

Cultures et politiques de l'évaluation en éducation et en formation

**RECHERCHE EN EDUCATION ET DECISION POLITIQUE :
CONDITIONS, APPORTS ET LIMITES DE L'EVALUATION**

Jean-Pierre Abbet* et Bruno Suchaut**

* Unité de recherche pour le pilotage des systèmes pédagogiques, jean-pierre.abbet@vd.ch

** Unité de recherche pour le pilotage des systèmes pédagogiques, bruno.suchaut@vd.ch

Mots-clés : évaluation, politique publique, acteurs, apports et limites

Résumé. Le but de cette contribution est de mettre en évidence les conditions dans lesquelles les politiques publiques relatives à l'éducation sont évaluées, en déterminant les apports et les limites inhérentes à une telle action. Pour ce faire, ces politiques, leurs objets, leurs processus dans le contexte suisse et en particulier vaudois sont pris en compte. Les apports différenciés de la recherche portant sur l'évaluation de dispositifs de politique publique versus des besoins de base des acteurs et usagers, d'une part, de la mise en œuvre de ces dispositifs par les pouvoirs publics ou des effets sur les bénéficiaires, d'autre part, sont également considérés en insistant particulièrement sur les types d'effets évalués. Sachant que les apports de l'évaluation doivent être reliés à des dispositions politiques, stratégiques et opérationnelles, les différents acteurs concernés par ces évaluations sont finalement considérés sous l'angle de leurs intérêts et de leurs regards particuliers sur l'éducation et les apports de la recherche.

1. Introduction

Le but de cette contribution est de mettre en évidence les conditions dans lesquelles les politiques publiques relatives à l'éducation sont évaluées, en déterminant les apports et les limites inhérentes à une telle action. Pour ce faire, ces politiques, leurs objets, leurs processus de mise en œuvre dans le contexte suisse et en particulier vaudois sont pris en compte. Les apports différenciés de la recherche visant l'évaluation des dispositifs de politique publique *versus* les besoins de base des acteurs et usagers sont tout d'abord envisagés. Dans le cadre de l'évaluation des politiques publiques, de tels apports sont ensuite considérés à propos de la distinction entre la mise en œuvre de ces dispositifs par les pouvoirs publics d'une part, les effets sur les bénéficiaires d'autre part, en insistant sur les types d'effets évalués. Puisque les évaluations sont destinées à des publics différents, les divers acteurs concernés sont passés en revue en fonction de leurs intérêts et de leurs regards particuliers sur l'éducation et les apports de la recherche.

En tenant compte de l'importance croissante accordée à l'évaluation des systèmes éducatifs, qui plus est dans un paysage en constante évolution, il importe de souligner l'origine politique et citoyenne d'une telle contribution. La demande originelle visait en effet à établir pour le monde politique vaudois l'état de l'évaluation des politiques publiques. On relèvera à cet égard que, en tant qu'organisme interne à l'administration de ce canton suisse, mais conduisant des recherches en éducation de façon externes, l'URSP (Unité de recherche pour le pilotage des systèmes pédagogiques) aide au pilotage des systèmes de formation par ses évaluations aux différents niveaux évoqués ici.

Cultures et politiques de l'évaluation en éducation et en formation

2. Question générale du choix et de la délimitation du champ et des objets d'évaluation

Dans le cas de la mise en œuvre d'une politique publique ou de la régulation ordinaire d'un système pédagogique, les questions initiales du choix d'évaluer, ainsi que de la détermination de ce qui fera ou non l'objet d'une évaluation, sont trop souvent escamotées. Elles ne sont pourtant ni évidentes ni inutiles, comme certains acteurs de l'éducation semblent tenté de le croire (Abbet, 2003). Deux écueils peuvent en effet favoriser un tel escamotage : pour le politique, le souci de parvenir rapidement à des décisions opérationnelles, pour le chercheur, celui de concentrer rapidement son savoir faire sur l'objet de recherche à construire (Abbet, *op. cit.*). Et pourtant, les objets, les niveaux et les types d'évaluation sont très divers, et donc les choix à opérer à cet égard fort nombreux. De plus, lorsqu'on se situe au plan de réformes scolaires, la plupart des innovations éducatives ou pédagogiques ont des répercussions sur l'ensemble du système, qu'il s'agisse de ses composantes « réformées » ou de celles qui sont censées subsister dans leur état antérieur.

Un tel état de fait devrait donc logiquement favoriser une réflexion préalable sur les choix de l'évaluation, sachant qu'aucun système de formation ou aucune réforme importante de celui-ci ne peuvent faire l'objet d'une évaluation exhaustive. Que l'on observe et analyse des dispositions institutionnelles, des pratiques professionnelles, des résultats scolaires, ou encore des vécus et opinions tels qu'exprimés par les différents acteurs du champ de l'éducation, il s'agit dès lors d'opérer des choix, si possible en connaissance de cause. Si l'on distingue les deux principaux niveaux que sont le cadre institutionnel et budgétaire d'une part, les contextes de l'enseignement que constituent l'établissement scolaire, la classe et l'enseignant d'autre part, différentes catégories qui peuvent faire l'objet d'analyses et d'évaluation émergent alors, en particulier dans le second cas, que ce soit du point de vue de l'organisation scolaire, de la composition des groupes classe, ou encore des situations pédagogiques et didactiques (Suchaut, 2008). Cette question préalable du choix et de la délimitation du champ et des objets d'évaluation doit ainsi faire l'objet d'une attention particulière. Il en va de même, ainsi qu'on va le voir, de la détermination du niveau et de la portée de l'évaluation en relation avec les intérêts différenciés des groupes d'acteurs considérés.

3. De quelles évaluations est-il question ?

Avant cela, il importe de souligner le fait que l'évaluation des politiques publiques relatives aux systèmes de formation se réfère nécessairement au cadre plus général de toute démarche d'évaluation. Une telle démarche doit en effet tenir compte des principales questions incontournables suivantes : quels sont les buts poursuivis, à savoir que cherche-t-on à évaluer et pourquoi ? Quels sont les moyens mobilisés pour y parvenir et par qui ? Finalement, quels sont les résultats attendus avec la démarche mise en œuvre et comment pourront-ils être perçus et si possible pris en compte par les acteurs ?

Les deux premières questions, en particulier, interpellent d'emblée. Cherche-t-on à évaluer la mise en œuvre par les pouvoirs publics d'une volonté politique qui se traduit en général par un dispositif législatif et réglementaire, voire un programme d'action, ou, *in fine*, veut-t-on connaître les effets de cette volonté sur les usagers et bénéficiaires ? Qu'il s'agisse de la mise en œuvre ou des effets, ou encore des deux, cherche-t-on, en outre, à mesurer l'amélioration de l'efficacité des politiques publiques seulement (ratios ressources/résultats) ou à mesurer l'amélioration de la situation des différents acteurs, c'est-à-dire des usagers et bénéficiaires, ainsi que des agents concourant à la mise en œuvre ? On peut également se demander si la seule satisfaction de ces acteurs constitue une mesure suffisante de cette amélioration de la situation ou si des critères plus objectifs doivent être dégagés ; dans ce cas lesquels, et comment les mesurer ?

Ne pas répondre à ces questions générales, tout au moins aux principales, revient, au bout du compte, à ne pas savoir vraiment ce que l'on évalue. Les éléments de réflexion qui suivent sont

Cultures et politiques de l'évaluation en éducation et en formation

ainsi l'occasion d'en discuter à partir d'un référentiel particulier, à savoir l'évaluation des politiques publiques relatives à l'enseignement dans le canton de Vaud, en Suisse. Ces éléments sont discutés du point de vue des activités conduites par l'URSP du Département de la formation, de la jeunesse et de la culture (DFJC) de ce canton francophone. Mentionnons, pour l'exemple, un autre référentiel cantonal, celui de Genève, qui a fait l'objet de l'analyse d'une politique publique dans le cadre circonscrit de l'enseignement secondaire obligatoire (Felouzis *et al.*, 2013).

3.1 Evaluation des politiques publiques et/ou des besoins des acteurs et usagers ?

Couvrir les besoins et les souhaits des citoyens/usagers/bénéficiaires constitue bien sûr le motif principal de la mise en œuvre des politiques publiques. Ces dernières répondent à des besoins avérés, identifiés par le monde politique, et qui se manifestent au moment où elles sont mises en œuvre. Les besoins qui ne répondent pas directement à ces caractéristiques, c'est-à-dire ceux qui ne sont pas explicitement couverts par des dispositions légales ou réglementaires, ou ne sont pas d'actualité tout en étant prévisibles, n'en sont pas moins réels et ne méritent pas moins d'être évalués. Il est bien sûr question ici des seuls besoins qui peuvent être identifiés, et non de besoins potentiels. C'est ce qui explique que, à côté et en complément de l'évaluation des politiques publiques, une évaluation des besoins des acteurs et usagers soit aussi nécessaire. Dans le domaine de l'enseignement, un tel complément se rapporte, par exemple, à l'identification et au traitement d'objets transversaux à la formation (les compétences sociales), ou parcourant celle-ci (les compétences disciplinaires dans les cursus de formation), pour prendre deux objets d'évaluation volontairement contrastés mais complémentaires. De telles évaluations se rapportent directement aux systèmes de formation, mais ne se réfèrent qu'indirectement à un dispositif particulier de politique publique. A l'inverse, évaluer la mise en œuvre de certains aspects de la nouvelle Loi sur l'enseignement obligatoire (LEO) du canton de Vaud, tels que la mise en filières ou en niveaux d'enseignement au secondaire obligatoire, ou encore l'intégration des enfants à besoins particuliers dans la scolarité ordinaire, sont deux cas illustratifs d'une évaluation au service d'une mesure spécifique de politique éducative.

Dans un certain sens, on pourrait presque dire que l'évaluation des besoins des acteurs, usagers et bénéficiaires devrait se situer en amont de tous dispositifs de politique publique, ces derniers devant répondre à de tels besoins et être évalués à l'aune de leur capacité à les satisfaire. Pratiquement, étant donné que les diverses manifestations des besoins et les diverses mises en œuvre de dispositifs ne se succèdent pas de façon schématique, ces évaluations de type différents sont plus ou moins conduites en parallèle. On peut noter par ailleurs que les besoins de base (éducation et instruction), qui évoluent sous diverses formes et concernant la satisfaction dans diverses dimensions, se situent le plus souvent dans le long terme, ce qui n'est pas toujours le cas des dispositifs de politiques publiques, qui peuvent être plus ou moins ponctuels et qui s'inscrivent souvent dans une temporalité connue et limitée.

On retiendra que, dans le cadre des missions qui lui sont dévolues, et compte tenu des moyens qui lui sont alloués, un centre cantonal dédié au pilotage des systèmes de formation tel que l'URSP évalue, de façon complémentaire, tant des dispositifs de politiques publiques mis en œuvre que des besoins des acteurs et usagers ; dans ce dernier cas, sous la forme de veille scientifique ou d'enquêtes parallèles et complémentaires à l'évaluation des dispositifs.

3.2 Dans le cas des politiques publiques, évaluation de la mise en œuvre par les pouvoirs publics ou des effets sur les usagers ?

En se concentrant plus strictement sur l'évaluation des politiques publiques, on relèvera que les niveaux d'impact des résultats des projets de recherche sont directement reliés aux différents niveaux des systèmes de formation auxquels on les réfère. On distingue en général quatre niveaux dans ces systèmes : (1) le niveau des décideurs politiques (2) celui des responsables et cadres de l'administration, (3) celui des directions régionales et enfin (4) celui des praticiens, en particulier les enseignants, mais aussi tous les professionnels qui gravitent autour de l'enseignement (plus ou

Cultures et politiques de l'évaluation en éducation et en formation

moins directement en contact avec les élèves, leurs parents, ainsi que les milieux professionnels ou sociaux).

Le *monde politique* cherche à mettre en œuvre les réformes nécessaires à la satisfaction des citoyens et au bon fonctionnement de la société, selon des choix qui parcourent l'échiquier politique. Il vise des intérêts variés qui sont rarement convergents et qui relèvent souvent d'un subtil équilibre. Les politiques sont ainsi intéressés à des résultats d'évaluation qui leur permettent – entre autres critères de choix – de légitimer leur action en intégrant les constats et opinions des divers acteurs auxquels s'adresse l'évaluation. Felouzis *et al.* (*op. cit.*) indiquent à cet égard que le rôle du politique par rapport à l'éducatif incite à penser que les systèmes de formation sont « pluri-régulés ». Les *responsables de l'administration* sont à l'interface des décisions politiques et des réalités du terrain. Ils constituent le premier échelon des professionnels qui, à ce titre, doivent non seulement décliner les décisions des politiques, mais aussi anticiper les conséquences prévisibles de ces décisions. Les résultats d'évaluation les intéressent donc, en principe, fortement. Les *directions des établissements scolaires*, proches des praticiens, doivent mettre en œuvre une politique pédagogique en contact avec les bénéficiaires, les élèves. Du point de vue de l'enseignement, ce niveau du système qui oriente directement l'action a besoin de sources diverses et l'évaluation devrait lui permettre de fonder son orientation en direction des praticiens. Les *praticiens* constituent le maillon de la chaîne le plus important en termes de résultats des bénéficiaires, voire celui qui est décisif. L'information fournie par la recherche qui concerne ce niveau provient généralement de deux sources distinctes : la relation à l'élève et à la matière enseignée au moyen de recherches provenant des instituts de formation des enseignants ; la connaissance plus globale du système, du fonctionnement et de ses résultats, provenant des instances d'évaluation du système. On voit donc bien, à ce niveau, l'intérêt d'une collaboration entre ces deux sources d'informations complémentaires, sans que l'on connaisse suffisamment l'intérêt des praticiens pour ces apports de l'évaluation.

En principe, ces niveaux sont hiérarchisés et les inflexions décidées au niveau politique devraient descendre en cascade jusqu'auprès des bénéficiaires. Ce point de vue, issu des principes démocratiques, et qui légitime cette hiérarchie, ne fait cependant pas ressortir ce qu'il y a de spécifique à chaque niveau. Par ailleurs, il est clair que l'on ne peut bâtir de la transparence, souvent invoquée dans le cadre des politiques publiques, de façon unilatérale. L'information doit donc, non seulement provenir, mais retourner aux différents niveaux du système qui en ont l'utilité. Ceci, dans une visée de régulation à plus ou moins long terme qui devrait être le propre des évaluations dans ce domaine, et qui constitue sa fonction principale. Les décideurs doivent donc faire parvenir des informations aux praticiens, mais ces derniers faire remonter des informations aux décideurs concernant leurs opinions et pratiques. Les chercheurs, pour leur part, transmettent des informations aux uns et aux autres dans le cadre de leurs évaluations.

On pourrait à première vue postuler que l'action de ces différents niveaux d'acteurs devrait, dans l'idéal, conduire à des effets optimaux auprès des usagers et bénéficiaires des prestations publiques. Il est alors entendu que la diversité interindividuelle de ces usagers pondère nécessairement ces effets en fonction de leurs caractéristiques et intérêts propres. Cela impliquerait donc qu'il suffise de mener des évaluations auprès des prestataires des politiques publiques pour avoir une bonne idée de la pertinence et de la qualité des actions conduites auprès des usagers et bénéficiaires. Comme un tel raisonnement n'est que théorique, en tous les cas dans le cadre de l'éducation et de l'enseignement, on ne peut faire l'économie d'évaluer la situation en termes de prestations effectives auprès des acteurs/usagers/bénéficiaires, et donc tenir compte de la position de ces derniers. La recherche évaluative couvre en général ces deux niveaux, prenant ses informations auprès des directions d'établissements, que ce soit au niveau de la formation obligatoire ou post obligatoire, auprès des enseignants et des autres professionnels intervenant dans les établissements, et aussi directement auprès des apprenants (élèves, apprentis, gymnasiens) et des parents d'élèves ou d'autres acteurs.

Cultures et politiques de l'évaluation en éducation et en formation

3.3 Les effets des politiques publiques et leur évaluation

L'évaluation des effets des politiques publiques auprès des acteurs concernés est une chose, la prise en compte effective des conclusions des évaluations en est une autre. Pour simplifier, et dans le contexte des éléments présentés ici, on peut toutefois considérer que les évaluations sont demandées pour être prises en compte, même si cela peut n'être que partiellement le cas pour diverses raisons d'opportunités politiques ou sociales que nous avons déjà discutées. Les effets des politiques publiques ne sont pas seulement dépendants des niveaux du système, mais aussi de leur plus ou moins grande prévisibilité dès l'origine du projet ; en d'autres termes, de l'anticipation des effets escomptés au moment de la mise en route des dispositifs. Ce n'est pas parce que des effets n'ont pas été spécifiquement prévus que leur identification ou mesure n'est pas utile, comme en témoigne une partie de nos travaux. L'inconvénient, lorsqu'on relève des effets dont le potentiel n'a pas été clairement circonscrit au départ, c'est qu'il est difficile de savoir s'ils correspondent à un besoin clairement identifié par les différents acteurs du système et s'ils peuvent permettre la mise en œuvre de potentiels changements.

Qu'il s'agisse des niveaux d'impacts, de leur anticipation ou des constats qui en découlent, il est clair que l'évaluation empirique pour l'aide au pilotage pédagogique ne peut se contenter d'avoir produit de nouvelles connaissances sur une situation investiguée ; encore faut-il les relier à des dispositions politiques, stratégiques et opérationnelles. Ces connaissances sont certes des préalables aux impacts souhaités, mais ceux-ci doivent en être distingués en fonction des critères retenus au préalable. Pour qu'une évaluation ait un impact réel sur les dispositifs et les programmes, un tel cheminement doit se faire dans deux directions : celle des producteurs d'évaluations vers les décideurs et praticiens (les utilisateurs) ; celle des utilisateurs vers les évaluateurs. C'est à la démarche de ces acteurs les uns vers les autres que l'on mesure l'intérêt qu'ils prennent à cette rencontre et à la mise en œuvre des résultats de l'évaluation Genelot et Suchaut, 2000 ; Suchaut, 2003, 2005). On touche là aux limites des effets de l'évaluation *stricto sensu*. Dans le cas de l'URSP, la réception de ses travaux et de ses recommandations peut être considérée généralement comme favorable, même si les conséquences qui en sont tirées lui échappent par définition. C'est à ce titre qu'une unité d'évaluation des systèmes de formation fondée sur des données empiriques, favorisant un débat serein et constructif, est pertinente en tant que structure d'interface entre différents niveaux d'interlocuteurs, pour lesquels elle constitue un outil média (et non immédiat) d'aide à la décision. Son rôle de tiers par rapport à ces différents acteurs est à ce titre primordial.

A cet égard, précisons encore que la place de l'organe d'évaluation par rapport aux politiques publiques qu'il évalue joue un rôle concernant son approche de la problématique et aussi relativement à la réception de ses évaluations. En d'autres termes, le fait que cet organe soit intégré à l'administration qui met en place ces politiques ou qu'il soit externe, temporairement mandaté pour telle ou telle évaluation, ne peut pas être ignoré puisque cette « intériorité » *versus* « extériorité » présente dans chaque cas des avantages et des inconvénients. Sans nous étendre sur cet aspect dans le cadre de notre contribution, on peut noter que l'extériorité présente l'avantage de l'impartialité à priori, bien que celle-ci ne soit que relative de la part d'un mandataire face à un mandant qui le choisit en fonction d'attentes plus ou moins explicites et précises. Cet avantage reste cependant très important face aux divers acteurs qui peuvent considérer que les résultats délivrés par un organe interne à l'administration le sont par une instance qui n'est pas neutre. Cette même extériorité présente cependant des inconvénients, en ce sens qu'elle est généralement ponctuelle et ne constitue pas une mémoire de l'évolution de la politique publique évaluée, n'en comprenant généralement qu'une partie des tenants et aboutissants sur la durée. De façon réversible, l'intériorité d'un organe d'évaluation pérennisé présente donc l'avantage d'une telle continuité et de la mémoire des mises en œuvre des politiques publiques, mais l'inconvénient de favoriser le fait que l'organisme qui évalue puisse être considéré comme juge et partie.

Cultures et politiques de l'évaluation en éducation et en formation

4. Les acteurs concernés, leurs intérêts et leurs regards

De façon tout à fait générale et simplifiée, mais correspondant à une réalité perçue par les différents acteurs de l'éducation et de la formation, on peut référer ces derniers aux trois mondes de la politique et de l'administration d'abord, des praticiens en relation avec les apprenants ensuite (enseignants et autres intervenants), et des chercheurs pratiquant l'évaluation enfin. Diverses difficultés peuvent rendre peu lisible l'influence de la recherche évaluative sur le cours des recadrages et des transformations des dispositifs qui sont opérés par les décideurs et les praticiens. Mentionnons en particulier le fait que ces changements interviennent de plus en plus de façon continue, ou en série, et non plus comme autrefois de façon ponctuelle, lors de grandes réformes, entrecoupées de périodes sans changements notables (Abbet, 2008). On ne peut ignorer non plus que les objets liés à l'éducation et à la formation sont la plupart du temps complexes, ni que les aspects d'opportunité sociale, politique, légale ou démocratique peuvent reléguer les résultats des recherches évaluatives centrées sur les aspects pédagogiques au second plan, ou à des jours meilleurs (*ibidem*).

Une méta analyse de la situation dans les champs de l'éducation et de la santé à l'école, portant sur des enquêtes de différents types (Abbet, 2008) a en outre montré que le statut de la demande d'évaluation, de façon ici tout à fait contre-intuitive, n'était pas déterminant. Si cette demande est explicite, mandatée pour un objectif localement identifié, que ce soit pour connaître des besoins ou pour évaluer un dispositif, on peut en effet s'attendre à priori à ce que la mise en œuvre des recommandations soit effective, alors que l'on s'attendra au contraire à un moindre investissement des pouvoirs publics dans ce sens lorsqu'ils se contentent de soutenir financièrement un projet d'évaluation dont ils ne sont pas à l'origine. Or, en fonction d'autres éléments d'appréciation tels que ceux mentionnés précédemment, une telle distinction peut devenir très relative, voire même inopérante. Il ressort des suivis de mise en œuvre de ces différents types de recherche qu'un problème à la fois plus général et plus prégnant peut être mis en évidence : celui relatif à la vision et au mode d'action des divers acteurs concernés. Que ce soit par rapport à la vision des tâches dans la durée ou à la façon de les aborder au quotidien, on peut dire du point de vue épistémique qu'il y a une faible concordance entre les modèles de fonctionnement de la recherche, du travail administratif et politique et de la pratique quotidienne des acteurs du terrain. Encore que le praticien puisse planifier son travail sur l'année scolaire, ou que le politicien ou le cadre administratif ne puisse mettre en route et faire aboutir un projet de loi sur quelques mois seulement, tout comme le chercheur ne peut normalement le faire pour ses projets de recherche. On peut plutôt penser qu'il s'agit d'habitudes différentes concernant le fait de se voir agir dans la durée et dans la façon de considérer les autres acteurs. Il est par exemple habituel de considérer que les politiques ont besoin de réponses claires et opérationnelles, alors que les chercheurs peuvent rendre compte de la complexité et d'intérêts divergents dans le cadre de leurs évaluations (Cros et Bon, 2006). Ces regards différents, ajoutés aux autres contraintes déjà mentionnées, ont donc des conséquences importantes puisque, quel que soit le type d'évaluation, on note qu'il faut une claire volonté ou des circonstances particulières pour faire correspondre les résultats des évaluations avec des mises en œuvre concrètes. Lorsque cette volonté ou ces circonstances font défaut, la prise en compte des résultats de la recherche évaluative visant des éléments de réforme institutionnelle ou des ajustements de pratiques peut devenir incertaine. Notons cependant qu'une telle situation est propre à la recherche évaluative dans le cadre des sciences humaines et sociales, contrairement à ce qui prévaut dans le cas des sciences techniques ou de la nature, dont les résultats sont au contraire instamment demandés pour être pris en compte.

En dépit de telles difficultés, on espère que les apports de la recherche puissent constituer *a minima* un outil de médiation et de mise en perspective pour les mondes des décideurs et des praticiens. Ce rôle de médiation par les résultats de la recherche permet en effet d'aller au-delà de la seule référence à l'expérience personnelle du praticien ou du décideur, de dépasser la seule comparaison intuitive avec des situations dont la comparabilité est plus ou moins fondée et, surtout, de démonter les préjugés ou les présupposés des uns et des autres. De façon complémentaire, la pratique de la décentration et de la prise de recul propre au chercheur pourrait

Cultures et politiques de l'évaluation en éducation et en formation

dans l'idéal devenir un bagage commun entre les différents acteurs des systèmes éducatifs, mais il s'agit là d'un travail difficile et exigeant, à reprendre constamment, pour lequel le chercheur a certainement un rôle à jouer, au moins dans le cadre de la recherche appliquée ou évaluative. La plupart du temps les chercheurs, lorsque cela est souhaité, sont prêts à le faire dans le prolongement des résultats de leurs travaux. Il ne faut cependant pas sous-estimer le fait que l'attitude critique qui fonde en particulier la décentration ne s'accorde pas facilement avec la position du décideur ou du praticien, qui souhaiteraient pouvoir se baser sur des évidences et des faits incontestables (Abbet, 2008). De façon plus abstraite, c'est sans doute ce qu'il faut entendre lorsque Vianou (2005) prône une conception systémique de la relation entre chercheurs et décideurs. N'oublions pas que, dans les milieux de la recherche, la controverse et le débat font partie intégrante de la démarche scientifique, puisqu'un résultat est « vrai » jusqu'à ce qu'on démontre qu'il est « faux ». La discussion des résultats est ainsi considérée comme une vertu dans le monde scientifique, puisque les faits sont par nature contestables, alors qu'elle est considérée comme une faiblesse dans le monde des décideurs.

5. Tension et articulation entre les intérêts des acteurs de l'évaluation

Si les évaluations externes basées sur des données empiriques et utilisant une méthodologie rigoureuse présentent une certaine garantie d'objectivité et produisent des résultats généralement fiables, leur appropriation par les acteurs du monde éducatif ne va pas de soi. L'attrait pour les conclusions des recherches peut en effet être très variable selon la catégorie d'acteurs considérée et selon la nature des résultats sachant que, par ailleurs, ces résultats prennent toujours place dans un contexte politique et social donné, celui-ci pouvant par nature varier au fil du temps. Il devient alors intéressant d'examiner les conditions selon lesquelles les intérêts des acteurs s'articulent, se rejoignent, divergent ou même s'opposent. Cela peut permettre d'apprécier les apports et les limites des recherches évaluatives, mais aussi les tensions que leurs résultats peuvent produire, sachant que les référentiels d'interprétation sont différents selon le groupe d'acteurs considéré parmi les trois que nous avons schématiquement retenus. Il est aisé de percevoir que, selon les acteurs considérés, les produits de la recherche évaluative ne seront pas considérés, appréciés, interprétés et utilisés de la même manière. Cela dépend en effet des référentiels mobilisés (Figari, 1995) et des paradigmes auxquels on peut se référer en matière d'évaluation (De Ketele, 1993). Ainsi, les résultats des recherches prendront un sens différent selon que l'on soit un administrateur (référentiel de gestion), un responsable des autorités du système (référentiel politique) ou un enseignant (référentiel pragmatique).

Le référentiel du chercheur peut tout d'abord être caractérisé par sa nature scientifique, avec une centration sur les questions méthodologiques et théoriques en référence avec le milieu de la recherche. Pour le praticien, l'enseignant par exemple, le référentiel d'interprétation est plus pragmatique et de nature pédagogique, puisque l'intérêt pour la recherche en éducation et ses résultats est principalement lié à la pratique professionnelle en termes d'apports didactiques ou dans le domaine de la gestion de la classe au sens large ; tout ce qui pourra contribuer à mieux répondre aux besoins des élèves et, plus largement, à améliorer les conditions d'enseignement. Enfin, pour le décideur, le référentiel est clairement politique, avec une situation toutefois complexe dans la mesure où l'intérêt des recherches est en partie déterminé par le contexte temporel. En effet, au-delà de l'intérêt même des résultats des travaux pour le pilotage, ceux-ci peuvent être lus et utilisés différemment du fait d'évènements particuliers. On se propose d'identifier de manière schématique différents cas de figure possibles illustrant la tension en termes de divergence dans le degré d'importance que revêt la recherche évaluative pour les trois groupes d'acteurs. Précisons que le terme générique d'« importance » peut recouvrir des sens différents (intérêt, utilité, pertinence, satisfaction,...) en fonction des situations analysées, mais il traduit globalement l'apport que peut avoir l'évaluation pour les acteurs dans leur contexte professionnel. Les graphiques suivants illustrent trois situations type possibles relatives à notre questionnement sachant, évidemment, que la réalité rend compte d'une plus grande diversité.

Cultures et politiques de l'évaluation en éducation et en formation

Un premier cas de figure représente la situation idéale selon laquelle les trois groupes d'acteurs mobilisent un intérêt élevé pour l'évaluation et ses résultats (figure 1). Le chercheur a ainsi la conviction que son travail apporte une contribution pertinente pour la connaissance de l'objet évalué, sur les plans théorique ou méthodologique, ou même simplement concernant l'originalité des résultats produits, ceux-ci pouvant alors être repérés par la communauté scientifique et faire l'objet de débats. Le praticien, quant à lui, considère que les résultats peuvent constituer un apport pour son métier et qu'il pourrait notamment mettre en pratique certaines conclusions mentionnées dans l'étude. L'intérêt du décideur peut se lire globalement dans la pertinence de la recherche en matière d'aide à la décision ou de justification de l'action politique. Certaines questions de politique éducative abordées par la recherche empirique ont pu, dans certains contextes, réunir l'intérêt de tous les acteurs considérés ici. On peut évoquer la question des modes de groupements des élèves dans les écoles et plus spécifiquement la taille de la classe. Au niveau politique tout d'abord, cette question, du fait de son fort impact sur les coûts, recueille toujours un fort intérêt de la part des décideurs (Meuret, 2001). Très souvent étudiée dans les recherches, la relation entre la taille de la classe et les progrès des élèves a fait l'objet de travaux relativement récents utilisant des méthodologies originales et robustes (Finn, 2002 ; Piketty, Valdenaire, 2006 ; Bressoux, Lima, 2011). Ces recherches ont permis d'avoir une connaissance plus fine du phénomène et des preuves plus nettes sur les bienfaits d'une forte réduction de la taille de la classe sur les progrès des élèves, notamment sur les plus faibles d'entre eux. Pour les enseignants et leurs représentants, il est facile de comprendre que les recherches mettant en évidence les effets positifs de la réduction de la taille de la classe soient positivement perçues et soulèvent un intérêt, puisque les résultats produits sont directement en lien avec l'amélioration de leurs conditions de travail.

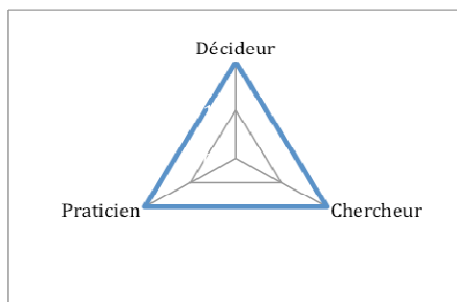


Figure 1

Un autre cas de figure possible (figure 2) est celui pour lequel l'évaluation présente un intérêt important pour le chercheur alors que le praticien ne perçoit dans les résultats qu'une utilité limitée et que le décideur ne voit pas l'intérêt de la recherche, pour diverses raisons possibles. Dans certains cas, la recherche explore des domaines ou des questions qui ne sont pas à l'ordre du jour au niveau de la politique éducative ou dont les résultats ne sont pas directement opérationnels pour les acteurs. On peut penser à certaines questions abordées par la psychologie cognitive touchant aux apprentissages des élèves. Les modélisations proposées dans certains de ces travaux sur l'apprentissage de la lecture ne débouchent pas immédiatement sur des préconisations concrètes pour le décideur et le praticien alors qu'elles sont importantes. On peut aussi citer des champs nouvellement explorés comme celui des compétences sociales développées au sein des systèmes de formation (Abbet et Moreau, 2012). Pour le chercheur, il y a un intérêt évident à travailler dans un domaine en plein développement, surtout si celui-ci se situe dans un champ où les besoins sont avérés. En revanche, l'attrait du décideur pour ce type de recherche peut être très limité, voir inexistant, notamment s'il n'a pas encore perçu les besoins ressentis ou exprimés par d'autres acteurs. Il aura alors plutôt tendance à mobiliser des caractéristiques plus traditionnelles des résultats scolaires (connaissances et compétences disciplinaires) pour instruire la politique. Il faut donc du temps pour que les nouvelles connaissances produites par les travaux puissent être intégrées dans des mesures ou dispositifs institutionnels. Quant au praticien, s'il peut percevoir l'intérêt de ce domaine de recherche, du fait notamment de sa confrontation quotidienne aux comportements des élèves, il ne voit pas forcément encore comment intégrer concrètement une

Cultures et politiques de l'évaluation en éducation et en formation

telle dimension dans son enseignement, en parallèle ou simultanément au travail sur les compétences académiques.

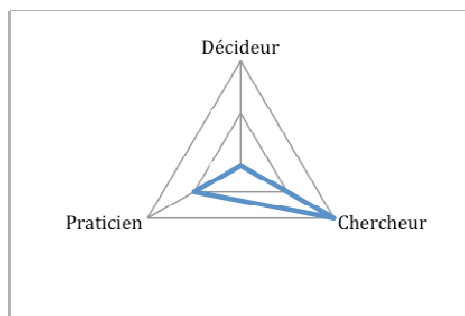


Figure 2

Une troisième situation générique (figure 3) est celle pour laquelle le décideur manifeste un grand intérêt pour les résultats produits, les considérant comme pertinents pour son action. En même temps, le chercheur ne voit pas dans les résultats un apport particulier dans le champ considéré et ne considère pas cette recherche comme efficiente au regard du temps mobilisé. Le praticien est dans la même posture que dans le cas précédent en manifestant un intérêt limité pour les résultats. Comparativement au contexte anglo-saxon, il est sans doute plus difficile de mobiliser des exemples pour ce cas de figure dans le contexte francophone, car la prise en compte des conclusions de la recherche en éducation y est moins systématique. On peut toutefois relever que les études qui touchent au pilotage du système sur la base d'indicateurs (les résultats des élèves notamment) peuvent potentiellement se rattacher à ce cas de figure. Cela peut aussi être le cas de dispositifs pédagogiques pour lesquels le décideur a besoin d'une évaluation spécifique, lui permettant de prendre une décision en termes de reconduction ou d'adaptation. Notons cependant que ce dernier cas de figure peut conduire les enseignants, par exemple, à être sensibles à l'impact d'une telle décision sur leurs pratiques, et donc à l'évaluation qui a permis de la fonder.

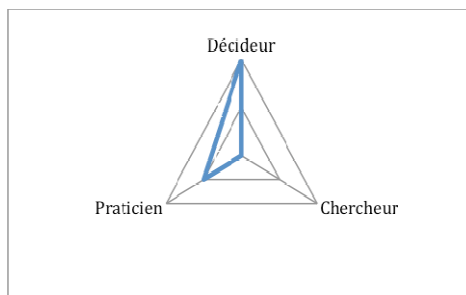


Figure 3

Nous n'ignorons pas que la réalité dépasse bien sûr de loin ces trois illustrations et il serait sans doute intéressant d'identifier et d'analyser de manière plus approfondie les causes qui contribuent à expliquer les variations des degrés d'intérêt des acteurs. Une telle analyse relève toutefois d'un travail spécifique, sans doute utile à mener dans l'espace romand de la recherche en éducation.

6. Conclusion

L'évaluation est nécessaire pour aider au pilotage des systèmes de formation. Les évolutions récentes concernant le développement des politiques publiques vont généralement dans le sens d'en rendre compte, dans un louable souci de transparence. L'idée même d'évaluation, et ses diverses déclinaisons dans le cadre de ces politiques, permet de mettre en évidence ce besoin de

Cultures et politiques de l'évaluation en éducation et en formation

« rendre compte » qui vise à partager de nouvelles informations et, au bout du compte, de « piloter » en connaissance de cause. Ces actions doivent tenir compte de la multiplicité des acteurs et des enjeux et ne peut seulement venir « d'en haut », des décideurs, même si ce sont souvent ces derniers qui donnent les impulsions pour mettre en œuvre une évaluation. Dans une perspective de mise en œuvre efficiente, parfois de mise en œuvre tout court, il y a un intérêt prépondérant à ce que les divers acteurs concernés, en particulier ceux des mondes politique et administratif d'un côté, éducatif et pédagogique de l'autre, puissent partager les enjeux de la formation sur la bases des évaluations proposées et de leurs résultats. Or ces acteurs ont souvent une visée, des perceptions et pratiques à ce point différentes que leurs interactions sont rendues extrêmement difficiles. Dans cette situation, le chercheur, ou évaluateur, peut jouer un rôle de médiation et d'aide à la décentration pour autant qu'il y investisse un intérêt au-delà des seuls résultats obtenus. La plupart des chercheurs du domaine de l'éducation sont prêts à jouer ce rôle si les autres acteurs le leur demandent, ce qui est loin d'être toujours le cas.

Notons finalement que, du point de vue de la recherche à strictement parler, les aspects méthodologiques associés à l'évaluation de dispositifs de politiques publiques sont à considérer avec toute la rigueur nécessaire. L'évaluation consiste souvent à déterminer si les améliorations éventuellement constatées sur les bénéficiaires peuvent être réellement le fait de la mesure mise en place (Givord, 2010). Pour diverses raisons de nature méthodologique, cette relation n'est pas toujours aisée à établir et doit faire l'objet de toute l'attention voulue.

7. Références bibliographiques

- Abbet J.-P. et Moreau, J. (2012). *Les compétences sociales et leur place dans la formation. Résultats d'une enquête auprès des élèves, apprentis et gymnasiens vaudois, ainsi que de leurs enseignants*. Lausanne : URSP, 12.1.
- Abbet J.-P. (2008). *Enquêtes dans les champs de la santé et de l'éducation à l'école. Types de recherches et acteurs concernés : quels résultats pour quels suivis ?* Lausanne : URSP.
- Abbet J.-P. (2003). Enjeux transversaux à l'éducation : quelles partitions pour quelle musique ? Symposium : Quelles régulations apporter à un système pédagogique en innovation ? *Actes du 15^e colloque international ADMEE/SSRE 2002*, Lausanne.
- Bressoux P., Lima L. (2011), La place de l'évaluation dans les politiques éducatives: le cas de la taille des classes à l'école primaire en France. *Raisons éducatives, gouverner l'éducation par les nombres? Usages, débats et controverses*, 99-123.
- Cros F. et Bon A. (coord.) (2006). Les évaluations en éducation au niveau international : impacts, contradictions, incertitudes. *Actes du séminaire international « L'évaluation au service de la qualité en éducation : pratiques et enjeux »*, CIEP, Sèvres.
- De Ketele J.-M. (1993). L'évaluation conjugée en paradigme. *Revue française de pédagogie*, N°103, 59-80.
- Felouzis G., Charmillot S. et Fouquet-Chauprade B. (2013). Comment organiser l'enseignement secondaire obligatoire ? Une politique publique et son analyse dans le canton de Genève, *Swiss Journal of Sociology*, 39 (2), 225-243.
- Figari G. (1995). *Evaluer: quel référentiel?* Bruxelles : De Boeck.
- Finn J.D. (2002), Small classes in American schools : research, practice, and politics. *Phi Delta Kappan*, 83(7), 551-560.
- Genelot S., Suchaut B. (2000) L'évaluation des innovations pédagogiques : quelles modalités de coopération entre les différents acteurs ? *Mesure et évaluation en éducation*, Vol. 23, N°1, 23-42.
- Givord P. (2010), *Méthodes économétriques pour l'évaluation de politiques publique*. Série des documents de travail de la Direction des Études et Synthèses Économiques de l'INSEE.
- Meuret D. (2001), *Les recherches sur la réduction de la taille des classes*. Rapport pour le Haut Conseil de l'Évaluation de l'École.
<http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/024000197/0000.pdf>

Cultures et politiques de l'évaluation en éducation et en formation

- Piketty T., Valdenaire M. (2006), L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français. *Les dossiers Evaluations et Statistiques*, N° 173.
- Piquée C., Suchaut B. (2002). Les opinions des acteurs confrontées à l'évaluation externe : le cas d'une action d'accompagnement scolaire au CP. *Les Sciences de l'Education pour l'ère nouvelle*, Vol. 35, N°3. 103-126.
- Suchaut B. (2003). De la nécessité d'évaluer les pratiques enseignantes : entre enjeux sociaux et obstacles méthodologiques. *Les Dossiers des sciences de l'éducation*, Presses Universitaires du Mirail, N°10, 17-30.
- Suchaut B. (2005). L'évaluation des dispositifs éducatifs et de formation. *Table ronde au colloque ADMEE, 24-26 octobre 2005 « Comment évaluer ? Outils, dispositifs, acteurs »*, I.U.F.M. Champagne-Ardenne, Reims.
- Suchaut B. (2008). Evaluation externe et recherche empirique en éducation : quels apports pour l'amélioration de l'école primaire ? *Les sciences de l'éducation : histoire, débats, perspectives*, 1 (2^e semestre), 89-104.
- Vianou K. (2005). De la recherche en éducation à la décision politique : l'expérience du PASEC en matière de communication pour favoriser l'utilisation de la recherche. Communication à l'Atelier de Ouagadougou sur l'éducation.