synthèse des façons de penser cette culture, ou, plutôt, ces cultures. L'auteur les ramène à deux catégories : celle « descendante » qui tente d'instaurer et de généraliser des formes d'« évaluation-contrôle » et celle « ascendante » qui tente, « à partir des pratiques », de favoriser « le développement des acteurs impliqués » par la co-construction d'un « certain statut de reconnaissance collective ».

Pourtant, on lit (ou on entend) des affirmations comme « il faut développer une culture de l'évaluation dans le système éducatif » ou bien « la France et les pays anglo-saxons n'ont pas la même culture de l'évaluation ». Cette notion doit donc certainement revêtir une signification importante et entraîner des caractéristiques différenciées des pratiques qui en découlent. C'est ce que nous allons chercher à repérer.

2 DISTINGUER ENTRE « LA », « UNE » ET « DES » CULTURE(S) DE L'ÉVALUATION

Clarifions, en effet, ce dont nous parlons :

- « La » culture de l'évaluation constitue une expression très synthétique qui ne peut que désigner le fait que, dans une institution donnée, la pratique de l'évaluation (sous quelque forme que ce soit) est généralement présente et, par conséquent, tolérée, sinon préconisée et organisée. Dans la mesure où on ne précise pas l'importance de son utilisation, il est difficile, aujourd'hui, de trouver des institutions d'où ce type de culture soit totalement absent.
- « Une » culture de l'évaluation revendique déjà l'appartenance à tel ou tel mouvement d'idées adhérant à une philosophie générale ou à une conception précise du statut et des méthodes qui en découlent. On parlera alors de culture des résultats (notion d'apparition récente), de culture des objectifs, de culture qualité, etc. On adhérera donc, selon les cas, à telle ou telle école de l'évaluation : socio-philosophique (l'évaluation à l'écoute des inégalités), politique (l'évaluation comme levier de pilotage du système éducatif) ou scientifique (l'évaluation comme application des travaux de recherche à l'amélioration des apprentissages et de la formation). Chacune de ces écoles fournira ses références (conception de l'éducation, réformes de société, choix pédagogiques) mais aussi ses méthodes et ses outils. L'adhésion à telle ou telle de ces écoles contribue ainsi au développement de l'une des « cultures » en question.
- « Des » cultures de l'évaluation coexistent donc dans les milieux de la formation et de l'éducation. Elles ne coïncident pas systématiquement avec les organisations, même s'il existe des connotations fortes autour d'une « culture d'entreprise » ou d'une « culture Éducation nationale », par exemple. Il apparaît très important de prendre acte de « ces » cultures différentes de l'évaluation. Cela veut dire, en effet, que plusieurs mondes possibles peuvent s'approprier des façons d'aborder l'évaluation, avec un système d'explications autonome (et, parfois, sans qu'il y ait préalablement choix, par les opérateurs eux-mêmes), des conceptions alternatives qu'ils impliquent, des modèles spécifiques qu'ils invoquent, des méthodologies captives qu'ils véhiculent.
- On adoptera, pour terminer, la formulation de Mottier Lopez (2013b) : une culture de l'évaluation est « une construction, dans et par un groupe social donné, des pratiques, normes, valeurs, discours, outils partagés en matière d'évaluation » (p. 203).

3 UN NOUVEL AVATAR : LA CULTURE DU RÉSULTAT ?

Une tendance institutionnelle ancienne dans les pays anglo-saxons mais récente en France préconise une *culture du résultat* par la mesure de la performance¹. Autant cette approche de l'évaluation par les résultats (opposée, souvent, à une *culture des moyens*) peut s'expliquer par une posture plus pragmatique en matière de modèles de pilotage des dispositifs éducatifs, autant il convient de s'arrêter un instant sur cette tendance qui semble poser problème lorsqu'elle introduit, par la force d'une loi, une vision déformante de l'évaluation. Nous nous trouvons, ici encore plus qu'ailleurs, devant un phénomène relevant de la fiction d'une « novlangue managériale », sorte d'idéologie utilisant à son profit des notions nobles pour en faire de véritables slogans censées la valoriser.

Expliquons-nous.

Historiquement, il s'agissait d'un principe de gestion qui, dans l'entreprise, tenait compte des résultats du salarié dans la fixation du montant de sa rémunération, principe sur lequel nous n'avons pas à discuter ici. C'est plus récemment que ce principe a été étendu à certaines administrations publiques (cf. *supra*). Mais, disons-le d'emblée : ce transfert s'est opéré avec une usurpation de concepts. En effet, le concept de *culture*² inspire le respect et appelle une reconnaissance universelle ; celui de *résultat* bénéficie d'un emploi quotidien depuis l'enfance (avec les résultats scolaires) qui lui donne un label d'évidence n'éveillant aucune méfiance. Pourtant, un seul exemple serait suffisant pour comprendre que les mots employés ont un autre sens : un établissement scolaire jugé selon ses résultats aux examens. L'opinion publique, alertée par les médias sur ce sujet, a déjà eu l'occasion de dénoncer les dangers de telles pratiques. On sait très bien, aujourd'hui, que ce type de résultats est conditionné par des facteurs souvent étrangers à la qualité éducative des prestations fournies par l'établissement, comme le milieu social et le type de recrutement de ses élèves. La confusion est introduite à deux niveaux :

- par l'emploi du concept de culture pour qualifier une directive politique ne se préoccupant que de faire pression sur une catégorie de personnels au détriment de la prise en compte de ce qui constitue véritablement la *culture* constituant le référent majeur de la mission du système éducatif d'un pays ;
- par l'utilisation du mot *résultat* pour désigner un chiffre (on a envie de dire : n'importe lequel) comme représentatif de l'ensemble des activités de direction, d'encadrement, d'accompagnement, de fonctionnement et, finalement, d'apprentissage de chaque élève.

1. LOLF : loi organique relative aux lois de finances (spécificité française), entrée en vigueur en 2006.

2. La culture pouvant être définie de manière limitée, ici, comme l'ensemble des traits distinctifs (tant intellectuels, moraux, techniques qu'artistiques) d'une société.

L'usurpation que constitue l'emploi de ces deux concepts de culture et de résultat est donc flagrante car on a confondu la mesure d'un indicateur arbitrairement choisi avec l'objectif à évaluer.

4 RECONNAÎTRE LES MODÈLES D'UNE CULTURE DE L'ÉVALUATION

En fait, l'emploi du terme *culture de l'évaluation* ne réussit qu'à noyer le poisson d'un ensemble de pratiques floues peu significatives d'une politique particulière. Il paraît plus intéressant de repérer, derrière ce langage, des modèles sous-jacents de l'évaluation. On peut en citer quelques-uns, non exhaustifs, inspirés des travaux sur l'évaluation des établissements (qui ont mobilisé les réseaux politiques et universitaires depuis les années 1960) mais qui peuvent aussi servir à décrypter les autres dispositifs de formation :

- *La culture du résultat*, déjà citée, habite les tenants de la mesure de l'efficacité des dispositifs d'éducation ou de formation. Le modèle sous-jacent est bien connu : il s'agit du courant « *school effectiveness* » (Scheerens et Bosker, 1997) qui cherche à appréhender les progrès pouvant être prédits par les caractéristiques des apprenants (milieu économique, socioculturel, familial, etc.). Dans ce modèle, l'évaluation est pratiquée à partir d'*indicateurs de performances* et a pour fonction d'analyser les corrélations entre ces caractéristiques et les performances obtenues par les dispositifs (notamment les écoles).
- *La culture de l'amélioration et du développement*, qui est née par réaction à la précédente. Le modèle sous-jacent à cette culture est celui de « *school improvement* » (Hopkins, 2001) qui s'intéresse davantage aux processus qu'aux résultats et qui s'inscrit dans un modèle plus large de changement. C'est l'auto-évaluation qui est, ici, la plus prisee.
- *La culture de l'évaluation des objectifs* imprègne les évaluateurs se situant dans la mouvance de la « pédagogie de la maîtrise » (alias pédagogie par objectifs) issue de *mastery learning*, concept inventé aux États-Unis dans les années 1960 et qui nous est parvenu par les travaux de Bloom³. Cette approche, ancienne et très exploitée, malgré les critiques ultérieures de morcellement excessif des séquences d'apprentissages dont elle a fait l'objet, s'articule autour de trois notions essentielles structurant le raisonnement évaluatif : a) un *objectif* (général ou secondaire) exprimant le résultat que l'apprenant doit atteindre ; b) un *comportement observable* répondant à la réalisation de cet objectif ; c) des *critères de réussite* définis-

3. Cf. sur cet auteur, les synthèses de De Landsheere (1992) ainsi que Hameline (1979).

sant les performances minimales exigées et, en même temps, le type d'évaluation qui leur sera attaché.

Cette *culture des objectifs* a dominé et domine encore les courants pédagogiques sensibles à la rationalité et à la pertinence de l'action éducative et induit une culture spécifique de l'évaluation qui ne manque pas de se conjuguer avec la culture du résultat dans certains dispositifs nationaux ou internationaux d'évaluation du niveau des écoles et des élèves.

5 QUAND LA CULTURE DE L'ÉVALUATION CACHE L'ÉVALUATION ELLE-MÊME

On a essayé de pointer, dans le développement qui précède, sinon la polysémie, du moins le caractère flou des discours sur la (ou les) culture(s) de l'évaluation et de suggérer de dégager les modèles sous-jacents de ceux qui s'en réclament. Dans cette optique, on gardera donc l'idée de repérer, plutôt, les influences exercées par des modèles d'évaluation comme ceux déjà évoqués (approche par les résultats, approche par l'amélioration, approche par les objectifs, etc.), y compris ceux inclus dans le recours aux méthodologies quantitatives *versus* qualitatives, dans les approches externes/internes, etc. Et on s'efforcera de fonder les pratiques évaluatives sur d'autres modélisations que celles pouvant se déduire des différentes « cultures » que l'on vient de visiter. Ainsi, il nous appartiendra d'énoncer une conception de l'évaluation qui ne serait pas dépendante de modes de pensée institutionnels, de stéréotypes, d'enfermements manichéens, d'habitudes de langage. Celle-ci proposera d'asservir cet ensemble de processus à une réflexion inspirée par l'approche philosophique de Dewey, tout particulièrement exprimée dans sa « théorie de l'enquête » et sur laquelle on s'appuiera lorsqu'on abordera la question méthodologique, en deuxième partie.

Chapitre 3

Des définitions de l'évaluation

Il est rare de trouver un ouvrage sur l'évaluation qui ne comporterait aucune définition de cette notion. Chaque auteur en propose une, si ce n'est la sienne propre. Devant une telle prolifération, on va, ici encore, tenter de voir clair, de voir large et d'avancer vers la perception de ce qui peut participer aux fondements de l'évaluation : les *définitions* font partie, par évidence, des points de passage obligés de cette quête, tout en posant, néanmoins, quelques problèmes.

1 AVANT DE DÉFINIR

1.1 Que fait-on en définissant l'évaluation ?

Attacher une sémantique à un concept, comme celui d'évaluation, ne se réduit pas à décliner une liste d'attributs qui prétendraient fixer la représentation de ce phénomène, comme, par exemple, déclarer qu'il s'agit d'un *jugement de valeur*, de *l'appréciation d'une performance* ou encore d'une simple *mise en relation entre référés et référents*, etc. Dans ce cas, définir se réduirait à affirmer ou à reproduire des allégations que l'évaluateur utiliserait comme assertions (sortes de vérités postulées).

Mais, alors, qu'est-ce que définir ? C'est, déjà, clarifier une chose sans passer par un raisonnement. Par exemple, ce serait dire que l'évaluation constitue, dans la plupart des programmes de formation, une transaction de jugement entre formateurs et apprenants.

Mais définir, c'est aussi démontrer qu'une chose existe : en expliquant la cause d'une chose, on dit ce qu'elle est. Par exemple, à partir de nombreuses

observations de phénomènes se déroulant dans les systèmes de formation, on pourrait montrer que l'évaluation éducative répond à un besoin social d'attribuer une valeur aux apprentissages et à leurs contextes : une valeur aux résultats des élèves, une valeur aux écoles qui les reçoivent, aux méthodes qui s'y développent et aux enseignants qui y exercent. Autrement dit, définir l'évaluation constituerait, à la fois, la recherche de connaissance la plus complète possible d'une réalité éducative en même temps que le fait de nommer un ensemble de processus inhérents à la régulation de l'action éducative.

1.2 Pour des définitions modestes

Nous postulons que l'évaluation constitue une forme de vie davantage qu'une procédure artificielle provoquée par un évaluateur : il existerait des *formes de vie évaluatives*, liées à des mondes différents (la justice, l'éducation, l'économie, notamment). Construire et proposer une définition de l'évaluation relève, selon nous, davantage de l'hypothèse, d'une part, et de la référenciation, d'autre part. De l'hypothèse, car il n'existe pas de vérité établie et on peut que proposer des définitions provisoires en tant qu'« hypothèses de travail », avant de procéder à la conception d'un dispositif d'évaluation. De la « référenciation », au sens de Bonniol et Vial (1997), définie comme l'action consistant à « s'étayer à un modèle antérieur » : l'évaluateur modeste en matière de définition présente au moins l'avantage de ne pas dresser d'emblée l'arbre qui cacherait la forêt.

1.3 Pour la prise en compte du discours et des jeux de langage dans la définition de l'évaluation

En effet, l'évaluation participe d'une activité de communication (Weiss, 1991) entre acteurs et, à ce titre, constitue une pratique sémiotique d'interprétation de signes. On peut très bien assimiler une procédure d'évaluation à un *jeu de langage*, une manière de produire de la signification par l'intermédiaire de règles, au sein d'un contexte particulier. Poser, en préalable, l'idée que *l'évaluation est un discours* entraîne une mise en doute de la légitimité de toute définition purement objectivante, surtout si l'on raisonne en termes *dialogiques*, avec le sens donné par Bakhtine (1970) au concept de « dialogisme », c'est-à-dire d'interaction qui se constitue entre le discours de l'énonciateur et les discours d'autres personnages. Cette analyse peut s'appliquer très pertinemment à la situation d'évaluation dans laquelle, en effet, l'être évalué (ou le groupe) ne peut être appréhendé qu'en tant que résultant d'interactions humaines.

1.4 La question de la scientificité de la définition

Le thème devient plus difficile et ne pourrait pas être longuement développé ici. Néanmoins, le professionnel (enseignant, formateur, expert) qui produit une définition de l'évaluation s'est-il interrogé, au préalable, sur sa

dimension scientifique ? S'est-il déjà forgé une représentation du caractère nomothétique⁴ de sa définition ? Il pourrait dire qu'évaluer, selon le modèle expérimental, c'est établir une relation de causalité entre des actions et des effets observés en accordant à cette définition une valeur de loi. Mais il pourrait aussi déclarer comme scientifique une définition d'inspiration herméneutique⁵ du type « évaluer, c'est interpréter les performances obtenues par un acteur (élève, apprenant en général, professionnel...) ».

On s'aperçoit déjà que définir l'évaluation ne relève pas de l'évidence, qu'il s'agit d'un acte qui structure d'emblée l'activité qui va suivre et qui procède, encore une fois, d'un fondement théorique à repérer et à assumer.

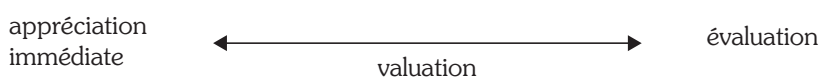
2 QUELLES DÉFINITIONS ?

En abordant le champ des *définitions* de l'évaluation, il faut rappeler une distinction porteuse de sens entre :

- la notion de *valuation* (Dewey, 2011) importée de la littérature anglophone, qui désigne, à la fois, le fait de priser (retenir pour précieux) et celui d'évaluer (attribuer une valeur à) ;
- la notion d'*évaluation*, davantage utilisée dans la littérature francophone, désignant essentiellement l'activité de mesure impliquant une comparaison et donnant lieu à une activité intellectuelle.

La notion de *valuation* étend le champ des postures évaluatives possibles tout en les posant sur un axe unifiant, allant de l'*appréciation immédiate*, proche du jugement de valeur à caractère émotif (du type de l'attribution du mot « bon » à une performance) à l'*évaluation*, donnant lieu à des mesures :

Figure 1. La valuation selon Dewey



Plutôt que de dresser une liste ou une analyse de définitions, ce qui a déjà été fait (Bonniol et Vial, 1997), on va les aborder d'un point de vue taxonomique afin de tenter d'en faire ressortir quelques orientations majeures. Pour éviter d'éventuelles querelles de typologies concurrentes, on a cherché une façon de classer les définitions de l'évaluation connues autour du modèle le plus large, le plus englobant qui soit et on a débouché sur deux catégories de choix structurantes : celle des choix *philosophiques* (qu'on peut appeler *paradigmes*) influençant les postures d'évaluation possibles et celle des choix de *finalités* assignées à l'évaluation.

4. Nomothétique : du grec *nomos* (loi), vise à obtenir des connaissances généralisables.

5. Herméneutique : qui a pour objet ou pour fonction d'expliquer, d'interpréter des textes.

Des définitions d'inspiration philosophique que Rodrigues (2006) classe autour de trois paradigmes⁶ :

- le paradigme *objectiviste*, au nom duquel on évalue le sujet (ou un ensemble de sujets) en tant qu'être social compris comme objet d'influences extérieures (l'évaluation sera alors définie comme une procédure de mesure objective de résultats, de comparaison entre sujets ou dispositifs, de classement de performances, etc.). C'est dans ce sens qu'il faut comprendre ce que Cardinet (1990, cf. *supra*) appelle *évaluation externe* : une mesure des performances d'un sujet, d'un groupe de sujets, d'un dispositif ou d'une institution, par un regard extérieur en comparaison avec d'autres sujets ou groupes de sujets ;
- le paradigme *subjectiviste* au nom duquel est postulée la complexité des situations à évaluer parce que construites par des acteurs : la compréhension des résultats requiert la connaissance des intentions des acteurs et de la manière dont ils interprètent les situations. L'évaluation sera alors définie comme un ensemble de procédures « autorégulatrices ». Cardinet parle, dans ce cas, d'*évaluation interne* : une analyse des résultats obtenus par un sujet ou un groupe de sujets du point de vue des évolutions de ces résultats ;
- le paradigme *interactif* (dialectique) : dans cette perspective, les acteurs de l'évaluation, selon Rodrigues, ne sont pas exclusivement envisagés comme des *objets déterminés* (approche objectiviste) ni comme des *sujets déterminants* (approche subjectiviste), mais comme des partenaires en co-évaluation. Cardinet appelle cette situation une *évaluation négociée* : une confrontation des analyses externes et internes en vue de rechercher la signification la plus complète des résultats d'un apprentissage ou d'une formation.

Des définitions selon les finalités assignées à l'évaluation. Trois finalités jalonnent la littérature sur l'évaluation : la formulation d'un *jugement de valeur*, *l'aide à la décision*, *l'attribution de sens*⁷. Faisons l'inventaire rapide des définitions relevant de la mise en avant de ces trois finalités.

2.1 L'évaluation définie comme jugement de valeur

Hadji (1990) définit l'évaluation comme l'acte par lequel on formule un jugement de valeur portant sur un objet déterminé (individu, dispositif) par

6. Paradigme : au sens de Kuhn (1983), un modèle cohérent de représentation du monde et d'interprétation de la réalité largement accepté dans un domaine particulier.

7. Pourquoi ne pas avoir intégré à cette liste la *mesure* ? À cette question posée par un membre cofondateur du livre, il a été répondu que la mesure intervenant, à sa place, dans la plupart des protocoles d'évaluation, elle ne devait pas être isolée comme définition alternative des autres...

le moyen d'une mise en relation entre deux types de données que certains auteurs appellent *référé* (données sur le sujet ou la situation à évaluer) et *réfèrent* (données de l'ordre de la norme, de l'objectif, de l'idéal). Il faudra revenir sur ces notions qui, dans le contexte de cet ouvrage, vont être revisitées et proposer des significations particulières.

Mais on s'arrêtera un instant sur le *jugement de valeur*. Une référence historique conforte les tenants inconditionnels de cette définition, celle du sociologue Max Weber (1965) : « Nous entendons par évaluation l'appréciation pratique d'un phénomène sur lequel notre activité peut exercer une influence en adoptant à son égard une attitude d'approbation ou de désapprobation » (Weber, 1965, p. 401). Mais les temps ont changé et, de nos jours, le concept d'évaluation s'est considérablement chargé d'attentes sociales, de recherche de transparence et d'équité, de protocoles marqués de rigueur et de scientificité. Le *jugement* devient suspect de subjectivité et *la valeur*, si elle reste largement invoquée, est devenue un concept d'utilisation très élargie.

Certes, il serait peut-être tentant de forger une définition en fonction de l'étymologie et de dire que dans la notion d'évaluation on retrouve la racine du mot valeur. Mais, en utilisant l'expression « jugement de valeur », on prend le risque d'introduire deux biais importants :

- celui de réduire l'évaluation au jugement d'une personne ou d'une institution sur une autre, limitant ainsi l'autonomie et la responsabilité de l'entité qui est jugée par un pouvoir arbitraire ;
- celui d'utiliser le concept de valeur qui introduit des risques évidents de manipulation et de confusion (valeur éthique, valeur économique, valeur sociétale, autre système de valeurs partagé ou non). Cardinet (1989) avait, en tout cas, répondu à l'avance à cette proposition par son article intitulé « Évaluer sans juger », dans lequel il justifie une évaluation descriptive *versus* une évaluation-jugement : « En ne prenant pas position au départ, l'école réserverait le jugement de valeur à celui qui devrait prendre effectivement plus tard une décision impliquant ces valeurs. Elle permettrait ainsi à diverses personnes de porter des regards différents sur la même réalité... » (Cardinet, 1989, p. 48).

Laissons à Dewey (1967) le rôle de mise au point : les jugements de valeur font partie de l'enquête – ils ont un statut opérationnel dans la formulation d'hypothèses – et peuvent lui être assujettis. La méthode par laquelle on contrôle les énoncés s'applique aux jugements de valeurs comme aux jugements de faits et les valeurs peuvent être ainsi éprouvées « à partir des conséquences qu'elles produisent dans l'application existentielle » (Dewey, 1967, p. 610). L'auteur souligne ici que le jugement de valeur est associé au « contrôle », c'est-à-dire qu'il doit être naturellement soumis à l'enquête car le jugement appréciatif, même s'il constitue une

étape normale et utile, exige un traitement rationnel. C'est le processus évaluatif dans son ensemble (inséré dans son contexte) qui est alors concerné et qui doit faire l'objet de l'enquête. Quoi qu'il en soit, on se rangera derrière la remarque de Dewey qui rappelle qu'il importe de distinguer entre l'acte non cognitif d'un jugement appréciatif et l'acte cognitif de la *valuation*. On va laisser ce débat ouvert pour n'en garder qu'une idée : une définition de l'évaluation invoquant le jugement de valeur ne peut se suffire à elle-même pour décrire l'ensemble du processus évaluatif et constitue, surtout, une interpellation.

2.2 L'évaluation définie comme aide à la décision

Toujours très prisee, la définition de Stufflebeam (1980) doit certainement remporter la palme du nombre de référencements : « un processus par lequel on traite des informations permettant de justifier des décisions possibles ». L'auteur s'appuie sur le recours à des modèles de décision appliqués à l'éducation et, se référant aux théories du changement (p. 48), il débouche sur la présentation de « types de décision » possibles, suite à une évaluation (planification, révision, structuration, mise en application).

Si l'on prend cette définition comme fondement de l'évaluation, on est tenté d'en déduire, tout naturellement, une conception très proche de l'évaluation par les objectifs dans la mesure où la *décision* n'est pas nettement réservée au sujet mais concerne plutôt, implicitement, l'évaluateur, assimilé au décideur ou, en tout cas, à son service. Et, on accepte, par ailleurs, d'appliquer une sorte de modèle méthodologique prescriptif traitant la démarche évaluative à travers le schéma systémique bien connu proposé par l'auteur (Stufflebeam, 1980), le CIPP⁸ (contexte, intrants, processus, produits), en la liant à une prise de décision en quelque sorte « pré-formatée ». On peut, en effet, à partir de la définition de l'auteur, se laisser conduire par ce cadre rassurant et justifier un protocole d'évaluation, depuis une conception (aide à la décision), en passant par une définition (citée *supra*), jusqu'à un modèle valant méthodologie (le CIPP et les schémas décisionnels). Mais, ce faisant, on aura appliqué un modèle pouvant servir de guide sans, pour autant, maîtriser l'ensemble d'un processus d'évaluation adapté à tel ou tel objet spécifique comme un apprentissage, une acquisition de compétence, l'impact d'une politique sur le fonctionnement d'un projet, etc.

8. Pour résumer et simplifier le modèle prescriptif d'évaluation proposé par Stufflebeam, retenons que l'évaluateur est invité à procéder, successivement, à une analyse du contexte (C) permettant de justifier le choix des intrants (I) à prendre en compte (comme les objectifs d'un programme), afin d'observer les processus (P) qui en découlent et de mesurer les résultats ou produits (P) du programme.

2.3 L'évaluation définie comme recherche ou changement de sens

2.3.1 Une recherche de sens

On a déjà vu (et on verra plus loin) comment l'évaluation cherche à créer du sens en comparant une situation avec des objectifs, une norme, etc. L'écart décelé entre les deux types de données ferait émerger le sens recherché par l'évaluation. Certaines définitions prises dans la littérature en font leur élément principal. Ainsi, Bonniol et Vial (1997) écrivent qu'évaluer, c'est poser des « problématiques du sens » (p. 23). En fait, plutôt qu'une définition, nous avons ici l'expression d'une fonction fondamentale de l'enquête évaluative : rechercher la réponse à une question et, par conséquent, trouver ou éclairer le sens d'un phénomène. Mais le passage de la définition à la réalisation constituera un long chemin à parcourir. Il n'en reste pas moins que la question du sens intéresse les évaluateurs qui, chacun dans son contexte, annoncent le sens qu'il convient de donner aux résultats de l'évaluation qu'ils ont organisée.

2.3.2 Un changement de sens

Autre chose est de dire que l'évaluation, en elle-même, consiste à modifier le sens des informations produites au sujet d'un apprenant ou d'un dispositif. Lesne (1984) illustre ce phénomène par la définition suivante : « L'évaluation d'un fait de formation consiste en la production d'un jugement de valeur ou d'utilité (d'une sorte d'étiquette) grâce auquel ce fait acquiert une signification différente, est perçu dans une nouvelle perspective, fait l'objet d'une nouvelle représentation, dès lors qu'un nouvel attribut, "ce qu'il vaut", lui est conféré » (p. 131). Voilà une définition qui peut interpeller l'évaluateur : l'idée que ce dernier est doté de pouvoirs importants de modification des caractéristiques recueillies sur l'apprenant ou sur le dispositif et qui, de par son intervention, deviendra une entité nouvelle, chiffrée, étiquetée, cataloguée. Parmi ces pouvoirs :

- celui de choisir, chez le sujet ou le groupe, les caractéristiques jugées représentatives sur lesquelles porteront les mesures ;
- celui d'interpréter des informations (première modification de la réalité) ;
- celui de les placer dans une échelle de notation, d'appréciation, etc. (deuxième modification de la réalité) ;
- celui de les reformuler, à travers des symboles chiffrés ou autres (troisième modification de la réalité) ;
- celui, final, le plus dangereux pour l'individu ou le groupe, de proposer un jugement porteur de toutes sortes de conséquences sur le curriculum, la réussite et la profession des évalués.

Cette définition de l'évaluation met en cause, selon nous, toutes les autres : elle montre bien, en tout cas, qu'il est difficile de bâtir une méthodologie dans ce domaine sur de simples définitions qui risquent de dénaturer un processus de régulation d'une formation ou d'un système en le transformant en instrument de contrôle, de pouvoir et, de manière plus générale, d'ajustement à la norme.

Ce rapide panorama de définitions de l'évaluation illustre bien l'idée selon laquelle celle-ci manipule en permanence le sens de l'école, de l'enseignement, de la formation et du pilotage des dispositifs institutionnels en organisant les protocoles d'évaluation à partir de définitions qui risquent d'être considérées comme fondatrices de modèles et de méthodologies alors qu'elles n'en constituent souvent que des avatars, des éléments de communication ou d'argumentation. Ces définitions gagneraient alors à n'être utilisées que comme des effets d'annonce d'un projet d'évaluation qui aura, auparavant, précisé ses fondements.

2.4 L'évaluation définie comme activité

Affirmer que l'évaluation est, fondamentalement, une *activité*, constitue une attitude certes déjà pressentie (Ardoino et Berger, 1989 ; Figari et Achouche, 2001 ; Tourmen, 2007) mais toujours téméraire, tant le concept même d'activité est difficile à définir et à utiliser sans en rester au flou qui caractérise son acception courante. Ce concept a, d'autre part, considérablement évolué grâce aux travaux d'ergonomie et de psychologie du travail (Leplat, 1997, 2000 ; Clot, 1999). En effet, le terme d'activité a pris un sens plus précis : il désigne alors un processus de réalisation d'une tâche, incluant toutes les activités mentales qui la rendent possible (la recherche d'informations, le diagnostic, le jugement, etc.). Pour approfondir la dimension ergologique et philosophique du concept d'activité, nous renvoyons aux écrits de Schwartz (2007) qui soutient l'idée que « chaque fois qu'il y a quelque chose comme *activité*... c'est toujours... en vue de résoudre un problème lié à l'unité de l'être humain : l'*activité* a pour fonction, pour signification ou pour objectif de recoudre des parties de l'être humain auparavant séparées » (p. 124). En bref, on utilisera le terme d'*activité* au sens attribué par l'ergonomie contemporaine qui y intègre, à la fois, l'action, c'est-à-dire ce qui se fait, mais aussi tous les processus mentaux qui la précèdent et l'accompagnent.

2.4.1 Une activité en situation

C'est dans ce sens que l'on décrira l'*activité d'évaluation*... En effet, on constate, aujourd'hui, que la pratique de l'évaluation est devenue un métier (l'évaluateur est reconnu en tant que tel dans plusieurs fonctions professionnelles comme celles de l'expert, de l'inspecteur, du consultant, du membre

de jury, du chercheur de même que celles de l'enseignant et du formateur). Mais, comme il s'agit d'éducation et de formation, l'activité professionnelle d'évaluation devient plus difficile du fait qu'elle est partagée avec les autres acteurs (apprenants, partenaires, notamment). C'est, enfin, une activité intégrée dans une situation particulière qui l'influence et la caractérise : une situation d'apprentissage ou de formation avec laquelle elle se trouve en interaction (Pastré, Mayen et Vergnaud, 2006.). En poursuivant dans ce sens, on peut dire que l'idée selon laquelle l'évaluation peut encore constituer un but en elle-même peut difficilement être soutenue. De nombreux travaux des vingt dernières années, portant sur les relations entre apprentissage et situation, mettent en avant la pertinence des concepts d'évaluation partagée, d'*évaluation située* (Mottier Lopez, 2006). Ce dernier auteur, s'appuyant sur une importante bibliographie, établit que, « dans une perspective située, évaluer est vu comme un acte de participation à une communauté de pratique qui s'incarne des significations, normes, valeurs... outils... socialement reconnus et partagés » (p. 88). L'évaluation serait, dans cette optique, une *activité* complexe, plurielle, partagée entre évaluateurs et évalués, ce que les ergonomes appellent une « co-activité ». Cette activité demande à être analysée dans les différentes opérations qui la concrétisent : ce sera l'objet de la deuxième partie de l'ouvrage.

2.4.2 Une expérience

À la lumière des travaux de P. Mayen (2005, 2009a) sur la place de l'expérience dans l'élaboration des compétences professionnelles et, plus récemment, dans la « validation des acquis de l'expérience » institutionnalisée dans un certain nombre de pays, on peut, enfin approcher l'évaluation comme constituant elle-même une *expérience*. Celle-ci comprend, en effet, un ensemble d'activités vécues par un sujet ayant à résoudre des tâches attendues dans l'exercice d'un acte professionnel qui organise des transactions entre évaluateurs et évalués. Selon Mayen, elles sont de deux ordres : des « *expériences processus* », dynamiques relevant du vivant, en train de se faire ou des « *expériences produits* », incorporées, en actes. D'ailleurs, on n'a pas encore assez exploité les apports de la didactique professionnelle (Pastré, Mayen et Vergnaud, 2006) à une certaine remise en question des théorisations de l'évaluation en éducation : on le voit déjà, en invoquant, ici, le lien entre validation des acquis de l'expérience et évaluation. D'autres apports apparaîtront plus loin, dans le domaine de l'évaluation des compétences (p. 163), par exemple.

Chapitre 4

Évaluation et théorie

1 POURQUOI ÉVOQUER LES LIENS ENTRE ÉVALUATION ET THÉORIE ?

Ce sont les auteurs nord-américains qui ont, les premiers, posé cette question et apporté quelques réponses. Les plus connus, Shadish, Cook et Leviton (1991), ont annoncé leurs positions dans leur titre même en parlant de *fondements* de l'évaluation de programme et en sous-titrant *théories d'une pratique*. Quant à Rossi (1991), il a présenté une synthèse des travaux allant dans ce sens. Il regrettait, en citant Wortman (1983), que « la recherche en évaluation » soit « une activité appliquée, largement (et malheureusement) athéorique » (p. 384), notant que si, suite aux nombreuses évaluations de programmes développées aux États-Unis, peu de changements majeurs ont pu intervenir, cela peut s'expliquer, en partie, par le fait que les programmes sociaux excèdent rarement les limites des idéologies politiques dominantes. Enfin, il déplore que de nombreux concepteurs de programmes ignorent les théories des sciences sociales qui leur permettraient pourtant d'estimer la complexité des processus ayant causé les problèmes qu'ils ont à évaluer.

Rossi formule alors le souhait d'une *évaluation conduite par la théorie*. Il suggère à l'évaluateur de construire un modèle théorique spécifique, mettant en perspective les intrants du programme, les processus de relation entre facteurs et les extrants, afin de mesurer tous ces aspects, en les rassemblant et en les analysant.

L'idée de lier évaluation et théorie n'est donc pas nouvelle, pas plus qu'elle n'est saugrenue. En revanche, dans certains pays, et

notamment en France, elle se heurte à des *a priori* disciplinaires et scientifiques qui tendraient à empêcher de dépasser le pragmatisme empirique dans lequel se noie souvent le discours sur l'évaluation sous prétexte que celle-ci ne serait qu'une pratique et que la théorie ne la concernerait donc pas.

2 PARLER THÉORIES À PROPOS DE L'ÉVALUATION ?

Si des théories sous-jacentes aux pratiques évaluatives existaient, de quoi s'agirait-il ? Lorsqu'on parle de théorie, on n'évoque pas simplement une construction intellectuelle non vérifiée par les faits ni une connaissance spéculative qui s'opposerait à une pratique. On peut s'intéresser à un sens plus noble de la *théorie* :

- dans le champ des sciences de la nature : un ensemble organisé de principes, de règles, de lois scientifiques visant à décrire et à expliquer des éléments en interaction (on pense à la théorie de la relativité) ;
- dans le champ des sciences humaines et sociales : un ensemble relativement organisé d'idées et/ou de concepts se rapportant à un domaine particulier (on pense à une théorie économique ou à la théorie des représentations) ;
- dans le champ des sciences de l'éducation, Laveault (1997) définit la théorie comme « un ensemble de propositions servant à unifier de façon logique des concepts afin d'expliquer et d'interpréter certains aspects de la réalité dont l'on cherche à rendre compte » (p. 2).

Pour ce qui nous intéresse, parler de théorie de l'évaluation, si l'on voulait bien risquer l'expression, pourrait consister à n'utiliser que cette deuxième définition : ce serait une construction intellectuelle utilisant un ensemble d'idées, de modèles (Mottier Lopez et Figari, 2012), de paradigmes cherchant à expliquer le fonctionnement des phénomènes liés à l'évaluation (attitudes, méthodes, interactions, effets...). Or ce que nous disent clairement les travaux nord-américains rassemblés par Shadish *et al.* (1991) c'est que des évaluations non fondées sur une vision théorique de l'objet à évaluer ainsi que de la méthodologie d'évaluation elle-même, ne peuvent être valides. C'est dans ce sens qu'on abordera, dans la deuxième partie de l'ouvrage, les possibilités d'intégration de cette théorisation dans l'élaboration des « référentiels d'évaluation » qui constitueront le pivot de la construction méthodologique, point d'orgue de cet ouvrage.

3 DIFFÉRENTES POSTURES CONCERNANT LA DIMENSION THÉORIQUE DE L'ÉVALUATION

3.1 L'évaluation ne serait qu'une pratique ?

Cette première affirmation s'avère encore courante : elle est tentante, en effet, puisqu'elle permet de se dispenser de réfléchir sur les fondements scientifiques de l'évaluation. Cette dernière resterait la seule pratique sociale et éducative non étudiable et non conceptualisable. Il s'agirait donc de fonctionner à l'aveugle en étant influencé par les injonctions de tel ou tel responsable administratif, du type : « Évaluez ! Il en restera toujours quelque chose ». Cette posture, à notre avis équivoque, peut inciter de nouveaux formateurs ou enseignants à recourir, sans trop y réfléchir, à des « fichiers » d'outils prêts à l'emploi, comprenant des protocoles préparés d'avance pour évaluer. En poussant un peu plus loin cette logique, on en arrive à la recherche sur internet de protocoles appelés souvent « modèles », promus par des auteurs qui prétendent aider à évaluer mécaniquement des apprentissages, fascinés par des batteries d'indicateurs déjà prêtes à l'emploi et applicables en toutes circonstances.

On voit bien que de telles pratiques, étant dépourvues de fondements, peuvent être difficilement crédibles.

3.2 L'évaluation ne serait pas à confondre avec l'activité scientifique ?

Barbier (2006), a lancé ainsi le débat : « Beaucoup de publications récentes assimilent activité scientifique et activité d'évaluation » (p. 358). Pour Barbier, ce n'est pas parce que l'évaluation produit de la signification qu'il faut « l'assimiler à une activité scientifique » et il voit, dans cette confusion, entretenue par les évaluateurs, d'une double appartenance de l'évaluation à la recherche et à la pratique, un désir de promotion de leur activité et de leur profession. Le débat peut continuer et laisser répondre que ces deux activités ne sont certes pas à confondre, mais qu'il est évidemment souhaitable que la recherche en sciences de l'éducation ou en sciences politiques, par exemple, soit utilisée comme source de référents majeure des choix méthodologiques opérés en évaluation. La position défendue ici est claire sur ce point : sans confondre évaluation et pratique scientifique, nous tentons de renforcer la dimension conceptuelle de l'activité d'évaluation et de traduire en procédures ce que les auteurs anglo-saxons déjà cités appellent évaluation *dirigée par la théorie*. C'est-à-dire que, pour nous, il s'agit de fonder scientifiquement des pratiques d'évaluation en les considérant comme des objets potentiels de recherche et de tenter d'en modéliser le processus afin de pouvoir lui appliquer un protocole méthodologique, celui de la « référentialisation » (cf. *infra*).

3.3 Une science de, par ou pour l'évaluation ?

Certes, il n'existe pas de consensus pour proclamer une *science de l'évaluation* : pas de savoir institué disponible, pas de théorie générale reconnue, pas de méthodologie inattaquable spécifique, pas de résultats nomothétiques¹ connus. En ce sens, on ne parlera pas d'« évaluation scientifique » en tant que telle. De même, il ne pourrait pas être question, non plus de *science par l'évaluation* ce qui reviendrait à tomber sous le coup de la critique précédente de Barbier : on ne peut pas soutenir longtemps qu'on *fait de la science* lorsqu'on fait, spécifiquement, de l'évaluation. Il resterait à tenter la description d'un cheminement scientifique (une *science pour l'évaluation*) qui apporterait à l'évaluation un certain nombre d'attributs : la conceptualisation préliminaire autour de l'objet à évaluer, la modélisation préalable du processus évaluatif, le recours à des méthodes de traitement d'information validées par les sciences sociales, l'utilisation du raisonnement argumentatif fondé sur la recherche de la preuve, respectant les étapes de l'exploration, de la problématisation, de la construction méthodologique suivie d'un traitement de données, le tout au service de la réponse aux questions de départ.

1. Nomothétique : se dit d'une science dont l'objet et la méthode permettent d'établir des lois générales ou universelles, représentées par des relations constantes entre les phénomènes observés.

Chapitre 5

L'évaluation, une activité langagière

1 LE PARLER ÉVALUATIF

Le *parler évaluatif* est en effet un langage particulier. Il prend parfois la forme d'un échange entre initiés ; il est codé (appréciations stéréotypées, échelles graduées, chiffres ou lettres...) et surcodé (signification différente d'un résultat selon son origine...) tout en présentant des possibilités de décodages très différentes.

Dans les situations éducatives données à évaluer, les processus explicatifs des activités de formation comme ceux relatifs à l'évaluation sont souvent observables et analysables à travers les traces qu'elles produisent. Ces traces deviennent alors des données sur lesquelles s'exercent les descriptions de la situation et sont intégrées dans ce que nous sommes tentés d'appeler le « discours évaluatif » dont les fonctions essentielles consisteront à nommer, à conceptualiser et à interpréter les relations pouvant être établies entre ces données. On a affaire, généralement, à deux types de discours évaluatifs : le discours courant et le discours scientifique, deux discours qui s'entremêlent parfois et qui, selon nous, mériteraient d'être davantage étudiés.

1.1 Le discours évaluatif courant

Les contenus et les raisonnements propres au discours utilisé dans les actes d'évaluation sont bien connus : promulgation et commentaire de résultats,

rapports, bilans, comptes rendus de jurys, édition de palmarès, etc.). Ce discours est constitué d'énoncés divers :

- il s'agit, souvent, d'énoncés de jugements étayés par des données appelées *résultats* ;
- chaque énoncé contient un raisonnement empirique du type : il est attribué telle valeur à telle performance, au vu de tel résultat et en référence à telle échelle de mesure ou de classement ;
- le raisonnement utilisé relève du bon sens (ou, parfois de l'expérience professionnelle), ne subit aucun contrôle et se voit validé par le consensus d'un milieu professionnel ou institutionnel donné.

C'est un discours reconnu et banalisé : il est souvent employé dans les échanges entre enseignants, parents et enfants autour de l'école, entre journalistes et public, entre décideurs et employés dans les organisations. Même s'il n'est pas incontestablement fondé, il est, en général, formulé avec conviction et peut faire autorité dans certaines situations (conseils de classe, comptes rendus d'audits, publications de classements ou de palmarès...).

Ces jugements peuvent appeler d'autres jugements ou contre-jugements ainsi que des contestations de résultats : cela confère parfois à l'évaluation une connotation incertaine, discutable, polémique et contribue quelque peu à brouiller l'image de l'évaluation.

1.2 Le discours évaluatif scientifique

C'est celui, nourri par les sciences sociales, le plus fréquemment utilisé dans les évaluations réalisées par des chercheurs et des experts et concernant des dispositifs (programmes, projets, curricula, politiques publiques) ou des situations éducatives de type apprentissages et acquisition de compétences. L'évaluateur utilise alors une valise de ressources conséquente, entraînant l'usage d'un vocabulaire spécialisé, plus technique et plus abstrait, en même temps :

- un arsenal conceptuel emprunté aux différentes disciplines-outils concernées (notions de cognition, de compétence, de représentation, d'opinion, d'image de soi, de conflit socio-cognitif, etc.) ;
- une boîte à outils particulièrement adaptée aux démarches d'évaluation (recueil et analyse qualitative et quantitative de données d'observation et d'opinions, analyse de contenu de textes et de discours, etc.).

L'imprégnation du discours évaluatif par les sciences sociales est telle que se dessine, depuis quelques années, une évolution dans le sens d'une certaine « confiscation » des activités d'évaluation par des disciplines de référence comme l'économie, la sociologie et les différentes sous-disciplines de la psychologie (cognitive, sociale expérimentale, du travail). On voit apparaître, en effet, des publications intitulées « évaluation de... » uniquement

construites sur des protocoles internes à ces disciplines, comme si l'évaluation n'existait pas, en elle-même, en tant que problématique complexe, interactive et pluridisciplinaire. On aperçoit alors le danger inverse à celui que signalait Barbier (2006, *supra*, p. 51) : celui de confondre une démarche scientifique à part entière (d'origine disciplinaire) avec une évaluation (même à dimension scientifique).

2 LE LANGAGE : UN PROBLÈME POUR L'ÉVALUATION

Qu'il soit courant ou de nature scientifique, le langage spécifique de l'évaluation pose problème. S'il reste général et approximatif, on a vu qu'il devenait peu crédible, sujet à beaucoup de subjectivité et d'arbitraire. Si, au contraire, il a recours à des notions spécifiques constituant un vocabulaire technique et conceptuel, il peut être perçu comme hermétique, abscons et, par conséquent, incompréhensible.

Et, pourtant, il ne peut y avoir de transactions d'évaluation sans communication entre acteurs de ces transactions. Plus encore : il ne peut y avoir communication évaluative sans recours à des concepts, des notions les plus précises possibles. Tout évaluateur (professionnel ou occasionnel) est amené à comprendre et à parler un langage spécifique. L'utilisation du langage fait partie du processus d'évaluation lui-même : il structure les échanges entre évaluateurs et évalués, la rédaction des rapports, les débats en jurys.

Nous allons, tout au long de cet ouvrage, utiliser et construire un langage adapté à la méthodologie de l'évaluation en convoquant les concepts utiles à la présentation des protocoles et en adaptant sinon de nouveaux concepts, du moins de nouvelles configurations de concepts parce qu'elles seront jugées indispensables à la conduite d'une opération d'évaluation.

Conclusion : une évaluation influencée par ses fondements théoriques

Cet ouvrage poursuit son objectif de constituer une ressource pour la construction de protocoles d'évaluation. Pour atteindre cet objectif, on a voulu, dans cette première partie, faire passer l'idée d'une priorité, d'un pré-requis incontournable selon nous : celui de pouvoir fonder théoriquement une méthodologie.

Nous partons de l'idée que l'évaluation n'est pas une pratique ordinaire ni l'étiquette que peut se donner toute étude disciplinaire portant sur un fait éducatif : pour nous, c'est un processus intégré à une technologie (Figari, à paraître) à part entière qui se développe dans un champ de pratiques et de concepts spécifiques et qui s'appuie sur des raisonnements scientifiques qui lui sont propres. Pour nous, il y a matière à théoriser l'évaluation et à réfléchir sur une épistémologie qui la concerne : nous avons écrit, ailleurs (Mottier Lopez et Figari, 2012), à l'appui de plusieurs auteurs, que « malgré le cloisonnement des sciences sociales et des sciences de l'éducation » qui auraient tendance à favoriser une sorte d'éclatement des approches scientifiques de l'évaluation, il y a « un espace » pour la « théoriser » en tant que phénomène « situé », dans le sens d'une « théorie pragmatique ».

Mais, sans aller si loin, ici, et pour résumer cette démarche, rappelons comment cette « fondation théorique » a été suggérée. On a voulu cheminer, à partir des discours sur l'évaluation (réalité vécue quotidiennement par les professionnels de cette activité) pour arriver aux significations théoriques sous-jacentes à ces discours. L'idée qui a entraîné ce cheminement consistait à montrer qu'il ne pouvait pas exister d'opposition entre pratique et théorie et qu'en même temps, toute pratique prenait une signification lorsqu'on décryptait le fondement théorique qu'elle supposait. Autrement dit, on peut réfléchir sur les fondements théoriques de l'évaluation sans passer par des prolégomènes savants et détaillés : on dispose, déjà, dans les discours habituels entre praticiens, par exemple, de traces évidentes de conceptualisation de cet objet.

Les débats tels que « mesure et contrôle *versus* évaluation », « quantitatif et qualitatif », « externe et interne », etc., ont fait apparaître des choix

de concepts féconds en matière de construction d'outils méthodologiques. La référence à des « cultures de l'évaluation » a mis à jour des modèles sous-jacents pouvant réinterroger les pratiques spontanées des évaluateurs. La réflexion sur les « définitions » de l'évaluation (du jugement de valeur à la recherche de sens en passant par l'aide à la décision) a révélé le statut téléologique de ces formulations : on a vu que « définir » l'évaluation c'est, en quelque sorte, annoncer les finalités qui lui ont été assignées. Enfin, pour faire la synthèse de cette revue préalable, on pourrait oser annoncer les liens indissociables entre évaluation et théorie : il faut les distinguer (l'évaluation n'est pas, à proprement parler, une activité de recherche) et aussi les relier (l'évaluation peine à produire du sens si elle n'est pas fondée théoriquement de même que la méthodologie ne pourra être reconnue que si elle s'appuie sur des modèles théoriques).

Ainsi fondée, située et annoncée, l'activité évaluative peut alors revêtir ses habits pratiques et instrumentaux.

Un principe organisateur apparaît : toujours des concepts avant les outils.

Mais, sans aller plus loin dans le débat épistémologique, on peut d'ores et déjà déduire de ce qui précède une idée fondamentale qui présidera à la présentation des chapitres suivants : les concepts précèdent les outils. Il a été montré, en effet, que le sens produit par une opération d'évaluation ne sera construit qu'en fonction de conceptualisations préalablement posées, autour de différentes formes comme :

- le modèle de l'objet à évaluer (autour de l'apprentissage, du dispositif, du programme...) ainsi que le modèle d'évaluation adapté (externe et objectiviste, interne et subjectiviste, négocié et interactif ?) ;
- le champ théorique de référence (sociologique, psychologique, éducationnel, économique, etc.) permettant de situer l'objet dans un domaine scientifique repérable et offrant les outils d'analyse correspondants ;
- la procédure de formulation de critères et de construction d'indicateurs adoptée ;
- le référentiel utilisé ou (re)construit pour l'occasion : c'est l'objet la deuxième partie.

Partie 2

MÉTHODOLOGIE
D'ENQUÊTE ÉVALUATIVE
PAR LA RÉFÉRENTIALISATION

Introduction

1 DES CONCEPTS À LA MÉTHODE

On vient de voir que le discours utilisé pour décrire les pratiques d'évaluation ne peut pas être dissocié des fondements évoqués (les différents courants qui ont marqué les conceptions diverses de l'évaluation, les cultures de l'évaluation, ses différentes définitions, les rapports entre évaluation et théorie). La deuxième partie aura donc pour but, tout naturellement, de tenir compte, dans l'élaboration d'un protocole à la fois stable et reproductible, des composantes théoriques énoncées dans la partie précédente.

Nous partons de l'idée que l'évaluation ne se réduit pas à un jugement ou à une appréciation personnelle (voire collective) mais qu'elle fonctionne comme un processus fortement impliqué dans des situations d'apprentissage, d'éducation, de formation et de communication. Nous prenons en compte, ensuite, la nécessité de procéder à un traitement d'informations appuyé sur des méthodes de sciences sociales (comme le questionnaire, l'entretien, l'analyse de documents et de textes, l'observation de situations). Nous tentons, enfin, d'asseoir la méthodologie présentée dans cet ouvrage sur une base théorique d'ensemble. Pour ne pas se cantonner dans une observation relevant du sens commun, fallait-il convoquer une procédure s'apparentant à la recherche scientifique ? Nous avons cherché, chez Dewey (1967) une possible réponse à cette question dans le cadre de sa « théorie de l'enquête » qui fournit de nombreux éléments de réflexion transposables, selon nous, à la problématique d'une méthodologie de l'évaluation.

2 UNE ENQUÊTE ÉVALUATIVE

L'auteur inscrit, en effet, sa conception de *l'enquête* dans une « logique » de recherche conjointe de vérité et de solutions aux problèmes humains et sociaux. Il en développe le caractère situé dans un environnement et inscrit dans l'expérience. Il définit enfin l'enquête, en réponse à une situation indéterminée, confuse, problématique : l'enquête « transforme une situation indéterminée en une situation déterminée » (Dewey, 1967, p. 22). C'est la démarche de l'être « en quête » de l'élucidation de « l'obscurité de son

environnement, de sa situation, à un moment précis de son développement » (p. 23). Dewey conçoit l'enquête comme un moyen pour penser et agir sur les problèmes que l'on rencontre dans son activité. Or l'évaluation peut être pensée comme une forme d'enquête au sens que lui donne l'auteur. C'est pourquoi nous avons cherché à inscrire la démarche méthodologique qui va suivre dans cette orientation.

Selon nous, l'évaluation participe de *l'enquête sociale*, telle que la définit Dewey : on la reconnaît en effet dans le fait qu'elle utilise l'observation de faits éducatifs (par exemple : des performances, des changements de comportements, des acquisitions, etc.) pour émettre une suggestion, une *idée* (par exemple : le poids du facteur *méthodes pédagogiques* ou celui de l'enseignant dans les progrès en matière d'acquisitions scolaires) qui, mise en relation avec ce que l'auteur appelle un « système de significations », peut devenir explication, connaissance, solution. L'évaluateur serait un enquêteur amené à observer et interroger une situation, a priori complexe, confuse et problématique, au moyen d'outils pertinents et reconnus, pour répondre à l'attente de solution d'un problème de sens posé par les acteurs de cette situation, avec le souci d'établir une forme de vérité, au sens pragmatique, c'est-à-dire en capacité de résoudre des problèmes ou de tirer des règles d'action.

Tout en empruntant aux techniques de l'enquête sociologique, ce que nous appellerons, dans le prolongement de ce que dit Dewey, *l'enquête évaluative* demande à être conçue avec sa finalité propre (« end-in-view ») consistant à répondre à un problème d'attribution de valeur (au sens anglo-saxon de « valuer »¹) à un objet préalablement mesuré et comparé. Il nous appartient donc, dans le développement de cette deuxième partie de l'ouvrage, de décrire ce qui nous paraît pouvoir constituer une méthodologie de conception d'une telle *enquête évaluative*. Cette procédure consistera, dans le sens de Dewey, non à *enquêter sur* un sujet, un dispositif ou un groupe, en restant à l'extérieur de ce sujet, ce dispositif ou ce groupe, mais à construire un protocole de « transformation d'une situation indéterminée en une situation déterminée » (déjà cité), « en connexion avec un tout contextuel ».

3 VERS UNE MÉTHODOLOGIE D'ENQUÊTE ÉVALUATIVE

Il s'agit donc d'aborder la construction de protocoles méthodologiques d'évaluation :

- en s'appuyant sur les bases définies jusqu'ici (conceptualisations qui éclairent et précisent ce qu'évaluer implique), exposées en première partie ;

1. Valuer : « attribuer une valeur à » selon Dewey (2011).

- en recourant à une modélisation du processus d'évaluation (telle qu'exposée *infra*) ;
- en procédant au repérage et/ou à la construction de ce que Dewey désigne par « système de signification » et qui s'apparente, selon nous, à un « référentiel », du moins à un artefact qui en tient lieu et sur lequel il conviendra de travailler en préalable.

Dans un premier temps, il paraît donc utile de traiter de la notion de *référentiel* et d'examiner la place que ce phénomène occupe dans l'évaluation, en relation avec d'autres formes du même lexème : référent, référé, référentialisation (Figari, 1994, p. 43). Ce sera l'objet du chapitre 7. Après quoi, on procédera à la description annoncée d'une méthodologie d'enquête évaluative qui prendra la forme d'une *méthodologie d'évaluation par la référentialisation*, en détaillant les différentes phases qui la développent (7-4).

4 MAIS QUI SONT LES ÉVALUATEURS ?

Dans la suite de cet ouvrage, nous emploierons souvent les expressions « l'évaluateur » ou « les évaluateurs ». Il serait légitime qu'on nous demande de qui nous parlons ; c'est le moment d'y répondre. De manière générale l'évaluateur est la personne qui évalue un objet quelconque, une situation, etc. Mais répondre par une telle évidence ne suffit évidemment pas. En effet, dans une opération d'évaluation, sait-on toujours repérer et reconnaître les personnes qui évaluent par différence avec celles qui n'évaluent pas ? Par exemple, dans une situation didactique, dirons-nous que l'évaluateur est l'enseignant ? L'élève ne participe-t-il pas, lui aussi, à l'évaluation de ses connaissances et de ses compétences ? Dans une situation d'inspection, dirons-nous que l'évaluateur est l'inspecteur ? L'enseignant n'est-il pas amené à exprimer, à son tour, l'appréciation de ses propres résultats ou performances ? Dans une situation d'évaluation de politique, de projet, de dispositif ou de programme, dira-t-on que l'évaluateur est l'expert, le consultant ou l'auditeur ? Les acteurs concernés ne sont-ils pas conviés, eux aussi, à discuter la pertinence d'une méthode ou la validité d'un résultat quand ils ne sont pas, eux-mêmes, membres du groupe de pilotage de l'évaluation ?

Les évaluateurs sont donc les personnes qui *participent* à une évaluation : commanditaires ou initiateurs, professionnels, sujets *soumis* à évaluation et, plus généralement, acteurs participant à un processus d'évaluation. Leur statut et leur pouvoir varient avec les institutions dans lesquelles l'opération se déroule.

Chapitre 6

La place des référentiels dans l'évaluation

1 L'OMNIPRÉSENCE DES RÉFÉRENTIELS

On remarque que les référentiels occupent une place importante dans le fonctionnement des programmes scolaires, de la formation professionnelle et des dispositifs éducatifs en général. Ils sont souvent utilisés comme des *allant de soi*, rarement définis par leurs concepteurs ou par leurs diffuseurs. La raison de cet engouement peut s'expliquer, selon Pastré (1999), par un besoin de se référer à des normes, dans une société marquée par l'anomie (absence d'organisation sociale résultant de la disparition des normes communément acceptées). Elle peut aussi répondre à une volonté d'harmoniser le système et de marquer l'expression d'un pouvoir institutionnel ou simplement disciplinaire. On trouve, par exemple :

- des *référentiels métier* (appelés parfois, *référentiels emploi*) : ceux-ci décrivent les activités, les compétences et les certifications caractérisant les différents métiers. En France, les fiches ROME¹, produites par le Pôle Emploi, en sont un premier exemple. Mais certains corps de métiers éditent leurs propres référentiels ;
- des *référentiels de compétences* : ceux-ci peuvent être détachés des précédents et formuler, de manière très précise, les capacités à répondre aux différentes tâches affectées à la pratique de tel ou tel métier. Ce type de référentiel est arrivé à l'école avec des « socles

1. ROME : Répertoire opérationnel des métiers et des emplois.

communs de connaissances et de compétences » (Ministère de l'Éducation nationale, 2005) constituant des « cadres de référence de la scolarité obligatoire » stipulant que chaque élève doit parvenir à la maîtrise du socle commun au terme de sa scolarité ;

- des *référentiels de formation* : ceux-ci proposent ou édictent un ensemble d'objectifs et de contenus d'une formation à concevoir dans le cadre d'une institution en charge d'organiser ce type d'activité ;
- des *référentiels d'évaluation* : ces derniers prenant souvent la forme
- de *référentiels de certification* comme, par exemple, en France, ceux de l'AFPA², qui, après avoir défini un emploi-type et une liste de compétences afférentes, énonce des critères d'appréciation et des seuils requis pour la tenue de l'emploi visé. Le document énonce, enfin, des modalités de mise en œuvre de l'évaluation.
- de *référentiels qualité*, comme celui de l'AERES (2009) qui organise un protocole d'évaluation de l'efficacité du système qualité mis en œuvre par les établissements d'enseignement supérieur, en France.
- et, enfin, de manière plus explicite, encore, de *référentiels d'évaluation* proprement dits, tel le référentiel des critères d'évaluation de l'AERES (2012). Ce document explicite la méthodologie d'évaluation définie, cette fois encore, par l'AERES : il précise les critères d'évaluation, indique pour chaque critère, le champ de l'évaluation couvert, les faits observables et les indices de qualité qui lui sont associés.

2 DES UTILISATIONS DIVERSES DE LA NOTION DE RÉFÉRENTIEL

Avant de les analyser et, par la suite, de les reconsidérer, il importe donc de passer par le détour de leur définition. Pour ce faire, on peut se reporter à des écrits antérieurs en rappelant (Figari, 1994, 2006) :

- l'idée qu'il s'agit d'un *système de référence(s)* soit « général », soit, de manière plus précise, d'un ensemble de référents organisé en fonction d'une opération d'évaluation particulière (Figari, 1994, p. 44) ;
- l'idée que le *référentiel* a pour fonction de rapporter les résultats de toute évaluation aux points de vue construits par l'évaluateur et par ses partenaires (Figari, 2006, p. 104). On se limitera donc à synthétiser, ici, quelques aspects du concept de *référentiel*. Tout d'abord, rappelons que ce dernier ne peut se comprendre sans recourir à celui de *référent* dont il est issu : le *référent* désignant l'élément à quoi un objet peut être rapporté ou comparé ou pouvant également

2. AFPA : Association de Formation Professionnelle des Adultes. Espace certification, exemple de la fiche présentant le métier de formateur.

constituer une information prise en référence pour expliquer une autre information ;

- la nécessité de différencier ce que nous venons d'appeler *référentiel* et l'équivalent d'idéal à atteindre, de norme par rapport à laquelle un jugement de valeur est porté : cette confusion engendre de fréquents malentendus que la « référentialisation » qui va être décrite par la suite aura pour but de dissiper ;
- le fait que, enfin, dans le monde de l'éducation et de la formation, le mot *référentiel* désigne souvent un ensemble de référents organisé autour d'un projet, d'un programme ou d'un dispositif de formation et/ou d'évaluation, souvent perçu comme « déjà là », comme nous allons le voir.

3 DES RÉFÉRENTIELS « DÉJÀ-LÀ » ?

On appellera ainsi les documents appelés *référentiels* ou en tenant lieu, émis par des institutions et invoqués par les acteurs comme normes de référence légitimes et non discutables : ils sont « déjà là » et, parfois acceptés sans remise en question, parfois contestés ou critiqués.

Ils sont formulés, le plus souvent, en termes de :

- prescriptions qui vouent la fonction éducative ou formative à l'application de modèles d'apprentissage implicites élevés au rang de marche à suivre ;
- normes qui sont omniprésentes et diversement perçues : assimilables à des référents de contrôle, elles peuvent être interprétées comme l'expression d'une soumission demandée par l'institution qui les produit (par exemple, le PER³ en Suisse). Comprises comme des règles, elles peuvent être intégrées... Perrenoud (1989) va même jusqu'à dire qu'« on ne peut évaluer qu'en fonction d'une norme, implicite ou explicite, qu'elle soit fixée d'avance ou se dégage de la comparaison même ». L'auteur distingue « deux types de normes : les règles de conduite d'une part, les normes d'excellence d'autre part » (p. 19).

On voit bien que les *référentiels déjà là* posent questions. Même si certains d'entre eux sont accompagnés de documents explicatifs, ils ne revêtent pas la même légitimité que les référentiels construits en situation, en collaboration avec les acteurs. Ils sont porteurs d'une légitimité institutionnelle alors que les seconds relèvent d'une légitimité locale, celle reconnue par les membres de la communauté éducative concernée. Le statut de ces référentiels demande, par conséquent, d'être revisité.

3. PER : Plan d'études romand (2001) définissant les missions, finalités et compétences attendues, discipline par discipline, des élèves suivant la scolarité obligatoire.

4 UNE DÉFINITION-PRINCEPS DU RÉFÉRENTIEL ?

Il n'est pas possible d'en rester là et la notion doit être clarifiée. Une référence-princeps nous semble pouvoir être convoquée ici : il s'agit de la définition-source qui reste celle du modèle cartésien : l'ensemble des axes (abscisses et ordonnées, par exemple) par rapport auxquels on situe un point... et qui permettent de définir les coordonnées de ce point. Si on généralise à l'espace ce modèle appliqué à un plan, on dispose d'une infinité d'axes possibles, ce qui permet de schématiser un référentiel autour de n axes correspondant à autant de questions, de dimensions structurant le mandat d'évaluation et de réponses à organiser. Par ailleurs, les axes du modèle sont indépendants et extérieurs au point dont on cherche les caractéristiques : cette extériorité par rapport à l'objet réel définit le statut de tout référentiel.

Avec un *référentiel* ainsi décrit, on est loin des normes et des prescriptions précédemment évoquées et constituant les fonctions perçues comme essentielles de ces objets.

5 RÉFÉRENTIELS ET ÉVALUATION

On a vu que l'évaluation est parfois définie comme la mesure de l'écart (ou, du moins comme l'appréciation de la relation) entre l'objet à évaluer (appelé « référé », Hadji, 1989) et un comparant (« le référent », notion connue et répandue dans la littérature spécialisée). Ce qui confirme bien que, même si on l'oublie c'est toujours le référent qui construit le sens de l'évaluation : par exemple, une note obtenue par un élève n'a pas la même signification si elle est comparée :

- aux objectifs du notateur ou à ceux des autres acteurs (sujets évalués, parents, recruteurs, etc.) ;
- à la moyenne et à l'écart-type des notes des autres sujets ou à des moyennes régionales ou internationales.

Un *référentiel d'évaluation* va alors revêtir plusieurs fonctions différenciées selon que cette évaluation interviendra *ex post* (après l'action, pour en apprécier l'impact) ou *ex ante* (synonyme d'étude de faisabilité préalable à l'action)⁴ ou encore *chemin faisant* (lorsqu'elle intervient en cours de projet, par exemple).

S'il s'agit d'une évaluation *ex post*, il servira de grille d'analyse de situations en place dans lesquelles le mandat d'évaluation intervient après l'action de formation : la reconstitution explicite d'un référentiel implicite sera alors nécessaire pour argumenter l'élaboration du protocole lui-même.

4. *Ex post*, *ex ante* : expressions souvent utilisées dans l'évaluation des politiques publiques au sein de la Communauté européenne.

Dans le cas d'une évaluation *ex ante*, il servira d'ensemble organisé de référents permettant d'offrir une taxonomie de comparants auxquels les acteurs (enseignants, formateurs, évaluateurs) seront invités à confronter une réalité observée. Le référentiel sera alors apte à offrir un choix d'éléments censés représentatifs de l'objet, assortis de la formulation de critères et/ou d'indicateurs⁵ proposés comme éléments d'un guide méthodologique.

Dans les deux cas, ces comparants (référents, critères, indicateurs) ne seront pas nécessairement des prescriptions, des normes ou des idéaux et pourront aussi être constitués d'informations choisies pour leur aptitude à éclairer une situation soumise à évaluation (par exemple : une moyenne indicative, des résultats extérieurs à ceux concernés...). En effet, on va pouvoir utiliser des informations comme un résultat obtenu lors d'une action de formation pour les comparer avec ceux obtenus par l'action en cours d'évaluation et repérer, ainsi, un écart, une différence qui deviendront signifiants. Leur signification sera d'autant plus précise qu'elle sera obtenue à partir de catégories issues de l'investigation préalable et non établies de façon arbitraire, comme c'est très souvent le cas. Le référentiel une fois conçu (reconstitué ou adopté après examen critique et/ou négociation, transaction, validation entre partenaires de l'évaluation) constituera un préalable à l'enquête évaluative.

5. Cf. pages 88-93.

Chapitre 7

Méthodologie d'évaluation par la référentialisation

1 SOURCES D'UN PROTOCOLE

1.1 Des notions préalables

1.1.1 *Référentiel et référentialisation*

On a constaté l'existence fréquente de référentiels « déjà là », c'est-à-dire préalablement publiés par l'institution qui les promulgue et perçus comme normatifs. D'autre part, on a suggéré que l'évaluation fonctionnant par mise en relation entre référés et référents, il paraissait utile de procéder, préalablement à une opération d'évaluation, à la formulation ou au décryptage d'un référentiel apte à conférer un sens aux résultats de cette opération. C'est cette procédure consistant à modéliser l'évaluation (cf. *infra*), autour d'un lien logique entre référé et référents par l'intermédiaire d'une formulation de critères et d'indicateurs, que nous avons appelée « référentialisation » (cf. Figari, 2006), pour marquer la différence entre un cadre de référence stabilisé (le référentiel) et l'activité continue d'élaboration et d'utilisation de ce document (la référentialisation). De manière générale, il s'agit de pouvoir se reporter à un référentiel, en disposant d'une vision d'ensemble du processus d'évaluation, vision partagée (ou répartie) entre les différents acteurs, sachant :

- qu'une évaluation ne pourra revêtir un sens quelconque qu'en fonction du référentiel explicite (annoncé, développé et argumenté en tant que tel) ou implicite (obéissant à une appartenance idéologique

- non consciente, à un courant d'idées dominant...) qui sera utilisé pour établir des diagnostics et poser des jugements ;
- que chaque évaluation dépend de son propre référentiel, les standards universellement applicables ne pouvant suffire, seuls, à l'attribution de signification ;
 - que, en conséquence, une « référentialisation » spécifique s'impose avant chaque évaluation.

1.1.2 La référentialisation : une notion à trois dimensions (conceptualisation, modélisation, instrumentation)

1.1.2.1 Première dimension : une conceptualisation

La *référentialisation* consiste, d'abord, en une conceptualisation par l'évaluateur du processus d'évaluation envisagé face à un objet, conceptualisation intégrée à la méthodologie elle-même et qu'elle va éclairer et justifier. Par exemple, selon que l'évaluateur regarde la situation à la lumière d'une conceptualisation objectiviste (1), subjectiviste (2) ou interactive (3) (cf. première partie, p. 42), il suivra plus particulièrement :

- un modèle quantitatif de comparaison entre une réalité et des standards (1) ;
- un modèle d'autorégulation, dans l'esprit d'une auto-évaluation mettant en œuvre des processus métacognitifs (2) ;
- un modèle de confrontation de jugements évaluatifs entre différents acteurs ou de recours à des référentiels communs ou négociés (3).

La première dimension de la *référentialisation* consiste donc en une conceptualisation de l'évaluation, en une construction d'un certain regard sur l'objet, ce qui entraînera un type d'évaluation conforme au regard qui aura été privilégié. Pour simplifier, on pourrait en conclure qu'il existera plusieurs façons d'opérer une référentialisation selon que l'on privilégiera l'un ou l'autre paradigme, à moins que les choix opérés par l'évaluateur ne consistent à réaliser une sorte de combinaison de plusieurs de ces paradigmes.

1.1.2.2 Deuxième dimension : une modélisation

La deuxième dimension de la référentialisation consiste en une possibilité de modélisation des processus de l'évaluation. Pour comprendre ce que nous entendons, ici, par *modèle*, il faut revenir à ce concept et se reporter à un écrit précédent (Mottier Lopez et Figari, 2012) : il s'agit d'une « représentation conceptuelle qui permet de rendre intelligible une situation... ou un phénomène... », « à des fins scientifiques mais aussi orientées vers l'action... » (p. 14). Un *modèle d'évaluation* prend, généralement, une forme préconstruite, validée par des recherches précédentes et peut représenter des niveaux de phénomènes (différents, selon les auteurs : par exemple, chez Stufflebeam, déjà cité) : c'est pourquoi l'emploi de la notion de modèle nous entraîne souvent vers des débats, certes respectables mais

peu utilisables en matière de méthodologie. On s'intéressera donc davantage à une *modélisation* plutôt qu'à un modèle, ceci pour deux raisons :

- le *modèle* désigne un schéma déjà établi et qu'il s'agit de suivre alors que la *modélisation* consiste à construire un *modèle* : peut-être moins nomothétique mais certainement plus adapté au déchiffrage du type d'objet à évaluer ;
- la *modélisation* constitue une procédure active faisant apparaître les relations stables de cause à effet ou de comparaison pouvant exister entre des éléments à observer : elle est donc un levier de l'évaluation.

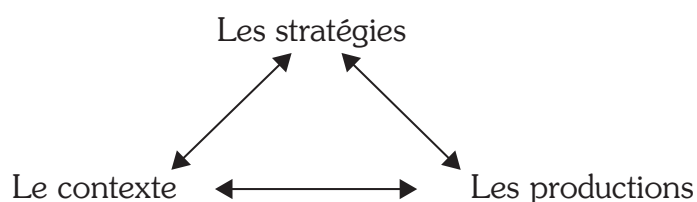
Arrivons à la modélisation envisagée. On a postulé, plus haut, que l'évaluation consiste à comparer différentes données décrivant des dimensions jugées représentatives de l'objet et que l'évaluateur va isoler et observer (à l'exclusion d'autres dimensions possibles) et qui nous semblent pouvoir être définies ainsi :

- *le contexte* de l'action à évaluer : ce qui relève des « déterminants » (éléments du contexte social, géographique, historique ; état des lieux et de la demande, état de la littérature en matière de travaux similaires disponibles) ;
- *les stratégies* des acteurs concernés : ce qui concerne les objectifs et les processus d'élaboration et de réalisation de projet, de programme, de séquence, de méthode, d'enseignement-apprentissage, etc. ;
- *les productions* : les résultats, les images, les effets, les documents produits (donnés à voir ou à utiliser).

L'évaluation, dans cette modélisation, consistera, par exemple, à confronter des données caractérisant le contexte avec celles décrivant les deux autres dimensions : stratégies et productions.

Le schéma suivant symbolise ce jeu de relations (et plus exactement d'interrelations) entre les données représentant successivement les éléments de contexte, de stratégies et de productions retenus comme significatifs du processus qui peut rendre compte de la méthodologie d'évaluation envisageable dans les situations à observer :

Figure 2. Première formulation du modèle d'évaluation



L'hypothèse de travail qui explique ce schéma consiste :

- à poser le *contexte* comme élément pouvant expliquer des *stratégies* (première forme d'évaluation) ;

- à suggérer que les *stratégies* orientent les types de *productions* (deuxième forme d'évaluation) qui s'ensuivront et que ces mêmes *productions* sont censées provoquer des effets sur le *contexte* pré-existant (troisième forme d'évaluation).

Proche du modèle CIPP de Stufflebeam (1980, p. 401), ce schéma s'en différencie cependant par sa caractéristique systémique plus prononcée. Ainsi, à la différence du profil linéaire du CIPP qui invite à « suivre » le processus dans l'ordre où il est esquissé (on s'occupe du contexte, ensuite des intrants, ensuite des processus, etc.), le modèle ci-dessus (qui sera développé *infra*) indique, d'emblée, qu'il n'y a pas de début et de fin et que, d'autre part, les relations entre les pôles sont interactives et itératives.

On retrouvera des traces de ce schéma dans des formes de modélisation plus élaborées (voir *infra*, p. 75-78) se rapportant aux relations entre ces trois pôles.

1.1.2.3 Troisième dimension : une instrumentation

L'instrumentation de l'évaluation, c'est la construction des outils d'enquête qui adoptent les formes classiques de l'observation, de l'interrogation et de l'analyse, dans une logique de référentialisation : les indicateurs, une fois construits, appellent la formulation de questions ou d'items qui sont eux-mêmes insérés dans des instruments adaptés. L'ensemble de ces instruments servira à récolter les informations que le référentiel aura convoquées. Leur analyse consistera à mettre en relation les informations obtenues pour chaque élément du modèle ICP avec les autres éléments.

Dans le cadre de la référentialisation précédemment décrite, les instruments destinés au recueil et au traitement des informations nécessaires à la mesure de ces relations seront choisis et élaborés, à propos de chaque *élément à évaluer*, en fonction des référents, critères et indicateurs figurant dans ce qu'on appellera, ici, un *référentiel d'évaluation*. Les instruments en question ne seront donc pas seulement importés d'une quelconque banque d'outils ou calqués sur des procédures habituellement empruntées à des disciplines de référence (la psychométrie ou la sociologie, par exemple) mais également conçus pour interroger une situation précise autour d'une problématique préexistante.

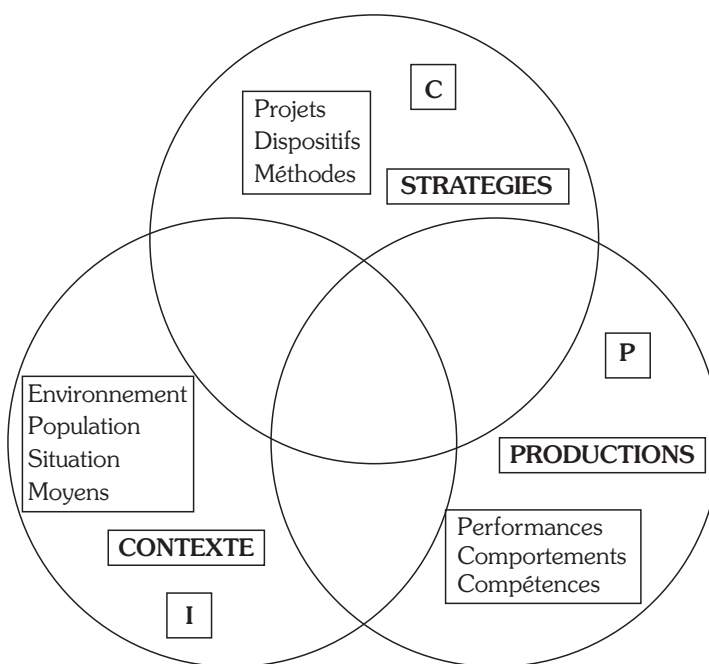
Avec la *référentialisation*, on sera ainsi passé d'une *conceptualisation* des processus d'évaluation en œuvre dans une situation donnée (à travers les définitions des processus constituant l'évaluation) à une *modélisation* des régularités du fonctionnement de cette dernière (à travers le regroupement des facteurs susceptibles de produire des effets autour de trois catégories : contexte, stratégies, productions), avant d'aborder une *instrumentation* (débouchant sur un ensemble d'outils) qui constituera la forme technique finale de *l'enquête d'évaluation*.

2 MODÉLISATIONS DE L'ÉVALUATION : SCHÉMAS PROGRESSIFS

Revenons sur la modélisation annoncée de l'évaluation car celle-ci préside au fonctionnement de l'ensemble du protocole instrumental qui est proposé.

Les schémas qui suivent tentent de la représenter. Elle a pour but de repérer les relations entre les trois dimensions I, C et P, (**I**nduit = contexte ; **C**onstruit = stratégies ; **P**roduit = productions) mises en jeu par les différentes formes d'évaluation (externe) ou de régulation (interne). On passe donc de la figure 2 schématisant les relations entre contexte, stratégies et productions à la figure 3 qui présente ces trois dimensions sous forme de trois espaces circulaires qui se croisent ainsi :

Figure 3. 2^e formulation du modèle d'évaluation



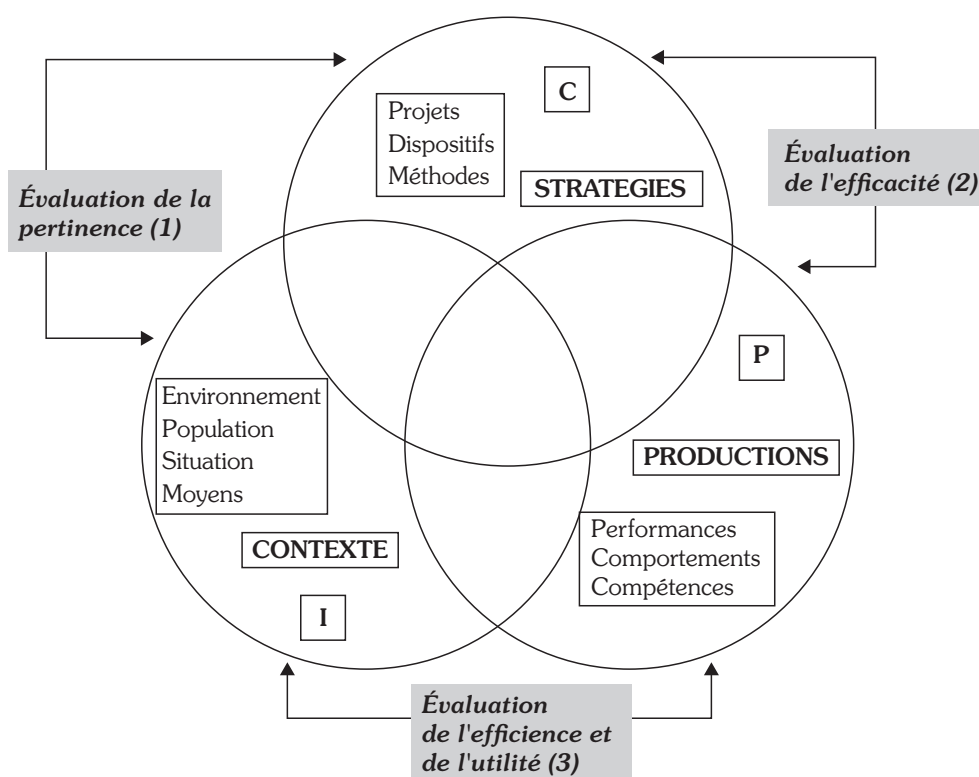
Sont posés ici, sous forme de cercles, les trois types de données que l'on suggère de construire afin de les croiser, avant d'entreprendre une démarche d'interprétation ou la production d'un jugement quelconque.

Les trois pôles du schéma précédent (destinées à devenir les matrices aptes à rassembler ces données) sont donc représentés par les trois zones identifiées comme constitutives d'une problématique d'évaluation courante (i.e. l'établissement des relations entre un contexte, des stratégies et des productions d'un objet – ou d'un dispositif – à évaluer qui seront posées dans la figure suivante). Trois zones dont il convient de préciser maintenant les interrelations, successivement, dans une perspective externe et dans une perspective interne (objet des figures suivantes).

2.1 Relations dans une perspective externe : évaluation de la qualité d'ensemble du dispositif (pertinence, efficacité, efficience)

Elles sont représentées par trois flèches à double sens reliant les trois cercles et qui symbolisent les trois formes prises alors par l'évaluation : celle de la pertinence, celle de l'efficacité, celle de l'efficience :

Figure 4. Une modélisation de l'évaluation externe



2.1.1 Évaluation de la pertinence (1)

Mettre en relation les données de contexte (I : environnement, population, situation, moyens) avec les choix stratégiques (C : projets, dispositifs, méthodes) revient à vérifier la *pertinence* de ces choix, au vu des résultats obtenus par la suite.

2.1.2 Évaluation de l'efficacité (2)

Mettre en relation les résultats attendus (C) avec les résultats obtenus (P) en termes de mesure des productions observées revient à évaluer l'*efficacité* des programmes, objectifs, moyens et méthodes déployés.

2.1.3 Évaluation de l'efficience (3)

Mettre en relation, selon un point de vue externe, les enjeux (I) et les résultats (P) revient à évaluer l'efficience d'un dispositif, en mesurant, ainsi, la rentabilité des investissements humains et techniques réalisés.

2.2 Relations dans une perspective interne : une évaluation de régulation

On se trouve, ici, dans des situations d'apprentissage, des séquences pédagogiques, des activités consécutives à la conduite de projets divers, supposant des interactions entre décideurs, acteurs de terrain, partenaires. D'où l'inévitable utilité d'une *régulation* des activités mises en place. Cette fonction de régulation est décrite par les auteurs comme une activité d'évaluation intégrée aux processus de formation et, en particulier, en matière d'apprentissage (par exemple, chez Allal et Mottier Lopez, 2007), mais elle est aussi en cause, de manière plus large, dans des situations d'évaluation à *mi-parcours* et permet alors de réorienter l'action. Cette fonction peut être utilisée, enfin, pour vérifier, au milieu du développement d'un projet, si les actions se déroulent comme prévu ou si elles appellent des correctifs et des améliorations. Ainsi, on peut distinguer :

- entre I et C : une régulation d'*ingénierie* consistant à réadapter les moyens, les supports et les autres ressources aux premiers diagnostics opérés ;
- entre C et P : une régulation *pédagogique* consistant à reconsidérer les méthodes d'intervention en fonction des réactions, difficultés ou autres phénomènes décelés en cours de fonctionnement ;
- entre P et I : une régulation de *feed-back* à partir de la confrontation entre la situation ayant fait l'objet de l'évaluation et la situation initiale.

Dans le schéma de modélisation, les relations de régulation figurent alors dans les zones d'intersection entre les trois cercles (voir Figure 5).

2.3 En synthèse : une modélisation de l'ensemble des regards d'évaluation

Il vient d'être procédé à la présentation successive de schémas modélisateurs des différents éléments à prendre en compte dans la construction d'une méthodologie de l'évaluation fondée sur la référentialisation. Mais il est important de noter que chacun des regards portés sur l'objet à évaluer n'acquiert de sens que dans une interaction avec les autres regards. C'est pourquoi il nous paraît pertinent de terminer cet exposé par le schéma d'ensemble des approches externes et internes dont il vient d'être question.

Figure 5. Une modélisation de l'évaluation interne

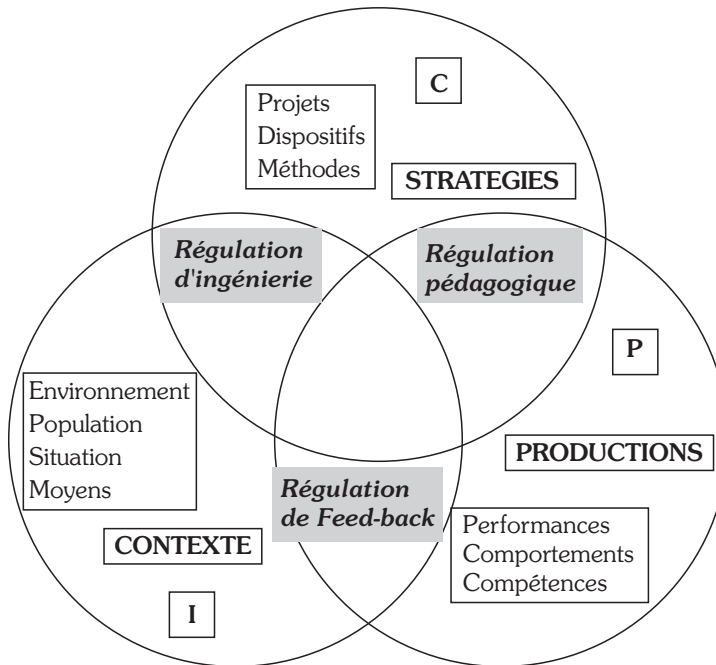
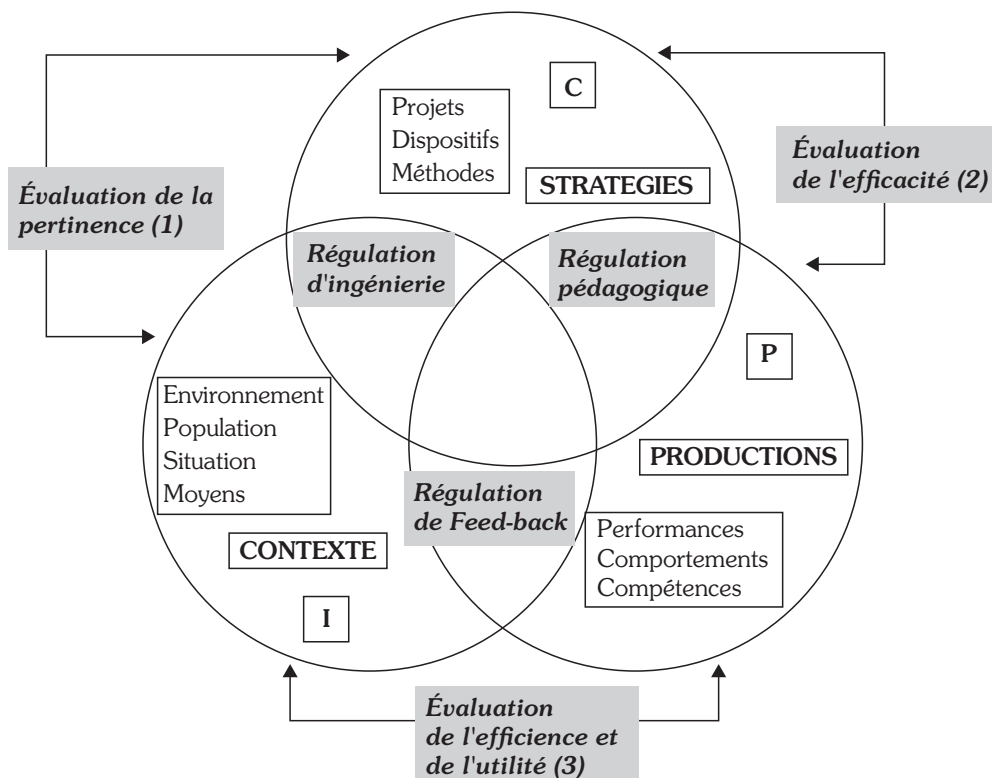


Figure 6. Modélisation de l'ensemble des processus d'évaluation



3 DÉVELOPPEMENT D'UN PROTOCOLE D'ÉVALUATION

3.1 Un protocole méthodologique

On ne va pas s'attarder sur le débat notionnel concernant la *méthodologie*. On n'en gardera que l'acception la plus simple : celle du discours sur la méthode, c'est-à-dire un raisonnement sur la construction d'un ensemble de méthodes appropriées à une investigation particulière constituant une évaluation. Alors qu'une *méthode* n'est constituée que par le recours à un outillage particulier afin de répondre à un type de question (perception par des individus ou une population, effets d'un changement, etc.), une *méthodologie* d'évaluation, plus large qu'une méthode, comprendra, à la fois, la recherche et la justification des méthodes pertinentes pour évaluer tel ou tel objet, et, également, la présentation du dispositif instrumental dans son ensemble.

La méthodologie, quelle qu'elle soit, aura donc le rôle primordial de chercher des techniques valides pour augmenter la compréhension d'un phénomène et produire des jugements argumentés et justifiés.

3.2 La référentialisation comme conception d'une méthodologie d'évaluation

Le but du travail entrepris ici consiste bien à *concevoir* (c'est-à-dire fonder et construire) une méthodologie d'évaluation sous forme de montage d'un dispositif d'*enquête* (au sens de Dewey). Cette méthodologie aura pour fonction d'élucider une situation problématique (par exemple : les effets d'un dispositif éducatif, les effets induits par une politique éducative, l'intérêt d'un outil pédagogique, les compétences acquises après le suivi d'une formation...). Cette méthodologie s'inspire de ce qui a été appelé précédemment la « référentialisation ».

3.3 En préalable, un constat et un postulat

Un constat : toute opération d'évaluation peut donner lieu à des démarches multiples (recours à des outils standardisés ou à des guides, élaboration d'instruments au coup par coup...), au choix de critères et d'indicateurs les plus divers, plus ou moins déjà établis. Dans de telles configurations, la subjectivité et l'arbitraire sont souverains. On connaît les risques de telles pratiques : désintérêt des acteurs, peur de l'évaluation (vécu d'insécurité et de déresponsabilisation, passivité, absence d'utilisation des résultats...).

Un postulat : toute opération d'évaluation concerne à la fois des acteurs et des enjeux. Les acteurs constituent, en effet, le matériau vivant de l'évaluation et, par conséquent, agissent et réagissent autour de cette évaluation. Quant aux enjeux, ils résident dans la construction, par

l'évaluation elle-même, de nouvelles significations des phénomènes, des comportements et des productions observés.

3.4 Enfin, l'organisation d'un protocole

On va pouvoir, à partir de ces préalables, proposer un protocole méthodologique de l'évaluation par la référentialisation.

Ce protocole comprend quatre phases :

La *phase 1* : exploration comportant :

- d'abord, une analyse du mandat (questionnement sur la situation à évaluer, approfondissement, décryptage) ;
- ensuite, en déduction de la démarche précédente, la formulation d'une problématique d'évaluation.

La *phase 2* : procédure centrale de référentialisation, reliant les éléments à évaluer (tels qu'ils ont été désignés par la problématique précédemment construite) avec les référents, les critères et les indicateurs nécessaires à toute évaluation.

La *phase 3* : instrumentation (choix et construction des outils destinés à renseigner les indicateurs).

La *phase 4* : analyse et interprétation des résultats produits par l'interrogation ou l'observation des acteurs et partenaires de l'évaluation.

Le tableau synoptique suivant met en perspective ces différentes phases en les illustrant uniquement par les questions auxquelles elles demandent d'apporter des réponses :

Tableau 1. Les différentes phases du protocole de référentialisation

Phase 1 Exploration	Analyse du mandat	– Quel contexte ? – Quelles stratégies ? – Quelles productions ?
	Problématique d'évaluation	À quelle(s) question(s) fondamentale(s) l'évaluation doit-elle répondre ?
Phase 2 Procédure centrale de référentialisation	Éléments à évaluer	Quelles dimensions de l'objet à observer ?
	Référents	À partir de quoi attribuer un sens à l'information ?
	Critères	Quels aspects représentatifs de l'objet à évaluer choisir pour porter un jugement ?
	Indicateurs	Quelles caractéristiques observables permettent la vérification du degré d'atteinte du critère ?
Phase 3	Instrumentation	Quels outils employer pour renseigner les indicateurs ?
Phase 4	Interprétation	Quelles significations attribuer aux résultats ?

Cette représentation en colonnes successives peut induire l'idée que l'on est invité à exécuter un travail vertical en suivant la succession des opérations. En fait, chemin faisant, on se rend compte assez vite que les catégories verticales ayant été distinguées, on a besoin d'éclairer les contenus de chacune d'elles par des allers-retours et des interactions entre catégories horizontales (dépendant davantage des aspects de la situation donnée à évaluer).

4 LES QUATRE PHASES DU PROTOCOLE D'ÉVALUATION ILLUSTRÉES PAR UN EXEMPLE-TYPE : L'ÉTABLISSEMENT SCOLAIRE

Il convient donc de dérouler, successivement, les phases annoncées du protocole. La formulation restera générale, pour s'adapter à des objets différents, mais, en même temps, elle aura besoin d'accrocher des représentations d'éléments descriptifs connus ou facilement repérables. C'est pourquoi il a été arbitrairement choisi d'évoquer, tout au long de l'exposé, un objet filigrané : ce sera l'établissement scolaire du fait qu'il a semblé rassembler un bon nombre de caractéristiques de situations éducatives plausibles, tant en matière d'apprentissage et de didactique qu'en matière de dispositifs, de contexte et de moyens.

4.1 Phase 1 : exploration (analyse du mandat et formulation de la problématique d'évaluation)

La première phase de la référentialisation constitue un moment essentiel, celui au cours duquel l'évaluateur prend connaissance d'une situation créée par une demande d'évaluation. De la compréhension qu'il tirera de ce premier travail dépend, en grande partie, la pertinence de l'opération qu'il conduira et, par conséquent, de sa pertinence et de son adéquation à la situation donnée. Deux opérations successives sont alors à distinguer : l'analyse du mandat et la formulation d'une problématique générale de l'évaluation réalisable à partir de cette analyse.

4.1.1 Analyse du mandat

Le mandat constitue une dimension fondatrice de toute évaluation. Il s'agit bien de le « décrypter », autrement dit, de le questionner, de l'approfondir, de chercher à en percevoir les spécificités, les enjeux qui l'expliquent, les incertitudes qui l'entourent, les contradictions qu'il peut déceler ainsi que les éléments essentiels qu'il comporte et qu'il est primordial de prendre en compte.

Il contient, en clair ou en implicite, des éléments nécessaires à la compréhension de ce qui va suivre : *l'annonce de l'objet formel* de

l'évaluation concernée (une politique, des apprentissages, un dispositif de formation, des méthodes de formation, des performances et des compétences) ; le *repérage des éléments de contexte* (situations appelant une opération d'enquête) et des *enjeux* qui conditionnent le projet d'évaluation (demandes des partenaires, pression éventuelle sur les résultats attendus) ; l'*analyse du questionnaire initial* contenu dans le document faisant office de mandat (interprétation et compréhension des questions et/ou des attentes des commanditaires) ; la *transformation du mandat*, par l'évaluateur, en contrat d'évaluation, tenant compte des démarches précédentes (c'est-à-dire les propositions de l'évaluateur quant à l'opération qu'il compte organiser, en réponse au mandat).

Tableau 2. Le traitement d'un mandat d'évaluation

Exemple : Traitement d'un mandat d'évaluation dans un établissement scolaire				
Mandat initial	Annonce de l'objet formel d'évaluation	Repérage du contexte	Analyse du mandat initial	Du mandat aux problématiques
Accompagnement à l'élaboration d'un projet d'établissement	Le projet d'établissement	Par ex : contexte social zones sensibles	Diagnostic de la situation	Questions d'évaluation propres à chaque élément à observer

En fait, le traitement du *mandat* demande de réaliser une véritable exploration de la situation donnée à évaluer qui en dépasse la simple compréhension. C'est, en effet, l'occasion d'opérer une étude préalable à partir d'une première analyse de données brutes obtenues auprès de l'institution commanditaire (documents descriptifs et prescriptifs, données sur les populations concernées, entretiens exploratoires). C'est à partir de ce travail que l'évaluateur peut émettre les éléments d'une problématique d'évaluation de cette situation et reformuler ainsi le *mandat* de départ devenant alors un véritable contrat de l'opération qui s'engage.

4.1.2 Élaboration d'une problématique

Le défaut fréquent des méthodes d'évaluation spontanées consiste à ne pas asseoir le questionnaire organisateur des futurs instruments sur un raisonnement interrogatif préalablement établi. En effet, l'évaluation porte rarement sur une question isolée (« mesurer tel ou tel élément », « vérifier telle

ou telle variable ») : elle découle, généralement, d'une situation complexe à partir de laquelle l'opérateur a réussi à dégager un ensemble de questions reliées entre elles, émanant d'une exploration des données attachées au mandat de départ : *la problématique*.

Pour définir de manière approfondie cette notion, nous renvoyons aux ouvrages de méthodologie de la dissertation (par exemple : Ferréol, 2000) ou de la recherche en sciences sociales (par exemple : Quivy et Van Campenhout, 2006), pour lesquels il s'agit, tout d'abord, de reformuler une *question de départ* en questionnement organisé autour d'un cadre théorique clairement défini. On peut, également emprunter aux travaux de Fabre (2009) sur la situation-problème en pédagogie qui étudie la « philosophie du problème » au service de l'apprentissage de la pensée : « Nous sommes souvent immergés dans des situations problématiques marquées par l'incomplétude et l'incertitude. C'est pourquoi la pensée ne peut être que problématique » (p. 18). Quant à Dewey (1967), il fonde le point de départ de sa Logique en affirmant que l'institution délibérée de problèmes devient un objectif de l'enquête.

On peut transposer ces réflexions au processus d'évaluation, en disant que la problématique consiste en une construction de questionnements issus d'un *mandat* et reformulés à travers un raisonnement conceptuel. Par exemple, s'il est demandé à un expert d'évaluer le projet d'une école, celui-ci analyse ce mandat et reformule alors la question de la recherche des effets du projet concerné en ayant recours au concept de « projet » (individuel, collectif, politique, éducatif, etc.) ainsi qu'à celui d'« organisation » (avec ses modèles explicatifs), aboutissant ainsi à la formulation des questionnements qui structureront le protocole d'évaluation. Ce questionnement intègre, pour finir, des hypothèses de réponses (comme la participation des acteurs).

4.2 Phase 2 : définition de l'objet à évaluer

La première phase d'élaboration d'une référentialisation consiste à définir un objet à évaluer, c'est-à-dire à transformer l'objet brut de départ en objet reconstruit. Il s'agit, alors, d'identifier et de définir les facettes pertinentes pouvant représenter le plus exactement l'objet, autrement dit, ce qui, par hypothèse, constitue les caractéristiques et les propriétés soumises à évaluation.

4.2.1 Éléments à évaluer

S'il paraît évident de commencer tout protocole d'évaluation par une définition de *l'objet à évaluer* (c'est-à-dire d'une situation, d'un apprentissage, d'un dispositif, d'un outil de formation, etc.), cette démarche ne peut consister seulement en une formulation globale. Il s'agit de faire émerger des

propriétés (ou des fonctionnalités) pouvant parfois ne pas être connues ou prises en compte par les acteurs et, par conséquent, d'analyser l'objet dans des dimensions et donc des caractéristiques observables pour en représenter toutes les facettes significatives.

Dans le cas de l'évaluation (ou de l'audit) d'un établissement scolaire, on va organiser la détermination des *éléments à évaluer* autour des catégories du modèle ICP, ce qui aboutit au tableau suivant :

Tableau 3. Décomposition de l'objet en éléments à évaluer

Les trois catégories d'ICP	Éléments à évaluer en tant que porteurs potentiels d'influences, d'effets, de résultats
I nduit Éléments facteurs d'influences (= déterminants)	I 1 : environnement I 2 : organisation I 3 : population I 4 : orientations institutionnelles
C onstruit Éléments facteurs de changement	C 1 : dispositifs C 2 : fonctionnement curriculaire C 3 : politique interne C 4 : stratégies et moyens
P roduit Éléments facteurs de résultats	P 1 : impact sur l'environnement P 2 : changements de comportements P 3 : réalisation des missions ; actions

Les trois dimensions du modèle d'évaluation ICP (Induit/Construit/Produit), présentées plus haut, peuvent se décomposer en « éléments à évaluer » :

- a) au niveau de I (Induit : contexte, situation, problèmes) : apparaissent les éléments à évaluer facteurs d'influences : l'environnement, l'organisation, la population, les missions attribuées. Ces *influences* peuvent, en effet, être attendues (espérées ou redoutées selon les cas) par les acteurs, en matière de contraintes, renforcements ou freins pouvant agir sur l'action soumise à évaluation ;
- b) au niveau de C (Construit : objectifs, projets, etc.), nous plaçons les éléments facteurs de changements (les dispositifs et les programmes censés provoquer des changements ou des améliorations par les moyens ou les contraintes qu'ils mettent en place), le fonctionnement curriculaire, c'est-à-dire l'organisation et le suivi des parcours des apprenants, les stratégies des acteurs (formateurs, évaluateurs, apprenants). On appellera *éléments facteurs de changements*, des catégories de stratégies ayant pour fonction d'entraîner tel ou tel

changement dans le fonctionnement d'un programme ou en termes d'acquisition de compétences ;

- c) au niveau de P (productions, décisions), on trouvera les éléments *porteurs de résultats* (impact sur l'environnement, changement de comportements, actions réalisées).

Ces éléments seront soumis à l'observation et c'est sur eux que toute l'information sera prélevée : d'où la nécessité qu'ils soient significatifs de l'objet à évaluer.

4.2.2 Identification des référents

Rechercher des référents (par exemple : le mandat et, par conséquent, les finalités de l'opération ou du dispositif, les données rapportées par la littérature ainsi que des données potentiellement comparatives) constitue la part centrale de tout raisonnement évaluatif. Ces référents peuvent être constitués par un mandat, des finalités et des objectifs, des orientations institutionnelles, des attentes des différents acteurs et partenaires, l'état de situations comparables, etc.). Ils peuvent être explicites mais aussi implicites : dans ce dernier cas, il s'agira de les repérer.

Comme cela a été développé précédemment (page 66), *les référents* de l'évaluation ont comme fonction de conférer du sens, un sens, aux informations relevées. Dewey (1967) avait déjà posé comme incontournable le rôle du recours à des faits extérieurs dans l'attribution d'un jugement : « Le jugement de valeur, comme tout jugement, ne peut être formé qu'en recourant à des faits extérieurs à l'appréciation immédiate, puisqu'on ne peut envisager les conditions et les conséquences d'une attribution de valeur qu'en prenant appui sur de tels faits » (p. 26).

C'est dire l'importance de leur explicitation et de leur prise en compte dans la formulation des conclusions, des jugements et des décisions. Ce travail intervient théoriquement à ce stade du protocole, mais, à la différence des autres catégories du protocole, *les référents* ne sont pas explicitement limités par les catégories précédentes. Un élément à évaluer ne provoque pas, nécessairement, la formulation d'un *référént* unique : celui-ci peut être pertinent pour plusieurs éléments à évaluer. Dans certaines situations d'évaluation, il se peut que les *référents* n'apparaissent pas nettement, ni dans le mandat, ni à l'issue de l'analyse des différentes informations fournies d'entrée de jeu. Cette éventualité peut laisser penser que l'évaluation en question n'est pas encore envisagée sur des bases claires et les évaluateurs sont alors amenés à reformuler ces référents sous contrôle des différents partenaires impliqués. Le tableau qui suit introduit des exemples de référents possibles en conservant, en exergue, la catégorisation des éléments à évaluer.

Tableau 4. Identification des référents

Éléments à évaluer en tant que porteurs potentiels d'influences, d'effets, de résultats	Référents possibles
I 1 : environnement I 2 : structures, organisation I 3 : population I 4 : missions	Existence repérée des déterminismes de la situation concernée (essentiellement socio-économiques) Règlement interne Modes de collaboration avec les populations et les partenaires concernés Textes officiels
C 1 : dispositifs C 2 : fonctionnement curriculaire C 3 : moyens engagés	Objectifs, projets Politique interne de développement de la réussite Stratégies conçues
P 1 : impact sur l'environnement P 2 : changements (de comportements) P 3 : réalisation des missions ; actions P 4 : résultats obtenus par les élèves	Attentes des acteurs Objectifs, projets, résultats espérés, Objectifs, projets, résultats espérés, Moyennes locales, régionales ou nationales

On trouve ici, en regard des *éléments à évaluer*, des *référents* éventuellement utilisables. D'où viennent-ils ? La réponse à cette question ne relève pas de l'évidence : ils ne sont pas des objets « donnés » comme tels et sont toujours à repérer et/ou à construire. Il convient de les chercher dans les documents où ils sont censés se trouver : projets, programmes, textes institutionnels, mais aussi comptes rendus d'entretiens. Une fois repérés, ils demandent à être vérifiés et validés en confrontation avec les différents acteurs impliqués dans l'évaluation (décideurs, professionnels, utilisateurs, partenaires, etc.). En effet, chacun d'entre eux disposant d'une clé ouvrant sur le sens éventuel d'un résultat, il est souhaitable de disposer du plus grand nombre de clés possibles.

L'évaluateur est donc amené à les formuler, aux trois niveaux du modèle ICP :

- au niveau de I (contexte, situations, problèmes), apparaissent les référents socio-économiques (déterminismes locaux en matière d'activités industrielles ou de services), les référents institutionnels (zones prioritaires, zones privilégiées) et les référents d'acteurs (constitués, notamment, par les attentes des parents et des acteurs socio-économiques) ;

- au niveau de C (dispositifs, projets, etc.), les *référents* sont tout naturellement constitués par les objectifs et les projets qui *prédisent* les effets de l'action ainsi que par les choix organisationnels et pédagogiques constituant les stratégies des acteurs. C'est, en général, la posture majeure spontanément adoptée par l'évaluateur, de comparer ce qui sera appelé résultats de l'évaluation aux objectifs des acteurs ;
- au niveau de P (productions, décisions), les *référents* peuvent être répartis en deux catégories essentielles de comparants : encore et toujours les objectifs de l'action, d'une part, et, d'autre part, les données extérieures à celle-ci, utilisées par comparaison afin de situer les résultats dans une échelle locale, régionale, nationale ou internationale.

Ces principaux *référents* une fois posés, il reste toujours à vérifier, lors de l'enquête exploratoire, qu'il n'en existe pas d'autres dans l'ensemble des informations recueillies. Certains référents peuvent même s'inviter, tout au long du déroulement de l'évaluation.

Pour résumer, on peut relever ici, notamment, l'existence de deux catégories de référents :

- les *référents externes* : déterminismes, orientations institutionnelles, confrontation des résultats avec d'autres (moyennes locales, régionales, nationales, internationales) ;
- les *référents internes* (construits par les acteurs) : les objectifs et les projets, les attentes, les stratégies conçues et déployées.

Les *référents* constituent bien le cœur du processus évaluatif : ils permettent, seuls, de justifier le sens qui sera attribué aux résultats de l'évaluation. Si l'on change des *référents* en gardant toutes les autres composantes du protocole, on risque d'obtenir des résultats d'évaluation différents. Ce qui amène à réfléchir sur les cas où un désaccord existe sur les référents entre des porteurs d'enjeux différents (par exemple, entre enseignant et élèves) : la question qui se pose alors est de savoir si les référents des uns sont plus légitimes que les référents des autres. On se trouve, alors, dans une situation à deux issues possibles (la troisième, celle de la prééminence des référents d'un acteur au détriment d'un autre étant exclue, par principe) :

- la première issue consiste à négocier, entre acteurs, la formulation des référents
- la seconde consiste à maintenir la présence de plusieurs référents, même contradictoires et de procéder à des évaluations différenciées en fonction des référents privilégiés.

4.2.3 *Élaboration des critères*

Mais si les *référénts* constituent le pôle de signification d'une procédure évaluative, ce sont les *critères* qui en sont le moteur méthodologique. Élaborer des critères (opération essentielle de toute méthodologie d'évaluation, établissant les étalons de valeur à partir desquels d'éventuels jugements d'évaluation pourront être portés) revient, en effet, à privilégier les aspects de l'objet qui serviront à apprécier et à juger sa valeur.

4.2.4 *Clarification de la notion de critère*

La notion de *critère* demande à être revisitée, ici, avant de l'instrumenter, car elle est parfois confondue avec celle d'*indicateur*. De nombreux évaluateurs emploient, même, l'un et l'autre, indistinctement. Certains commanditaires se satisfont, même, d'une catégorie mixte de « critères/indicateurs », voire d'« indicateurs généraux ».

Dans les années 1980, Bonniol et Genthon (1989) avaient clairement établi qu'un « critère d'évaluation... consiste en une dimension abstraite... » (efficacité, pertinence, efficience, etc.) « utilisée pour évaluer un objet ». Les auteurs ajoutaient que ces critères, si les conditions favorables sont installées, « peuvent devenir critères de réalisation » (d'une tâche relative à l'apprentissage, par exemple). Bonniol et Genthon précisaient enfin que « l'objet » que les critères concernent peut être :

- un *produit* : dans ce cas, le critère (plus formel que fonctionnel) peut être « la précision », « l'exactitude » ;
- une *procédure* : le critère porte alors sur des « démarches », des « techniques », des « façons de faire » (et peut devenir plus fonctionnel que formel) et il peut être formulé en termes de « pertinence des méthodes » ;
- un *processus* : le critère porte alors sur des « intentions », des « manières d'aborder une situation »...

Par ailleurs, Ardoino et Berger (1989) ont également publié un lexique précis des termes les plus employés dans les pratiques d'évaluation, lexique dans lequel le *critère* tient une place centrale. Il y est défini comme un « étalon de valeur fondant le jugement et, éventuellement, la décision » et, enfin, comme « une interface entre les visions du monde, d'une part, et, d'autre part, les indicateurs, plus opératoires ». Des listes de critères généraux, plus ou moins standardisés, existent (par exemple, l'efficacité, l'efficience, la pertinence, la cohérence, la conformité, etc.) et sont utilisables pour évaluer tel ou tel objet, s'ils correspondent à l'intention (ou l'obligation) de l'évaluateur.

Mais d'autres dimensions éclairant des aspects de l'objet à évaluer peuvent aussi être choisies comme critères, si elles se prêtent à une formulation abstraite suffisamment représentative de cet objet, comme la valeur ajoutée, l'impact, l'effet-maître, la progression, etc.

4.2.5 Formulation des critères

Dans ce protocole, on va utiliser la notion de critère pour formuler ce sur quoi les performances de l'objet pourront être appréciées, compte tenu, dans chaque élément prévu comme étant à évaluer, des référents qui auront été repérés et validés dans la phase précédente. L'établissement pris en exemple, ici, objet collectif et dispositif complexe, s'il en est, va ainsi être regardé à travers le prisme des critères, devenus, chacun à son tour, des regards posés sur une zone qu'ils vont spécifiquement éclairer.

Les *critères* peuvent maintenant être formulés, tels que présentés dans le tableau suivant :

Tableau 5. Élaboration des critères

Éléments à évaluer En tant que porteurs potentiels d'influences (I), d'effets (C), de résultats (P)	Référents mobilisables	Critères possibles
I 1 : environnement	– Déterminismes (sociaux, économiques, etc.) pesant sur les situations scolaires	– Pertinence de la prise en compte de l'environnement
I 2 : structures, organisation	– Conceptions institutionnelles (textes réglementaires)	– Conformité aux normes institutionnelles
I 3 : attentes des populations	– Politique interne de prise en compte des populations de partenaires	– Correspondance entre projet et attentes des acteurs
I 4 : missions	– Orientations, projets	– Cohérence entre missions et choix
C 1 : dispositifs	– Objectifs, projets	– Pertinence des objectifs, projets
C 2 : fonctionnement curriculaire	– Politique interne de développement de la réussite	– Qualité de l'accompagnement des parcours
C 3 : moyens engagés	– Stratégies conçues	– Efficience (rapport coût/ investissement)
P 1 : impact sur l'environnement	– Attentes des acteurs	– Prise en compte des orientations du groupe de pilotage

Éléments à évaluer En tant que porteurs potentiels d'influences (I), d'effets (C), de résultats (P)	Référénts mobilisables	Critères possibles
P 2 : changements (de comportements) P 3 : réalisation des missions ; actions P 4 : résultats obtenus par les élèves	– Objectifs, projets – Idem – Résultats espérés, moyennes locales, régionales ou nationales	– Fonctions remplies par le projet ; atteinte des objectifs – Efficacité des actions – Comparaisons diachroniques et synchroniques

Ce moment du protocole d'évaluation (= la formulation des critères) constitue l'étape-charnière au cours de laquelle on prépare le passage d'une logique réflexive (que faut-il évaluer et avec quels référents ?) à une logique instrumentale (avec quels indicateurs et à travers quels items de questionnaire ?). La jonction entre les deux logiques s'opère au moyen du choix des *critères* qui constitue le terme de la réflexion préalable ayant pour fonction de clarifier les buts de l'évaluation, autour desquels il restera, par conséquent, à fonctionner de manière technique (i.e. à construire un recueil de données).

On voit bien, dans la colonne regroupant des critères potentiels d'évaluation, que ceux-ci sont encore exprimés avec des termes abstraits (pertinence, cohérence, efficacité, mise en œuvre, comparaison, etc.), ce qui constitue l'aboutissement de la phase réflexive. Ces critères (qui annoncent les choix que les évaluateurs ont opérés lorsqu'ils ont sélectionné les dimensions de l'objet qu'ils ont considérées comme représentatives) deviennent alors les organisateurs de l'instrumentation qui va suivre.

4.2.6 Construction des indicateurs

Un emploi multiple de la notion d'indicateur. On a beaucoup glosé sur les *indicateurs* d'évaluation, beaucoup usé et abusé de cette notion, en lui attribuant des sens différents. Comme pour la notion de *critère*, celle d'*indicateur* mérite assurément d'être revisitée ici. En effet, de nombreuses publications, parmi lesquelles des rapports officiels d'institutions nationales et internationales, construisent, manient, commentent toutes sortes d'*indicateurs* spécifiquement dédiés aux systèmes éducatifs. On n'en rappellera que deux, pour exemples :

4.2.6.1 Sens n° 1 : une information (généralement statistique) jugée représentative d'un phénomène

- Les *indicateurs* de l'OCDE (2012) qui portent, entre autres, sur le niveau de formation et les taux d'obtention d'un diplôme, l'égalité entre les sexes et l'équité, les retombées de l'éducation sur l'économie, le marché du travail et la société. Par exemple, le programme PISA (2012)¹ qui vise à fournir aux pays membres de cette organisation des informations (obtenues par enquêtes) sur les performances des jeunes de 15 ans en matière de lecture, de mathématiques et de sciences. Il est question, dans les rapports triennaux qui en découlent, d'*indicateurs de performance*, de *contexte*, de *tendances*. La publication de ces *indicateurs* fait connaître les points forts et les points faibles des systèmes éducatifs de chacun des pays enquêtés.
- Les *indicateurs statistiques* du ministère de l'Éducation nationale (2012) qui proposent une analyse synthétique des coûts, des activités et des résultats du système éducatif français, déclinée en 31 indicateurs statistiques. La publication de chiffres, appelés *indicateurs*, se propose d'analyser les évolutions et les comparaisons qui permettent d'actualiser les jugements pouvant être portés sur l'état de l'École en France.

On pourrait rappeler ou citer de nombreux autres documents présentant ou analysant des *indicateurs* de toutes sortes, économiques, sociaux, éducatifs, etc. Ce sont des statistiques, des chiffres qui donnent une indication, une information privilégiée voire représentative d'un phénomène. Mais on l'aura bien compris : ce n'est pas ce type d'*indicateurs* dont il est question ici, même s'ils peuvent nourrir, ici et là, certaines évaluations.

4.2.6.2 Sens n° 2 : les dimensions concrètes d'un critère

Dans le cadre d'une « méthodologie de l'évaluation », on utilisera plutôt la notion d'*indicateur* dans son sens instrumental : celui de dimension concrète d'un critère.

En effet, alors que le critère, comme nous l'avons déjà souligné, est toujours présenté sous forme abstraite, l'*indicateur*, lui, veut et doit être concret. Il s'agit d'un élément observable constituant le signe de la présence d'un phénomène et permettant d'en mesurer le niveau. Dans ce qui nous occupe, il renvoie donc, toujours, même implicitement, à un critère. Ainsi, le *critère de conformité* du fonctionnement d'un dispositif d'aide aux élèves en difficulté pourra être mesuré au moyen de l'*indicateur des écarts* entre une norme donnée (circulaire, plan d'action, etc.) et une réalisation

1. Programme International pour le Suivi des Acquis des élèves

qui en a été inspirée. De la même manière, un *critère d'efficacité* d'un enseignement (renforcé par tel ou tel dispositif) sera mesuré par un *indicateur de taux de réussite* à telle ou telle épreuve ou tel ou tel examen.

Dans le protocole de référentialisation, les indicateurs trouveront naturellement leur place en regard de chaque critère (catégorie abstraite) comme les éléments observables et mesurables (catégories concrètes) nécessaires à l'implémentation² de ce dernier. C'est ce qui apparaît dans le tableau suivant :

Tableau 6. La déclinaison (?) des indicateurs

Éléments à évaluer en tant que porteurs potentiels d'influences, d'effets, de résultats	Référents possibles	Critères	Indicateurs
I 1 : environnement	– Importance comparée des déterminismes	– Pertinence de la prise en compte de l'environnement	– Existence d'une réflexion sur les données – éléments retenus
I 2 : structures, organisation	– Conceptions institutionnelles (textes réglementaires)	– Conformité aux normes institutionnelles	– Confrontation des faits ou des actions avec les normes
I 3 : attentes des populations	– Politique interne de prise en compte des populations de partenaires	– Évolution des populations de partenaires	– Comparaison avant et après action
I 4 : missions	– Orientations, projets	– Cohérence entre missions et choix	– Expression par les acteurs des éléments ayant contribué aux réalisations
C 1 : dispositifs	– Objectifs, projets	– Mise en œuvre des objectifs, projets	– Existence et visibilité de formes du projet

2. Qu'on nous excuse cet anglicisme assumé (pouvant être compris comme « actualisation » ou « instanciation ») qui permet, ici, avec un seul mot, d'indiquer le processus de mise en œuvre technique du développement d'un projet (ce projet étant, en l'occurrence, l'énoncé d'un critère).

C 2 : fonctionnement curriculaire	– Politique interne de développement de la réussite	– Accompagnement des parcours	– Traces des moments-clés du parcours et du devenir des apprenants
C 3 : moyens engagés	– Stratégies mises en oeuvre	– Efficience	– Rapport entre ressources disponibles et ressources mobilisées ; entre investissement et résultats
P 1 : impact sur l'environnement	– Problèmes et besoins	– Réponses aux besoins	– Opinions et représentations des acteurs
P 2 : changements (de comportements)	– Attentes des acteurs	– Prise en compte des orientations du groupe de pilotage	– Idem
P 3 : réalisation des missions ; actions	– Objectifs, projets	– Fonctions remplies par le projet ; atteinte des objectifs	– Comportements observés avant/ après
P 4 : résultats obtenus par les élèves	– Idem – Résultats espérés, moyennes locales, régionales ou nationales	– Efficacité des actions – Comparaisons diachroniques et synchroniques – Compétences acquises – Performances scolaires	– Opinions des acteurs – Quantité, nature et qualité des actions – Appréciations des enseignants – Résultats aux tests

Le cas pris en référence, ici, ne pouvait être détaillé et exhaustivement traité, mais il convient de lire ces indicateurs comme des exemples.

Le travail de l'évaluateur en matière de formulation d'indicateurs s'avère exigeant et précis. Exigeant dans le respect d'une logique rigoureuse : celle de la déclinaison de chaque critère en indicateurs observables et mesurables. Précis, dans la mesure où la formulation de chaque indicateur va, en quelque sorte, formater et découper les informations à obtenir, de telle manière qu'elles ne pourront être utilisées à d'autres fins que celles pour lesquelles l'indicateur les aura prévues. Une fois validée par les acteurs

concernés par l'évaluation envisagée, cette liste d'indicateurs constitue une base formelle sur laquelle peuvent alors se construire différents instruments destinés à les renseigner.

La phase 2, ainsi organisée, a contribué à façonner les différentes facettes de ce qui est devenu une batterie d'indicateurs pouvant être présentée à des acteurs (décideurs, formateurs, enseignants, apprenants, stagiaires, etc.), et dorénavant discutée, modifiée, selon des processus participatifs.

4.3 Phase 3 : construction d'instruments

La construction d'instruments d'évaluation dans le cadre d'une « référentialisation » se développe dans cette logique d'enchaînement en marche depuis le début du protocole.

4.3.1 Des indicateurs à la rédaction des items de questionnement

La troisième phase devient donc essentiellement technique ; elle consiste en la conception et la passation d'outils de questionnement, d'observation et d'analyse de données. Il s'agit, maintenant, de décliner chaque indicateur en autant d'items que nécessaire pour obtenir les informations qui le renseigneront.

Qu'est-ce qu'un item ? C'est une proposition comme « un nombre d'occurrences », ou bien « un taux de satisfaction » constituant une catégorie d'une grille ou d'un questionnement. La formulation de l'item, entre un indicateur qui nomme ce que l'on cherche à savoir et la réponse que l'on attend, permet de préciser et de limiter le type d'information à récolter. L'item peut, enfin, être utilisé sous forme de question.

Tableau 7. La formulation des items

Éléments à évaluer En tant que porteurs potentiels d'influences, d'effets, de résultats	Indicateurs	Exemples d'items de questionnement ou d'observation
I 1 : environnement	– Existence d'une réflexion sur les données et éléments retenus	– Nombre d'occurrences d'unités de discours sur les facteurs environnementaux
I 2 : structures, organisation	– Existence des normes réglementaires dans le projet formalisé – Évolution des attentes	– Mention des normes dans le projet
I 3 : attentes des populations		– Données avant et après action – Taux d'intérêt ressenti

I 4 : missions	– Expression par les acteurs des éléments ayant contribué aux réalisations	– Avis des utilisateurs
C 1 : dispositifs	– Traces d'évolution et adaptation du projet	– Données, avant et après projet
C 2 : fonctionnement curriculaire	– Suivi des parcours et du devenir des apprenants	– Taux de réussite et taux d'insertion professionnelle
C 3 : moyens engagés	– Rapport entre ressources disponibles et ressources mobilisées ; entre investissement et résultats	– Ratios des deux sortes de ressources
P 1 : impact sur l'environnement	– Opinions et représentations des acteurs	– Indice de satisfaction des acteurs – Données comparées entre avant et après – Avis des acteurs
P 2 : changements (de comportements)	– Évolution des comportements observés	– Taux, inventaire, classement – Identification – Classement et pourcentages
P 3 : réalisation des missions ; actions	– Quantité, nature et qualité des actions	
P 4 : résultats obtenus par les apprenants	– Appréciations des enseignants – Résultats aux tests – Appréciations des enseignants – Résultats aux tests	

4.3.2 L'intégration des items dans des instruments appropriés

On commence par répartir les items déclinant les indicateurs dans des matrices d'instruments aptes à les renseigner :

- on peut regrouper, par exemple, les items constituant des questions dans des instruments de type questionnaire ou grille d'entretien de même que des items de type « occurrences » dans des grilles d'analyse de contenu ;
- il sera aisé, en utilisant les indicateurs et les items qui les déclinent, d'élaborer des questionnaires portant sur les attentes et les besoins

d'une population particulière, des grilles d'observation d'activités d'apprentissage ou de phénomènes de groupe, des grilles d'entretien à destination de publics directement impliqués (parents, élèves ou étudiants, décideurs, etc.).

On n'entrera pas, ici, dans le détail de la construction de chaque type d'instrument : chacun d'entre eux pourra suivre les règles établies dans la méthodologie de l'enquête en sciences sociales. Les ouvrages spécialisés sur ce domaine ne manquent pas (Quivy et Campenhoudt, 2006 ; Ghiglione et Matalon, 1978 ; Ghiglione *et al.*, 1980).

Il restera à tester et améliorer les instruments ainsi construits auprès de panels représentatifs de la population à traiter.

4.3.3 Le recueil des informations

Entre la phase 3 et la phase 4 s'interpose le travail de recueil des informations enclenché par la mise en œuvre des instruments. On intégrera leur passation dans le temps de travail ou d'apprentissage, ceci afin d'associer l'évaluation au processus de formation et, par conséquent de l'aborder davantage comme complice de ce dernier.

Là encore, il s'agit d'un travail couramment pratiqué en sciences sociales : entretiens de formes variées, questionnaires divers, analyse de contenus de documents et de discours, observation de situations sous techniques multiples dont les procédures n'ont pas à être exposées ici.

Il restera à remplir une matrice de résultats (tableaux à double entrée comprenant, en ligne, les catégories d'éléments à évaluer et, en colonne, les données recueillies par les indicateurs).

La collecte des informations une fois réalisée, on suppose la matrice exhaustivement remplie. La « valuation », au sens de Dewey, peut alors commencer.

4.4 Phase 4 : Analyse et interprétation

Il s'agit en effet, à ce stade, de « faire parler » les données obtenues, de leur attribuer un sens. Dans le cadre d'une démarche scientifique, on s'attacherait à mettre en perspective les résultats et l'hypothèse formulée au départ. Dans le cadre d'une opération d'évaluation, avec le même esprit, on mettra en relation ces mêmes résultats avec les catégories de la référentialisation précédemment développées. Chaque *élément à évaluer*, regardé au moyen de *critères* construits par l'évaluateur et examiné au travers d'*indicateurs* renseignés par des informations dûment recueillies, aura obtenu des résultats qui seront alors confrontés aux référents prévus. De cette confrontation devrait naître une interprétation attributive de sens.

En matière d'évaluation, on veillera particulièrement à distinguer *analyse* et *interprétation* ou, du moins à fonder des interprétations éventuellement productrices de jugements sur les analyses préalables des résultats. C'est ce qui va être développé.

4.4.1 Analyse des résultats

Comme le résumant Tremblay et Perrier (2006), trois formes d'analyse s'offrent à l'évaluateur :

- a) *l'analyse descriptive* dressera un portrait de la situation observée ou étudiée, suite à une simple compilation des données obtenues. Par exemple, on pourra dresser les caractéristiques d'un parcours, d'un profil, d'un groupe, etc. ;
- b) *l'analyse explicative* s'attachera à argumenter une prise de décision en fonction de la mise à jour de liens causaux entre variables comme, par exemple le lien entre l'existence d'un dispositif et un changement de comportement ou de performance ;
- c) *l'analyse compréhensive*, enfin, s'imposera s'il s'agit d'étudier des rapports entre les résultats obtenus et les perceptions des acteurs. Par exemple, une évaluation pourra prendre en compte l'implication des sujets dans un projet éducatif ou les choix opérés en matière de stratégie curriculaire pour expliquer telle réussite ou tel échec.

Au terme de l'analyse, apparaît alors un nouveau paysage de l'opération évaluative : on ne dispose plus de plusieurs registres possibles de traitement d'informations sur lesquelles il serait encore possible de porter des jugements plus ou moins éclatés, mais d'un système global d'organisation des données mettant en relation les unes avec les autres à travers les catégories de la référentialisation.

4.4.2 Interprétation des résultats

Selon Lahire (2005), cité par Soulet (2011), une enquête en sciences sociales est ponctuée d'actes d'*interprétation*, à tout moment du processus, d'indices, de traces, de sélections, de codage, de corrélations statistiques. Nous rejoignons Soulet (2011), lorsqu'il explique *l'interprétation* comme une mise en relation de faits entre eux : « *L'interprétation*, dans ses multiples registres, de donation de sens, de dévoilement de signification [...], à des fins d'élucidation, de production et de maîtrise du sens, est au cœur de l'activité des sciences sociales. Elle se centre sur la mise en relation de faits, de phénomènes et du sens de cette relation » (p. 2). Dans le domaine de l'évaluation, c'est précisément ce que la référentialisation précédemment décrite permet de faire : mettre en relation, par exemple, des faits de contexte avec des phénomènes de l'ordre de

l'action ou du résultat (relations entre I, C et P, décrites en début de deuxième partie), ce qui infère un sens au compte-rendu des données produites par l'ensemble de l'opération (et, donc, constitue une *interprétation*, au sens de Soulet).

Dans l'exemple suivi tout au long de cette partie, interpréter les résultats de l'évaluation d'un établissement scolaire ou de formation reviendra donc à extraire la signification des écarts, dans le cadre de chaque questionnement problématique : entre les informations obtenues et celles issues des différentes catégories du référentiel précédemment établi ; entre les différentes catégories elles-mêmes (issues de I, C et P).

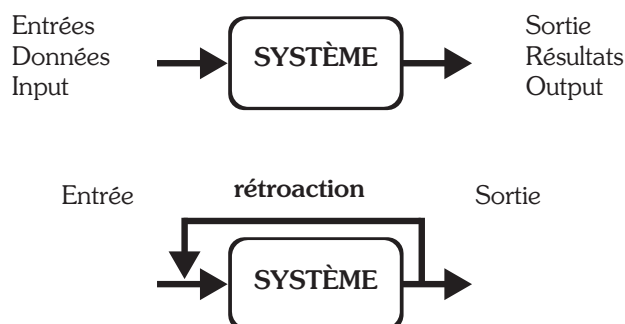
Cette procédure amène à formuler des conclusions sous forme de diagnostics, de prédictions ou de scénarios prospectifs, de réponses à des questions de départ, de préconisations.

4.4.3 *Exploitation et restitution*

Lorsque l'évaluateur a élaboré l'interprétation des résultats de l'opération qu'il a menée, peut-on dire qu'il a terminé l'évaluation ? Poser la question ainsi revient, évidemment, à répondre par une prudente négative. En effet, si l'on se demande où en est le processus lui-même, celui qui consiste en une activité de régulation de l'action éducative, on doit convenir que, si régulation il y a, la boucle de rétroaction qui la caractérise n'est pas aboutie tant que les acteurs de l'évaluation ne sont pas confrontés aux conclusions des différents diagnostics qui les concernent à quelque chef que ce soit (apprenants mais aussi éducateurs, formateurs, enseignants, décideurs).

Lorsque l'on *rend compte* d'une évaluation, les résultats promulgués ne comportent pas seulement une dimension informative générale. Ils présentent au moins, deux autres fonctions utilisables par les évaluateurs :

- a) le bilan d'une action, d'un apprentissage, d'un projet : le rapport d'évaluation (ou présentation synthétique des résultats) constitue la forme traditionnelle de compte-rendu attendue par les institutions qui ont émis le mandat fondateur. Il comprend, souvent, des résultats suivis, généralement accompagnés de comparaisons, d'interprétations, de jugements de valeur et de préconisations ;
- b) la réutilisation corrective : il s'agit de terminer la boucle de rétroaction évoquée plus haut, à la manière d'un système qui utilise son résultat (output ou grandeur de sortie ou encore extrant) comme input (grandeur d'entrée ou intrant), notamment à des fins de correction : autrement dit, le résultat d'une action est à nouveau renvoyé dans le processus qui l'a engendré, soit pour l'accentuer (rétroaction positive) soit pour le freiner (rétroaction négative).

Figure 7. La boucle de rétroaction du modèle systémique

Ce modèle existe en cybernétique, en automatique, en systémique et s'utilise aussi en sciences sociales. Il peut aider, par exemple, à illustrer le fonctionnement d'un système d'apprentissage, alimenté par un *input* (c'est-à-dire un programme, un projet, des objectifs, etc.) favorisant un processus A (c'est-à-dire un apprentissage) qui, à son terme, *output*, peut être réutilisé après l'évaluation (B) pour réalimenter le système (en devenant un nouvel input) qui va le réutiliser comme un renforcement ou un frein. L'intérêt de ce modèle, lorsqu'on le projette dans le processus évaluatif, réside dans l'explication qu'il donne du rôle de l'évaluation dans le fonctionnement d'un système d'enseignement/apprentissage : celle-ci ne s'arrête pas à l'exposé des sorties du système (notes, appréciations, jugements, résultats statistiques, par exemple) mais elle contribue à relancer le processus qu'elle a accompagné, et ici, principalement, l'apprentissage lui-même, en lui indiquant ce qu'il doit améliorer pour le réussir. On voit, alors, comment la *restitution* de l'évaluation peut s'enrichir par l'intégration de cette fonction.

Au moment de terminer le déroulement du protocole de référentialisation, il paraît opportun de signaler la relation de sens existant entre ce dernier et ce qui, dans la littérature anglo-saxonne, est appelé « cadre logique » ou « modèle logique » (McLaughlin et Jordan, 2004) qui a été largement utilisé dans l'évaluation de programme. Ce cadre, très pragmatique, consiste à discerner, lors de l'évaluation d'une situation, les liens entre les *inputs* (=ce que l'on investit), les *outputs* (=ce que l'on fait) et les *outcomes* (= les changements que l'on obtient sur le court, le moyen et le long terme). Ce cadre, appelé aussi « modèle »³, constitue, en fait, une technique de rationalisation appliquée à certaines formes de l'évaluation, privilégiant une approche linéaire et quantitative de la causalité. Dans la modélisation qui est présentée ici, il n'a pas été question de prescription rationalisante adressée aux évaluateurs, mais d'une procédure d'explicitation d'un

3. On se reportera à l'ouvrage de Mottier Lopez et Figari (2012), pour inventorier les différentes façons d'évoquer des « modèles » d'évaluation.

processus d'évaluation. Le « modèle logique » pourrait servir de contrôle ou de contrepoint ainsi que de support à un dispositif méthodologique : il ne remplace pas la référentialisation.

4.4.4 Expliquer les résultats en regard de la référentialisation

Au terme de l'exposé d'une méthodologie de l'évaluation par la référentialisation, il doit maintenant apparaître comme évident que cette dernière, ayant accompagné les différentes phases de la procédure (1 – définition de l'objet à évaluer ; 2 – construction d'instruments ; 3 – interprétation), est devenue organisatrice du raisonnement évaluatif dans son ensemble. C'est-à-dire que, pour rester dans l'exemple de l'évaluation d'un établissement scolaire, les différentes phases suggérées ont suivi une logique de construction de sens du matériau informatif plutôt qu'une succession temporelle et linéaire de procédures. Les évaluateurs concernés, disposant d'un référentiel à la formulation duquel ils ont eux-mêmes participé, peuvent élaborer puis renseigner leurs outils de questionnement et d'observation selon une vision d'ensemble sans cesse disponible avec le seul souci de chercher, dans l'ensemble des catégories de données envisagées, les croisements qui s'avèrent pertinents. Ils peuvent, enfin, justifier tel ou tel résultat présenté par des choix, des événements, des actions pris en compte par le référentiel.

5 EN CONCLUSION : UTILISATIONS D'UNE MÉTHODOLOGIE DE L'ENQUÊTE ÉVALUATIVE

On a proposé, tout au long du chapitre, la mise en place d'un protocole d'*enquête évaluative*. Ce protocole, selon nous, suggère plusieurs utilisations possibles : une vision d'ensemble de ce que mobilise une opération d'évaluation ; un guide méthodologique pour construire des instruments d'évaluation ; une grille d'analyse pour étudier un dispositif éducatif. Ensuite, il ne sera pas inutile de rappeler, à l'issue d'un tel travail, le statut d'une telle méthodologie, entre qualitatif et quantitatif, ancrée dans expérience et l'intérêt que l'on peut trouver à « l'enquête évaluative ».

5.1 Une vision d'ensemble de ce que mobilise une opération d'évaluation

Le problème qui se pose, d'emblée, à tout évaluateur, est toujours celui de savoir comment commencer à élaborer un dispositif d'évaluation adapté à une demande, une commande ou une obligation, à partir d'une situation sociale généralement complexe, comprenant des acteurs (décideurs, évaluateurs, évalués, partenaires, utilisateurs...), des contraintes (en termes d'environnement et de moyens), des problématiques (conflits, difficultés

de communication, échecs déjà enregistrés...) ainsi que des projets de changements (nouvelle politique, nouveaux projets...). La référentialisation suggère de dresser, à partir de telles situations, un tableau synoptique mettant en relation la question d'évaluation posée à un objet avec des hypothèses de réponse qui se construisent à travers la recherche des référents, l'élucidation ou l'élaboration des critères et la construction d'indicateurs. Ensuite, par la modélisation ICP envisagée, la référentialisation invite à une seconde mise en relation entre les données obtenues et un système de significations, pour parler comme Dewey, comprenant, dans sa structure même, une recherche des interactions entre contexte et stratégies, entre stratégies et résultats, entre contexte et résultats (cf. modélisation, p. 73-78).

5.2 Un guide méthodologique pour construire une opération d'évaluation

Certes, nous ne disposons pas, ici, d'un véritable guide pouvant servir de livre de recettes ou de recueil d'instruments préfabriqués applicables à des catégories de situations. Néanmoins, il s'agit bien d'une proposition de marche à suivre pour la conduite d'une opération d'évaluation dans son intégralité.

On a pu, en effet, répondre aux questions comme :

- quelle posture adopter face à un mandat d'évaluation (observateur ou participatif, par exemple) ?
- quels concepts mobiliser pour justifier une vision globale d'évaluation (en privilégiant quelle définition, quel paradigme) ?
- quels travaux préalables effectuer pour disposer d'un diagnostic sur lequel programmer l'ensemble d'un protocole d'évaluation (à partir de l'élucidation d'un mandat) ?
- comment construire une problématique et un questionnement d'évaluation ?
- comment utiliser et/ou construire un référentiel d'évaluation ? Et, dans ce cadre, comment définir des référents, déterminer les éléments à évaluer, établir des critères et construire des indicateurs ?
- quel outil privilégier en fonction du questionnement choisi ?
- comment traiter et interpréter les données ainsi obtenues ?
- comment gérer et réguler les relations entre les acteurs (évaluateurs et évalués, commanditaires et opérateurs, acteurs et partenaires) ?

5.3 Une grille d'analyse pour étudier un dispositif éducatif

Un tel protocole, une fois rédigé, constitue, en effet, un dispositif d'observation d'une situation, un questionnement systématique d'un objet, bref, une grille d'analyse utilisable par le décideur, l'expert, le consultant,

le formateur ou l'enseignant, pour mener une investigation particulière comme l'observation d'un groupe en formation (classe, stage, dispositif), d'un projet éducatif ou comme l'audit d'une institution éducative (par exemple, un établissement d'enseignement ou de formation), une étude sur un programme.

5.4 Méthodologie qualitative avant méthodologie quantitative

On a remarqué que le protocole déroulé tout au long de cette troisième partie n'a pas comporté de volet quantitatif : pas de comptage, pas de statistiques, pas même de chiffres. Il s'agissait donc bien de présenter une méthodologie d'évaluation et non une méthodologie de la mesure. Cette dernière, d'ailleurs, n'existe pas en tant que telle : il faut plutôt chercher des méthodes ou des instruments de mesure pouvant quantifier et comparer des informations obtenues à partir de descriptifs qualitatifs des objets donnés à évaluer. On dispose, pour ce faire, d'une abondante ressource dans la consultation de la revue *Mesure et Évaluation en Éducation*, éditée par l'ADMEE Canada, en association avec l'ADMEE Europe⁴. Mais ce que nous avons développé ici consiste en une modélisation, d'essence qualitative, d'un processus d'évaluation qui ne demande, dans son application sur le terrain, qu'à se doter des caractéristiques quantitatives nécessaires.

5.5 Une enquête évaluative ancrée dans l'expérience

Chaque activité évaluative est une expérience : une expérience en cours, mais aussi une séquence parmi d'autres expériences passées et à venir. Dewey (1967) va même jusqu'à postuler que « l'éducation est une reconstruction ou réorganisation de l'expérience qui augmente la signification de l'expérience même et l'habileté à diriger le cours des expériences successives » (p. 103). *L'expérience évaluative* demande, quant à elle, à être caractérisée :

- évaluer c'est vivre une expérience avec des évalués (si on veut encore recourir à cette dualité *évaluateur/évalué* au lieu de parler d'évaluants pour les deux partenaires). On l'a vu tout au long de la référentialisation, à l'occasion du déroulement de la réflexion spécifique aux différentes phases dans laquelle intervient la recherche commune des référents, l'élaboration partagée du choix des critères, la discussion autour de l'interprétation des résultats et de la validation des produits de l'évaluation par les acteurs concernés ;

4. Association pour le Développement des Méthodologies de l'Évaluation en Éducation (Europe et Canada).

-
- l'enquête évaluative se déroule, comme le préconise Dewey, dans un continuum expérientiel entre un environnement (une situation) et un (ou des) individu(s) ;
 - une expérience évaluative ne suffit pas pour résoudre les problèmes d'une autre expérience : il y a nécessité de revenir, chaque fois à l'enquête. Autrement dit le *savoir évaluer* ne suffit pas et ne remplace pas l'enquête évaluative.

Conclusion : le sens et l'intérêt d'une « enquête évaluative »

La lecture de Dewey nous conforte dans l'idée que l'évaluation par la référentialisation constitue une activité très proche de la conception pragmatique de l'enquête exposée dans la *Logique*. Selon l'auteur, « l'enquête est générale... elle n'est pas réservée au savant ou au logicien, mais elle est utilisée par tout homme affrontant une difficulté... » (Dewey, 1967, p. 33). Appliquée à l'évaluation, l'enquête ainsi abordée consiste à élucider, comme le préconise l'auteur, une problématique annoncée en tant que telle ou formulée par l'évaluateur (par exemple : l'efficacité d'une politique éducative, la résolution d'une difficulté d'apprentissage, la recherche des causes d'une situation d'échec, le dysfonctionnement d'un dispositif éducatif...). On peut encore emprunter à Dewey (1967) une définition des fonctions générales de l'enquête pour l'appliquer à celles d'une *enquête évaluative* : « la situation qui provoque l'enquête » (ici, l'évaluation) « est obscure, troublée, inconnue : il est impossible d'agir à partir d'elle, d'où l'institution de l'enquête » (ici, l'évaluation). « L'enquête éclairant la situation, celle-ci devient connue, vraie... : la vérité apparaît comme la fin (end-in-view) d'une enquête instituée pour élucider une situation indéterminée » (p. 35). C'est en ce sens que la décomposition du mot *enquête* en *en-quête* rejoint la démarche évaluative : l'*enquête évaluative* intègre bien le rôle des valeurs dans l'orientation des activités humaines et, par conséquent, dans les effets produits par ces activités sur le contexte social étudié. L'évaluation guidée par la référentialisation reste *en quête* du sens des actions éducatives.

Partie 3

MÉTHODOLOGIE APPLIQUÉE
À DIFFÉRENTS CHAMPS
DE L'ÉDUCATION ET
DE LA FORMATION

Introduction

On a exploré, dans la première partie, les questions théoriques que pose l'évaluation et éclairé les choix conceptuels préalables qui la gouvernent, afin de pouvoir attribuer un sens clair aux appréciations que la suite des procédures aura construites. On a proposé, dans la deuxième partie, un protocole d'« enquête évaluative » appuyé sur un modèle méthodologique général que nous avons appelé « référentialisation ». Le but de l'ouvrage consistant à établir une démarche complète, son aboutissement logique était donc attendu dans des exemplifications. C'est pourquoi l'application concrète à des situations de terrain, soumises à évaluation fait l'objet de la troisième partie.

1 L'ACTIVITÉ D'ÉVALUATION DANS LES MISES EN SCÈNE DE L'ÉDUCATION, DE LA FORMATION ET DES APPRENTISSAGES

Qu'elle soit intégrée aux processus d'apprentissage dans la classe comme dans la séquence de formation (Laveault, 2012) ou qu'elle soit plus spécifiquement appliquée au *programme*, au sens anglo-saxon (Mesure et Évaluation en Éducation, 2006), l'activité évaluative s'insère toujours dans un dispositif¹ donné, qu'il soit établi par l'évaluateur ou imposé. C'est cette mise en scène de protocoles évaluatifs, quelles que soient les situations auxquelles ces derniers s'appliquent, que nous avons cherché à mettre en valeur dans les exemples qui vont suivre, avec l'hypothèse que chaque exemple auquel on tente d'appliquer une méthodologie générale, constitue, néanmoins un objet unique et complexe. D'où l'intérêt d'adapter toute modélisation aux problématiques spécifiques de cet objet. L'application d'une même méthodologie (avec des formulations de consignes identiques) à des situations et des objets différents devrait favoriser la maîtrise d'une approche plus large (avec des types de réponses diversifiés) de l'enquête évaluative.

1. On distinguera *dispositif éducatif* et *dispositif d'évaluation* de toute situation éducative.

2 LA RÉFÉRENTIALISATION : UNE COMPARAISON ET UNE MISE EN RELATION

L'approche par la référentialisation a mis en valeur deux processus, au moins, qui caractérisent une évaluation et que l'on va retrouver dans certains épisodes des exemples qui vont suivre :

- l'évaluation résulte souvent d'une *comparaison* entre différentes données (qui revêtent des fonctions différenciées, par exemple des données décrivant l'existant face à des données de référence). Autrement dit, le *travail* de l'évaluateur, selon le modèle qui s'esquisse ici, consiste bien à *comparer* des données décrivant des dimensions différentes de l'objet (ou de l'action) à évaluer. Par exemple, on compare un *réfèrent* et un *référé*, un *état initial* et un *état final*, etc. ;
- l'évaluation consiste à étudier les *relations* entre les différentes observations issues de l'activité, c'est-à-dire à les analyser : les interpréter, les comprendre. Or un objet à évaluer comporte, naturellement, des dimensions (des aspects, des facettes, des caractéristiques) innombrables qui ne pourraient jamais être intégralement observées, au risque de perdre le sens. D'où l'intérêt de disposer d'un modèle comme ICP, présenté en seconde partie, qui suggère d'analyser les relations entre contexte (Induit), stratégies (Construit) et productions (Produit).

L'évaluateur construit ainsi un regard sur l'objet (ou l'action) au moyen d'un modèle explicite du jugement qu'il pense devoir porter sur cet objet.

3 APPLICATION À DES PRATIQUES ÉVALUATIVES

Parmi les champs de pratiques de l'évaluation dans le domaine de l'éducation et de la formation, on va proposer d'examiner, ici, l'évaluation des politiques éducatives (chapitre 8), des projets éducatifs (chapitre 9), des dispositifs et des programmes de formation et d'apprentissage (chapitre 10), des outils pédagogiques (chapitre 11), de l'acquisition et de la mise en œuvre des compétences (chapitre 12). Mais on n'entrera pas directement dans des exemples d'activités évaluatives insérées dans des séquences didactiques d'enseignement-apprentissage (comprenant l'étude des phénomènes cognitifs proprement dits) car il conviendrait de mobiliser un autre appareillage plus spécifique, relatif à l'évaluation formative et certificative, notamment. Ce champ d'exercice de l'évaluation en éducation comporte, par ailleurs, une nombreuse littérature. Il n'est besoin que de consulter, parmi d'autres, pour en retrouver le chemin, des auteurs comme : Abernot (1996), Allal (1991), Cardinet (1986), De Ketele (1986), Grégoire (1996), Postic (1977), Scallon (1988), Tousignant (1990), etc.

4 MODE DE PRÉSENTATION DES CAS

Nous proposons cinq exemples de pratiques évaluatives ayant eu recours à tout ou partie du protocole de référentialisation précédemment exposé. Il n'était pas envisageable de les décrire dans leur intégralité : c'est pourquoi nous avons cherché à retrouver, dans chaque opération, les traces de la procédure en question dans un plan et au moyen de tableaux récapitulatifs empruntant les mêmes catégories. Ici, les cinq pratiques présentées comportent une origine réelle mais ont été « formatées » pour constituer des cas d'école.

Chaque exemplification est traitée de la manière suivante : (1) une introduction, (2) l'exemple présenté de deux manières différentes : texte et tableau.

(1) Une introduction

Elle définit la situation particulière examinée en la rangeant dans une classe d'objets à évaluer et inclut des rubriques descriptives :

- *le concept sous-jacent* à cet objet est explicité afin d'établir un lien entre dimensions théoriques et méthodologiques de l'élaboration qui a été opérée ;
- *les différents cas de figure* que peut prendre l'évaluation de cet objet sont cités, de manière à relativiser l'application choisie par rapport à d'autres cas du même type ;
- il s'agit alors de repérer *les fonctions* apparemment remplies par l'opération afin de donner à voir l'influence des buts réels de l'évaluation sur le choix des protocoles et des outils.

(2) *L'exemple* lui-même est présenté de deux façons successives, à chaque phase de la référentialisation :

- *l'analyse du cas*, sous forme d'exposé des modalités utilisées, suivies d'exemples d'informations extraites de chaque opération
- *la mise en tableau* reproduisant les mêmes contenus, sous forme synoptique et destinée à mettre en perspective l'organisation de la procédure adoptée, avec la mise en relation des correspondances entre les différentes catégories du protocole et qui prend la forme déjà proposée page 80 :

Tableau 8. Rappel du protocole de référentialisation

Phase 1 Exploration	Analyse du mandat	– Quel contexte ? – Quelle stratégie ? – Quelles productions ?
	Problématique d'évaluation	À quelle(s) question(s) fondamentale(s) l'évaluation doit-elle répondre ?
Phase 2 Procédure centrale de référentialisation	Éléments à évaluer	Quelles dimensions de l'objet à observer ?
	Référents	À partir de quoi attribuer un sens à l'information ?
	Critères	Quels aspects représentatifs de l'objet à évaluer choisir pour porter un jugement ?
	Indicateurs	Quelles caractéristiques observables permettent la vérification du degré d'atteinte du critère ?
Phase 3	Instrumentation	Quels outils employer pour renseigner les indicateurs ?
Phase 4	Interprétation	Quelles significations attribuer aux résultats ?

Ce protocole (proposé et explicité dans le chapitre précédent) va servir de grille de lecture des différents cas choisis comme exemples d'application dans les domaines d'éducation et de formation qui vont être présentés dans ce chapitre. Cependant, pour éviter le maniement d'un tableau synoptique trop lourd, il a été procédé à son découpage en sous-tableaux traitant, chaque fois, une phase du protocole.

Enfin, on remarquera que la construction des « items » n'est pas reprise, en phase 3, dans les applications décrites dans cette partie, leur construction n'étant pas assez clairement explicitée dans les exemples d'origine. Enfin, chaque phase observée se terminera par une synthèse des apports qu'elle aura contribué à fournir à la démarche d'évaluation.

Chapitre 8

Première application : Évaluation de politiques éducatives

Un exemple : Évaluation d'une politique territoriale d'éducation

INTRODUCTION : LES POLITIQUES ÉDUCATIVES COMME OBJETS À ÉVALUER

Les politiques éducatives sont d'abord des politiques publiques et elles nous intéressent ici, plus précisément, dans leur dimension locale. L'évaluation occupant désormais une place importante dans la gestion des politiques publiques, celle des politiques éducatives n'en est devenue que plus présente dans la gestion des organisations en général et des systèmes scolaires, en particulier. Évaluer une politique publique (éducative) consistera donc à en mesurer les effets en termes d'atteinte d'objectifs, d'efficacité, d'efficience, de pertinence et de cohérence.

Mais la pratique n'est pas si simple. Solaux (2000) rappelle la complexité de toute évaluation de politique éducative, située au carrefour de plusieurs disciplines de référence, parmi lesquelles les sciences politiques (Faure *et al.*, 1995) et les sciences de l'éducation (Tourmen, 2007). Quant à Henriot (2009), citant des enquêtes sur l'utilisation de ces évaluations, elle constate même que « celles-ci ont peu d'effets en termes de décisions » (p. 39), pour

des raisons diverses (soit elles sont mal comprises ou mal reçues, soit elles entraînent des coûts non supportables, soit elles sont contestées par une partie des acteurs). Le développement de méthodologies adaptées, fondées et participatives pourrait, alors, s'avérer d'autant plus pertinent.

Concept de politique publique éducative

Politique publique

Une politique publique est souvent définie, dans les documents nationaux et internationaux, comme un système d'action collective planifié (un programme d'action) visant à provoquer un changement dans les conduites d'acteurs, individuelles ou organisées, dans le cadre d'un système prescriptif établi par une institution reconnue. Il s'agit de ce que Henriot (2009) qualifie d'*action publique*, nationale ou locale et de ce que Muller (1990) définit, par ailleurs, comme « un mécanisme de fabrication... d'idées, de valeurs qui vont constituer une vision du monde » (p. 157).

Éducative

Appliquée à l'éducation, cette définition concerne, à la fois, des programmes nationaux ou locaux, constituant des dispositifs souvent périscolaires, destinés à accompagner l'enseignement scolaire ainsi qu'à conforter l'éducation familiale, sociale et culturelle.

Il faut rappeler que la terminologie « évaluation de politiques publiques » reflète les discours ayant cours particulièrement en France alors que partout ailleurs, c'est l'expression anglo-saxonne de *program evaluation* qui désigne la même pratique, la notion de *programme* étant conçue comme un ensemble de finalités et d'objectifs organisé dans le cadre d'un dispositif prévoyant une planification, le recours à des méthodes adaptées et à des protocoles d'évaluation intégrés.

Différents cas de figure

En termes de domaines d'application, on trouve, par exemple :

- l'évaluation de l'aménagement des rythmes scolaires
- l'évaluation des zones d'éducation prioritaire ou équivalents
- l'évaluation de la lutte contre l'échec scolaire
- l'évaluation d'établissements, d'innovations éducatives...

Situations d'évaluation de politiques éducatives

L'évaluation des politiques éducatives peut répondre à des préoccupations diverses et nombreuses :

- une régulation des systèmes éducatifs comme le remarque Demailly (2001) ;
- une mise en lumière des pratiques de pilotage et de formation des organisations éducatives comme l'explique Verdière (2001) ;
- une mobilisation autour de projets de changement comme le signale Petitpierre (2001).

Fonctions attribuables à l'évaluation des politiques publiques

Sous une forme typologique le « petit guide de l'évaluation des politiques publiques » publié par le Conseil Scientifique de l'Évaluation, en France (1996), propose de regrouper ainsi les fonctions de l'évaluation des politiques publiques :

- une fonction déontologique consistant à rendre compte aux responsables politiques et aux citoyens de la manière dont une politique a été mise en œuvre et des résultats obtenus... (*accountability*) ;
- une fonction gestionnaire consistant à répartir plus rationnellement les ressources humaines et financières entre différentes actions, améliorer la gestion des services chargés de les mettre en œuvre ;
- une fonction décisionnelle consistant à préparer des décisions concernant la poursuite, l'arrêt ou la refonte d'une politique ;
- une fonction d'apprentissage et de mobilisation voulant contribuer à la formation et à la motivation des agents publics et de leurs partenaires en les aidant à comprendre les processus auxquels ils participent et à s'en approprier les objectifs.

On ne peut être sûr que ces fonctions se réalisent dans le cas qui va suivre, mais on peut toujours invoquer l'une ou l'autre d'entre elles comme référent explicatif de certains résultats de l'évaluation réalisée.

MÉTHODOLOGIE REPÉRÉE

Du protocole méthodologique utilisé dans cet exemple, on retiendra les éléments suivants :

1 PHASE 1 : EXPLORATION (MANDAT ET PROBLÉMATIQUE)

On a vérifié qu'elle s'est déroulée de la manière suivante, consultable sous deux versions successives (par une analyse cursive du cas puis par une version synoptique sous forme de tableau).

1.1 Analyse du cas

1.1.1 Le mandat

La collectivité commanditaire (une communauté d'agglomération), inscrite dans le dispositif « Politique de la ville » (Conseil scientifique de l'évaluation, 1996), a souhaité établir un état des lieux du fonctionnement et des effets du volet éducatif de ce dispositif et procéder à une évaluation globale de sa propre politique éducative. L'analyse du mandat, des attentes et intentions des prescripteurs fait ressortir la volonté d'identifier les différentes fonctions attribuées aux dispositifs éducatifs et institutionnels ainsi que de construire une évaluation participative qui permette de passer d'une politique d'accumulation de dispositifs à une politique globale.

Par rapport aux effets éventuels de contexte, les mêmes prescripteurs ont souhaité identifier les spécificités démographiques, socio-économiques et scolaires caractérisant les publics touchés par les dispositifs, les activités développées dans le cadre des dispositifs ainsi que les pratiques et les attentes des jeunes dans leurs temps libres.

En ce qui concerne les stratégies, ils ont décelé des problématiques à résoudre relatives :

- aux méthodes à adopter pour formuler un projet
- aux axes du projet à choisir prioritairement pour sa formulation

En matière de productions visées, ils attendaient des preuves de :

- l'impact des dispositifs sur les publics
- la faisabilité du passage entre des objectifs de dispositifs institutionnels (épars) et la formulation d'un projet global

1.1.2 La problématique

Cette analyse du mandat, en mettant en lumière les attentes spécifiques vis-à-vis de l'évaluation, fait ressortir une question essentielle : en quoi les dispositifs éducatifs, constitutifs de la politique de la ville, deviennent-ils les éléments de ce que cette politique nomme « projet éducatif local » ?

1.2 La mise en tableau

L'analyse précédente peut gagner à être présentée sous forme synoptique, dans un tableau récapitulatif de la procédure ayant guidé la conduite de la phase 1 :

**Tableau 9. Évaluation d'une politique territoriale d'éducation
Phase 1**

Exploration (mandat et problématique) (exemples)	
Analyse du mandat	Problématique d'évaluation
– Par rapport au contexte <i>Réaliser un état des lieux des dispositifs éducatifs développés sur les quartiers prioritaires</i> – Par rapport aux stratégies <i>Formaliser un projet éducatif intégrant des dispositifs éducatifs épars</i> – Par rapport aux productions <i>Étudier l'impact d'un échantillon d'actions éducatives conduites sur les quartiers prioritaires</i>	À quelle question fondamentale l'évaluation doit-elle répondre ? <i>En quoi les dispositifs éducatifs, constitutifs de la politique de la ville, participent-ils de l'élaboration d'une politique éducative territoriale ?</i>

1.3 Les apports de cette phase à la conduite de l'évaluation

On a pu apercevoir la façon avec laquelle les évaluateurs en charge de cette opération ont débuté leur travail : ils ont examiné les actes de commande comme l'appel d'offres, le cahier des charges ainsi que les documents pouvant décrire l'ensemble des enjeux de l'évaluation attendue. À partir d'une compréhension de ces enjeux, ils ont reformulé ces attentes par une « problématique » d'évaluation qui a ainsi transformé le pouvoir du commanditaire en nouveau pouvoir, celui de l'évaluateur, consistant à traduire en catégories méthodologiques neutres ce qui représentait, au départ, une volonté ou une attente d'acteurs impliqués dans une situation particulière. Cette problématique a, apparemment, constitué la base de questionnement qui a gouverné le fonctionnement de la phase suivante.

2 PHASE 2 : PROCÉDURE CENTRALE DE RÉFÉRENTIALISATION

Dans les différentes actions qui constituent cette phase, on pourra relever, à partir des dispositifs éducatifs annoncés dans cet exemple, des éléments définis comme étant à évaluer ainsi que des exemples de choix de référents, critères et indicateurs adaptés.

2.1 Analyse du cas

2.1.1 Définition des éléments à évaluer

Il s'agissait bien de repérer les *facettes de l'objet à observer* qui devraient constituer l'ossature de l'ensemble de la référentialisation. Parmi ces éléments, on trouve, par exemple :

En termes d'effets éventuels de contexte : les caractéristiques de la population de 0-16 ans sur le territoire.

En ce qui concerne les stratégies : les objectifs implicites du projet d'éducation ainsi que le recours aux programmes et dispositifs contractuels.

En matière de productions espérées : la participation des acteurs éducatifs.

2.1.2 Choix des référents

Pour attribuer un sens aux informations, les évaluateurs du projet éducatif ont été amenés à privilégier certains référents :

Pour évaluer les effets éventuels de contexte, les données fournies par l'Insee², par l'organisme qui gère les allocations familiales et par l'inspection académique ont constitué des référents systématiquement consultés.

Pour évaluer les effets stratégiques, le concept de projet tel qu'il a été théorisé dans la littérature (Boutinet, 1990) ainsi que les documents d'orientation locaux, régionaux, nationaux sur les « projets éducatifs locaux » ont été utilisés comme référents.

Pour évaluer les effets attendus en matière de participation des acteurs, les évaluateurs se sont référés à la notion de « participation symbolique à la citoyenneté » dont s'inspire Hart (1992) pour proposer une échelle de mesure de ce phénomène.

2.1.3 Élaboration des critères

Pour porter un jugement, les évaluateurs ont choisi comme aspects représentatifs du projet éducatif les critères suivants :

- *Au niveau des effets de contexte* : le critère d'évolution sociodémographique de la population des individus de zéro à seize ans.
- *Au niveau des effets stratégiques* : le critère de pertinence des objectifs du projet.
- *Au niveau des productions attendues* : la cohérence des dispositifs.

2. INSEE : Institut national de la statistique et des études économiques (France)

2.1.4 Construction des indicateurs

Les critères d'évaluation du projet éducatif ont été vérifiés au moyen d'indicateurs censés décrire leurs caractéristiques observables. On relève, parmi d'autres, des indicateurs comme :

- *En matière d'effets de contexte* : la progression et la répartition de l'effectif des enfants et des jeunes, dans la population globale du territoire, depuis 5 ans.
- *En matière d'effets stratégiques* : l'existence (et la nature) d'objectifs éducatifs opérationnels et d'actions prioritaires ; le type de perception et les modes d'utilisation des différents dispositifs par les opérateurs (comme les associations).
- *En matière de productions espérées* : les modalités de participation des parents proposées par les structures dans les activités elles-mêmes ou dans des espaces de concertation spécifiques.

2.2 La mise en tableau

Comme pour la phase 1, on présente ici, sous forme synoptique, la procédure ayant guidé la conduite de la phase 2 :

**Tableau 10. Évaluation d'une politique territoriale d'éducation
Phase 2**

Procédure centrale de la référentialisation (exemples)			
Éléments à évaluer	Référents	Critères	Indicateurs
Quelles dimensions de l'objet à observer ?	À partir de quoi attribuer un sens à l'information ?	Quels aspects représentatifs de l'objet à évaluer choisir pour porter un jugement ?	Quelles caractéristiques observables permettent la vérification du degré d'atteinte du critère ?
<i>Les caractéristiques de la population 0-16 ans sur le territoire</i>	<i>Données Insee Données Caisse d'allocations familiales Données Inspection académique</i>	<i>Évolution socio- démographique de la population 0-16 ans</i>	<i>Effectif des enfants et des jeunes sur le territoire et sa progression depuis 5 ans Répartition de ces effectifs dans la population globale et sa progression depuis 5 ans</i>

Procédure centrale de la référentialisation (exemples)			
Éléments à évaluer	Référents	Critères	Indicateurs
<i>Les objectifs implicites du projet d'éducation</i>	<i>Concept de projet</i>	<i>Pertinence des objectifs du projet</i>	<i>Existence et nature des objectifs éducatifs, stratégiques, opérationnels et d'actions prioritaires Existence d'initiatives de formation/réflexion des acteurs aux connaissances sur l'enfance et la jeunesse et de leurs besoins, dans l'élaboration des projets</i>
<i>Les programmes et dispositifs contractuels</i>	<i>Données nationales Agence Nationale pour la cohésion sociale et l'égalité des chances Discours des acteurs associatifs</i>	<i>Cohérence des dispositifs</i>	<i>Spécificités des activités conduites grâce aux différents dispositifs (publics cibles, type d'activité, modes d'intervention individuels/ collectif...) Types de perception et modes d'utilisation des différents dispositifs par les opérateurs (associations notamment) Répartition par tranches d'âges des bénéficiaires du programme de réussite éducative sur la ville par rapport à la France</i>

<i>La participation des acteurs éducatifs</i>	<i>Projets des opérateurs</i>	<i>Pertinence de la prise en compte des besoins des publics</i>	<i>Modalités de participation des parents proposées par les structures dans les activités elles-mêmes ou dans des espaces de concertation spécifiques</i>
---	-------------------------------	---	---

2.3 Les apports des procédures de la phase 2 à la conduite de l'évaluation

Cette phase du protocole a permis aux évaluateurs de passer d'une posture passive d'observateurs invités à répondre à des questions à une posture active responsable d'une opération consistant à produire une batterie d'indicateurs représentatifs des problèmes qu'une évaluation est chargée d'élucider. Les indicateurs, construits dans une logique déductive (éléments à évaluer, référents, critères, indicateurs) ont pu alors constituer la structure de questionnement des différents instruments qui restaient à construire.

3 PHASE 3 : INSTRUMENTATION

Quels outils les évaluateurs de la politique éducative pris en exemple ont-ils employés pour renseigner les indicateurs précédemment construits ? En fait, il s'agit d'une succession de procédures développant, chacune, selon les besoins, une interrogation, une observation, un questionnement.

3.1 En guise d'analyse du cas

Dans cette opération d'évaluation, on repère trois types de procédures qui se sont succédé en matière d'instrumentation :

- des choix préalables ont été effectués (lignes 1,2,3) :
 1. Deux sortes de documents-cadres ont été privilégiées : d'une part, des prescriptions de dispositifs et des bilans d'actions antérieures, d'autre part, des analyses statistiques produites par des institutions publiques.
 2. Deux catégories de populations ont été ciblées : les élus et les acteurs éducatifs concernés.
 3. Toutes les activités se déroulant dans les dispositifs à évaluer ont été interrogées.

- *des instruments spécifiques ont été construits et exploités (lignes 4, 5) :*
- 4. Il a été procédé à la construction de grilles d'analyse de documents et de grilles d'entretiens individuels et collectifs, d'enquêtes, de focus groupe, auprès des élus, des professionnels et des jeunes.
- 5. Les informations recueillies avec les instruments ont été confrontées entre elles.
- *des actions d'exploitation des résultats ont été menées (ligne 6) :*
- 6. Les résultats ont été exploités par des présentations devant un comité de pilotage chargé d'intégrer les résultats de l'évaluation dans la conduite du projet global. Ils ont été mis en valeur par l'élaboration d'un rapport validé par le groupe de pilotage et diffusé aux différents partenaires engagés dans la politique de la ville.

3.2 Mise en tableau

**Tableau 11. Évaluation d'une politique territoriale d'éducation
Phase 3**

Instrumentation (exemples)	
Quels outils employer pour renseigner les indicateurs ?	
1 – Quels documents descriptifs et prescriptifs choisis ?	– <i>Documents juridiques sur les dispositifs, sur les bilans des actions</i> – <i>Données statistiques produites par des institutions (Caisse d'allocations familiales, Inspection académique, Insee, données locales...)</i>
2 – Quelles populations interrogées ?	<i>Élus et acteurs éducatifs des villes de la communauté d'agglomération</i>
3 – Comment ont été sélectionnés les activités et les acteurs concernés ?	<i>Toutes les activités des dispositifs ont été identifiées par la réalisation d'un tableau de bord</i>
4 – Quels instruments construits ?	– <i>Grilles d'analyse de documents</i> – <i>Grilles d'entretiens individuels et collectifs, questionnaires d'enquêtes, focus groupe auprès des élus, des professionnels et des jeunes</i>
5 – Comment ont été traitées les informations ?	<i>Analyse des documents produits, des enquêtes et des réponses fournies par les acteurs interrogés</i>

6 – Comment les résultats ont-ils été exploités et mis en valeur ?	– <i>Par des présentations devant un comité de pilotage chargé d'évaluer les résultats et les conséquences sur le projet global</i> – <i>Par l'élaboration d'un rapport validé par le groupe de pilotage et diffusé aux différents partenaires engagés dans la politique de la ville</i>
--	---

3.3 Les apports de cette phase au processus d'évaluation

Il est clair que l'évaluation est souvent perçue comme l'observation instrumentée d'une situation donnée. Dans cette veine, il est souvent admis que l'instrumentation constitue elle-même, en soi, l'évaluation et qu'il suffit d'en utiliser les résultats. Mais nous avons tenté de montrer, dans les deux parties précédentes, que les résultats étaient, en fait, très dépendants du choix des éléments sélectionnés, des référents privilégiés, des critères choisis et des indicateurs qui en découlaient : il en résulte un changement majeur de statut de l'instrumentation en tant que telle. Il convient donc, selon nous, de vérifier si l'instrumentation est pertinente au regard de ce qui vient d'être rappelé et si elle est méthodologiquement valide. L'examen du cas en cours de présentation ne laisse pas apparaître de problème particulier à ce stade.

4 PHASE 4 : INTERPRÉTATION

Les données ayant été recueillies selon le protocole précédent, les évaluateurs ont disposé, alors, de résultats classés selon les indicateurs construits par la référentialisation entreprise. L'interprétation consistait à attribuer du sens à ces résultats.

Pour en rendre compte, on intégrera l'interprétation directement dans la mise en tableau, de manière à synthétiser, dans un seul support, l'opération consistant à élaborer cette interprétation évaluative en cohérence avec les indicateurs ayant provoqué les informations à analyser. La colonne *instrumentation* n'y est conservée que pour mémoire.

**Tableau 12. Évaluation d'une politique territoriale d'éducation
Phase 4**

Interprétation (exemples)		
Indicateurs	Instrumentation	Interprétation
Quelles caractéristiques observables permettent la vérification du degré d'atteinte du critère ?	Quels outils employer pour renseigner les indicateurs ?	Quelles significations attribuer aux résultats ?
Éléments contextuels : <i>Effectif des enfants et des jeunes sur le territoire et sa progression depuis 5 ans</i>	<i>Analyse de tableaux de bord</i>	Exemple : <i>Les enfants de 0 à 5 ans, au nombre de 1 064 représentent une proportion supérieure (24 %) à celle des autres communes et de l'agglomération (22 %).</i> <i>Les 6-10 ans, 20 %, sont aussi un peu plus représentés que sur l'Agglomération (19 %).</i>
Éléments stratégiques : <i>Existence et nature des objectifs éducatifs, stratégiques, opérationnels et d'actions prioritaires</i>	<i>Analyse documentaire</i> <i>Entretiens individuels</i> <i>Entretiens collectifs</i>	<i>Initiée dès la fin des années 90, formalisée au début des années 2000, la politique éducative de la ville concernée est restée sur les mêmes orientations depuis lors. Elle est donc considérée malgré plusieurs caractéristiques et orientations de travail encore pertinentes, comme « démodée ».</i> <i>Cette politique a toutefois le mérite d'exister et de poser une base de travail pour son renouvellement. Cette politique est constituée de 6 orientations.</i>

Productions : <i>Type de perception et modes d'utilisation des différents dispositifs par les opérateurs (associations notamment)</i>	<i>Analyse documentaire</i> <i>Entretiens individuels</i>	<i>Le dispositif a ses limites (travail avec les « grands » absenteïstes), mais des résultats pédagogiques ont été pointés sur l'amélioration de l'organisation de l'élève, son comportement en classe, ainsi que sur l'amélioration de ses résultats scolaires.</i>
--	--	--

4.1 Les apports de cette phase au processus d'évaluation

L'interprétation reste l'activité la plus polémique, la plus incertaine et la plus lourde de conséquences de l'ensemble du processus d'évaluation. Le reconnaître ne suffit pas pour se dispenser d'en assurer l'objectivation la plus importante possible. La méthode utilisée ici a consisté à confronter les résultats (obtenus par l'administration des instruments d'interrogation, d'analyse et d'observation) avec les référents attachés aux éléments à évaluer définis dans les phases précédentes. Ici la confrontation a amené une interprétation mitigée en termes d'atteintes d'objectifs et de définition de politique locale d'éducation.

Chapitre 9

Deuxième application : Évaluation de projets éducatifs

*Un exemple :
Évaluation d'un projet éducatif local*

INTRODUCTION : LES PROJETS ÉDUCATIFS COMME OBJETS À ÉVALUER

Les projets éducatifs ne sont pas des dispositifs comme les autres : ce ne sont ni des programmes, ni des actions de formation mais des processus particuliers qu'il convient de définir avec précision avant d'en entreprendre l'évaluation. S'ils s'inscrivent souvent dans des politiques d'ensemble, comme dans le cas précédent, ils n'en demeurent pas moins un mode très particulier d'organisation de l'éducation et de la formation.

Concept de projet

Derrière cette notion, « figure emblématique de notre modernité », selon Boutinet (1990), le concept nomade et polysémique de *projet* décrit bien une démarche d'ensemble empreinte, à la fois, d'anticipation et de volontarisme. C'est en même temps un concept *empirique* (désignant la gestion des adaptations quotidiennes), un concept *théorique* (en sciences humaines, le projet définit des processus de comportements individuels et collectifs

anticipant des situations futures), et un concept *opératoire* (permettant, dans des conditions précises, de guider l'action). L'évaluation d'un tel processus ne pourra donc pas consister, uniquement, à comparer un état A (de début de l'action projetée) à un état B (de fin d'action) mais aura également à prendre en compte les phénomènes de participation, de communication, d'implication, etc.

Projet éducatif

Les actions éducatives appelées « projets » sont toujours inscrites dans des orientations qui sont explicitement ou implicitement préétablies. L'évaluateur est donc amené à repérer ces dernières avant d'aborder l'élaboration d'une quelconque méthodologie. Les approches de cette activité peuvent varier : celle, *philosophique* qui fait appel à des valeurs invoquées comme justifications des objectifs privilégiés ; celle, *psychologique* qui considère le projet comme une réponse à l'incertitude de l'avenir, justifiant des nécessités de changement et d'anticipation (Boutinet, 2004) ; celle, *sociologique*, influencée soit par le modèle déterministe représenté par Bourdieu (1980) présentant le projet comme *socialement conditionné*, soit par le modèle de l'action représenté par Boudon (1979), concevant le projet comme une *agrégation de conduites individuelles* motivées par un espoir de gain des acteurs ; celle, *gestionnaire*, enfin, qui conçoit la conduite de projet comme « l'art de diriger et de coordonner des ressources humaines et matérielles pour atteindre des objectifs fixés, dans des limites de temps, de budget et de satisfaction des partenaires » (Guide to the Project Management Body of Knowledge, 2000). Dans le fonctionnement des organisations, la plupart des projets recoupent plusieurs de ces approches et demandent, par conséquent, d'être évalués dans une perspective pluraliste.

Différents cas de figure

Depuis les années 1980, la notion de projet a envahi le monde de l'éducation et de l'enseignement. Les dispositifs relevant du projet constituent, aujourd'hui, un mode courant d'organisation de l'activité humaine, en général, et de l'action éducative en particulier.

On trouve des projets d'établissement, des projets pédagogiques de classes, de disciplines, des projets éducatifs d'associations, de villes (Projets éducatifs locaux¹), de régions...

1. Projet éducatif local (PEL) : en France, ce sigle désigne un projet d'éducation élaboré à l'échelle d'un territoire

Les situations d'évaluation de projet

Lorsqu'une action appelée *projet* est donnée à évaluer, trois doutes apparaissent d'emblée et demandent à être résolus avant la conception d'un quelconque protocole :

- toute action éducative s'inscrit dans un projet : cette évidence ne peut pas être occultée. Celui-ci peut être formulé mais il peut aussi rester implicite et connu seulement de quelques initiés ou d'observateurs attentifs. La question peut alors se poser de déterminer ce qui relève de la programmation institutionnelle d'une action et ce qui relève de la mise en œuvre d'un projet d'acteurs locaux : chacune de ces deux situations demandera le déploiement de procédures d'évaluation particulières, soit isolément, soit en interaction ;
- tout projet écrit ne recouvre pas automatiquement un processus de projet : il peut être l'œuvre d'un décideur ou d'un groupe de pression et ne pas représenter une communauté tout entière. L'inverse apparaît aussi fréquemment : un collectif de formateurs, de pédagogues, d'enseignants peut organiser une activité éducative autour d'un idéal commun sans que celle-ci soit officialisée dans le cadre d'un dispositif connu et appelé *projet* ;
- le projet peut être analysé comme objet à évaluer à partir du moment où il correspond à une configuration précise qui en justifie l'appellation, soit au nom d'une approche classique par objectifs (Hameline, 1979), soit en référence à une conception anthropologique de l'anticipation d'une action humaine, comme le propose Boutinet (1990).

MÉTHODOLOGIE REPÉRÉE

Du protocole méthodologique utilisé dans cet exemple, on retiendra les éléments suivants.

1 PHASE 1 : EXPLORATION (MANDAT ET PROBLÉMATIQUE)

On a vérifié qu'elle s'est déroulée de la manière suivante, consultable sous deux versions successives (par une analyse cursive du cas puis par une version synoptique sous forme de tableau).

1.1 Analyse du cas

1.1.1 Le mandat

Une collectivité locale de 7 000 habitants souhaite évaluer chemin faisant les objectifs, le pilotage et les résultats du Projet Éducatif Local installé sur son territoire.

L'exploration du mandat, des attentes et intentions des prescripteurs a permis aux évaluateurs de conserver, dans la volonté exprimée par les prescripteurs, les éléments suivants :

- évaluer la situation intermédiaire de l'avancée des actions éducatives produites dans le cadre d'un Projet Educatif Local
- soumettre ces avancées à un regard externe dans le cadre d'une évaluation participative.

En matière de prise en compte du contexte, les évaluateurs ont souhaité :

- apprécier des conditions du développement de la politique éducative locale
- utiliser la procédure comme un moyen de « relance » et de remobilisation des acteurs

En ce qui concerne les stratégies, ils ont décelé des problématiques à résoudre relatives :

- à l'animation et au développement du projet (pilotage)
- à l'existence de plusieurs types de représentations et d'appropriations du projet chez les acteurs éducatifs

Pour évaluer les productions, ils ont cherché des preuves sur :

- les effets du projet auprès des publics visés (enfance, jeunesse, familles)

1.1.2 La problématique

Cette brève analyse permet déjà de dégager la problématique de l'évaluation entreprise : il s'agit, en effet, de découvrir en quoi le dispositif « projet éducatif local » a entraîné des dynamiques éducatives en direction des publics visés.

1.2 Mise en tableau

**Tableau 13. Évaluation d'un projet éducatif local
Phase 1**

Exploration (mandat et problématique) (exemples)	
Analyse du mandat	Problématique d'évaluation
– Par rapport au contexte <i>apprécier les conditions du développement de la politique éducative locale</i> – Par rapport aux stratégies <i>analyser l'appropriation des objectifs par les « porteurs du projet » et les associations</i>	– <i>À quelle(s) question(s) fondamentale(s) l'évaluation doit-elle répondre ?</i>

– Par rapport aux productions <i>étudier les effets du projet auprès des publics visés (enfance, jeunesse, familles)</i>	<i>En quoi le projet éducatif local a-t-il :</i> – <i>fait progresser l'appropriation des activités éducatives par les acteurs locaux</i> – <i>développé de nouvelles activités</i> – <i>contribué à fédérer des dynamiques éducatives</i>
--	---

2 PHASE 2 : PROCÉDURE CENTRALE DE RÉFÉRENTIALISATION

Dans les différentes actions qui constituent cette phase, on pourra relever, en examinant les caractéristiques du projet figurant dans cet exemple, des éléments définis comme étant à évaluer ainsi que des exemples de choix de référents, critères et indicateurs adaptés.

2.1 Analyse du cas

2.1.1 Définition des éléments à évaluer

Il s'agissait bien de repérer, comme dans l'exemple précédent, les *facettes de l'objet à observer* qui devaient constituer l'ossature de l'ensemble de la référentialisation. Parmi ces éléments, on peut repérer, par exemple, les choix suivants :

- *En termes d'effets éventuels de contexte* : les modes d'accompagnement et de soutien de l'enfant et du jeune tout au long de son parcours
- *En ce qui concerne les stratégies* : les actions tendant à renforcer l'éducation à la citoyenneté et favoriser la participation
- *En matière de productions espérées* : un dispositif de pilotage et un suivi en continu

2.1.2 Choix des référents

Afin d'attribuer un sens aux informations, les évaluateurs du projet éducatif ont été amenés à privilégier certains référents :

- *Pour évaluer les effets éventuels de contexte* : les données fournies par l'Insee², par l'organisme qui gère les allocations familiales et par l'inspection académique ont constitué des référents systématiquement consultés.
- *Pour évaluer les effets stratégiques* : le concept de projet tel qu'il a été théorisé dans la littérature (Boutinet, 1990) ainsi que les documents d'orientation locaux, régionaux, nationaux sur les « projets éducatifs locaux » ont été utilisés comme référents.

2. INSEE : Institut national de la statistique et des études économiques (France).

- Pour évaluer les effets attendus en matière de participation des acteurs : les évaluateurs se sont référés à la notion de « participation symbolique à la citoyenneté » dont s'inspire Hart (1992) pour proposer une échelle de mesure de ce phénomène.

2.1.3 Élaboration des critères

Pour porter un jugement, les évaluateurs ont choisi comme aspects représentatifs du projet éducatif les critères suivants :

- Au niveau des effets de contexte : le critère d'évolution sociodémographique de la population des individus de zéro à seize ans.
- Au niveau des effets stratégiques : le critère de pertinence des objectifs du projet
- Au niveau des productions attendues : la cohérence des dispositifs

2.1.4 Construction des indicateurs

Les critères d'évaluation du projet éducatif ont été vérifiés au moyen d'indicateurs censés décrire leurs caractéristiques observables. On relève, parmi d'autres, des indicateurs comme :

- En matière d'effets de contexte : la progression et la répartition de l'effectif des enfants et des jeunes, dans la population globale du territoire, depuis 5 ans.
- En matière d'effets stratégiques : l'existence (et la nature) d'objectifs éducatifs opérationnels et d'actions priorisées ; le type de perception et les modes d'utilisation des différents dispositifs par les opérateurs (comme les associations).
- En matière de productions espérées : les modalités de participation des parents proposées par les structures dans les activités elles-mêmes ou dans des espaces de concertation spécifiques.

2.2 Mise en tableau

**Tableau 14. Évaluation d'un projet éducatif local
Phase 2**

Procédure centrale de la référentialisation (exemples)			
Éléments à évaluer	Référents	Critères	Indicateurs
Quelles dimensions de l'objet à observer ?	À partir de quoi attribuer un sens à l'information ?	Quels aspects représentatifs de l'objet à évaluer choisir pour porter un jugement ?	Quelles caracté- ristiques obser- vables permettent la vérification du degré d'atteinte du critère ?

<i>Éléments stratégiques (1,2,3)</i> 1. Accompagnement et soutien de l'enfant et du jeune tout au long de son parcours	– Confrontation avec les données statistiques initiales – avec les données statistiques des mêmes activités éducatives organisées dans des communes de population équivalente	Renforcement de l'offre éducative	– Évolution de la fréquentation des activités extrascolaires sur la commune par rapport à l'année N – 1 – Évolution du nombre d'enfants touchés sur les différents dispositifs de soutien scolaire par rapport à N – 1
2. Actions visant à renforcer l'éducation à la citoyenneté et à favoriser la participation des enfants et des jeunes	Données statistiques initiales Discours des acteurs associatifs	Développement des démarches de participation et de responsabilisation chez les jeunes en difficulté	Évolution du nombre d'enfants par année de naissance fréquentant une association sportive par rapport à N-1 Nature de l'évolution des projets pédagogiques des structures sur ce sujet
3. Dispositif de pilotage et de suivi en continu	Les fiches « Évaluation annuelle des projets » de la CAF du département	Qualité de la gestion de l'information des acteurs éducatifs	Types de médias utilisés pour communiquer avec les acteurs éducatifs Nature des contenus médiatiques transmis aux publics visés

2.3 Les apports de cette phase au processus d'évaluation

Le mandat ayant consisté à demander une évaluation, chemin faisant, d'un dispositif déjà conçu et en cours d'exercice, les évaluateurs avaient donc

minimisé la prise en compte du contexte (première dimension du modèle d'évaluation développé dans la deuxième partie de l'ouvrage) et concentré la méthodologie d'évaluation autour des *objectifs* affichés par le projet en examinant la dimension « productions » à travers l'état actuel (au moment du déroulement de l'évaluation) des activités et des évolutions repérables. Les éléments à évaluer ainsi que les référents étaient en quelque sorte prédéterminés. Les indicateurs qui ont été retenus représentaient la partie visible des critères construits par les évaluateurs à partir des représentations des commanditaires et des acteurs éducatifs interrogés.

3 PHASE 3 : INSTRUMENTATION

Quels outils les évaluateurs du projet éducatif local pris en exemple ont-ils employés pour renseigner les indicateurs précédemment construits ?

3.1 En guise d'analyse du cas

En fait, il s'agit d'une succession de procédures développant, chacune, selon les besoins, une interrogation, une observation, un questionnement.

On retrouve les trois types de procédures utilisées dans le cas précédent :

- *Des choix préalables ont été effectués (lignes 1, 2,3) :*
 1. Le document-cadre privilégié a été le « Projet Éducatif Local » rédigé cinq ans auparavant et qui pouvait servir de « pré-référentiel » dans la mesure où les objectifs qu'il contenait avaient eux-mêmes été construits à partir de l'analyse d'une situation locale et en collaboration avec des citoyens acteurs éducatifs.
 2. Les populations visées étaient les décideurs (élus), les enseignants et les responsables de structures sociales et éducatives participant au projet éducatif.
 3. Le mandataire avait repéré les acteurs à interroger (en fonction de sa connaissance du terrain) et l'évaluateur avait établi la liste des activités à observer (en fonction des objectifs figurant dans le document de référence précédemment repéré) et à partir d'échantillons élaborés selon la méthode des quotas.
- *Des instruments spécifiques ont été construits et exploités (lignes 4, 5) :*
 4. des entretiens semi-directifs et individuels auprès des acteurs (promoteurs et bénéficiaires du projet) s'imposaient
 5. Une analyse des documents et des réponses fournis par les acteurs interrogés a permis de mesurer les différents degrés d'atteinte des objectifs du projet.
- *Des actions d'exploitation des résultats ont été menées (ligne 6) :*
 6. Les informations recueillies ont été mises en valeur par des présentations devant un comité technique chargé d'évaluer la vraisemblance

et les retombées possibles de l'évaluation et par l'élaboration d'un rapport validé par le groupe de pilotage et diffusé aux différents partenaires engagés dans le projet.

3.2 Mise en tableau

**Tableau 15. Évaluation d'un projet éducatif local
Phase 3**

Instrumentation (exemples)	
Quels outils employer pour renseigner les indicateurs ?	
1 – Quels documents descriptifs et prescriptifs choisis ?	<i>Le document « Projet Educatif Local » rédigé cinq ans auparavant</i>
2 – Quelles populations interrogées ?	<i>Les décideurs (élus), les enseignants, les responsables de structure</i>
3 – Comment ont été sélectionnés les activités et les acteurs concernés ?	<i>Acteurs : repérage, par le mandataire, Activités à observer : échantillons élaborés selon la méthode des quotas</i>
4 – Quels instruments construits ?	<i>Entretiens semi-directifs (individuels et de groupe)</i>
5 – Comment ont été traitées les informations ?	<i>Analyse des documents et des réponses fournis par les acteurs interrogés</i>
6 – Comment les résultats ont-ils été exploités et mis en valeur ?	<i>Par des présentations devant un comité technique chargé d'évaluer la vraisemblance et les retombées possibles de l'évaluation Par l'élaboration d'un rapport validé par le groupe de pilotage et diffusé aux différents partenaires engagés dans le Projet</i>

3.3 Les apports de cette phase au processus d'évaluation

Il est à remarquer, dans cet exemple, la part importante prise par les acteurs dans la maîtrise et le contrôle des instruments d'évaluation :

- l'analyse des documents produits par les porteurs du projet a été réalisée sous contrôle des auteurs
- l'interrogation des responsables éducatifs a dû suivre les informations fournies par les commanditaires (sujets à interroger préalablement repérés)
- les rapports successifs ont été présentés et discutés devant une instance constituée par des acteurs du projet (groupe de pilotage).

4 PHASE 4 : INTERPRÉTATION

Les données ayant été recueillies selon le protocole précédent, les évaluateurs ont disposé, alors, de résultats classés selon les indicateurs construits par la référentialisation entreprise. L'interprétation consistant à attribuer du sens à ces résultats, pour en rendre compte, on l'intégrera, comme dans les cas qui précèdent, directement dans la mise en tableau, de manière à synthétiser, dans un seul support, la mise en cohérence avec les indicateurs ayant provoqué les informations à analyser. La colonne *instrumentation* n'y est conservée que pour mémoire.

4.1 Analyse du cas mise en tableau

**Tableau 16. Évaluation d'un projet éducatif local
Phase 4**

Interprétation (exemples)		
Indicateurs	Instrumentation	Interprétation
Quelles caractéristiques observables permettent la vérification du degré d'atteinte du critère ?	Quels outils employer pour renseigner les indicateurs ?	Quelles significations attribuer aux résultats ?
Élément stratégique 1 : <i>Évolution de la fréquentation des activités extrascolaires sur la commune par rapport à l'année N – 1</i>	<i>Analyse de tableaux de bord</i>	<i>Concernant les activités extra scolaires (mercredis, petites et grandes vacances), nous pouvons constater une progression de la fréquentation en nombre de journées enfants du centre de loisirs, de l'accueil multisports, du pôle jeunesse, etc.</i>
Élément stratégique 2 : <i>Nature de l'évolution des projets pédagogiques des structures sur ce sujet</i>	<i>Analyse de contenus des projets pédagogiques des structures</i>	– <i>Évolution des projets pédagogiques en matière de règles de vie</i> – <i>L'absence de relevés précis sur la construction ou la mise en œuvre de ces règles de vie n'autorise pas à s'exprimer sur un renforcement de l'éducation à la citoyenneté dans les espaces éducatifs.</i>

Élément stratégique 3 : <i>Types de médias utilisés pour communiquer avec les acteurs éducatifs</i>	<i>Analyse documentaire</i> <i>Analyse de discours des acteurs</i>	– <i>Existence de 11 médias construits par les acteurs éducatifs.</i> – <i>Communication globale sur le projet à renforcer.</i> – <i>Visibilité et lisibilité de l'action à renforcer par la création d'un outil de communication rassemblant l'ensemble des propositions de loisirs faites aux enfants et aux jeunes sur la commune</i>
---	---	--

4.2 Les apports de cette phase au processus d'évaluation

Le mandat ayant porté essentiellement sur la dimension stratégique du projet donné à évaluer, il s'ensuit, assez naturellement, une interprétation des données recueillies, en termes de conformité aux objectifs affichés et, de manière générale, dans un registre stratégique. Il s'agissait, en effet, de répondre à des interrogations telles que « comment tel objectif a-t-il été opérationnalisé ? », « pourquoi tel dispositif prévu n'a-t-il pas été réalisé ? », « comment s'explique telle ou telle dynamique perceptible plutôt ici que là ? », etc.

Chapitre 10

Troisième application : Évaluation des dispositifs et programmes

Un exemple : Évaluation d'un dispositif de classes-passerelles vers des écoles d'ingénieurs

INTRODUCTION : LES DISPOSITIFS ÉDUCATIFS ET LES PROGRAMMES COMME OBJETS À ÉVALUER

Concepts de dispositif et de programme

Nous avons volontairement maintenu le double titre qui rassemble et distingue en même temps *dispositif* et *programme* du fait que les deux notions sont utilisées pour désigner un même type d'objet : un mode d'organisation de l'action éducative, sous forme de découpage d'un contenu d'apprentissage ou de formation au service de finalités et d'objectifs. On l'appelle plus couramment *programme* dans le contexte anglo-saxon et *dispositif* dans la culture francophone. Mais le *programme* sous-entend plus facilement l'existence de contenus explicites alors que le *dispositif* désigne plus souvent le mécanisme mis en place pour favoriser des apprentissages. Ainsi, on parle de *programme d'enseignement* et de *dispositif d'alternance*, en France) et de *programme de réduction des conflits chez les enfants*, au Canada.

On choisira cependant, ici, d'utiliser le terme *dispositif* dans une acception générale de *mise en situation* (Altet, 1997) d'un projet éducatif, englobant la notion de programme. Une traduction anglophone du *dispositif* éducatif émerge cependant dans les documents internationaux, celle de *educational arrangements* (UNESCO, 2008).

Nature et conceptualisation d'un dispositif éducatif

Contrairement à ce que l'on pourrait croire, le dispositif n'est pas un simple organigramme mais un objet complexe qui demande à être conceptualisé avant d'être construit ou évalué. Il existe même une théorie du dispositif (Peeters et Charlier, 1999) qui le définit comme la concrétisation d'une intention au travers de la mise en place d'environnements aménagés (au moyen de mécanismes finalisés constituant des instruments pour l'action). Le dispositif constitue, en général, un instrument de projet et organise, dans cette perspective, un espace de médiation entre les acteurs.

Différents cas de figure

On désigne comme dispositifs de nombreuses configurations destinées à faire fonctionner des activités éducatives : des dispositifs de lutte contre l'échec scolaire, de veille éducative, de réussite éducative, d'alternance travail/études, de zones d'éducation prioritaire, des mises en situation de groupe, d'expérimentation, de découverte, et la liste n'est pas close, comme l'illustre l'exemple qui va suivre.

Évaluation des dispositifs éducatifs

Les modes d'évaluation des dispositifs éducatifs sont très variés du fait de la diversité des regards qui sont portés sur ces objets. L'approche sociologique, avec ses outils d'observation et d'enquête ainsi que ses analyses quantitatives, y joue un rôle important. Mais des problématiques souvent complexes et localisées amènent des évaluateurs à construire des protocoles plus spécifiques. C'est alors qu'intervient l'intérêt de mobiliser un protocole de référentialisation.

MÉTHODOLOGIE REPÉRÉE

Du protocole méthodologique utilisé dans cet exemple, on retiendra les éléments suivants :

1 PHASE 1 : EXPLORATION (MANDAT ET PROBLÉMATIQUE)

On a vérifié qu'elle s'est déroulée de la manière suivante, consultable sous deux versions successives (par une analyse cursive du cas puis par une version synoptique sous forme de tableau).

1.1 Analyse du cas

1.1.1 Le mandat

Il s'agissait, à la demande du ministère français de l'Agriculture, d'évaluer le devenir des étudiants issus d'un type de classes préparatoires à l'entrée des grandes écoles d'ingénieurs agronomes. Ces classes sont destinées à offrir à des sujets issus (initialement) de l'enseignement technique (par le biais des filières de formation des techniciens supérieurs¹), l'opportunité de rejoindre des grandes écoles d'ingénieurs en agronomie. Il s'agissait, par la même occasion, d'évaluer les effets du dispositif des classes passerelles entre BTS(A) et Grandes Écoles d'agronomie, cette opération intéressant aussi l'institution commanditaire : en effet, les prescripteurs ont la volonté de repérer et mettre en valeur les atouts de ce dispositif dans une logique externe, comparative, et, autant que se peut, longitudinale.

En matière de prise en compte du contexte, les évaluateurs ont souhaité observer le passage entre études universitaires courtes et longues et étudier les profils différenciés des candidats.

Concernant les stratégies, les observations ont porté principalement sur celles des élèves consistant à utiliser un passage par l'approche professionnelle pour réussir un concours à dominante académique ; puis sur celles des responsables du dispositif et des enseignants consistant à s'appuyer tout particulièrement sur des motivations et des compétences construites, en partie, sur des expériences de terrain.

Quant à l'évaluation des productions, elle a porté notamment sur l'impact de ce dispositif sur la réussite au concours final et sur l'insertion professionnelle des anciens étudiants.

1.1.2 La problématique

L'évaluation entreprise devait répondre à un défi important, celui de l'hétérogénéité du public de ces classes face à l'enjeu d'une entrée unique et aménagée dans un système de « Grandes Écoles » à la française privilégiant un modèle de recrutement très sélectif autour de matières principalement scientifiques. La deuxième interrogation dominant l'ensemble de l'opération consistait à élucider les stratégies d'utilisation par les étudiants d'un dispositif curriculaire particulier.

1. Il s'agit des Brevets de Techniciens Supérieurs de l'Agriculture (BTSA), en France.

1.2 Mise en tableau

**Tableau 17. Évaluation d'un dispositif de classes-passerelles
vers des écoles d'ingénieurs
Phase 1**

Exploration (mandat et problématique) (exemples)	
Analyse du mandat	Problématique d'évaluation
Par rapport au contexte <i>Identifier les attentes différenciées des étudiants face à une voie de réussite</i> Par rapport aux stratégies <i>Mettre en lumière les stratégies intrinsèques et extrinsèques des étudiants</i> Par rapport aux productions <i>Déterminer l'impact sur la réussite au concours final et sur l'insertion professionnelle</i>	À quelle(s) question(s) fondamentale(s) l'évaluation doit-elle répondre ? – <i>Hétérogénéité du public face à un dispositif d'accès à un système de Grandes Écoles</i> – <i>Élucider les effets de stratégies d'utilisation d'un dispositif curriculaire</i>

1.3 Les apports de cette phase au processus d'évaluation

Là encore, cette première phase montre l'intérêt d'utiliser l'analyse du mandat et de la situation qu'il met en lumière pour opérer une transformation : celle d'une demande initiale en problématique d'évaluation (c'est-à-dire d'une attente du commanditaire en projet d'intervention par les évaluateurs). Ici, l'institution demandait une étude sur la pertinence d'un dispositif spécifique mis en place pour favoriser l'accès d'un public rural au système des « Grandes écoles ». La problématique d'évaluation qui en a été dégagée a consisté à élucider les effets des stratégies d'utilisation de ce dispositif par un public diversifié.

2 PHASE 2 : PROCÉDURE CENTRALE DE RÉFÉRENTIALISATION

On a relevé, dans les procédures qui constituent la phase 2 de cet exemple, des éléments à évaluer ainsi que des choix de référents correspondant bien à des problématiques curriculaires (motivations d'études, stratégies d'utilisation des dispositifs particuliers...). Quant aux critères

et aux indicateurs, ils ont été volontairement confondus en une seule catégorie².

2.1 Analyse du cas

2.1.1 Définition des éléments à évaluer

- *En termes d'effets éventuels de contexte, ont été notamment mis en exergue* : les milieux de vie ruraux ou urbains des élèves, les informations contenues dans les dossiers d'entrée, la vie scolaire dans les Écoles.
- *En ce qui concerne les stratégies* : les stratégies curriculaires des élèves (contournements des difficultés d'entrée ou de réussite dans des classes préparatoires traditionnelles, projets à moyen ou à long terme des élèves).
- *En matière de productions espérées* : des connaissances générales et des compétences enrichies par l'expérience de la classe-passerelle (disciplinaires pour les techniciens et professionnelles pour les élèves issus d'études générales).

2.1.2 Choix des référents

En ce qui concerne la prise en compte du contexte, pour comprendre les choix et les réussites curriculaires des étudiants, les évaluateurs se sont référés aux milieux de vie (ruraux ou urbains) dont ils étaient issus. En effet, un parcours de candidat citadin passant par une classe (fût-elle « passerelle ») pour accéder à une Grande École d'Agronomie ne devait pas s'analyser de la même façon que celui d'un candidat d'origine rurale. On pouvait également mieux expliquer les choix des étudiants en se référant aux pratiques de vie scolaire significativement différentes en fonction de leur histoire scolaire (figurant dans les dossiers d'entrée).

Au niveau stratégique, le référent jugé essentiel a été le « projet » des étudiants, recueilli essentiellement par entretiens semi-directifs et à la lumière duquel les évaluateurs ont pu comprendre les particularités des choix *intrinsèques* (dans une logique reliant BTS agricole, Grande École d'Agronomie et profession adaptée) ou *extrinsèques* (dans une logique de recherche de réussite par des moyens détournés).

En termes de productions attendues, l'hypothèse des évaluateurs portait sur l'acquisition de connaissances générales approfondies pour le public issu de l'enseignement agricole et de connaissances professionnelles enrichies pour le public issu de l'enseignement général. Les

2. L'exemple a été reconstruit sur un cas réel au cours duquel, en effet, les commanditaires ont demandé aux évaluateurs, pour l'occasion, de renoncer à la distinction entre critère et indicateur.

résultats attendus portaient également sur une reconnaissance du profil spécifique et sur une plus grande réussite des élèves de type « passerelle ».

2.1.3 Indicateurs généraux (critères et indicateurs volontairement confondus)

Quels aspects représentatifs (ou critères) de l'objet à évaluer (c'est-à-dire le dispositif des « classes-passerelles ») ont-ils été sélectionnés pour porter une appréciation ? Et quelles caractéristiques visibles (ou indicateurs) pouvaient-elles en être observées ?

En ce qui concerne la prise en compte du contexte, on peut citer, par exemple, des indicateurs comme « établissement scolaire d'origine », « expérience des stages », « performances des élèves dans les disciplines scientifiques ».

Au niveau stratégique, parmi de nombreux indicateurs retenus, on a relevé ceux portant sur les motifs des choix, les stratégies de contournement, les options internes à l'École, les recherches d'emploi.

En termes de productions attendues, ont été choisis comme indicateurs principaux les performances comparées aux concours des élèves à profils d'origine différents, les types d'emploi et les réussites professionnelles après l'École.

2.2 Mise en tableau

**Tableau 18. Évaluation d'un dispositif de classes-passerelles
vers des écoles d'ingénieurs
Phase 2**

<i>Procédure centrale de référentialisation</i>		
Éléments à évaluer	Référents	Indicateurs généraux (critères et indicateurs volontairement confondus)
Quelles dimensions de l'objet à observer ?	À partir de quoi attribuer un sens à l'information ?	Quels aspects représentatifs de l'objet à évaluer choisir pour porter un jugement ? Et Quelles caractéristiques observables permettent la vérification du degré d'atteinte du critère ?

Éléments contextuels : <i>profils d'origine</i>	– <i>Milieux de vie ruraux ou urbains</i> – <i>Dossiers d'entrée</i> – <i>Vie scolaire dans les Écoles</i>	– <i>Établissement scolaire d'origine</i> – <i>Expérience des stages</i> – <i>Performances des élèves dans les disciplines scientifiques</i>
Éléments stratégiques : <i>stratégies curriculaires</i>	– <i>Projets des élèves</i>	– <i>Motifs des choix</i> – <i>Stratégies de contournement</i> – <i>Choix des options internes à l'École</i> – <i>Recherches d'emploi</i>
Productions : – <i>connaissances générales et professionnelles</i> – <i>compétences acquises</i>	– <i>Reconnaissance du profil spécifique des élèves « passerelle » dans les grandes Écoles</i> – <i>Réussites aux concours</i>	– <i>Performances comparées des élèves à profils d'origine différents</i> – <i>Types d'emploi et réussites professionnelles après École</i>

2.3 Les apports de cette phase au processus d'évaluation

Les profils d'origine, les stratégies curriculaires ainsi que les compétences et connaissances attendues ont constitué un exemple équilibré d'ingrédients utilisables pour une évaluation de dispositif : ces trois dimensions combinées ont permis de construire des indicateurs ayant suivi les parcours des étudiants-candidats : ces indicateurs, une fois renseignés, sont porteurs d'explication des effets de ces classes dans une logique d'ensemble plutôt que dans une observation morcelée mettant en valeur tel ou tel facteur.

3 PHASE 3 : INSTRUMENTATION

Quels outils les évaluateurs du dispositif de classes-passerelles pris en exemple ont-ils employés pour renseigner les indicateurs généraux précédemment construits ?

Ici, on remarque une prédilection pour des analyses de profils individuels (diachroniques et synchroniques), des entretiens et des questionnaires destinés à des populations relativement nombreuses (nombre total de réponses : 460), une enquête de terrain auprès d'une population dispersée de professionnels anciens élèves.

3.1 En guise d'analyse du cas

On retrouve les trois types de procédures d'instrumentation utilisées dans le cas précédent.

- *Des choix préalables ont été effectués (lignes 1, 2, 3) :*
 1. Les évaluateurs ont d'abord procédé à une analyse des programmes officiels, référentiels, tableaux de réussite disponibles dans la documentation de l'institution commanditaire.
 2. Les populations interrogées étaient les enseignants, étudiants (candidats, élèves en cours d'étude, lauréats), professionnels anciens élèves.
 3. Un repérage des groupes de populations concernés (élèves, enseignants, professionnels) a été réalisé par l'institution.
- *Des instruments spécifiques ont été construits et exploités (lignes 4, 5) :*
 4. Les instruments les plus adaptés à ces choix méthodologiques étaient, à juste titre, ceux procédant de l'enquête classique et employés ici : questionnaires et entretiens semi-directifs.
 5. À partir de l'analyse des informations recueillies, il a été procédé à une reconstitution des parcours des étudiants.
- *Des actions d'exploitation des résultats ont été menées (ligne 6) :*
 6. Elles ont été conformes au cahier des charges de l'évaluation commanditée, c'est-à-dire la présentation d'un rapport validé par le groupe de pilotage et diffusé aux différents partenaires de l'évaluation (ministère mandataire, enseignants, étudiants).

3.2 Mise en tableau

**Tableau 19. Évaluation d'un dispositif de classes-passerelles
vers des écoles d'ingénieurs
Phase 3**

Instrumentation	
Quels outils employer pour renseigner les indicateurs ?	
1 – Quels documents descriptifs et prescriptifs choisis ?	<i>Programmes, référentiels, tableaux de réussite</i>
2 – Quelles populations interrogées ?	<i>Enseignants, étudiants (candidats, en cours d'étude, lauréats, professionnels anciens élèves)</i>
3 – Comment ont été sélectionnés les activités et les acteurs concernés ?	<i>Repérage par l'institution des populations concernées (élèves, enseignants, professionnels)</i> – <i>Étudiants : populations entières</i> – <i>Enseignants : échantillon élaboré selon la méthode des quotas</i> – <i>Professionnels : sujets repérés par le mandataire</i>

4 – Quels instruments construits ?	<i>Enquête :</i> – questionnaires – entretiens semi-directifs
5 – Comment ont été traitées les informations ?	<i>Reconstitution des parcours à partir de l'analyse des informations recueillies</i>
6 – Comment les résultats ont-ils été exploités et mis en valeur ?	– Par des présentations devant un comité de pilotage – Par l'élaboration d'un rapport diffusé aux différents partenaires de l'évaluation (ministère mandataire, enseignants, étudiants)

3.3 Les apports de cette phase au processus d'évaluation

L'évaluation de ce dispositif a apparemment mobilisé un appareillage méthodologique important (analyse de documents, entretiens préparatoires, enquête auprès de populations entières, recherche de professionnels épars avec entretiens biographiques, mise en discussion des rapports...). On pourrait se demander si de tels moyens étaient indispensables à la mise en lumière du rôle et des effets des classes-passerelles entre enseignement technique et Grandes Écoles d'Agronomie. La réponse est à chercher, selon nous, dans la relation avec le projet né de la commandite posée : il a été choisi de réaliser, non pas seulement une étude d'impact mais aussi et surtout une évaluation partagée dans laquelle les acteurs voulaient comprendre l'ensemble des facteurs agissant sur la dynamique du dispositif à observer. Dans ce sens, l'appareillage méthodologique a favorisé une réflexion plus complète des acteurs.

4 PHASE 4 : INTERPRÉTATION

Les données ayant été recueillies selon le protocole précédent, les évaluateurs ont disposé, alors, de résultats classés selon les indicateurs construits par la référentialisation entreprise. L'interprétation consistant à attribuer du sens à ces résultats, pour en rendre compte, on l'intégrera, comme dans les cas qui précèdent, directement dans la mise en tableau, de manière à synthétiser, dans un seul support, la mise en cohérence avec les indicateurs ayant provoqué les informations à analyser. La colonne *instrumentation* n'y est conservée que pour mémoire parce qu'elle fait partie d'une continuité *indicateurs-instrumentation-interprétation*.

4.1 Mise en tableau

**Tableau 20. Évaluation d'un dispositif de classes-passerelles
vers des écoles d'ingénieurs
Phase 4**

Interprétation (exemples)		
Indicateurs	Instrumentation	Interprétation
Quelles caractéristiques observables permettent la vérification du degré d'atteinte du critère ?	Quels outils employer pour renseigner les indicateurs ?	Quelles significations attribuer aux résultats ?
Éléments contextuels : <i>profils d'origine</i>	<i>Analyse des programmes, référentiels, tableaux de réussite</i>	<i>Pertinence des programmes par rapport aux profils d'origine rurale et professionnelle</i> – <i>impertinence avec les autres profils</i>
Éléments stratégiques : <i>stratégies curriculaires</i>	<i>Enquête :</i> – <i>questionnaires</i> – <i>entretiens semi-directifs</i>	– <i>Stratégies « intrinsèques » cohérentes et porteuses de réussite</i> – <i>Stratégies « extrinsèques » aux résultats mitigés</i>
Productions : – <i>connaissances générales et professionnelles</i> – <i>compétences acquises</i>	<i>Élaboration d'un rapport validé par le groupe de pilotage et diffusé aux différents partenaires de l'évaluation (ministère mandataire, enseignants, étudiants)</i>	<i>Mise en évidence :</i> – <i>d'une filière pertinente et productrice de compétences sur un axe professionnalisant</i> – <i>de l'incohérence de son utilisation à des seules fins de contournement de la difficulté des concours d'entrée dans les « Grandes Écoles »</i>

4.2 Les apports de cette phase au processus d'évaluation

L'interprétation des différents résultats obtenus après passation des instruments précédemment indiqués confirme la problématique construite après analyse du mandat initial : les évaluateurs ont disposé, en effet, d'explications plus complexes qu'attendu des effets produits par le dispositif donné à évaluer. En analysant les réponses des sujets aux différents stades

du curriculum concerné (étudiants en BTSA³, étudiants hors BTSA, étudiants participant aux classes passerelles (d'origines diverses), lauréats aux concours d'entrée ainsi que non-lauréats, étudiants en grandes Écoles (d'origines diverses) et professionnels après École, il a été confirmé que le dispositif en cause donnait libre cours à de nombreuses stratégies d'utilisation et à plusieurs formes de réussite possibles. Il a été vérifié, également, que certains parcours, surtout ceux guidés par une stratégie « intrinsèque » (public d'origine rurale, passé par un BTSA, utilisateur de la classe-passerelle, lauréat ou non au concours d'entrée en Grande École) obtenaient des insertions professionnelles réussies et variées.

3. Il s'agit des Brevets de Techniciens Supérieurs de l'Agriculture (BTSA), en France.

Chapitre 11

Quatrième application : Évaluation des outils pédagogiques

*Un exemple : Évaluation des supports
pédagogiques d'une formation de formateurs*

INTRODUCTION : LES OUTILS PÉDAGOGIQUES COMME OBJETS À ÉVALUER

Évaluation des outils pédagogiques : référence à l'ingénierie éducative (ou pédagogique)

Qu'est-ce qu'un outil, un instrument, ou un support pédagogique ? On a coutume de ranger ces objets dans la catégorie des « méthodes » pédagogiques comme étant l'ensemble des moyens d'organisation d'une situation de formation utilisant l'espace, le temps, les matériaux, les ressources, les mises en relation choisis pour permettre et/ou favoriser des apprentissages. Il existe, néanmoins, peu de conceptualisations de cette notion, en soi. Si l'on veut mobiliser des concepts capables de fonder le choix et le maniement de ces outils, il faut alors chercher, à l'amont de leur élaboration, dans les modèles pédagogiques qu'ils illustrent explicitement ou implicitement. Cette recherche peut effectivement s'entreprendre à l'occasion de l'évaluation de tel ou tel outil. Mais on en restera, ici, à l'évocation d'une approche plus générale des « outils » comme éléments d'ingénierie

éducative (ou pédagogique). Appelée également, au Québec, *design pédagogique*, cette activité consiste à concevoir et à mettre en œuvre un dispositif ou un programme de formation en faisant appel, ou non, à ce que les tenants des technologies d'information et de communication, appellent, aujourd'hui, l'ingénierie éducative (Tardif, 1998 ; Depover *et al.*, 2003). Cette dernière notion peut, en tout cas, être revendiquée dans le cadre plus large de l'ensemble des techniques, méthodes et moyens conçus et/ou utilisés pour organiser des séquences d'apprentissage, avec ou sans recours aux médias informatisés. Évaluer des outils pédagogiques consistera donc à apprécier leur pertinence et leur efficacité, au regard du processus d'ingénierie développé en réponse à des objectifs précis et clairement identifiés.

Différents cas de figure

Dans ce domaine, les objets donnés à évaluer sont nombreux, qu'ils s'appellent outils ou instruments, supports, manuels ou fichiers, didacticiels ou documents *on line* (comme ce sera le cas dans l'expérience qui va être relatée ci-dessous). En ce qui concerne l'évaluation d'outils informatisés, on renverra aux auteurs précédents tout en mettant en garde contre l'apparente facilité des protocoles programmés en liaison avec des séquences didactiques (comme c'est le cas, par exemple, dans des dispositifs de e-learning), présentant des formes techniques rigoureuses ainsi que ludiques, mais ne répondant pas, systématiquement, à des modèles d'évaluation préalablement invoqués.

MÉTHODOLOGIE REPÉRÉE

Du protocole méthodologique utilisé dans l'opération prise en exemple, on retiendra, en suivant toujours le même plan, les éléments suivants.

1 PHASE 1 : EXPLORATION (MANDAT ET PROBLÉMATIQUE)

Comme dans les exemples précédents, on a cherché à savoir comment les évaluateurs ont analysé le mandat reçu et exprimé la problématique d'évaluation qui s'en dégageait.

1.1 Analyse du cas

1.1.1 Le mandat

Il était demandé de vérifier la qualité des outils pédagogiques créées par les membres d'un réseau national de formateurs ayant participé à des sessions de formation continue organisées et financées par un programme européen ADAPT. Ces outils (manuels, fichiers, mallettes, supports informatisés *on line* et *off line*) devaient être évalués par un groupe d'évaluateurs affectés

à l'opération, avant leur diffusion dans les organismes de formation où ils devaient être améliorés et utilisés. En conséquence, l'institution commanditaire avait besoin de disposer d'une méthode d'évaluation applicable à l'ensemble de ces outils et pouvant être praticable par des évaluateurs dispersés dans l'ensemble du pays.

En matière de prise en compte du contexte, les concepteurs du dispositif d'évaluation avaient à intégrer les exigences du programme européen soucieux de contrôler la conformité des produits subventionnés à ses orientations et à ses priorités. Ils avaient, aussi, à considérer la reconnaissance attendue par les auteurs des outils ainsi que par les formateurs de formateurs qui avaient encadré ces auteurs.

En ce qui concerne les stratégies d'acteurs, il fallait prendre en considération, à la fois, la dimension formative de l'opération (avec une forme d'évaluation adaptée au suivi des processus de construction de compétences ainsi qu'à la possibilité d'amélioration des productions) et la situation de contrôle qui lui était clairement associée (entraînant l'élaboration d'une mesure de conformité et d'écarts).

En ce qui concerne l'anticipation du type de résultats espérés, les concepteurs avaient, enfin, à garder à l'esprit deux enjeux. Le premier résidait dans l'importance que tout programme européen accorde à la reproductibilité d'une telle opération en matière d'aide au développement de modèles de formation d'adultes et de méthodologies de formation. Ceci, d'autant plus qu'il s'agissait, ici, d'une double possibilité de réinvestissement : en termes de production d'outils et de supports de formation, d'une part, et en termes de méthodologies d'évaluation, d'autre part. Le second enjeu consistait en l'utilisation prévisible des résultats pour des améliorations d'outils.

1.1.2 La problématique

Les questions posées à l'évaluation étaient très spécifiques à l'objet à traiter. Il s'agissait :

- de définir les caractéristiques et les fonctions attendues de chaque outil,
- de disposer, pour ce faire, d'un « référentiel » représentatif des options, objectifs et compétences des partenaires du programme au sein duquel devait se mettre en place l'ensemble du dispositif d'évaluation envisagé,
- de concevoir des instruments d'évaluation laissant le moins de place possible aux jugements appréciatifs du genre « bon », « médiocre », etc. et répondant davantage à des règles de conformité à des critères les plus objectifs possibles et des indicateurs laissant peu de place aux approximations.

1.2 Mise en tableau

Tableau 21. Évaluation des supports pédagogiques d'une formation de formateurs
Phase 1

Exploration (mandat et problématique) (exemples)	
Analyse du mandat	Problématique d'évaluation
Par rapport au contexte – Besoin de disposer d'une méthode d'évaluation applicable à l'ensemble de ces outils et pouvant être praticable par des évaluateurs dispersés dans l'ensemble du pays Par rapport aux stratégies – À la fois : dimension formative et situation de contrôle Par rapport aux productions – <i>Reproductibilité</i> – <i>Résultats à prendre en compte pour des améliorations d'outils</i>	À quelle(s) question(s) fondamentale(s) l'évaluation doit-elle répondre ? – <i>Disposer d'un « référentiel » représentatif des options, objectifs et compétences des partenaires du programme au sein duquel devait fonctionner l'ensemble du dispositif d'évaluation d'outils pédagogiques.</i> – <i>Qu'est-ce qu'un « bon » outil pédagogique ?</i>

1.3 Les apports de cette phase au processus d'évaluation

Devant un mandat si ambitieux (consistant à construire des instruments d'évaluation de supports pédagogiques divers utilisables dans des activités professionnelles diverses, des régions diverses et des organismes de formation tout aussi divers), une première réponse aurait pu consister à renoncer à une telle entreprise. Cependant, une analyse approfondie de ce mandat a, dans un premier temps, éclairé cette ambition : plus qu'un jugement de valeur à porter sur chaque outil, ce qui était attendu était plutôt la mise en place d'un dispositif d'évaluation de supports pédagogiques selon un référentiel implicitement utilisé par l'organisation commanditaire. Un référentiel construit avec la collaboration des professionnels de formation et des didacticiens concernés par les contenus disciplinaires des différents outils et que les futurs évaluateurs devaient s'approprier. Il apparaissait essentiel, alors, qu'une définition consensuelle du « bon » outil pédagogique puisse être adoptée par l'ensemble du réseau producteur et utilisateur de ces ressources. Dans cette optique, l'évaluation pouvait, alors, commencer.

2 PHASE 2 : PROCÉDURE CENTRALE DE RÉFÉRENTIALISATION

Dans un tel contexte (programme européen à visées complexes, production d'outils pédagogiques pour des domaines de formation diversifiés), que fallait-il évaluer et comment pouvait-on justifier et formuler des référents, critères et indicateurs pertinents ?

2.1 Analyse du cas

2.1.1 *Choix des éléments à évaluer*

Les effets éventuels du contexte sur la qualité des outils (influence du cadre européen, spécificité des disciplines de formation, culture des organismes de formation participant à l'opération) ne devaient pas être évalués car cela aurait consisté à introduire la reconnaissance d'une variabilité fonctionnelle des outils d'une même catégorie. Le dispositif de formation voulant être « national », un même outil (manuel, fichier, support informatique, etc.) devait comporter des caractéristiques unifiées, reconnaissables et échangeables d'un organisme à l'autre.

La dimension stratégique de l'outil à évaluer a fait l'objet, en revanche, d'un traitement particulier. Il a été demandé de tenir compte des spécificités de chaque outil selon qu'il était destiné à la formation individuelle, à la formation accompagnée par tutorat ou à la formation collective. L'évaluation devait faire ressortir soit la spécialisation de chacun d'entre eux dans l'une de ces situations soit son adaptabilité à l'une ou à l'autre.

Quant au « produit », il a constitué l'élément à évaluer dominant ; il s'agissait, en effet, d'apprécier la qualité de supports de formation classés en « écrits », « audio », « vidéo », « multimédia » (off line et on line), « autres ». Même si les processus de construction des outils étaient potentiellement considérés, les outils eux-mêmes en tant que produits finis, livrables, reproductibles, restaient, aux yeux des partenaires de l'opération, les véritables « objets » à évaluer.

2.1.2 *Référents*

Les éléments contextuels ayant été très peu pris en compte dans cette phase de référentialisation, examinons la suite du protocole, c'est-à-dire :

Les éléments stratégiques : comme principaux référents, les évaluateurs ont pointé les objectifs généraux du programme européen, les objectifs généraux de la formation de formateurs (incluant une production d'outils), et, enfin, les attentes d'un public de formateurs, consultants et experts.

Les productions ayant constitué, dans cet exemple, la partie la plus visible de l'objet à évaluer (voir *supra*), les référents invoqués relèvent surtout des normes et des techniques éditoriales comme les modèles de communication audio-visuelle et informatique, les modèles didactiques spécifiques aux disciplines concernées, etc.

2.1.3 Critères

Quels aspects représentatifs de chacun des outils les évaluateurs ont-ils sélectionnés pour porter un jugement ?

Au niveau stratégique, les critères jugés les plus adéquats ont été ceux de conformité aux objectifs et de pertinence pédagogique.

Au niveau des productions, on trouve de nombreux critères descriptifs des caractéristiques spécifiques de chacun des outils, comme ceux de complétude du produit, lisibilité du plan formel, homogénéité entre les parties, adéquation à l'utilisation, facilité de consultation, etc.

2.1.4 Indicateurs

Les caractéristiques observables de ces critères étaient, elles aussi, nombreuses et aisées à formuler comme, par exemple, la présence des éléments suivants : titre, sommaire, objectifs, mode d'emploi, apports de connaissances, ressources internes, activités d'évaluation, etc.

2.2 Mise en tableau

Tableau 22. Évaluation des supports pédagogiques d'une formation de formateurs
Phase 2

Procédure centrale de référentialisation			
Éléments à évaluer	Référents	Critères	Indicateurs
Quelles dimensions de l'objet à observer ?	À partir de quoi attribuer un sens à l'information ?	Quels aspects représentatifs de l'objet à évaluer choisir pour porter un jugement ?	Quelles caractéristiques observables permettent la vérification du degré d'atteinte du critère ?
Éléments contextuels	<i>Référents décontextualisés</i>	<i>Critères décontextualisés</i>	<i>Indicateurs décontextualisés</i>

Éléments stratégiques – Formation individuelle – Formation accompagnée par tutorat – Formation collective	– Objectifs généraux du programme européen – Objectifs généraux de la formation de formateurs avec production d'outils – Attentes d'un public de formateurs, consultants et experts	– Conformité aux objectifs – Conformité aux objectifs – Pertinence pédagogique	– Objectifs visibles – Objectifs visibles – Présence de guides d'usage, de modes d'emploi...
Productions – Caractéristiques des produits finis : « écrits », « audio », « vidéo », « multimédia » (off line et on line), « autres » : Structure, présentation, illustrations, utilisation, etc.	– Normes éditoriales – Modèles de communication audio-visuelle et informatique – Etc. Modèles didactiques spécifiques aux disciplines concernées, etc.	– Complétude du produit – Lisibilité du plan formel – Homogénéité entre les parties – Adéquation à l'utilisation – Facilité de consultation – Etc.	– Présence des éléments : titre, sommaire, objectifs, mode d'emploi, apports de connaissances, ressources internes, activités d'évaluation etc.

2.3 Les apports de cette phase au processus d'évaluation

De manière curieuse, les données de contexte n'ont pas été formellement convoquées, ici, même si elles auraient pu tempérer sensiblement les choix instrumentaux. On voit, en effet, que seules les dimensions horizontales vouées aux *stratégies* et aux *productions* ont été réellement traitées. En fait, le contexte, dans ce cas, était porteur d'unification plutôt que de diversification : les outils pédagogiques devaient pouvoir être utilisés à un niveau national, répondre à des normes édictées dans le cadre d'un programme européen et pouvoir être évalués de la même manière par des formateurs diversement situés. En revanche les deux autres dimensions (stratégies et productions attendues) étaient pleinement concernées : les critères et les indicateurs ont été formulés, alors, de manière la plus précise et la moins équivoque possible.

3 PHASE 3 : INSTRUMENTATION

3.1 En guise d'analyse du cas

Construire des outils d'évaluation pour évaluer des outils pédagogiques : c'est bien ce qu'ont tenté de faire les évaluateurs dans cette opération. Il faut rappeler qu'ici, ces outils devaient être utilisés par d'autres évaluateurs que ceux ayant contribué à les construire. Ce qui demandait une réduction à l'extrême des marges de manœuvre possibles laissées aux opérateurs finaux : le référentiel devait être très précis, les indicateurs très formels et les items très simplifiés. Malgré tout, la procédure d'instrumentation a suivi les mêmes modalités que dans les autres situations et on peut constater que les indicateurs, construits selon le même protocole, conservent ici aussi le même statut. Ils désignent les aspects précis que l'évaluateur regarde et vérifie lorsqu'il manipule le manuel, le fichier ou la séquence filmée.

- *Des choix préalables ont été effectués (lignes 1, 2, 3) :*
 1. Les évaluateurs ont d'abord procédé à une lecture exploratoire d'échantillons de supports pédagogiques par discipline et par catégorie, afin de disposer, pour chaque catégorie, d'un modèle d'objet à évaluer.
 2. Les populations interrogées étaient les formateurs et les didacticiens concernés par chaque type de document.
 3. Un repérage des outils produits et un panel de formateurs ont été proposés par les responsables du dispositif de formation.
- *Des instruments spécifiques ont été construits et exploités (lignes 4, 5) :*
 4. Des grilles d'analyse des outils pédagogiques ont été élaborées et des focus-groupes sur la création de ces outils ont été installés.
 5. Une analyse de contenu d'outils pédagogiques et de fiches d'entretiens a été produite.
- *Des actions d'exploitation des résultats ont été menées (ligne 6)*
 6. Elles ont suivi les règles prévues : évaluations des outils adressées aux auteurs et entretiens avec ces derniers dans une perspective correctrice et formative.

Après corrections, les outils pédagogiques évalués ont été présentés en forum public avant leur diffusion dans les réseaux de formation auxquels ils étaient destinés.

3.2 Mise en tableau

**Tableau 23. Évaluation des supports pédagogiques d'une formation de formateurs
Phase 3**

Instrumentation	
Quels outils employer pour renseigner les indicateurs ?	
1 – Quels documents descriptifs et prescriptifs choisis ?	<i>Cahier des charges spécifique pour la création d'outils pédagogiques</i>
2 – Quelles populations interrogées ?	<i>Formateurs, didacticiens</i>
3 – Comment ont été sélectionnés les activités et les acteurs concernés ?	<i>Échantillon d'outils et panel de formateurs proposés par les concepteurs du dispositif de formation de formateurs</i>
4 – Quels instruments construits ?	<i>Grille d'analyse des outils pédagogiques Focus-groupes sur la création d'outils</i>
5 – Comment ont été traitées les informations ?	<i>Analyse de contenu de documents et d'outils entretiens</i>
6 – Comment les résultats ont-ils été exploités et mis en valeur ?	<i>Jurys ayant délibéré à partir des résultats de l'analyse des outils ; modifications apportées aux outils Forum public de présentation des outils accompagnés de leur évaluation</i>

3.2.1 Aperçu d'un outil d'évaluation des manuels pédagogiques (extrait)

On a, ici, le résultat d'un travail de référentialisation qui prend la forme d'un outil d'évaluation de manuels pédagogiques. On a gardé, pour l'exemple, un élément à évaluer : la structure, comportant, dans cet extrait, les éléments de la structure et l'organisation des informations. Ces deux éléments sont proposés à l'évaluation au moyen de critères (complétude du produit, etc.) déclinés, chacun en indicateurs dont les items numérotés appellent une réponse par oui ou par non assortie, éventuellement, d'un commentaire utilisable au moment de l'interprétation.

Tableau 24. Exemple d'un élément à évaluer du point de la structure et l'organisation des informations du manuel.

Instrument de Contrôle et de Validation de supports pédagogiques	
Forme du support : Écrit	Type : Manuel
I – STRUCTURE	

I-1 – ELEMENTS DE LA STRUCTURE

Critère : Complétude du produit

Indicateurs : présence des éléments suivants :	OUI	NON	Commentaires
1. – Titre			
2. – Sommaire			
3. – Présentation des objectifs			
4. – Mode d'emploi			
5. – Apport de connaissances			
6. – Activités pédagogiques			
7. – Activités d'évaluation			
8. – Ressources internes			
9. – Ressources externes			
TOTAUX / 9 :	0	0	

I-2 – ORGANISATION DES INFORMATIONS

Critère : Lisibilité du plan formel

Indicateurs : caractéristiques du sommaire	OUI	NON	Commentaires
1. Découpage précis en parties et sous parties			
2. Fonctions différentes des parties et sous parties			
3. Présentation aérée du sommaire			
TOTAUX / 3 :	0	0	

Critère : Cohérence de l'organisation

Indicateurs : Caractéristiques des parties :	OUI	NON	Commentaires
1. Équilibre des parties en nombre de pages			
2. Identité de structure des parties			
TOTAUX / 2 :	0	0	

Critère : Séquenciation des contenus

Indicateurs : Présentation des contenus :	OUI	NON	Commentaires
1. Suggestion du parcours d'apprentissage			
2. Découpage en séquences			
TOTAUX / 2 :	0	0	

Il faut rappeler que ces outils étant destinés à être utilisés par des évaluateurs occasionnels et dispersés, il a fallu limiter, autant que faire se pouvait, les marges d'interprétation et de jugement. Seules restaient à déterminer par les instances locales d'évaluation, les seuils d'acceptabilité et les préconisations d'amélioration.

3.3 Les apports de cette phase au processus d'évaluation

Le tableau précédent situe la construction des instruments dans le continuum méthodologique institué par la référentialisation mais le résultat de ce travail n'est concrètement visible que dans les grilles d'analyse des outils qui ont été produites et dont il vient d'être donné un aperçu, ici, dans un extrait de l'instrument d'évaluation des manuels pédagogiques. On peut remarquer la recherche de marges d'interprétation subjective minimales des réponses attendues par la formulation des indicateurs.

4 PHASE 4 : INTERPRÉTATION

4.1 En guise d'analyse du cas

Cette phase revêt, dans cette situation d'évaluation d'outils pédagogiques, un statut et des modalités très spécifiques. Il s'agit moins d'interpréter les résultats produits par l'ensemble de l'opération que de clore une phase d'évaluation dans l'ensemble du modèle de vie d'un outillage pédagogique. Ce modèle de vie d'un outil pédagogique se trouve clairement représenté par

Gerard et Rogiers (1993) qui le schématisent sous la forme : « conception-édition-évaluation-utilisation » (p. 7 ; c'est nous qui soulignons). L'entreprise consiste bien, en effet, à suivre l'intégration du processus d'évaluation ainsi construit dans la vie des outils, leur utilisation par les formateurs et par les apprenants. L'évaluation d'un outil pédagogique telle qu'elle a été présentée dans le cadre des exemples d'application d'une méthodologie de référentialisation constitue, en effet, comme l'indiquent les auteurs précités, une phase de mise à l'épreuve de l'outil avant son utilisation en situation de formation. Ici, ce dernier a été soumis à une analyse semblable à celle appliquée aux autres cas éducatifs ou de formation, ce qui invite à une interprétation plus formative que critique. L'outil pédagogique est évalué comme une situation pédagogique mettant en scène un formateur et un apprenant, une procédure d'information ou d'apprentissage, des ressources et des règles autour d'un objectif partagé de formation.

4.2 Mise en tableau

**Tableau 25. Évaluation des supports pédagogiques
d'une formation de formateurs
Phase 4**

Édition	Évaluation	Interprétation
– Règles, normes fixées par l'institution (= référents de l'évaluation)	– Utilisation d'indicateurs précis et univoques.	– Existence de caractéristiques essentielles de l'outil – Co-présence de scores positifs indiquant la qualité de l'outil – Remarques éventuelles à interpréter par les auteurs et destinées aux corrections ou aux améliorations

4.3 Les apports de cette phase au processus d'évaluation

Cet exemple voulait montrer, en plus de l'application d'une méthodologie d'évaluation à des outils ou supports pédagogiques, comment la référentialisation pouvait modifier profondément le statut d'un instrument d'évaluation « à distance », selon qu'il est produit (par des experts) en dehors de l'espace d'utilisation et imposé à des utilisateurs passifs ou selon qu'il est produit à l'intérieur de cet espace par l'intermédiaire de la co-construction d'un référentiel validé par tous les acteurs concernés.

Chapitre 12

Cinquième application : Évaluation de compétences

*Un exemple : l'évaluation des acquis
de l'expérience en jury de VAE*

INTRODUCTION : LES COMPÉTENCES COMME OBJETS À ÉVALUER

Un objet nouveau pour l'évaluation

Les compétences semblent constituer, depuis les années 2000, un nouvel objet à évaluer dans les mondes de l'éducation et de la formation (De Ketele, 2001). On a en effet assisté à l'émergence de l'Approche Par Compétences (au niveau international), des dispositifs de Validation des Acquis de l'Expérience en France (VAE, depuis la loi de 2002). Ainsi, l'évaluation des compétences constitue une problématique nouvelle dans le champ de l'évaluation, la notion de compétence étant d'introduction plutôt récente dans le champ des sciences de l'éducation.

Ce nouvel objet soulève toutefois de nombreuses questions : les compétences sont pour partie invisibles, composites, complexes et sujettes à des définitions variables. Malgré la prégnance – controversée – des classiques découpages en « savoirs/savoir-faire/savoir-être », on observe l'émergence de définitions, inspirées des recherches en psychologie du travail, qui

lient compétences et situations de travail (ou de vie). Être compétent est ainsi de plus en plus souvent défini comme la capacité à traiter une classe de situations professionnelles quel que soit son état (normal/dégradé) et son degré de complexité (Vergnaud, 2001 ; Mayen, Métral et Tourmen, 2010) – autrement dit, être compétent, c'est « résoudre des situations complexes » (Gérard, 2005).

Différents cas de figure

Les professionnels de l'enseignement et de la formation sont appelés à évaluer des compétences pour différents objectifs : on pourra ainsi distinguer des évaluations de compétences à visée certificative (type VAE, épreuves pratiques, note à l'issue d'un stage...), d'autres à visée formative (évaluations en cours de formation, bilans de compétences...), pour reprendre la célèbre définition de Scriven (1967). Les professionnels de l'éducation et de la formation évaluent aussi des compétences de différentes manières :

- seuls, ou en jury – souvent avec des professionnels, comme dans le cas de la VAE ou des suivis de stage,
- de manière directe : sur base d'observations de l'activité d'une personne, comme dans des épreuves pratiques en situation de travail réelle ou simulée, ou de manière indirecte : sur base d'un discours ou de traces de l'activité d'une personne, comme en VAE, en retour de stage, en bilan de compétences...
- à différents moments : comme le constatent Tardif et Dubois (2010), l'évaluation des compétences en formation peut être faite selon une « logique photographique » (estimation d'un niveau à un moment M) ou selon une « logique vidéographique », en prenant en compte une progression sur un temps long (par le biais de portfolios et carnets de compétences, par exemple).

Plusieurs recherches – basées sur des analyses du travail en didactique professionnelle – ont permis d'avancer sur la connaissance de l'activité des évaluateurs quand ils évaluent des compétences, de manière directe ou indirecte (Mayen, Mayeux et Savoyant, 2006 ; Tourmen, 2009, 2012). Le protocole de référentialisation peut s'avérer pertinent pour étudier la méthodologie sous-jacente et, comme nous allons le voir, souvent implicite dans l'activité évaluative qui s'intéresse aux compétences, telle que celle d'un jury de VAE. Le cas présenté ici est celui d'un jury de VAE (Validation des Acquis de l'Expérience) de niveau de BTS, regroupant des enseignants de différentes disciplines et un professionnel du secteur concerné. Son déroulement a été enregistré et intégralement retranscrit¹, puis analysé avec un cadre de didactique professionnelle. L'analyse des interactions a permis de mettre à jour les critères et indicateurs utilisés par le jury pour estimer le

1. Données fournies par le Cereq.

niveau de compétence acquis par les candidats dans leur expérience professionnelle (Mayen, Tourmen, Mayeux et Chausse, 2008 ; Tourmen, Mayen et Figari, 2009), mais aussi le protocole de référentialisation implicite et sous-jacent qui était à l'œuvre, comme nous allons le montrer à présent.

MÉTHODOLOGIE REPÉRÉE

Du protocole méthodologique utilisé dans cet exemple, on retiendra les éléments suivants :

1 PHASE 1 : EXPLORATION (MANDAT ET PROBLÉMATIQUE)

On a vérifié qu'elle s'est déroulée de la manière suivante, consultable sous deux versions successives (par une analyse cursive du cas puis par une version synoptique sous forme de tableau).

1.1 Analyse du cas

1.1.1 Le mandat

Bien que nous n'ayons pas eu accès à l'observation de l'activité qui a précédé le jury, nous pouvons faire l'hypothèse que tous ses membres connaissaient, à des degrés divers et au moins à minima, l'existence du cadre législatif et réglementaire de la Validation des Acquis d'Expérience². La raison d'être de leur activité de jury, ce qu'il leur est possible et impossible de faire est en effet fixé par ce cadre qui propose de certifier des compétences aux vues d'une expérience professionnelle. Comme il est stipulé dans la loi, « toute personne engagée dans la vie active est en droit de faire valider les acquis de son expérience, notamment professionnelle, en vue de l'acquisition d'un diplôme, d'un titre à finalité professionnelle ou d'un certificat de qualification figurant sur une liste établie [...] ». L'objet de l'évaluation est précisé : « Peuvent être prises en compte, au titre de la validation, l'ensemble des compétences professionnelles acquises dans l'exercice d'une activité salariée, non salariée ou bénévole, en rapport direct avec le contenu du diplôme ou du titre. » Ses modalités sont également prescrites. Le jury peut attribuer la totalité du diplôme ou du titre. À défaut, il se prononce sur l'étendue de la validation et, en cas de validation partielle, sur la nature des connaissances et aptitudes devant faire l'objet d'un contrôle complémentaire. Le jury se prononce au vu d'un dossier constitué par le candidat, à l'issue d'un entretien. Les membres du jury de VAE dont il est ici question

2. Il s'agit, en France, de la loi de Modernisation Sociale (n° 2002-73, articles 133 et 134, du 17 janvier 2002).

se sont donc insérés dans ce cadre réglementaire qui leur prescrivait une finalité, des conditions de validité, des objets à évaluer (l'expérience globale d'un candidat) et des modalités d'évaluation (dossier, entretien et/ou mise en situation). Le jury a procédé à la lecture des dossiers remis par les candidats plusieurs jours auparavant et à des entretiens de 30 à 45 minutes avec eux. Dans ce chapitre, nous ne traiterons que de l'exemple du traitement d'une candidate³.

Par rapport aux effets éventuels de contexte, les membres du jury se sont intéressés à l'origine de la candidate à la VAE, à savoir sa formation initiale, son milieu d'origine (rural ou non) et ses premières activités professionnelles, ainsi qu'à l'origine de sa demande de VAE à cette étape de son parcours.

En ce qui concerne les stratégies, les membres du jury ont aussi souhaité mieux connaître le déroulement de la carrière professionnelle de la candidate et examiner l'ensemble des activités réalisées dans son parcours personnel et professionnel, en lien avec le diplôme demandé par la candidate.

En matière de productions visées, les membres du jury ont cherché à estimer si la candidate avait développé, dans son parcours, des compétences, aptitudes et capacités utiles à un bon exercice du métier.

1.1.2 La problématique

L'évaluation devait donc permettre d'estimer le niveau de compétences atteint par une candidate après plus d'une dizaine d'années d'expérience dans la gestion d'une exploitation agricole, avec son mari. Autrement dit, d'évaluer un niveau de compétences à partir du seul récit d'une expérience passée.

1.2 La mise en tableau

L'analyse précédente peut gagner à être présentée sous forme synoptique, dans un tableau récapitulatif de la procédure ayant guidé la conduite de la phase 1 :

3. Pour une présentation plus systématique du traitement de ce cas, se référer à Mayen, Tourmen, Mayeux et Chausse (2008) ou Tourmen (2009).

**Tableau 26. Évaluation des compétences
Phase 1**

Exploration (mandat et problématique) (exemples)	
Analyse du mandat	Problématique d'évaluation
– Par rapport au contexte <i>Origine de la candidate (formation, milieu, premières expériences) et de la demande de VAE</i> – Par rapport aux stratégies <i>Activités menées dans son parcours et diplôme souhaité par la candidate</i> – Par rapport aux productions <i>Compétences développées dans son parcours au regard de celles attendues par le diplôme</i>	À quelle(s) question(s) fondamentale(s) l'évaluation doit-elle répondre ? – <i>Quel est le niveau de compétences de la candidate à ce jour ?</i> – <i>Son expérience a-t-elle permis de les acquérir/maintenir à un niveau suffisant au regard du diplôme délivré ?</i>

1.3 Les apports de cette phase au processus d'évaluation

Si le mandat évaluatif n'est pas systématiquement re-questionné ni ré-examiné en VAE, car considéré comme connu des membres du jury voire évident (même si parfois méconnu), les membres d'un jury ré-examinent l'origine et la nature de la demande de VAE de chaque candidat, au vu de son parcours spécifique.

2 PHASE 2 : PROCÉDURE CENTRALE DE RÉFÉRENTIALISATION

Les membres du jury ont référentialisé l'expérience spécifique de la candidate au regard de leur propre cadre de référence, pour partie constitué des critères et indicateurs officiels de compétences contenus dans le référentiel de diplôme, pour partie fondé sur leur propre expérience et connaissance du métier concerné. Ils se sont ainsi intéressés à trois dimensions de l'expérience de la candidate :

1. les situations vécues (c'est-à-dire le I du modèle) ;
2. les activités menées (c'est-à-dire le C du modèle) ;
3. les acquis développés (c'est-à-dire le P du modèle).

2.1 Analyse du cas

2.1.1 Définition des éléments à évaluer

Il s'agissait bien de repérer les *facettes de l'objet à observer* qui devraient constituer l'ossature de l'ensemble de la référentialisation. Parmi ces éléments, on trouve, par exemple :

- *En termes d'effets éventuels de contexte* : les situations vécues par la candidate : les formations reçues, les lieux de travail successifs et leurs caractéristiques (degré d'autonomie, nature des missions confiées, niveau de poste...), la variété des situations rencontrées...
- *En ce qui concerne les stratégies* : les activités menées par la candidate dans ces situations, les périmètres d'actions effectivement réalisées, leur richesse et leur complexité, leur réussite...
- *En matière de productions espérées* : les acquis développés par la candidate au cours de ces activités, à savoir les connaissances développées et/ou intégrées dont il a été fait la preuve, les modes de raisonnement utilisés et considérés comme maîtrisés, les capacités à prendre sa propre activité comme objet de pensée (ou capacité de recul)...

2.1.2 Choix des référents

Pour évaluer les effets éventuels de contexte, les évaluateurs se sont référés à ce qu'ils connaissent de l'origine habituelle des élèves candidats à ce diplôme (milieu urbain ou rural, par exemple) et de l'origine habituelle des professionnels, ainsi qu'aux filières de formation et situations les plus courantes ou recommandées dans ce métier.

Pour évaluer les effets stratégiques, les évaluateurs se sont référés aux champs d'activité qu'ils connaissent et qui sont ceux attendus d'un professionnel du secteur, ainsi qu'aux résultats habituellement atteignables.

Pour évaluer les effets attendus, les évaluateurs se sont enfin intéressés à ce que la candidate a tiré de son expérience, des savoir-faire, des connaissances et manières de raisonner (par exemple, un raisonnement économique et comptable qui est évalué en cours de jury) dont ils ont tenté d'estimer la validité, la complétude et la pertinence.

2.1.3 Élaboration des critères

Au niveau des effets de contexte : les évaluateurs ont cherché à évaluer selon les critères suivants : 1) adéquation des situations rencontrées par rapport au référentiel de certification, 2) complexité des situations rencontrées au regard du référentiel, 3) adaptabilité du parcours.

Au niveau des effets stratégiques : les évaluateurs ont cherché à évaluer selon les critères suivants : 1) richesse des activités menées par rapport au référentiel, 2) réussite des activités et 3) complexité des activités menées.

Au niveau des productions attendues : les évaluateurs ont cherché à évaluer selon les critères suivants : 1) capacité à mobiliser des connaissances du niveau du diplôme, 2) capacité de recul sur l'activité, 3) étendue des acquis au regard du diplôme et des fonctions visées.

2.1.4 Construction des indicateurs

Pour estimer l'atteinte des critères énoncés ci-dessus, les membres du jury se sont appuyés sur une partie des indicateurs suivants, dont nous avons fait la synthèse à l'issue de l'observation de plusieurs jurys (Mayen, Tourmen, Mayeux et Chausse, 2008). A nouveau, ces éléments n'ont pas été énoncés de manière exhaustive au préalable, mais ils sont apparus au fur et à mesure des débats du jury :

- *En matière d'effets de contexte* : les évaluateurs prélèvent des informations sur le type de situations vécues (secteur, domaine d'activité, type d'entreprise, niveau de poste, fonctions et missions confiées, niveau de responsabilité et d'autonomie...), les caractéristiques du poste de travail (objets transformés, techniques manipulées...), la variété des situations rencontrées, les degrés de performance attendus, la nature des problèmes à traiter et la disponibilité des solutions, enfin la diversité des expériences vécues.
- *En matière d'effets stratégiques* : les évaluateurs prélèvent des informations sur l'étendue des activités réalisées, leur degré d'autonomie et part dans le collectif, les résultats et effets produits, leur reconnaissance par des employeurs, pairs ou clients (l'absence d'échec manifeste faisant souvent office d'indicateur indirect de réussite dans les jurys de VAE), la nature des activités menées (exécution, diagnostic, outils mobilisés...) et leur complexité (raisonnements de diagnostic, d'anticipation ou d'analyse impliqués).
- *En matière de productions espérées* : les évaluateurs prélèvent des informations sur les acquis de la candidate, à savoir le niveau de connaissance des objets du travail, la justesse des raisonnements techniques et l'usage de concepts scientifiques et termes techniques, le recours à des ressources (livres...) pour agir, la clarté de la description des activités dans le dossier et à l'oral, la capacité à en dépasser la simple description et à en analyser les conditions et les résultats, les leçons tirées et les prises de recul, des formes de généralisation et de prise de conscience, enfin le nombre et l'importance des acquis au regard du diplôme visé.

2.2 La mise en tableau

Comme pour la phase 1, on présente ici, sous forme synoptique, la procédure ayant guidé la conduite de la phase 2 :

**Tableau 27. Évaluation des compétences
Phase 2**

Procédure centrale de la référentialisation (exemples)			
Éléments à évaluer	Référents	Critères	Indicateurs
Quelles dimensions de l'objet à observer ?	À partir de quoi attribuer un sens à l'information ?	Quels aspects représentatifs de l'objet à évaluer choisir pour porter un jugement ?	Quelles caractéristiques observables permettent la vérification du degré d'atteinte du critère ?
<i>Situations vécues</i>	<i>Parcours habituels ou souhaités</i>	<i>Adéquation des situations vécues / référentiel, complexité, variété et adaptabilité</i>	– Type de situations vécues – Caractéristiques du poste de travail – Variété des situations rencontrées – Degrés de performance attendus, nature des problèmes à traiter et disponibilité des solutions, enfin diversité des expériences vécues...
<i>Activités menées</i>	<i>Activités habituelles ou souhaitées</i>	<i>Richesse, réussite, complexité des activités menées</i>	– Étendue des activités réalisées, – Degré d'autonomie – Part dans le collectif, – Reconnaissance par des employeurs, pairs ou clients – Nature des activités menées (exécution, diagnostic, outils mobilisés...) – Complexité (raisonnements de diagnostic, d'anticipation ou d'analyse impliqués)...

<i>Acquis de l'expérience</i>	<i>Acquis nécessaires pour valider le diplôme</i>	<i>Capacité à mobiliser des connaissances du niveau du diplôme, recul sur l'activité, étendue des acquis</i>	– Niveau de connaissance des objets du travail – Justesse des raisonnements techniques et usage de concepts scientifiques et termes techniques – Recours à des ressources – Clarté de la description des activités dans le dossier et à l'oral – Capacité d'analyse des résultats, de prise de recul, de généralisation – Nombre et importance des acquis au regard du diplôme visé...
-------------------------------	---	--	---

2.3 Les apports de cette phase au processus d'évaluation

Il n'a pas été réalisé d'enquête spécifique sur les référents – si ce n'est les compétences attendues dans le référentiel de diplôme qui ont été rappelées à plusieurs reprises en cours de jury –, ceux-ci étant censés être connus et partagés par les membres du jury, même s'ils peuvent susciter des débats. Les critères et indicateurs de compétences n'ont pas non plus été pré-définis ni pré-discutés avant le jury, si ce n'est la liste de compétences du référentiel de certification qui a été rappelée. La plupart des critères et indicateurs utilisés pour évaluer les acquis de l'expérience de la candidate ont ainsi émergé au fur et à mesure des débats du jury, sans ordre prédéfini, et c'est l'observation des interactions qui en a permis la synthèse présentée ici. Nous sommes donc dans un cas particulier d'évaluation où le référentiel agissant se révèle chemin faisant, en cours d'activité, et se discute au même moment. Nous faisons l'hypothèse qu'il est tellement incorporé à l'expérience professionnelle des membres du jury – et donc seulement partiellement conscient – que son identification/discussion avant jury serait certainement discutable, si ce n'est à disposer d'une liste basée sur l'analyse de jurys précédents.

3 PHASE 3 : INSTRUMENTATION

3.1 En guise d'analyse du cas

Dans cette opération d'évaluation, on repère trois types de procédures qui se sont succédé en matière d'instrumentation :

- *Peu de choix préalables ont été effectués, mais beaucoup se sont imposés (lignes 1, 2, 3) :*
 1. le cadre réglementaire de la VAE s'est imposé, ainsi que le référentiel du diplôme concerné ;
 2. la population à interroger n'a pas été choisie par le jury mais lui a été envoyée par l'institution certificatrice ;
 3. les modalités d'évaluation ont aussi été imposées, même si le cadre a pu être rediscuté et ajusté au moment du jury.
- *Le jury a eu recours à des instruments classiques en éducation qui n'ont pas nécessité d'élaboration spécifique au préalable (lignes 4, 5) :*
 4. le jury a eu recours à des instruments de recueil ou d'analyse tels qu'utilisés en sciences sociales (lecture des dossiers, débat contradictoire, entretien collectif, prise de notes), ainsi qu'à un référentiel de certification ;
 5. les informations ont été traitées en cours de jury par l'ensemble des membres et discutées oralement lors des délibérations ;
- *Des actions d'exploitation des résultats ont été menées (ligne 6) :*
 6. les résultats du jury ont été notés dans un procès-verbal qui a stipulé le résultat de la VAE (validé totalement, partiellement, refusé), des recommandations et des commentaires de la part du jury.

3.2 Mise en tableau

**Tableau 28. Évaluation des compétences
Phase 3**

Instrumentation (exemples)	
Quels outils employer pour renseigner les indicateurs ?	
1 – Quels documents descriptifs et prescriptifs choisis ?	<i>Cadre réglementaire de la VAE Référentiel de diplôme</i>
2 – Quelles populations interrogées ?	<i>Candidats imposés par l'institution</i>
3 – Comment ont été sélectionnés les activités et les acteurs concernés ?	<i>Imposés</i>
4 – Quels instruments construits ?	<i>Aucun</i>

5 – Comment ont été traitées les informations ?	<i>En cours de jury par un débat contradictoire</i>
6 – Comment les résultats ont-ils été exploités et mis en valeur ?	<i>Dans un procès-verbal</i>

3.3 Les apports de cette phase au processus d'évaluation

Cette phase n'a pas donné lieu à des innovations ni à des discussions spécifiques car elle est très normée dans le cadre d'une VAE.

4 PHASE 4 : INTERPRÉTATION

4.1 En guise d'analyse de cas

Les membres du jury ont analysé, en cours de jury, des données prélevées sur le parcours de la candidate et son expérience, même s'ils étaient arrivés avec une première conception construite à la lecture préalable du dossier (Tourmen, 2009). L'interprétation de l'expérience et des acquis de la candidate a ainsi eu lieu chemin faisant, en temps réel, tout au long des débats du jury et de l'entretien avec la candidate, un avis commun s'étant construit et débattu avant d'être finalement stabilisé.

4.2 Mise en tableau

**Tableau 29. Évaluation des compétences
Phase 4**

Interprétation (exemples)		
Indicateurs	Instrumentation	Interprétation
Quelles caractéristiques observables permettent la vérification du degré d'atteinte du critère ?	<i>Quels outils employer pour renseigner les indicateurs ?</i>	<i>Quelles significations attribuer aux résultats ?</i>
Éléments contextuels : situations vécues par la candidate	<i>Lecture du dossier Entretien avec la candidate Comparaison avec le référentiel de certification</i>	<i>Compétence validée totalement/partiellement/non acquise</i>
Éléments stratégiques : activités menées par la candidate	<i>Lecture du dossier Entretien avec la candidate Comparaison avec le référentiel de certification</i>	<i>Compétence validée totalement/partiellement/non acquise</i>

Interprétation (exemples)		
Indicateurs	Instrumentation	Interprétation
Productions : <i>acquis de la candidate</i>	<i>Lecture du dossier Entretien avec la candidate Comparaison avec le référentiel de certification</i>	<i>Compétence validée totalement/partiellement/ non acquise</i>

4.3 Les apports de cette phase au processus d'évaluation

Durant la délibération, l'observation de plusieurs jurys a montré que les membres n'examinent pas tous les critères avec le même degré de détail. « Ainsi, ils procèdent par « sondages » et se centrent sur les critères qui semblent poser problème chez un candidat ou sur ceux pour lesquels il manque des informations » comme l'ont constaté Mayen, Tourmen, Mayeux et Chausse (2008). De plus, « les discussions du jury se centrent sur l'examen des points qui posent problème, font débat ou posent question, les acquis manifestes étant plus rapidement validés ». Le jugement final est le fruit d'une pondération entre différents critères et les acquis sont pondérés au regard des situations rencontrées et des activités menées par un candidat. Toujours selon Mayen, Tourmen, Mayeux et Chausse (2008), « un candidat à qui il manque une compétence, par exemple en langue, mais qui a témoigné d'une expérience riche et de la maîtrise des autres compétences du diplôme peut se voir valider le diplôme » (p. 21). Enfin, le jugement final se construit tout au long du jury : les membres énoncent un premier diagnostic sur l'expérience du candidat dès le début du jury, sur base des lectures individuelles du dossier, diagnostic qui est explicité, affiné, testé, débattu puis stabilisé au moment de la délibération finale (Tourmen, 2009).

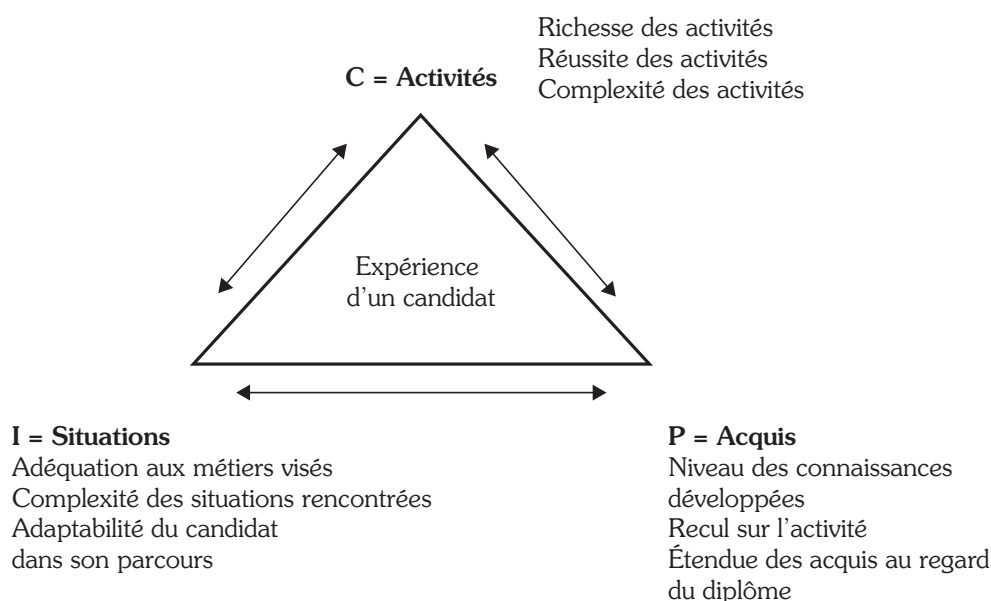
L'examen d'un jury de VAE à l'aide du protocole de référentialisation a donc permis de mettre à jour les caractéristiques suivantes. Tout d'abord, les évaluateurs prélèvent des informations sur l'objet à évaluer (les acquis de l'expérience d'un candidat, que l'on peut aussi appeler compétences) à travers la prise en compte de trois dimensions que le schéma ICP précédemment décrit permet de modéliser de manière pertinente et dynamique :

- I : les situations vécues par un candidat (autrement dit le contexte de son expérience) ;
- C : les activités menées dans ces situations (autrement dit les stratégies adoptées) ;
- P : enfin, les acquis retirés (ce que l'activité en situation a permis de produire/construire comme connaissances et capacités).

Nous avons ainsi reformulé le schéma ICP suivant comme modélisation possible de l'expérience telle qu'elle est évaluée en jury de VAE

(Mayen, Tourmen, Mayeux et Chausse, 2008 ; Tourmen, Mayen et Figari, 2009). Le modèle systémique restitue la dynamique de l'expérience, envisagée comme un processus ayant un contexte et des déterminants (I), un déroulement (C) et des effets (P).

Figure 8. Application du modèle ICP à l'évaluation des acquis de l'expérience



L'examen du jury avec la grille de la référentialisation a aussi montré que si le cadre réglementaire est davantage contraint que dans d'autres formes d'évaluation – où le choix des méthodes de collecte de données seront plus ouverts, comme en évaluation de programmes –, les membres du jury peuvent toutefois creuser tel ou tel point lors de leurs questions aux candidats (en cas d'entretien), revenir à la lecture de parties du dossier et ainsi procéder à un recueil de données complémentaire. Ils prélèvent ainsi un faisceau d'indices qui sont ensuite croisés, mis en relation et pondérés pour aboutir à un jugement final à partir d'une somme de jugements provisoires et partiels émis tout au long de l'activité collective. On peut donc dire que le *référentiel d'évaluation implicite* du jury se révèle, se construit et se débat en cours de jury, de manière dynamique dans le cours de l'activité. L'observateur du jury (par exemple, le chercheur) peut utilement utiliser le modèle de la référentialisation pour reconstruire après coup le référentiel en actes du jury (questions posées, référents mobilisés, critères et indicateurs prélevés et discutés, instruments utilisés, interprétations réalisées), dans une sorte de « référentialisation à l'envers » (Tourmen et Droyer, 2011). Le modèle de la référentialisation devient alors un modèle d'analyse de l'activité évaluative. Il permet d'en reconstruire et d'en comprendre la logique sous-jacente et souvent non explicite pour les acteurs eux-mêmes.

Conclusion : référentialisation et exemplification

On a examiné cinq situations d'évaluation différentes qui présentaient la caractéristique commune d'avoir été traitées selon une méthodologie se réclamant de la référentialisation. On a pu voir, ainsi, comment cette dernière s'adapte aux différences de publics, de problématiques et d'enjeux et, en même temps, comment elle oriente de manière homogénéisante les choix successifs d'une chaîne de procédures constitutive de toute évaluation.

Au fil des exemples analysés, cette chaîne est apparue comme immuable : mandat – problématique – définition des éléments à évaluer – repérage des référents – élaboration des critères et des indicateurs – instrumentation – interprétation.

L'épreuve de l'exemplification a fait son œuvre, celle de pouvoir conclure, après expérience, que ce qui paraît différent sur le terrain s'apparente cependant, si on veut bien opérer la conceptualisation nécessaire, à un schème unique apte à rendre compte d'une démarche – toujours complexe – d'évaluation.

Conclusion

Selon le mode de lecture adopté pour prendre connaissance de cet ouvrage sur la méthodologie de l'évaluation en éducation (cf. les « chemins » proposés en introduction), on aura pu nourrir, à son propos, des impressions différentes.

On aura pu découvrir, par exemple, une méthodologie de l'évaluation traditionnelle du type « question-exploration-instrumentation-résultat » ; ou bien une autre façon de dérouler un enchaînement de procédures mettant en relation les différents facteurs de résultats constituant les formes données à voir à l'issue d'une évaluation. Ces deux façons d'appréhender le protocole proposé pourraient aussi apparaître dans leur combinaison et leur complémentarité au sein d'une opération utilisant la *référentialisation*.

Quant aux auteurs eux-mêmes, peut-être souhaiteraient-ils que soit principalement retenu de ce travail trois idées directrices : une *méthodologie générale de l'évaluation*, une *posture d'évaluateur réflexif*, et, d'une manière plus large, une *enquête* à la recherche du sens d'une situation problématique soumise à évaluation.

1 UNE MÉTHODOLOGIE « GÉNÉRALE » D'ÉVALUATION

Il s'agit, en effet, d'une méthodologie de conduite d'une opération d'évaluation accessible à des acteurs éducatifs. Cette approche les invite à inclure leur opération dans une vision d'ensemble, dans un modèle holistique leur permettant de clarifier les fondements du processus qu'ils déclenchent ou auquel ils participent. Cette méthodologie a permis de construire une forme de tableau de bord consistant à rappeler d'où l'on part (contexte, mandat ou demande, situation problématique) et où l'on va (type de résultats, mode d'utilisation) en passant par la construction d'un questionnement et/ou d'une observation particulièrement adapté(e), le tout organisé au moyen d'une modélisation de l'évaluation (distinction et comparaison des données relatives au contexte, aux stratégies et aux productions : le modèle que nous avons schématisé autour du sigle ICP : induit, construit, produit).

Les auteurs ont eux-mêmes expérimenté la possibilité d'entrer dans le protocole de manière alternative, à des stades différents du modèle :

- soit en commençant par observer la façon dont les acteurs ont traité les questions issues du décryptage d'un mandat inséré dans un contexte donné (le I du modèle), en continuant par repérer les stratégies mises en place (le C du modèle) en évaluant, au passage, la pertinence de ces dernières, pour en arriver à comprendre et expliquer les résultats obtenus par l'action éducative (le P du modèle) ;
- soit en commençant par étudier le projet et les stratégies en cours (le C du modèle) pour remonter vers le contexte, les attentes et les moyens qui les ont influencés (le I du modèle), ou vers les résultats ou produits de ces stratégies (le P du modèle), en se limitant à vérifier leur efficacité ;
- soit en commençant par prendre acte des résultats d'une action (le P du modèle) pour en chercher les causes éventuelles soit dans le contexte et ses déterminismes (le I du modèle) soit dans les stratégies utilisées pour traiter le problème (le C du modèle).

Une telle méthodologie reste *générale* dans la mesure où elle se propose comme modulable, flexible, adaptable, négociable entre les acteurs participant à l'évaluation.

2 UNE POSTURE D'ÉVALUATEUR RÉFLEXIF

2.1 Une posture

En empruntant d'abord à Jorro (2000) le sens donné au terme de *posture*, on introduit l'ensemble des profils, des attitudes, des gestes que peuvent prendre non seulement les enseignants, mais aussi, par extension, l'ensemble des évaluateurs. Pour l'auteur, les gestes de l'évaluation reflètent, en effet, les interactions entre enseignants et apprenants (ou entre décideurs et acteurs de terrain) dans leurs transactions successives. Pour nous, les différentes phases que nous avons distinguées se traduisent par autant de postures :

- l'exploration du mandat (ou demande) pour en définir la problématique et les éléments qui seront donnés à évaluer se traduit par des modalités contractualisées de dialogue avec les décideurs ou les commanditaires ;
- le repérage des référents des différents partenaires ainsi que la construction d'un appareillage de critères et d'indicateurs constituent un parti pris systématique annoncé et négocié avant le début de l'opération d'évaluation ;
- l'instrumentation et le recueil d'informations en découlant sont présentés comme postures méthodologiques induites par la référentialisation ;
- le retour à la confrontation avec les acteurs pour en formuler collectivement les interprétations possibles fait partie des gestes inhérents à la conduite de toute évaluation.

2.2 Un évaluateur réflexif

En adoptant l'approche méthodologique présentée dans cet ouvrage, on joue, effectivement, sur des attitudes particulières supposées être développées par les acteurs, attitudes reliables par la notion de *réflexivité*. Cette notion a été largement traitée en sciences de l'éducation dans le domaine spécifique de la formation des enseignants, suite aux travaux de Schön (1993), déjà nommés, sous la forme de publications très nombreuses qui ne nous concernent, ici, que parce qu'ils rejoignent le concept d'*expérience réflexive* de Dewey, liée à la notion d'*enquête* que nous lui avons empruntée. Comme le remarquent Saussez *et al.* (2001), le « praticien qui réfléchit sur sa pratique... conduit l'élaboration conceptuelle de l'expérience... en progressant dans l'abstraction et la formalisation ». Par ailleurs, nous avons déjà rappelé que Dewey (1993) définit *l'enquête* comme « la transformation contrôlée d'une situation indéterminée... en situation déterminée... et unifiée » (p. 165) : c'est à cette *réflexion* constituée par une transformation, via la conceptualisation et la modélisation qui ont été engagées, que travaille l'évaluateur tout au long de la *référentialisation* dont il a déroulé le protocole.

3 UNE EN-QUÊTE À LA RECHERCHE DU SENS D'UNE SITUATION PROBLÉMATIQUE

3.1 La recherche de résolution d'une énigme

Une évaluation, quelle qu'elle soit, est toujours la résolution d'une énigme. En effet, il s'agit, par exemple, de chercher à savoir si des objectifs sont atteints, si des attentes (et lesquelles) ont été comblées, si des situations difficiles peuvent être expliquées, si des apprentissages ont été réussis ou pourquoi ils ne l'ont été que partiellement, etc. Dès qu'on analyse la situation sur laquelle l'évaluation va porter, on s'aperçoit toujours qu'elle est multiforme, complexe, ambiguë.

Devant ce problème, l'évaluateur est toujours amené à choisir entre plusieurs démarches comme, par exemple, réduire la question à un élément de problème connu et adopter alors une instrumentation existante, utiliser son expérience et adopter une méthodologie déjà éprouvée dans des situations jugées identiques, invoquer les ressources offertes par les sciences sociales en matière de questionnement, d'observation ou d'analyse et les appliquer à la population concernée par l'évaluation. L'expérience nous apprend que ces différentes postures sont possibles, souvent employées, parfois même combinées dans ce que d'aucuns appelleraient du « bricolage » permettant de produire des résultats plus ou moins attendus, plus ou moins appréciés, plus ou moins exploités. La doxa de l'évaluation est ainsi apparemment respectée et chaque acteur de l'évaluation (l'évaluateur comme

l'évalué) quitte le champ des opérations pour reprendre ses activités habituelles et... recommencer.

Mais, dans son approche à la fois philosophique et éducative, Dewey nous a invités à aller plus loin dans *une quête de la vérité et de la connaissance* ainsi que dans une *quête de la valeur* : ce sont ces deux orientations qui ont guidé l'organisation de notre réflexion sur l'*en-quête* (plutôt que l'enquête) évaluative.

3.2 Une quête de la vérité et de la connaissance

Pour Deledalle (1993), Dewey conçoit la vérité comme une construction de l'enquête : il s'agit de la situation rétablie après une reconstruction du réel. L'auteur lui préfère l'expression d'« assertibilité garantie », constituée d'assertions contrôlées *expérimentalement* (au sens large de « éprouvé », « mis à l'épreuve ») et *expérieniellement* (c'est-à-dire en continuité avec l'expérience de la situation). Selon Dewey (1993), « la connaissance est la fin de l'enquête ». En effet, vérité et connaissance se rejoignent dans la mesure où cette dernière est toujours vraie car elle est « la relation des moyens (méthodes) employés et des conclusions atteintes à titre de conséquences de ces moyens » (p. 64). En suivant Dewey, on peut postuler que, dans sa dimension scientifique, l'évaluation est en quête de connaissance (on cherche à connaître des relations de cause à effet) et que, dans sa construction du jugement (à l'instar de la procédure judiciaire), elle est à la recherche de la vérité.

3.3 Une quête de la valeur

C'est effectivement l'ensemble du processus évaluatif qui est ici en cause. La *valuation* ou *attribution de valeur*, la formulation d'un *jugement de valeur* constituent la forme finale prise par l'évaluation, comme cela a été noté en première partie. Nous disons bien *noté*, c'est-à-dire remarqué et référé : il restait à tenter d'introduire cette modalité dans un protocole méthodologique.

En utilisant la *référentialisation*, l'évaluateur a été invité à chercher, d'abord, puis à formuler, en second lieu, les fondements de ses choix ainsi que de ses interprétations et de ses exploitations finales. C'est-à-dire qu'on lui a demandé de tenir deux postures fondatrices de valeur :

- celle de dire en fonction de quoi le jugement pouvait être proféré : en fonction de quel(s) référent, de quel(s) critère(s) et de quel(s) indicateur(s), serait-il procédé à l'annonce d'un résultat ?
- celle de procéder à une confrontation des valeurs portées par les acteurs de l'évaluation : évaluateurs, évalués, partenaires. Le jugement devenant alors un « jugement de valeurs ».

L'enquête évaluative par la référentialisation constitue une entreprise ne comportant ni début ni fin : elle est un moment réflexif sur la formation, une expérience partagée de recherche de reconnaissance, de validation, de mise en lumière. Elle puise son raisonnement dans l'immensité des travaux antérieurs sur l'évaluation en éducation, elle ouvre des perspectives et appelle des débats et des compléments souhaités par les auteurs.

Bibliographie

- Abernot, Y. (1996). *Les méthodes d'évaluation scolaire*. Paris : Dunod.
- Aeres (2009). *Le référentiel qualité de l'AERES*. Consulté le 12 septembre 2013 dans <http://www.aeres-evaluation.fr/Publications/Methodologie-de-l-evaluation/Documents-generaux-de-reference>.
- Aeres (2012). *Le référentiel des critères d'évaluation des entités de recherche*. Consulté le 12 septembre 2013 dans <http://www.aeres-evaluation.fr>.
- Allal, L. (1991). *Vers une pratique de l'évaluation formative*. Bruxelles : De Boeck.
- Allal, L. (2007). Évaluation dans le contexte de l'apprentissage situé : peut-on concevoir l'évaluation comme un acte de participation à une communauté de pratiques. In M. Behrens (Éds), *La qualité en éducation : pour réfléchir à la formation de demain* (pp. 39-56). Québec : Presses de l'Université du Québec.
- Allal, L. (2008). Conceptualiser les outils d'évaluation des apprentissages. In G. Baillat, J.-M. De Ketele & L. Paquay (Éds), *Évaluer pour former* (pp 71-81). Bruxelles : De Boeck.
- Allal, L. & Mottier Lopez, L. (2007). *Régulation des apprentissages en situation scolaire et en formation*. Bruxelles : De Boeck.
- Altet, M. (1997). Professionnaliser les métiers d'enseignants. Savoirs professionnels et formation. *Cahiers du CREN*, 16-17.
- Ardoino, J. & Berger, G. (1989). D'une évaluation en miettes à une évaluation en actes. ANDSHA-Matrice.
- Bakhtine, M.-M. (1970). *La poétique de Dostoïevski*. Paris : Seuil.
- Barbier, J.-M. (2006). Quelques questions pour la recherche dans le domaine de l'évaluation. In G. Figari, G. & M. Achouche (Éds), *L'activité évaluative réinterrogée* (pp. 351-360). Bruxelles : De Boeck.
- Blais, J.-G. et al. (1999). Le palmarès des écoles en question. *Mesure et évaluation en éducation*, 22(1), 1-20.

- Bonboir, A. (1972). *La docimologie*. Paris : PUF.
- Bonniol, J.J. & Genthon, M. (1989). L'évaluation et ses critères : les critères de réalisation. *Repères pour l'enseignement du français*, 79, 107-115.
- Bonniol, J.-J. & Vial, M. (1997). *Les modèles de l'évaluation*. Bruxelles : De Boeck.
- Boudon, R. (1979). *La logique du social*. Paris : Hachette.
- Bourdieu, P. (1980). *Le sens pratique*. Paris : Minuit.
- Boutinet, J.-P. (1990). *Anthropologie du projet*. Paris : PUF.
- Boutinet, J.-P. (2004). *Psychologie des conduits à projet*. (4^e éd.) Paris : PUF.
- Cardinet, J. (1986). *Évaluation scolaire et pratique*. Bruxelles : De Boeck.
- Cardinet J. (1989) Évaluer sans juger. *Revue française de pédagogie*, 88(1), 41-52.
- Cardinet, J. (1990). L'évaluation externe, interne et négociée. In *Hommage à Jean Cardinet*. Genève : Del Val.
- Clot, Y. (1999). *La fonction psychologique du travail*. Paris : PUF.
- Conseil scientifique de l'évaluation (1996). *Petit guide de l'évaluation des politiques publiques*. Paris : La Documentation française.
- Crahay, M. (2006) Qualitatif-Quantitatif : des enjeux épistémologiques convergents ? In L. Paquay (Éds), *L'analyse qualitative en éducation* (pp. 33-52). Bruxelles : De Boeck.
- De Ketele, J.-M. (1986). *Méthodologie de l'observation*. Bruxelles : De Boeck.
- De Ketele, J.-M. (2001). Place de la notion de compétence dans l'évaluation des apprentissages. In G. Figari & M. Achouche (Éds), *L'activité évaluative réinterrogée. Regards scolaires et socioprofessionnels* (pp. 39-43). Bruxelles : De Boeck Université.
- Deledalle, G. (1993) La théorie de l'enquête et le problème de la vérité (1967) Présentation. In J. Dewey, *Logique. Une théorie de l'enquête*. Paris : PUF.
- De Lansheere, G. (1972). *Évaluation continue et examens. Précis de docimologie*. Bruxelles : Labor.
- De Landsheere, V. (1992). *L'éducation et la formation*. Paris : PUF.
- Del Rey, A. (2013) *La tyrannie de l'évaluation*. Paris : La Découverte.

- Demailly, L. (Éds). (2001). *Évaluer les politiques éducatives*. Bruxelles : De Boeck.
- Demeuse, M. (2004). *Introduction aux théories et aux méthodes de la mesure en [...] sciences de l'éducation*. Liège : Les Éditions de l'Université de Liège.
- Depover, C., De Lièvre, B., Quintin, J.-J., Porco, F. & Floquet, C. (s.d.). Les modèles d'enseignement et d'apprentissage. Consulté le 12 septembre 2013 dans <http://ute.umh.ac.be/dutice/uv6a/>
- Depover, C., Quintin, J.-J. & De Lièvre, B. (2003). Le design et la production des supports de formation, In C. Depover & L. Marchand (Éds), *E-learning et formation des adultes en contexte professionnel* (pp. 95-124). Bruxelles : De Boeck.
- Dewey, J. (1967) *Logique. La théorie de l'enquête*. (2^e éd.) Paris : PUF.
- Dewey, J. (2010). *Le public et ses problèmes*. Paris : Gallimard Folio essais.
- Dewey, J. (2011). *La formation des valeurs*. Paris : La Découverte.
- Eduscol (2011). *Bilan des expérimentations académiques 2008-2009*. Consulté le 12 septembre 2013 dans <http://eduscol.education.fr/cid50828/enseigner-et-evaluer-par-competences.html>
- Fabre, M. (2009). *Philosophie et pédagogie du problème*. Paris : Vrin.
- Faure, A., Pollet, G. & Warin, P. (1995). *La construction du sens dans les politiques publiques : débats autour de la notion de référentiel*. Paris : L'Harmattan.
- Figari, G. (1994). *Évaluer : quel référentiel ?* Bruxelles : De Boeck.
- Figari, G. (2006). Les référentiels entre théorie et méthodologie. In G. Figari & L. Mottier Lopez (Éds), *Recherche sur l'évaluation en éducation : problématiques, méthodologies et épistémologie. 20 ans de travaux autour de l'ADMEE-Europe* (pp. 101-108). Paris : L'Harmattan.
- Figari, G. (2009, mars). *L'évaluation : construction et distorsion du sens de l'éducation*. III^e Conférence de l'ERCILIS : La crise du sens, Université de Sfax, Tunisie.
- Figari, G. (2012). Configurations pour une épistémologie de l'évaluation. In L. Mottier Lopez & G. Figari (Éds), *Modélisations de l'évaluation en éducation* (pp. 211-224). Bruxelles : De Boeck.
- Figari, G. (sous presse). L'évaluation entre technicité et théorisation ? *Mesure et évaluation en Éducation*, 36(3).

- Figari, G. & Achouche, M. (Éds) (2001). *L'activité évaluative réinterrogée. Regards scolaires et socio-professionnels*. Bruxelles : De Boeck.
- Figari, G. & Mottier Lopez, L. (Éds) (2006). *Recherche sur l'évaluation en éducation : problématiques, méthodologies et épistémologie. 20 ans de travaux autour de l'ADMEE-Europe*. Paris : L'Harmattan.
- Ferréol, G. (2000). *La dissertation sociologique*. Paris : Armand Colin.
- Gerard, F.-M. (2002). L'indispensable subjectivité de l'évaluation. *Antipodes*, 156, 26-34.
- Gérard, F.-M. (2005). *L'évaluation des compétences à travers les situations complexes*. Actes du colloque de l'Admee Europe, Reims
- Gerard, F.-M. & Rogiers, X. (1993). *Concevoir et évaluer des manuels scolaires*. Bruxelles : De Boeck.
- Ghiglione, R. & Matalon, B. (1978). *Les enquêtes sociologiques. Théories et pratiques*. Paris : Armand Colin.
- Ghiglione, R., Beauvois, J.-L., Chabrol, C. & Trognon, A. (1980). *Manuel d'analyse de contenu*. Paris : Armand Colin.
- Project Management Institute (2000). *Guide to the Project Management Body of Knowledge*. Bruxelles : PMI.
- Grégoire, J. (Éds). (1996). *Évaluer les apprentissages*. Bruxelles : De Boeck.
- Gros, A. (2011). Les formes de l'enquête historique : John Dewey & Max Weber. *L'Atelier du Centre de recherches historiques*, 7. Consulté le 12 septembre 2013 dans <http://acrh.revues.org/3668>
- Hadji, C. (1989) *L'évaluation, règles du jeu*. Paris : ESF
- Hameline, D. (1979). *Les objectifs pédagogiques en formation initiale et en formation continue*. Paris : ESF
- Hart, R. (1992). *Children's Participation from Tokenism to citizenship*. Florence : Centre de recherche Innocenti de l'UNICEF.
- Henriot, M. (2009). *L'évaluation participative de l'action publique*. Paris : L'Harmattan.
- Hopkins, D. (2001). *School improvement for real*. London : Falmer Press.
- Jorro, A. (2000). *L'enseignant et l'évaluation. Des gestes évaluatifs en questions*. Bruxelles : De Boeck.
- Kuhn, T. (1983). *La structure des révolutions scientifiques*. Paris : Éditions Champs Flammarion.

- Lahire B. (2005). *L'Esprit sociologique*. Paris : La Découverte.
- Laveault, D. (1997). *Introduction à la recherche*. Ottawa : Edu 5590, Université d'Ottawa.
- Laveault, D. (2012). Autorégulation et évaluation soutien d'apprentissage. In L. Mottier Lopez & G. Figari (Éds), *Modélisations de l'évaluation en éducation* (pp. 115-130). Bruxelles : De Boeck.
- Leplat, J. (1997). *Regards sur l'activité en situation de travail*. Paris : PUF.
- Leplat, J. (2000). *L'analyse psychologique de l'activité en ergonomie*. Toulouse : Octarès.
- Lesne, M. (1984). *Lire les pratiques de formation d'adultes*. Paris : Edilig.
- Machado, A. (2007). *Évaluation et participation. Une étude sur le rôle des acteurs dans la formation continue*. Thèse de doctorat en Sciences de l'éducation, Universidade do Minho.
- Mayen, P. (2009a). Dix savoirs sur l'expérience, acquis à travers l'expérience de la validation des acquis de l'expérience. *Cahiers du travail social*, 123, 8-19.
- Mayen, P. (2009b). Expérience et formation des adultes. In J.-M. Barbier, E. Bourgeois, G. Chapelle & J.-C. Ruano-Borbalan (Éds), *Encyclopédie de la formation* (pp. 763-780). Paris : PUF.
- Mayen, P., Mayeux, C. & Savoyant, A. (2006). Construire la référence : une activité permanente des jurys de VAE. In G. Figari, P. Rodrigues, M. Palmira Alves & P. Valois (Éds), *Évaluation des compétences et apprentissages expérientiels* (pp. 23-47). Lisbonne : Admee Europe et Educa.
- Mayen, P., Tourmen, C., Mayeux, C. & Chausse, C. (2008). *Les conditions de réussite de l'activité d'un jury de VAE*. Rapport remis au Ministère de l'Agriculture, Dijon : Enesad-Eduter
- Mayen, P. & Savoyant, A. (2009). Élaboration et réduction de l'expérience dans la validation des acquis de l'expérience. *Relief*, 28, 59-69.
- Mayen, P., Métral, J.-F. & Tourmen, C. (2010). Les situations de travail : références pour les référentiels. *Recherche et formation*, 2, 31-46.
- McLaughlin, J. A. & Jordan, G. B. (2004). Using Logic Models. In J. Wholey, H. P. Hatry & Newcomer, K. E. (Éds), *Handbook of Pratical Program Evaluation* (pp. 7-32). Hoboken : Jossey-Bass.
- Merle, P. (1996). *L'évaluation des élèves. Enquête sur le jugement professoral*. Paris. PUF.

- Ministère de l'Éducation nationale (2005). Loi d'orientation et de programme pour l'avenir de l'école. Consulté le 18 septembre 2013 dans <http://www.education.gouv.fr/bo/2005/18/MENX0400282L.htm#loi>.
- Ministère de l'Éducation nationale (2012). Les 31 indicateurs sur le système éducatif français. L'état de l'école No. 22. Consulté le 18 septembre 2013 dans http://cache.media.education.gouv.fr/file/etat22/44/3/DEPP-EE-2012_231443.pdf.
- Mons, N. (2004). *L'évaluation quantitative internationale des systèmes éducatifs : apports, limites et renouvellement*. IREDU-CNRS/ Université de Bourgogne. Consulté le 18 septembre 2013 dans http://www.prisme-asso.org/IMG/pdf/L_evaluation_quantitative_internationale_des_systemes_educatifs.pdf.
- Mottier Lopez L. (2006). Évaluation située des apprentissages : le rôle fondamental de la médiation sociale. In G. Figari & L. Mottier Lopez (Éds), *Recherche sur l'évaluation en éducation* (pp. 82-89). Paris : L'Harmattan.
- Mottier Lopez, L. (2013a). Pour une culture partagée de l'évaluation. Du retour à des pratiques d'évaluation traditionnelles vers la construction d'un nouveau rapport entre la recherche, le terrain et le politique. *Revue Éducation Canada*. Consulté le 18 septembre 2013 dans <http://www.cea-ace.ca/education-canada/article/pour-une-culture-partagée-de-l'évaluation>
- Mottier Lopez, L. (2013b). Vers une culture de l'évaluation conjointement construite. In J.-F. Marcel & H. Savy (Éds), *Évaluons, évoluons. De l'évaluation dans l'enseignement agricole* (pp. 197-207). Dijon : Educagri Éditions Recherche.
- Mottier Lopez, L. & Figari, G. (Éds). (2012). *Modélisations de l'évaluation en éducation. Questionnements épistémologiques*. Bruxelles : De Boeck.
- Muller, P. (1990). *Les politiques publiques. Que sais-je ?* Paris : PUF.
- Mutelesi, E. (1998). Subjectivité comme auto-organisation. Une étude du constructivisme radical au départ de Husserl. Dissertation doctorale à l'Institut Supérieur de Philosophie, Université Catholique de Louvain.
- Noël, B. (1997). *La métacognition*. Bruxelles : De Boeck.
- OCDE (2012). *Regards sur l'éducation 2012 : les indicateurs de l'OCDE*. Paris : OCDE.
- Orwell, G. (1950). 1984. Paris : Gallimard.
- Pastré, P. (1999). Travail et compétences : un point de vue de didacticien. *Formation emploi*, 67, 109-126.

- Pastré, P., Mayen, P. & Vergnaud, G. (2006). La didactique professionnelle. *Revue française de pédagogie*, 154, 145-198.
- Peeters, H. & Charlier, P. (1999). Contributions à une théorie du dispositif. *Hermès*, 25, 15-23.
- Petitpierre, C. (2001) Pour une régulation formative du changement. In L. Demailly (Éds), *Évaluer les politiques éducatives* (pp. 163-180). Bruxelles : De Boeck
- Pfauwadel, A. (2010). *Évaluer tue*. Forum des psys. Évaluation, une culture de mort, Saison 1, Épisode 12. Consulté le 12 septembre 2013 dans <http://forumpsy.wordpress.com/>
- PISA (2012). *Évaluation des compétences des jeunes de 15 ans, en mathématiques, lecture et sciences*. Consulté le 19 Septembre 2013, dans https://www.educa.ch/sites/default/files/20111205/brochure_pisa-fra-2012.pdf.
- Postic, M. (1977). *Observation et formation des enseignants*. Paris : PUF.
- Quivy, R. & Van Campenhoudt, L. (2006). *Manuel de recherche en sciences sociales*. Paris : Dunod.
- Rodrigues, P. (2006) Le recours à des paradigmes de l'évaluation. In G. Figari & L. Mottier Lopez (Éds), *Recherche sur l'évaluation en éducation* (pp. 193-201). Paris : L'Harmattan.
- Rossi, P. H. (1991). Comprehensive, tailored, theory-driven evaluations. In W. R. Shadish, T. D. Cook & L. C. Leviton (Éds), *Foundations of program evaluation. Theories of practice* (pp. 377-440). Newbury Park : Sage.
- Saussez, F., Ewen, N. & Girard, J. (2001). Au cœur de la pratique réflexive, la conceptualisation ? *Recherche et Formation*, 36, 69-87.
- Scallon, G. (1988). *L'évaluation formative des apprentissages*. Vol. 2. *L'instrumentation*. Québec : Les Presses de l'Université Laval.
- Scheerens, J. & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford : Pergamon.
- Schön, D. (1993). *Le praticien réflexif. À la recherche du savoir caché dans l'agir professionnel*. Montréal : Éditions Logiques.
- Schwartz, Y. (2007). Un bref aperçu de l'histoire culturelle du concept d'activité. *Activités*, 4(2), 122-133. Consulté le 18 septembre 2013 dans <http://www.activites.org/v4n2/v4n2.pdf#page=123>.
- Scriven M. S. (1967). *The methodology of evaluation (AERA Monograph Series in Curriculum Evaluation No. 1)*. Chicago : Rand McNally.

- Shadish, W. R., Cook, T. D. & Leviton, L. C. (1991). *Foundations of program evaluation. Theories of practice* (pp. 377-440). Newbury Park : Sage.
- Solau, G. (2000). *L'évaluation des politiques d'éducation. Documents (actes et rapports) pour l'éducation*. Dijon : CRDP de Bourgogne.
- Soulet, M.-H. (2011). Interpréter, avez-vous dit ! *SociologieS, Expériences de recherche, Régimes d'explication en sociologie*. Consulté le 18 septembre 2013 dans <http://sociologies.revues.org/3471>.
- Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. L., Merriman, H. O. & Provus, M. M. (1980). *L'évaluation et la prise de décision en éducation*. Ottawa : NHP.
- Suchaut, B. (2008). *La loterie des notes au bac. Un réexamen de l'arbitraire de la notation des élèves*. Dijon : IREDU, Université de Bourgogne, CNRS.
- Tardif, J. (1998). *Intégrer les nouvelles technologies de l'information. Quel cadre pédagogique ?* Paris : ESF.
- Tardif, J. & Dubois, B. (2010). Construire des dispositifs en vue de l'évaluation et du développement des compétences : comment ? In L. Paquay, C. Van Nieuwenhoven & P. Wouters (Éds), *L'évaluation, levier du développement professionnel ?* (pp. 133-146) Bruxelles : De Boeck.
- Torres, J.-C. (2010). *L'évaluation dans les établissements scolaires. Théories, objets et enjeux*. Paris : L'Harmattan.
- Tourmen, C. (2007). *Les compétences des évaluateurs. Le cas des évaluateurs de politiques publiques*. Thèse de doctorat en Sciences de l'éducation, LES, Université de Grenoble Alpes.
- Tourmen, C. (2009). L'activité évaluative et la construction progressive du jugement. *Les Dossiers des sciences de l'éducation*, 22, 101-119.
- Tourmen, C. (2012). *Comment voir la pensée dans l'action ?* Actes du 24^e colloque International de l'Admee-Europe (Luxembourg, 11-13 janvier).
- Tourmen, C. & Droyer, N. (2011). Une méthode pour analyser et outiller les jurys d'évaluation. Le cas d'un jury de recrutement d'allocataires de recherche. *Actes du 23^e colloque de l'Admee Europe* (Paris, 12-14 janvier).
- Tourmen, C., Mayen, P. & Figari, G. (2009). Quels critères et indicateurs pour évaluer les acquis de l'expérience ? *Actes du 21^e colloque International de l'Admee-Europe* (Louvain-La-Neuve, 21-23 janvier).

- Tousignant, R. (1990). *Les principes de la mesure et de l'évaluation des apprentissages*. Québec : Gaëtan Morin.
- Tremblay, R.-R. & Perrier, Y. (2006). *Savoir plus : outils et méthodes de travail intellectuel* (2^e éd.). Montréal : Les Éditions de la Chenelière.
- UNESCO (2008). *Inclusive education : the way of the future*. Geneva : International Conference Centre.
- Verdière, J. (2001). Les difficultés de l'évaluation de l'éducation prioritaire. In L. Demailly (Éds), *Évaluer les politiques éducatives* (pp. 77-94). Bruxelles : De Boeck.
- Vergnaud, G. (2001). Psychologie du développement cognitif et évaluation des compétences. In G. Figari & M. Achouche (Éds), *L'activité évaluative réinterrogée. Regards scolaires et socioprofessionnels* (pp. 43-51). Bruxelles : De Boeck.
- Vergnaud, G. (2008). De la didactique des disciplines à la didactique professionnelle, il n'y a qu'un pas. *Travail et Apprentissages*, 1, 51-58.
- Weber, M. (1965). *Essai sur la théorie de la science*. Paris : Plon.
- Weiss, J. (1991). *L'évaluation : problème de communication*. Neuchâtel : Del Val.
- Wortman, P. M. (1983). Evaluation research : A methodological perspective. *Annual Review of Psychology*, 34, 223-260.

Table des figures

Figure 1. La valuation selon Dewey (Chapitre 3. Des définitions de l'évaluation, 2. Quelles Définitions ?)	41
Figure 2. Première formulation du modèle d'évaluation (Chapitre 7. Méthodologie d'évaluation par la référentialisation, 1. Sources d'un protocole, 1.1 Des notions préalables)	73
Figure 3. 2^e formulation du modèle d'évaluation (Chapitre 7. Méthodologie d'évaluation par la référentialisation, 2. Modélisations de l'évaluation : schémas progressifs)	75
Figure 4. Une modélisation de l'évaluation externe (Chapitre 7. Méthodologie d'évaluation par la référentialisation, 2.1 Relations dans une perspective externe : évaluation de la qualité d'ensemble du dispositif (pertinence, efficacité, efficience)	76
Figure 5. Une modélisation de l'évaluation interne (Chapitre 7. Méthodologie d'évaluation par la référentialisation, 2.2 Relations dans une perspective interne : une évaluation de régulation)	78
Figure 6. Modélisation de l'ensemble des processus d'évaluation (Chapitre 7. Méthodologie d'évaluation par la référentialisation, 2.3 En synthèse : Une modélisation de l'ensemble des regards d'évaluation)	78

- Figure 7. La boucle de rétroaction du modèle systémique**
(Chapitre 7. Méthodologie d'évaluation
par la référentialisation, 4.4 Phase 4 : Analyse
et interprétation) 99
- Figure 8. Application du modèle ICP à l'évaluation des acquis
de l'expérience**
(Chapitre 12. Cinquième application :
Évaluation de compétences,
4.3. Les apports de cette phase au processus
d'évaluation) 175

Table des matières

Sommaire	5
Remerciements	7
Préface	9
Introduction	13
1. Origines croisées d'une façon de penser la méthodologie de l'évaluation en éducation	13
2. Méthodologie de l'évaluation : pas seulement une tautologie	14
3. Un livre de « méthodologie générale sur l'évaluation en éducation »	15
4. Une volonté de « fonder » sur des concepts toute méthodologie de l'évaluation	16
5. Une approche influencée par la théorie de l'enquête de Dewey	16
6. D'autres influences à assumer	17
7. Une exposition en trois moments	17
8. Trois chemins de lecture	18
 PARTIE 1. LES FONDEMENTS DE L'ACTIVITÉ ÉVALUATIVE	 19
Introduction	21
1. Une grande diversité de pratiques d'évaluation	21
2. Des pratiques pas toujours réfléchies	22
3. Un besoin de visions claires et globales	22
4. La nécessité d'appuyer les méthodes d'évaluation sur des conceptualisations génératrices de sens	23
 Chapitre 1. L'évaluation en débats	 25
1. Mesure/évaluation	25
2. Contrôle/évaluation	27

3.	Qualitatif/quantitatif	27
4.	Externe/interne	28
5.	Objectif/subjectif	29
5.1	<i>Objectif</i>	30
5.2	<i>Subjectif</i>	30
6.	Des oppositions réductrices ou des amorces de théorisation ?	31
Chapitre 2. Des cultures de l'évaluation ?		33
1.	Culture de l'évaluation : un thème fréquent	33
2.	Distinguer entre « la », « une » et « des » culture(s) de l'évaluation	34
3.	Un nouvel avatar : la culture du résultat ?	35
4.	Reconnaître les modèles d'une culture de l'évaluation	36
5.	Quand la culture de l'évaluation cache l'évaluation elle-même	37
Chapitre 3. Des définitions de l'évaluation		39
1.	Avant de définir	39
1.1	<i>Que fait-on en définissant l'évaluation ?</i>	39
1.2	<i>Pour des définitions modestes</i>	40
1.3	<i>Pour la prise en compte du discours et des jeux de langage dans la définition de l'évaluation</i>	40
1.4	<i>La question de la scientificité de la définition</i>	41
2.	Quelles définitions ?	41
2.1	<i>L'évaluation définie comme jugement de valeur</i>	43
2.2	<i>L'évaluation définie comme aide à la décision</i>	44
2.3	<i>L'évaluation définie comme recherche ou changement de sens</i>	45
2.3.1	Une recherche de sens	45
2.3.2	Un changement de sens	45
2.4	<i>L'évaluation définie comme activité</i>	46
2.4.1	Une activité en situation	47
2.4.2	Une expérience	47
Chapitre 4. Évaluation et théorie		49
1.	Pourquoi évoquer les liens entre évaluation et théorie ?	49
2.	Parler théories à propos de l'évaluation ?	50

3.	Différentes postures concernant la dimension théorique de l'évaluation	51
3.1	<i>L'évaluation ne serait qu'une pratique ?</i>	51
3.2	<i>L'évaluation ne serait pas à confondre avec l'activité scientifique ?</i>	51
3.3	<i>Une science de, par ou pour l'évaluation ?</i>	52
	Chapitre 5. L'évaluation, une activité langagière	53
1.	Le parler évaluatif	53
1.1	<i>Le discours évaluatif courant</i>	53
1.2	<i>Le discours évaluatif scientifique</i>	54
2.	Le langage : un problème pour l'évaluation	55
	Conclusion : une évaluation influencée par ses fondements théoriques	57
	PARTIE 2. MÉTHODOLOGIE D'ENQUÊTE ÉVALUATIVE PAR LA RÉFÉRENTIALISATION	59
	Introduction	61
1.	Des concepts à la méthode	61
2.	Une enquête évaluative	61
3.	Vers une méthodologie d'enquête évaluative	62
4.	Mais qui sont les évaluateurs ?	63
	Chapitre 6. La place des référentiels dans l'évaluation	65
1.	L'omniprésence des référentiels	65
2.	Des utilisations diverses de la notion de référentiel	66
3.	Des référentiels « déjà-là » ?	67
4.	Une définition-princeps du référentiel ?	68
5.	Référentiels et évaluation	68
	Chapitre 7. Méthodologie d'évaluation par la référentialisation	71
1.	Sources d'un protocole	71
1.1	<i>Des notions préalables</i>	71
1.1.1	Référentiel et référentialisation	71
1.1.2	La référentialisation : une notion à trois dimensions (conceptualisation, modélisation, instrumentation)	72

2.	Modélisations de l'évaluation : schémas progressifs	75
2.1	<i>Relations dans une perspective externe : évaluation de la qualité d'ensemble du dispositif (pertinence, efficacité, efficience)</i>	76
2.1.1	Évaluation de la pertinence (1)	76
2.1.2	Évaluation de l'efficacité (2)	76
2.1.3	Évaluation de l'efficience (3)	77
2.2	<i>Relations dans une perspective interne : une évaluation de régulation</i>	77
2.3	<i>En synthèse : une modélisation de l'ensemble des regards d'évaluation</i>	77
3.	Développement d'un protocole d'évaluation	79
3.1	<i>Un protocole méthodologique</i>	79
3.2	<i>La référentialisation comme conception d'une méthodologie d'évaluation</i>	79
3.3	<i>En préalable, un constat et un postulat</i>	79
3.4	<i>Enfin, l'organisation d'un protocole</i>	80
4.	Les quatre phases du protocole d'évaluation illustrées par un exemple-type : l'établissement scolaire	81
4.1	<i>Phase 1 : exploration (analyse du mandat et formulation de la problématique d'évaluation)</i>	81
4.1.1	Analyse du mandat	81
4.1.2	Élaboration d'une problématique	82
4.2	<i>Phase 2 : définition de l'objet à évaluer</i>	83
4.2.1	Éléments à évaluer	83
4.2.2	Identification des référents	85
4.2.3	Élaboration des critères	88
4.2.4	Clarification de la notion de critère	88
4.2.5	Formulation des critères	89
4.2.6	Construction des indicateurs	90
4.3	<i>Phase 3 : construction d'instruments</i>	94
4.3.1	Des indicateurs à la rédaction des items de questionnaire	94
4.3.2	L'intégration des items dans des instruments appropriés	95
4.3.3	Le recueil des informations	96
4.4	<i>Phase 4 : Analyse et interprétation</i>	96

4.4.1	Analyse des résultats	97
4.4.2	Interprétation des résultats	97
4.4.3	Exploitation et restitution	98
4.4.4	Expliquer les résultats en regard de la référentialisation	100
5.	En conclusion : utilisations d'une méthodologie de l'enquête évaluative	100
5.1	<i>Une vision d'ensemble de ce que mobilise une opération d'évaluation</i>	100
5.2	<i>Un guide méthodologique pour construire une opération d'évaluation</i>	101
5.3	<i>Une grille d'analyse pour étudier un dispositif éducatif</i>	101
5.4	<i>Méthodologie qualitative avant méthodologie quantitative</i>	102
5.5	<i>Une enquête évaluative ancrée dans l'expérience</i>	102
	Conclusion : le sens et l'intérêt d'une « enquête évaluative »	105
	PARTIE 3. MÉTHODOLOGIE APPLIQUÉE À DIFFÉRENTS CHAMPS DE L'ÉDUCATION ET DE LA FORMATION	107
	Introduction	109
1.	L'activité d'évaluation dans les mises en scène de l'éducation, de la formation et des apprentissages	109
2.	La référentialisation : une comparaison et une mise en relation	110
3.	Application à des pratiques évaluatives	110
4.	Mode de présentation des cas	111
	Chapitre 8. Première application : Évaluation de politiques éducatives	113
	Introduction : Les politiques éducatives comme objets à évaluer	113
	<i>Concept de politique publique éducative</i>	114
	Politique publique	114
	Éducative	114
	<i>Différents cas de figure</i>	114
	<i>Situations d'évaluation de politiques éducatives</i>	114
	<i>Fonctions attribuables à l'évaluation des politiques publiques</i>	115

Méthodologie repérée	115
1. Phase 1 : Exploration (mandat et problématique)	115
1.1 Analyse du cas	116
1.1.1 Le mandat	116
1.1.2 La problématique	116
1.2 La mise en tableau	116
1.3 Les apports de cette phase à la conduite de l'évaluation	117
2. Phase 2 : Procédure centrale de référentialisation	117
2.1 Analyse du cas	118
2.1.1 Définition des éléments à évaluer	118
2.1.2 Choix des référents	118
2.1.3 Élaboration des critères	118
2.1.4 Construction des indicateurs	119
2.2 La mise en tableau	119
2.3 Les apports des procédures de la phase 2 à la conduite de l'évaluation	121
3. Phase 3 : Instrumentation	121
3.1 En guise d'analyse du cas	121
3.2 Mise en tableau	122
3.3 Les apports de cette phase au processus d'évaluation	123
4. Phase 4 : Interprétation	123
4.1 Les apports de cette phase au processus d'évaluation	125
Chapitre 9. Deuxième application :	
Évaluation de projets éducatifs	127
Introduction : Les projets éducatifs comme objets à évaluer	127
<i>Concept de projet</i>	127
<i>Projet éducatif</i>	128
<i>Différents cas de figure</i>	128
<i>Les situations d'évaluation de projet</i>	129
Méthodologie repérée	129
1. Phase 1 : Exploration (mandat et problématique)	129
1.1 Analyse du cas	129
1.1.1 Le mandat	129
1.1.2 La problématique	130
1.2 Mise en tableau	130

2.	Phase 2 : Procédure centrale de référentialisation	131
2.1	<i>Analyse du cas</i>	131
2.1.1	Définition des éléments à évaluer	131
2.1.2	Choix des référents	131
2.1.3	Élaboration des critères	132
2.1.4	Construction des indicateurs	132
2.2	<i>Mise en tableau</i>	132
2.3	<i>Les apports de cette phase au processus d'évaluation</i>	133
3.	Phase 3 : Instrumentation	134
3.1	<i>En guise d'analyse du cas</i>	134
3.2	<i>Mise en tableau</i>	135
3.3	<i>Les apports de cette phase au processus d'évaluation</i>	135
4.	Phase 4 : Interprétation	136
4.1	<i>Analyse du cas mise en tableau</i>	136
4.2	<i>Les apports de cette phase au processus d'évaluation</i>	137
Chapitre 10. Troisième application : Évaluation des dispositifs et programmes		139
Introduction : Les dispositifs éducatifs et les programmes comme objets à évaluer		139
<i>Concepts de dispositif et de programme</i>		139
Nature et conceptualisation d'un dispositif éducatif		140
<i>Différents cas de figure</i>		140
<i>Évaluation des dispositifs éducatifs</i>		140
Méthodologie repérée		140
1.	Phase 1 : Exploration (mandat et problématique)	141
1.1	<i>Analyse du cas</i>	141
1.1.1	Le mandat	141
1.1.2	La problématique	141
1.2	<i>Mise en tableau</i>	142
1.3	<i>Les apports de cette phase au processus d'évaluation</i>	142
2.	Phase 2 : Procédure centrale de référentialisation	143
2.1	<i>Analyse du cas</i>	143
2.1.1	Définition des éléments à évaluer	143
2.1.2	Choix des référents	143
2.1.3	Indicateurs généraux (critères et indicateurs volontairement confondus)	144

2.2	<i>Mise en tableau</i>	144
2.3	<i>Les apports de cette phase au processus d'évaluation</i>	145
3.	Phase 3 : Instrumentation	145
3.1	<i>En guise d'analyse du cas</i>	146
3.2	<i>Mise en tableau</i>	146
3.3	<i>Les apports de cette phase au processus d'évaluation</i>	147
4.	Phase 4 : Interprétation	147
4.1	<i>Mise en tableau</i>	148
4.2	<i>Les apports de cette phase au processus d'évaluation</i>	148
Chapitre 11. Quatrième application :		
	Évaluation des outils pédagogiques	151
Introduction : Les outils pédagogiques comme objets à évaluer		151
	<i>Évaluation des outils pédagogiques : référence à l'ingénierie éducative (ou pédagogique)</i>	151
	<i>Différents cas de figure</i>	152
Méthodologie repérée		152
1.	Phase 1 : Exploration (mandat et problématique)	152
1.1	<i>Analyse du cas</i>	152
1.1.1	Le mandat	152
1.1.2	La problématique	153
1.2	<i>1.2. Mise en tableau</i>	154
1.3	<i>Les apports de cette phase au processus d'évaluation</i>	154
2.	Phase 2 : Procédure centrale de référentialisation	155
2.1	<i>Analyse du cas</i>	155
2.1.1	Choix des éléments à évaluer	155
2.1.2	Référents	155
2.1.3	Critères	156
2.1.4	Indicateurs	156
2.2	<i>Mise en tableau</i>	156
2.3	<i>Les apports de cette phase au processus d'évaluation</i>	157
3.	Phase 3 : Instrumentation	158
3.1	<i>En guise d'analyse du cas</i>	158
3.2	<i>3.2. Mise en tableau</i>	159
3.2.1	Aperçu d'un outil d'évaluation des manuels pédagogiques (extrait)	159
3.3	<i>Les apports de cette phase au processus d'évaluation</i>	161

4.	Phase 4 : Interprétation	161
4.1	<i>En guise d'analyse du cas</i>	161
4.2	<i>Mise en tableau</i>	162
4.3	<i>Les apports de cette phase au processus d'évaluation</i>	162
Chapitre 12. Cinquième application : Évaluation de compétences		163
Introduction : Les compétences comme objets à évaluer		163
<i>Un objet nouveau pour l'évaluation</i>		163
<i>Différents cas de figure</i>		164
Méthodologie repérée		165
1.	Phase 1 : Exploration (mandat et problématique)	165
1.1	<i>Analyse du cas</i>	165
1.1.1	Le mandat	165
1.1.2	La problématique	166
1.2	<i>La mise en tableau</i>	166
1.3	<i>Les apports de cette phase au processus d'évaluation</i>	167
2.	Phase 2 : Procédure centrale de référentialisation	167
2.1	<i>Analyse du cas</i>	168
2.1.1	Définition des éléments à évaluer	168
2.1.2	Choix des référents	168
2.1.3	Élaboration des critères	168
2.1.4	Construction des indicateurs	169
2.2	<i>La mise en tableau</i>	170
2.3	<i>Les apports de cette phase au processus d'évaluation</i>	171
3.	Phase 3 : Instrumentation	172
3.1	<i>En guise d'analyse du cas</i>	172
3.2	<i>Mise en tableau</i>	172
3.3	<i>Les apports de cette phase au processus d'évaluation</i>	173
4.	Phase 4 : Interprétation	173
4.1	<i>En guise d'analyse de cas</i>	173
4.2	<i>Mise en tableau</i>	173
4.3	<i>Les apports de cette phase au processus d'évaluation</i>	174
Conclusion : référentialisation et exemplification		177
Conclusion		179
1.	Une méthodologie « générale » d'évaluation	179

2.	Une posture d'évaluateur réflexif	180
2.1	<i>Une posture</i>	180
2.2	<i>Un évaluateur réflexif</i>	181
3.	Une enquête à la recherche du sens d'une situation problématique	181
3.1	<i>La recherche de résolution d'une énigme</i>	181
3.2	<i>Une quête de la vérité et de la connaissance</i>	182
3.3	<i>Une quête de la valeur</i>	182
	Bibliographie	185
	Table des figures	195