

La recherche évaluative

Christian Depover

Université de Mons-Hainaut (Belgique)

Thierry Karsenti

Université de Montréal

Vassilis Komis

Université de Patras (Grèce)

Résumé

Introduction

1. Peut-on concilier recherche et évaluation ?
2. Les formes de recherche évaluative
3. Validité et complémentarité des approches en matière de recherche évaluative
4. La recherche évaluative orientée vers la conception (*design based research*)
5. La recherche évaluative orientée vers la prise de décision

Conclusion

Activités d'appropriation

Concepts importants

Lectures complémentaires

Résumé

Ce texte présente la recherche évaluative comme un mode de recherche à part entière, notamment lorsqu'elle est placée dans une double perspective de rigueur et de production de connaissances. Il a également pour objet de démarquer la recherche évaluative par rapport à certaines formes d'évaluation. Deux modalités de recherche évaluative sont distinguées. La première est orientée vers l'amélioration d'un programme ou d'un dispositif alors que la seconde vise plutôt à généraliser ou à poursuivre un programme mis en place.

Introduction

Comme le fait remarquer Pinch (2009), la recherche évaluative doit occuper une place importante parmi les formes de recherches susceptibles d'être mobilisées en éducation. En effet, contrairement à d'autres domaines scientifiques qui se consacrent surtout à la description des phénomènes, l'éducation est avant tout une science de l'action. Or, une action, pour être efficace et se donner un maximum de chances d'atteindre son but, doit être régulée. Cette régulation exige le recours à l'évaluation sous différentes formes.

Lorsque l'évaluation comme moyen d'améliorer l'action est menée d'une manière systématique à partir de données empiriques, certains auteurs comme Patton (1990) ou Jones (2000) parlent de recherche évaluative, ou encore de recherche d'évaluation. Toutefois, pour considérer qu'il s'agit d'un processus de recherche au sens strict, il convient d'ajouter une seconde condition, à savoir que les techniques mises en œuvre conduisent également à la production de nouvelles connaissances susceptibles d'enrichir, de compléter, de nuancer ce que l'on sait à propos d'un domaine.

Placée dans cette double perspective de rigueur et de production de connaissances, l'évaluation peut, comme nous allons tenter de le montrer dans ce texte, être considérée comme une méthode de recherche à part entière.

1

PEUT-ON CONCILIER RECHERCHE ET ÉVALUATION ?

Entre les processus mis en œuvre en matière de recherche et d'évaluation, il existe un grand nombre de similitudes, notamment en matière de récolte

et de traitement des informations ainsi que d'interprétation des résultats. Sur ces aspects, la différence est surtout une question d'appréciation personnelle. Ainsi, on se montrera généralement plus exigeant quant à la qualité des données, la fiabilité des outils utilisés, la rigueur de l'interprétation lorsqu'on aura choisi de s'inscrire dans un processus de recherche plutôt qu'un autre. Une différence plus fondamentale apparaît toutefois lorsqu'on s'intéresse aux résultats produits, puisque, pour considérer qu'il s'agit de recherche au sens plein du terme, il faut que le processus mis en œuvre conduise à la production de nouvelles connaissances qui puissent être considérées comme originales. Pour satisfaire cette dernière exigence, il est particulièrement important que la recherche évaluative s'appuie sur une analyse approfondie de la littérature et que, sur cette base, soient formulées des questions de recherche précises et détaillées, ce qui est plus rarement le cas en matière d'évaluation.

La production de résultats originaux qui caractérise la recherche va souvent de pair avec la liberté laissée au chercheur de choisir ses orientations de recherche. Or, en matière de recherche évaluative, cette liberté peut être entravée par le fait que souvent cette forme de recherche répond à une commande et que, dès lors, le choix des questions auxquelles elle s'intéressera risque d'être, plus ou moins profondément, infléchi par la demande de l'organisme qui financera les travaux.

Le fait que la recherche soit menée en réponse à une commande, plus ou moins explicite, soulève d'autres difficultés, notamment liées au calendrier d'exécution. En effet, une recherche évaluative commanditée par une autorité se voit obligée de fournir une réponse dans un délai plutôt bref, de manière que le commanditaire puisse prendre une décision à l'échéance fixée, par exemple généraliser certaines pratiques, poursuivre une mise à l'essai, adopter de nouveaux outils, etc. Or, un calendrier de recherche s'inscrit plutôt dans le moyen, voire le long terme, car la production de connaissances originales comporte une part d'aléas qui s'accommode fort mal d'un calendrier trop rigide.

Le rapprochement entre recherche et évaluation, que l'on observe à l'occasion de la recherche évaluative, permet d'inscrire la recherche dans l'action. En effet, les recherches servent souvent à éclairer une problématique, à mieux comprendre un phénomène ou à démontrer l'efficacité d'une méthode, mais elles n'aboutissent que rarement à une décision qui influe directement sur les pratiques éducatives. En matière de recherche évaluative, au contraire, il existe souvent un lien étroit entre l'objet de la recherche et la décision à prendre. Ce lien est assez évident lorsqu'il s'agit d'une recherche faisant suite à une demande externe, mais c'est aussi le cas, comme nous le montrerons, lorsque la recherche répond au choix d'une équipe de développement d'associer une évaluation systématique à la conception d'un nouvel outil ou d'un nouveau programme. Lorsqu'il s'agit d'élaborer des outils et des programmes, la recherche évaluative

ne peut se limiter à apprécier globalement leur efficacité ou leur pertinence, elle doit s'attacher, à chacune des étapes de leur élaboration, à fournir les informations qui permettront d'éclairer les décisions de l'équipe de conception.

Ce lien entre recherche et action conduit, parfois, à évoquer le terme de *recherche appliquée* à propos de la recherche évaluative. Parler de recherche appliquée amène à distinguer ce type de recherche d'autres formes qui seraient à caractère plus fondamental, avec le risque de faire prédominer la première sur les autres, du fait qu'il est généralement plus facile de susciter une mobilisation pour des projets qui apporteront des réponses à court terme aux problèmes imminents d'ordre éducatif que pour des projets à long terme aux résultats parfois plus incertains.

C'est ainsi qu'on observe, au grand regret de certains chercheurs qui y voient une forme de limitation de leur autonomie, une tendance de plus en plus affirmée de la part des agences de financement (UNESCO, Banque Mondiale, ACIDI, etc.) à privilégier la recherche appliquée et notamment la recherche évaluative plutôt d'autres formes de recherche qui valoriseraient davantage la liberté, la créativité et surtout l'indépendance du chercheur.

2

LES FORMES DE RECHERCHE ÉVALUATIVE

2.1 La recherche évaluative visant à améliorer un programme ou un dispositif (orientée vers la conception)

Cette forme de recherche s'effectue généralement durant l'action afin de permettre un meilleur ajustement aux buts recherchés. Elle fait appel à des personnes externes à la conception et à l'exécution de l'action, mais peut aussi mobiliser des acteurs qui prennent une part active dans l'action.

Lorsque la recherche évaluative concerne la mise au point d'outils ou de programmes, on parle généralement de «recherche conception» (*design research*). Il peut s'agir, par exemple, de compléter l'élaboration d'un logiciel éducatif par une série de mises à l'essai rigoureuses dont les résultats auront pour effet d'infléchir la mise au point de l'outil et d'accroître les chances d'aboutir à un produit qui satisfasse les différentes parties concernées.

En matière de programme, on peut s'intéresser à la mise au point d'une méthodologie d'enseignement particulière, mais aussi à des interventions à orientation sociale ou sanitaire visant, par exemple, à réduire la propagation du VIH/SIDA par une campagne d'éducation bien ciblée. La recherche évaluative qui vise à l'élaboration de programmes s'inscrit généralement dans

le cadre d'une politique publique plus globale, comme réduire l'échec scolaire, pour le premier exemple, ou améliorer l'accès aux soins préventifs, pour le second.

À la suite de Brown (1992), un certain nombre de chercheurs s'intéressant essentiellement à l'enseignement des sciences ont utilisé l'expression *design research* pour caractériser une approche de recherche qui consiste à soutenir la mise au point de dispositifs pédagogiques par des prises d'informations systématiques. Comme le rappellent Collins et coll. (2004), le but de la recherche centrée sur le design est d'améliorer la manière dont un dispositif fonctionne dans la réalité (plutôt qu'en laboratoire) et, lorsqu'un problème apparaît, d'analyser son origine pour remédier aux problèmes dans le cadre d'un processus de raffinement continu.

2.2 La recherche évaluative visant à valider un programme ou un dispositif (orientée vers la décision)

Il s'agit d'une forme de recherche qui survient nécessairement à la fin de l'action ou une fois le dispositif terminé et qui est généralement confiée à une personne ou à une organisation externe qui n'est pas directement engagée dans l'élaboration de l'action ou du dispositif.

La recherche évaluative de ce type répond généralement à une commande des autorités et est donc directement liée aux questions qu'elles se posent à un moment déterminé ou aux décisions qu'elles seront amenées à prendre à court terme.

Elle dépend souvent d'enjeux politiques, en particulier lorsque les résultats de la recherche contribuent à fonder une décision qui peut avoir un impact sur les orientations en matière de politique éducative à l'échelle locale, régionale ou nationale. C'est le cas notamment lorsqu'il s'agit de valider une nouvelle approche pédagogique sur la base d'un échantillon d'établissements avant de généraliser la réforme à l'ensemble du territoire national.

Ainsi, certaines recherches évaluatives pourraient, par exemple, contribuer à orienter une politique en fournissant des informations importantes pour déterminer quels programmes mériteraient d'être soutenus. Toutefois, comme le soulignent pertinemment Cronbach et coll. (1981), la véritable mission de l'évaluation n'est pas de décider à la place des autorités, mais plutôt d'éclairer et d'informer tous les citoyens dans le cadre d'un processus démocratique de prise de décision.

La recherche évaluative peut aussi porter sur la comparaison entre plusieurs modalités de mise en œuvre d'une action, comme lorsqu'il s'agit de choisir entre deux méthodes d'apprentissage de la lecture celle dont il est préférable de préconiser l'usage auprès des enseignants de première année primaire.

3

**VALIDITÉ ET COMPLÉMENTARITÉ
DES APPROCHES EN MATIÈRE
DE RECHERCHE ÉVALUATIVE**

La recherche évaluative mobilise aussi bien des méthodologies qualitatives centrées sur la compréhension en profondeur des phénomènes (approche inductive) que des méthodologies quantitatives orientées vers la mise en évidence de relations de causalité (approche déductive). Pour Pinch (2009, p. 393), il est nécessaire de « considérer une variété de modèles, d'approches, de méthodologies et de sources de données », et ce, en fonction des objectifs de la recherche évaluative menée. Pour évoquer les qualités attendues de la recherche évaluative, nous serons donc amenés à faire référence à ces deux méthodologies.

Le contrôle de la validité dans une recherche quantitative fait appel à des procédures bien connues et largement traitées dans le présent ouvrage (voir le chapitre 7). Tout d'abord, dans le cas particulier de telles recherches évaluatives, il s'agit d'assurer les conditions qui permettront d'attribuer la causalité des phénomènes aux variables indépendantes mises en évidence dans le plan expérimental. Ainsi, si l'on veut établir que la disponibilité des manuels scolaires améliore les performances d'apprentissage, il faut être certain qu'aucune autre cause ne peut expliquer ces résultats. Pour obtenir cette certitude, il convient d'être attentif à la représentativité des échantillons constitués et à la pertinence du plan expérimental adopté.

Rappelons que la manière d'établir la validité d'une recherche est assez différente dans une méthodologie qualitative (voir Patton, 2002) puisque, dans le cadre de celle-ci, il ne s'agit plus de manipuler la situation conformément à un plan expérimental défini, mais d'appréhender la dynamique des événements et de mettre en évidence leur intentionnalité à partir d'observations effectuées le plus souvent en contexte naturel.

Pour cette raison, la recherche qualitative (voir le chapitre 6) insistera, pour établir la validité de ses résultats, sur des critères comme l'honnêteté, la richesse et la profondeur des données récoltées, ou encore elle mettra en avant l'importance de la triangulation des informations prélevées sur le terrain. Il n'en demeure pas moins, comme l'indiquent Patton (2002) ou encore Pinch (2009, p. 393), que les chercheurs en éducation se « doivent d'être plus ouverts à la recherche qualitative et aux techniques de collecte et d'analyse de données qui les sous-tendent ».

Des auteurs comme Gorard et Taylor (2004) insistent sur l'intérêt, en matière de recherche éducative, de combiner les approches en recourant au principe de la triangulation méthodologique afin d'articuler recherches

quantitatives et qualitatives. Dans une approche de ce type, on ne se contentera pas, pour étudier un phénomène, de mobiliser à la fois des données qualitatives et quantitatives, mais on s'efforcera de croiser les méthodologies en injectant notamment des éléments propres à la recherche qualitative, comme la compréhension en profondeur des phénomènes à la faveur de prises d'informations multiples, dans un dispositif expérimental basé sur la comparaison de plusieurs groupes.

Dans un souci d'appréhension la plus complète possible du phénomène étudié, on peut également faire intervenir d'autres formes de triangulation en recherche évaluative, comme la triangulation des observateurs. Celle-ci consiste à utiliser plusieurs chercheurs qui interviennent de manière indépendante pour observer ou analyser une situation, un contexte ou un phénomène particuliers.

Cohen et coll. (2007) parlent de triangulation spatiale en recherche évaluative pour désigner une approche qui permet de dépasser les études menées uniquement dans un contexte ou une culture particuliers en diversifiant les lieux de prise d'information. Cette forme de triangulation est mise en œuvre notamment dans le cadre des études transculturelles ou encore à l'occasion d'une méthodologie d'investigation proposée par Miles et Huberman (2003) sous le nom d'analyse inters-sites, sur laquelle nous reviendrons plus loin dans ce texte.

La triangulation temporelle en recherche évaluative, pour sa part, consiste à prendre en compte l'évolution d'un phénomène, liée notamment à la maturation des sujets par des prises d'informations répétées et plus ou moins distantes dans le temps (études longitudinales).

D'une manière générale, la triangulation dans la recherche évaluative repose sur l'idée d'utiliser plusieurs méthodes de collecte de données pour étudier un phénomène afin de le saisir dans toute sa richesse et sa complexité, mais c'est aussi un moyen d'améliorer la validité interne d'une recherche qualitative en confrontant les données issues de plusieurs sources. Par exemple, une étude sur l'opinion des enseignants à propos d'une réforme peut donner lieu à une enquête au moyen d'un questionnaire écrit dont les résultats seront confirmés et complétés par une série d'entrevues. Des données issues de l'observation du comportement des élèves peuvent être confrontées avec celles issues d'entretiens semi-directifs, etc.

Une difficulté inhérente à la recherche évaluative est liée au fait que le recours à une méthode d'investigation particulière peut conduire le chercheur à se construire une représentation particulière du phénomène à l'étude. C'est comme si la méthode jouait un rôle de filtre à travers lequel l'objet d'étude était examiné. Comme le rappelle Lin (1976), le chercheur doit s'assurer que les données rassemblées ne sont pas de simples artefacts de la méthode d'investigation employée.

Le fait d'utiliser plusieurs méthodes d'investigation et de montrer que ces méthodes conduisent à des données qui convergent dans l'explication d'un phénomène permet de se prémunir contre ce type de distorsion de la réalité.

4

LA RECHERCHE ÉVALUATIVE ORIENTÉE VERS LA CONCEPTION (*DESIGN BASED RESEARCH*)

Cette forme de recherche s'intéresse à la mise au point de programmes ou de dispositifs pédagogiques dans le cadre d'un processus itératif où alternent conception, développement et prise de mesure sur le terrain, de manière à assurer le meilleur ajustement possible de ces programmes ou de ces outils aux besoins des bénéficiaires. Une telle approche est fort populaire dans les milieux préoccupés par la mise au point d'environnements d'apprentissage qui exploitent les TIC, mais aussi par l'élaboration de programmes d'intervention à caractère social ou pédagogique.

À titre d'illustration, nous présenterons succinctement un modèle de design pédagogique sur la base duquel nous travaillons depuis une dizaine d'années dans le cadre de la conception d'environnements d'apprentissage exploitant les TIC (Depover et Marchand, 2002). Ce modèle, qui est de type sériel, en ce sens qu'il décompose le processus de conception d'un environnement d'apprentissage en une série d'étapes strictement définies, prévoit des prises d'information systématiques à l'issue de chacune des étapes du processus. Ces prises d'information font appel à des méthodologies déterminées dont le choix est fonction de la décision à prendre à un moment fixé. Ainsi, à l'issue de la phase de mise au point de l'interface, on entreprendra des études de terrain pour valider la pertinence ergonomique de l'outil en organisant des mises à l'essai sur la base d'un échantillon de sujets issu de la population cible. Au moment du choix de certaines stratégies pédagogiques, on validera ce choix non seulement auprès d'experts, mais aussi en analysant les réactions de certains groupes d'apprenants. Lorsqu'une partie significative du dispositif aura été mise au point, on appliquera des épreuves permettant d'effectuer des mesures d'efficacité pédagogique.

En référence à ce modèle, nous mettrons en évidence deux éléments qui constituent, selon nous, des traits caractéristiques d'une approche de type recherche évaluative. D'une part, le fait de procéder à des mesures répétées tout au long du processus d'élaboration et, d'autre part, l'utilisation des informations issues de ces mesures pour infléchir le processus d'élaboration. Ainsi, contrairement à ce qui a longtemps été pratiqué en matière de design pédagogique, il ne s'agit pas d'attendre que le processus de production soit

arrivé à son terme pour confronter les produits au terrain, mais d'inclure la prise de mesure au sein même du processus d'élaboration. Pour attirer l'attention sur l'importance de la prise en compte des réactions ou des opinions des bénéficiaires, on parle souvent à ce propos de design incrémental, de manière à souligner que chaque étape s'appuie sur les précédentes, mais aussi, qu'à chaque étape, l'ensemble du processus peut être remis en cause en fonction des résultats des essais effectués.

Une autre question importante qui découle directement du design des environnements d'apprentissage est celle de leur implantation dans différents contextes d'apprentissage. Même si on a pris soin de tenir compte des réactions du public tout au long de la mise au point d'un dispositif et que son efficacité pédagogique a été établie dans un contexte particulier, il n'est pas pour autant certain que celui-ci sera bien accueilli au-delà du contexte qui lui a donné naissance.

À cet égard, l'analyse inter-sites proposée par Miles et Huberman (2003) peut utilement être mobilisée en vue d'étendre à d'autres contextes la portée des observations qui ont été le plus souvent effectuées à partir d'un site d'étude particulier. Parmi les variables à prendre en compte au moment du passage d'un site particulier à d'autres sites choisis afin de mieux représenter les différents contextes d'application de l'outil, il importe particulièrement de considérer les difficultés dues au phénomène de résistance à l'innovation. En matière de résistance au changement, Huberman, dans un ouvrage publié par l'Unesco en 1974, dénombre plus de vingt sources de résistance, parmi lesquelles figurent, par ordre d'importance, l'insuffisance de matériel, le manque de préparation des enseignants et la difficulté pour les acteurs de terrain de percevoir les avantages réels de l'innovation proposée.

Collins et coll. (2004) insistent également sur l'importance de prises d'informations multiples dans une recherche orientée vers le design. Ainsi, ils distinguent trois catégories de variables par rapport auxquelles ils estiment important d'effectuer des mesures : les variables de climat (motivation, engagement personnel, coopération...), les variables d'apprentissage (acquisitions, stratégies cognitives et métacognitives...) et les variables systémiques (soutien, stratégies de diffusion et d'adoption, coûts...). Pour chacune de ces catégories de variables, des outils de mesures particuliers seront mobilisés, comme les outils d'observation (grilles, enregistrements vidéo) pour les variables de climat, les tests de rendement scolaire pour les acquisitions et les interviews ou les enquêtes pour les variables systémiques. Pour ces auteurs, la prise d'informations par rapport aux variables systémiques est particulièrement importante, car l'étude d'un dispositif se fait toujours par rapport à un contexte d'implantation particulier, et le fait d'avoir prouvé l'efficacité d'un dispositif dans un contexte déterminé ne garantit en rien celle-ci dans un autre contexte.

Pour ce qui est de la mise au point de programmes d'intervention, l'approche incrémentale que nous venons de décrire est également applicable, dans la mesure où la conception ainsi que l'implémentation du programme respectent un certain nombre d'étapes strictement définies et que les personnes qui prennent part à la démarche de recherche évaluative sont mobilisées dès les phases initiales du projet. Dans le cas contraire, ce sera plutôt en termes de recherche évaluative orientée vers la prise de décision qu'il s'agira d'envisager l'analyse des programmes d'intervention.

En matière de recherche évaluative orientée vers le design, le risque est grand de privilégier l'évaluation et de négliger la recherche. En effet, ce qui justifie la prise d'informations tout au long de l'élaboration d'un outil ou d'une méthode particulière, c'est avant tout la possibilité d'ajuster de manière continue le processus de conception et de production. Pour souligner l'importance de cet aspect, on parle aussi d'évaluation formative, afin de caractériser cette forme d'évaluation. Rien d'étonnant, dès lors, qu'on puisse être tenté, à l'occasion d'un processus de ce type, de faire l'économie d'une réflexion qui permettrait d'inscrire les informations recueillies dans une démarche de recherche plus globale en faisant appel à un cadre conceptuel et à des questions de recherche susceptibles de conduire à la production de résultats originaux.

5

LA RECHERCHE ÉVALUATIVE ORIENTÉE VERS LA PRISE DE DÉCISION

5.1 Une forme de recherche en phase avec la prise de décision

Comme nous l'avons déjà vu précédemment, cette forme de recherche peut, dans certains contextes, reposer sur une démarche quantitative et mettre en œuvre un dispositif expérimental plus ou moins complexe, même si cette approche n'exclut pas que des données de nature qualitative puissent être mises à contribution.

Certains travaux actuellement en cours, qui portent sur le choix des politiques les plus efficaces en matière d'aide au développement, nous paraissent être des exemples intéressants qui illustrent parfaitement l'intérêt porté à ce type d'approche en matière de recherche, mais aussi d'aide à la décision. En quelques mots, il s'agit de choisir entre différentes stratégies d'investissement dans le secteur de l'éducation. Est-ce que le fait d'améliorer l'accès à certaines ressources comme les manuels et les tableaux ou d'accroître la proportion des enseignants améliore les performances scolaires ? Si oui, dans

quelle mesure ? Certaines de ces recherches peuvent conduire à des résultats surprenants, parfois même contraires à ce qu'on a toujours tenu pour acquis en éducation. Ainsi, selon des données rapportées par Banerjee et Duflo (2009), le fait d'augmenter la proportion des manuels scolaires jusqu'à un pour deux élèves ne modifie pas les résultats scolaires, non plus que le fait d'améliorer de manière substantielle le ratio maître-élèves. Par contre, la mise sur pied, au Kenya, d'un programme de santé publique portant sur le traitement des vers intestinaux a permis de réduire l'absentéisme scolaire de 25 %, alors que l'effet de l'engagement d'un maître supplémentaire, ce qui représentait un coût beaucoup plus important, a eu un effet bien moindre sur la participation scolaire. Évidemment, il est important de faire appel aux démarches de la recherche qualitative pour à la fois nuancer et contextualiser de tels résultats.

A contrario, certains usages des technologies en classe, comme le recours au langage LOGO, ont progressivement disparu, faute d'avoir pu démontrer leur efficacité pédagogique. L'ouvrage célèbre de Cuban (2001), *Oversold and Underused*, témoigne de manière remarquable de cette désillusion due à la difficulté de mettre en évidence l'efficacité de certains outils ou de certaines approches pédagogiques mobilisant les TIC.

Les résultats de ce genre vont parfois à l'encontre des croyances intuitives en mettant au jour les répercussions que l'absence de résultats positifs peut avoir sur le maintien de certains usages, voire l'implantation de nouvelles méthodes. Ces résultats qui nous amènent à nous interroger montrent selon nous encore la nécessité d'éprouver, par des recherches évaluatives faisant appel à de multiples méthodologies (qualitatives et quantitatives), le bien-fondé des changements que nous envisageons d'apporter au fonctionnement de l'école, et ce, avant de les généraliser à l'ensemble des classes susceptibles d'en bénéficier.

5.2 Comparer les performances de plusieurs groupes pour tenter d'établir des relations causales

Une autre approche utile pour apprécier l'efficacité d'une action ou d'un dispositif pédagogique consiste à recourir à une mesure *avant-après*, en d'autres termes, à appliquer une épreuve d'évaluation de la performance scolaire avant (prétest) et après (post-test) l'application d'un traitement pédagogique donné (voir le chapitre 7). La mise en évidence de relations causales, qui constitue une des caractéristiques importantes de la recherche quantitative, procède selon une approche déductive sur la base d'hypothèses ou de questions de recherche. Cette forme de recherche fait également appel à la notion de plan expérimental, c'est-à-dire à une manipulation plus ou moins complexe de la situation à partir de laquelle les mesures *avant-après* seront effectuées.

En effet, pour créer les conditions permettant d'établir l'existence d'une relation de causalité entre une variable indépendante (le traitement appliqué) et une variable dépendante (les résultats observés), il est nécessaire de comparer les performances enregistrées par différents groupes d'élèves.

5.3 La mesure des effets des programmes et des politiques

L'un des principaux freins à l'évaluation objective de politiques ou de programmes réside dans la difficulté de préciser, voire de quantifier les effets qui y sont associés. Par exemple, la plupart des politiques en matière d'aide au développement visent d'une manière ou d'une autre à lutter contre la pauvreté et à améliorer le bien-être des populations. Un tel objectif fait aisément consensus sur le plan international, mais que signifie-t-il concrètement et surtout comment peut-on vérifier s'il est effectivement atteint ? Pour pouvoir être évaluée, une politique doit être définie en objectifs clairs et déclinée en indicateurs de suivi qui soient suffisamment mesurables. C'est sur cette base qu'un programme déterminé pourra être mis en œuvre et que son impact possible pourrait être réellement apprécié.

Les effets d'un programme, lorsque les variables à considérer sont complexes ou que leur nombre est important, doivent être évalués au moyen de méthodologies de recherche à la fois qualitatives et quantitatives.

5.4 Une approche plus qualitative basée sur l'implication des acteurs

Une approche proposée dans les années 1990 par Fetterman et ses collaborateurs (Fetterman et coll., 1996) insiste sur la valeur émancipatrice (*empowerment*) des procédures de recherche-évaluation mises en œuvre. Pour Fetterman (2001), l'évaluation émancipatrice est conçue pour « aider les personnes à améliorer les programmes auxquels elles participent en mettant en œuvre une sorte d'autoévaluation » (p. 81). Il s'agit de comprendre ce qui se passe dans une situation donnée à partir de la perspective des participants.

À cette fin, Fetterman (2001) décrit trois étapes. La première consiste à demander aux personnes prenant part à un programme de définir leur propre vision de ce programme avec le soutien d'un spécialiste de l'évaluation. Cette phase se termine habituellement par un atelier où les différentes visions sont mises en évidence. La deuxième étape consiste à décrire les activités à mener dans le cadre du programme et à les classer par ordre de priorité lors d'une activité de groupe. Enfin, la troisième étape consiste à engager le groupe dans un processus de recherche-évaluation basé sur l'autoréflexion et la construction d'une représentation commune. Ainsi, selon Fetterman, le groupe s'organise progressivement sous la forme d'une communauté de pratiques qui partage

un certain nombre de règles et de valeurs en vue d'une tâche à mener en commun. Les personnes qui constituent cette communauté intègrent la logique de l'évaluation en menant à bien leur propre évaluation en suivant un processus décrit par Patton (1997) comme une immersion.

Il est clair que, placée dans une telle perspective, la recherche évaluative dépasse le cadre d'une approche purement diagnostique préparant la prise de décision, mais intègre aussi la possibilité d'agir sur la situation. Par exemple, dans un projet d'éducation sanitaire, le fait d'engager les participants dans une réflexion par rapport au programme conduit à une amélioration de leurs connaissances dans le domaine et les rend plus critiques vis-à-vis des programmes qui leur seront proposés par la suite. Par rapport à cet aspect, cette forme de recherche évaluative s'apparente à une recherche-action dont les effets porteraient en priorité sur les acteurs plutôt que sur d'autres composantes du dispositif.

Certains chercheurs comme Wandersman (2005) inscrivent plus résolument l'évaluation émancipatrice dans une perspective de recherche-action en insistant sur le fait que ses effets peuvent non seulement concerner les personnes, mais aussi les programmes ou les organisations qui ont part. Ainsi, Miller et Lennie (2005) montrent, à travers l'analyse d'un programme soutenu par la Croix-Rouge australienne, que l'approche émancipatrice a permis d'améliorer globalement la qualité du programme ainsi que sa durabilité à long terme.

5.5 Pour conclure à propos de la recherche évaluative orientée vers la prise de décision

Même si les principes décrits dans les sections précédentes peuvent paraître complexes et parfois lourds à mettre en œuvre, ils sont les seuls à garantir des conclusions valides. Qu'on opte pour une approche basée sur des plans expérimentaux, qu'on privilégie l'implication des acteurs, ou encore qu'on combine les approches quantitative et qualitative, il est essentiel de respecter les principes méthodologiques propres à chacune des approches.

Lorsqu'on hésite à se lancer dans une recherche qui risque de s'avérer onéreuse, il faut avoir à l'esprit les conséquences négatives (en matière de coûts, mais aussi de capital humain) des réformes injustifiées ou de l'abandon de réformes qui auraient pu conduire à des progrès considérables pour l'ensemble de la nation.

Beaucoup de pays dans le monde mettent en œuvre des réformes ou l'ont déjà fait pour les harmoniser avec une approche par compétences, mais très peu d'entre eux ont eu la prudence de mettre ces réformes à l'essai et de contrôler rigoureusement les résultats de ces mises à l'essai. Aujourd'hui, on commence à percevoir les conséquences de cette négligence, qui se traduit,

chez certains, par une remise en cause globale de leur système éducatif et, chez d'autres, par un repli frileux sur des approches très classiques qui ont le mérite de rassurer les enseignants et les parents. Pour les autorités responsables de l'éducation, c'est souvent difficile de résister aux modes du moment et de prendre le temps de la réflexion et surtout de la mise à l'essai, au risque de rater le train de la nouveauté et de l'innovation.

Dans le chef du chercheur, le principal écueil à éviter par rapport à ce genre de recherche qui répond généralement à une commande externe, c'est d'accepter, consciemment ou non, l'idée que la recherche puisse être au service des autorités et de se plier ainsi trop facilement aux contraintes qu'elles imposent aux équipes de recherche : échéanciers déraisonnables, restrictions sur le choix des méthodologies ou, plus grave, pressions pour que les conclusions aillent dans le sens attendu par les autorités. Quand on s'engage dans une recherche évaluative de ce type, il faut être armé pour résister à des pressions multiples émanant des parties en présence, mais aussi être capable de rédiger un rapport qui ne froissera pas exagérément les susceptibilités de certains sans pour autant masquer une réalité que tous ne sont pas toujours prêts à accepter.

Conclusion

Dans ce chapitre, nous avons essayé de montrer que la recherche évaluative n'est pas une recherche au rabais, mais que, bien au contraire, si le chercheur s'efforce de mettre en œuvre des méthodologies rigoureuses – quelles soient de nature qualitative, quantitative ou mixte – et s'il se ménage un espace de liberté et d'autonomie suffisant pour exprimer sa créativité, cette forme de recherche peut conduire à faire avancer de manière appréciable la compréhension des phénomènes éducatifs.

Toutefois, pour donner à ce type d'approche son statut de recherche à part entière, il est essentiel pour le chercheur de respecter un certain nombre de principes qui garantissent non seulement la validité des résultats qu'il produira, mais aussi leur honnêteté et leur probité.

Plus particulièrement, lorsqu'il s'agit d'une recherche qui répond à une commande et fournit des éléments utiles à une prise de décision, la tentation pour certains d'orienter les résultats risque d'enlever toute valeur à cette forme d'investigation. Même si elles concernent la recherche en général, les questions d'éthique professionnelle occupent une place particulière en matière de recherche évaluative. S'il est vrai que toute vérité n'est pas toujours bonne à dire, il est difficilement acceptable pour un chercheur de ne pas rendre compte, publiquement ou non, des résultats qu'une démarche rigoureuse de recherche a mis en évidence. Si une certaine latitude existe pour le chercheur, c'est seulement à l'étape de l'interprétation des données, à l'occasion de laquelle la responsabilité lui incombera de donner plus ou moins de poids à certains aspects.

C'est aussi au chercheur d'estimer dans quelle mesure il accepte d'être contraint dans sa liberté d'expression. Ainsi, il n'est pas rare que certains commanditaires assortissent leur financement d'une clause de confidentialité qui empêche ou limite la diffusion des résultats par le chercheur mandaté. Si ces clauses sont bien définies au départ, ce sera au chercheur de décider si la commande est compatible avec l'idée qu'il se fait de sa liberté universitaire ainsi qu'avec les règles en vigueur dans son institution. Par contre, si rien n'a été défini au départ et que le commanditaire refuse qu'on publie certains résultats parce qu'ils le desservent, ce serait une faute déontologique pour le chercheur d'accepter le dictat du bailleur de fonds.

Une particularité intéressante de la recherche évaluative réside dans les buts qu'elle poursuit. Ainsi, non seulement elle vise à comprendre, à l'instar de la recherche qualitative (paradigme d'intelligibilité), et à dégager des lois et des principes (paradigme nomothétique), comme le fait la recherche quantitative, mais aussi elle a des visées pragmatiques et politiques en contribuant à orienter des décisions. C'est probablement cette richesse qui fait sa popularité, mais aussi sa principale faiblesse, car, en poursuivant tous ces buts à la fois, il arrive que le chercheur oublie qu'ils ne pourront être atteints que dans la mesure où la rigueur méthodologique sera strictement respectée, ce qui est loin d'être toujours le cas.

Activités d'appropriation

1. Nommez au moins trois critères qui permettent de distinguer recherche évaluative et évaluation.
2. Proposez un plan de recherche évaluative pour vérifier l'efficacité d'une politique éducative. Précisez les étapes à suivre, les actions à mettre en œuvre et les précautions méthodologiques à prendre.
3. Indiquez les points de convergence et de divergence entre :
 - recherche évaluative et recherche-action ;
 - recherche évaluative et recherche appliquée ;
 - recherche évaluative et recherche fondamentale ;
 - recherche évaluative et recherche conception (*design research*).
4. Pour chacun des contextes de recherche décrits brièvement ci-dessous, indiquez quelle forme de recherche évaluative (orientée vers la conception ou orientée vers la décision) vous paraît le mieux convenir (dans certains cas, les deux peuvent faire l'affaire) et justifiez votre réponse.

Situation 1 : Une équipe pluridisciplinaire s'est donné pour objectif de concevoir un logiciel de simulation destiné à des élèves du début de l'enseignement secondaire. L'équipe souhaite prendre toutes les précautions pour que le logiciel produit réponde le mieux possible aux besoins du milieu.

Situation 2 : Les autorités du ministère de l'Éducation envisagent une réforme de grande ampleur des programmes d'études. À cette fin, elles mettent en place une équipe chargée de suivre l'application de la réforme.

Situation 3 : Après trois années de difficultés dues à une modification en profondeur des approches pédagogiques préconisées au début de l'enseignement primaire, le ministre de l'Éducation s'interroge sur l'opportunité de poursuivre la réforme.

Concepts importants

Vous trouverez la définition des mots clés suivants dans le glossaire de cet ouvrage :

- Recherche évaluative orientée vers la conception
- Recherche évaluative orientée vers la prise de décision

Lectures complémentaires

Jones, R. A. (2000). *Méthodes de recherche en sciences humaines*, Bruxelles, De Boeck Université.

Cet ouvrage présente différentes méthodes de recherche en sciences humaines et consacre un chapitre à la recherche d'évaluation.

Miles, M. B., et A. M. Huberman (2003). *Analyse des données qualitatives*, Bruxelles, De Boeck Université.

Cet ouvrage présente des approches originales en matière de recherche qualitative et d'étude des innovations qui sont susceptibles d'être mises en œuvre dans le cadre de la recherche évaluative.

Cronbach, L. J. (1983). *Designing Evaluations of Educational and Social Programs*, San Francisco, Jossey-Bass.

Ouvrage de base sur l'évaluation des programmes sociaux et sur la recherche évaluative.