

# L'évaluation d'impact

---

INTERVENTION DEVANT LE GROUPE « ACTION SOCIALE » DE LA SFE — 15/10/19 – THOMAS DELAHAIS

Bonjour à tous,

Merci de m'avoir invité à cette séance pour vous parler d'évaluation d'impact. Je suis Thomas Delahais, je suis évaluateur dans une coopérative appelée Quadrant Conseil, et je travaille en particulier sur les questions d'impact.

Ce qui nous a intéressé dans cette question, c'est ce qui s'est passé dans les années 2000, en France et dans le monde, lorsque tout d'un coup des économistes, Mme Duflo notamment qui a reçu son prix Nobel hier, sont venus sur le champ de l'évaluation et ont expliqué que les évaluations menées jusque-là n'étaient pas scientifiques parce qu'elles ne traitaient pas de l'impact ; et qu'en conséquence ces évaluations ne pouvaient pas décentement soutenir la décision publique. Il fallait donc tout changer.

Comme beaucoup d'évaluateurs, je pense, nous avons été piqués, on a d'abord dit que c'était n'importe quoi, que ce que ces économistes faisaient avait beaucoup de défauts, que l'impact c'était complexe, que la vision de l'action publique qui sous-tendait cette approche de l'impact était problématique. Et tout cela est toujours vrai, j'y reviendrai tout à l'heure.

Mais nous nous sommes aussi dit que cette injonction à l'impact elle n'était pas inintéressante non plus. Que peut-être nous nous étions détachés par exemple des débats sur les méthodes, que finalement nous nous détournions des questions les plus difficiles sous prétexte qu'il était impossible d'y répondre. Bref, ce qui s'est passé ces 15 dernières années, disons, c'est que les évaluateurs se sont ressaisis des questions d'impact et ont essayé d'avoir non seulement un regard critique sur la notion, mais aussi de l'embrasser.

Une chose importante c'est que je ne connais pas la protection de l'enfance, j'ai appris beaucoup de choses ce matin, mais ce n'est pas ma pratique. Ce que je vais dire, je vais le faire en m'appuyant sur ce que je connais sur les politiques de la jeunesse, de l'éducation, et sur le travail social, mais je compte sur la discussion pour qu'on puisse voir ce qui est pertinent dans le cadre de la protection de l'enfance.

---

Ce que je vais aborder donc ici, c'est d'abord la notion d'impact, ce qu'elle recouvre en théorie et en pratique (I). Puis nous parlerons d'une notion de l'impact en particulier, l'impact contrefactuel, et en particulier la méthode expérimentale par assignation aléatoire, et en quoi il faut faire attention en particulier à cette approche lorsque l'on travaille dans l'action sociale (II). Et puis je terminerai en essayant de vous donner un aperçu de la diversité des approches de l'impact (III), et en vous disant que vous pouvez vous saisir des méthodes et des outils de l'évaluation d'impact, au service de l'action sociale.

## LA NOTION D'IMPACT

L'impact d'abord c'est une analogie balistique, comme l'impact d'une balle, pour parler des conséquences de l'action publique. C'est important d'avoir en tête que c'est une notion qui n'est pas stabilisée, et qu'elle est utilisée avec des sens très différents.

J'en propose 3 ici, qui sont proches, mais ne se recoupent pas totalement :

- l'impact au sens des effets qui justifiaient l'intervention initialement.
- l'impact au sens des conséquences de l'intervention évaluée sur les bénéficiaires finaux de l'intervention
- l'impact au sens de l'ensemble des conséquences d'une intervention, en particulier les effets inattendus, désirables ou indésirables.

Enfin il y a la notion de mesure d'impact social ou d'évaluation d'impact social, qui en France est liée au secteur de l'économie sociale et solidaire, mais qui globalement est plutôt un type d'évaluation qu'une évaluation d'impact au sens des points mentionnés ci-dessus.

Dans tous les cas, l'impact va porter sur des choses plus difficiles à estimer que dans une évaluation classique, c'est-à-dire un rapport de causalité, d'abord, et ensuite des conséquences plus lointaines, ou initialement inconnues, et donc il faut une ou des méthodes spécifiques pour traiter de ces questions.

Donc une évaluation d'impact, **c'est une évaluation qui vise à estimer des impacts, et qui pour le faire s'appuie sur des méthodes spécifiquement conçues pour cela.** J'insiste au moins sur une chose à cette étape, c'est qu'une évaluation d'impact ce n'est jamais une simple mesure du changement. Par contre ça peut être pas mal de choses différentes, on y reviendra aussi.

## POURQUOI EST-CE QU'ON FAIT DES ÉVALUATIONS D'IMPACT ?

Si je réponds basiquement à cette question, pourquoi est-ce qu'on fait des évaluations d'impact, je le ferais de la même façon que pour une évaluation classique, c'est-à-dire potentiellement pour apprendre, juger et décider. L'évaluation d'impact oblige à se poser la question des impacts que l'on veut véritablement avoir et peut-on raisonnablement les avoir. Elle oblige à réfléchir aux bons critères de jugement ; elle aide à faire des choix... comme n'importe quelle (bonne) évaluation.

Alors la question ce serait plutôt « **pourquoi est-ce qu'on ferait des évaluations d'impact** » ? pourquoi a-t-on une injonction assez généralisée à ce sujet, qu'on retrouve notamment dans les documents officiels, dans les discours de tous ?

D'abord, il ne faut pas sous-estimer le fait que dans beaucoup de cas, quand on parle d'évaluation d'impact, on parle d'autre chose. On parle d'évaluation classique, parfois on parle même de performance un peu basique.

Ensuite il y a l'idée qu'il y a besoin de preuves robustes, « scientifiques » pour mettre en place de bonnes politiques publiques. Ça, c'est quelque chose d'ancien, on remonte au moins aux années 1960 avec la société expérimentale de Donald Campbell. Que dit Campbell ? Il dit que dans la vraie vie on ne peut pas savoir si les interventions « marchent », si elles sont bonnes pour le bien public, et que donc il vaut mieux tester des choses dans un cadre expérimental et les évaluer avec une méthode en particulier, qui est la méthode expérimentale par assignation aléatoire.

Cette méthode consiste à tester une intervention auprès d'individus et de les comparer, au regard d'un certain nombre d'indicateurs d'impact, avec d'autres individus, qui ont les mêmes caractéristiques, mais qui ne sont pas soumis à l'intervention. L'impact est la différence constatée entre la situation avec et la situation sans.

Cette méthode expérimentale va gagner beaucoup en importance dans les années 1970 et 1980, avant de tomber un peu en désuétude, notamment pour des raisons méthodologiques (est-ce que ce qui marche dans une expérimentation marchera une fois déployé, par exemple ?), mais aussi parce qu'on se rend compte assez vite que les politiques ne donnent pas une importance si importante que cela à la preuve. Dans les années 1980 et 1990, l'évaluation va connaître beaucoup de

mutations, dans des directions très différentes, d'un côté la montée de l'évaluation participative, les évaluations émancipatrices, la montée en puissance du registre de la pertinence, et de l'autre côté la montée de la question de *value for money*, de performance.

Dans les années 1990 revient cependant, à l'initiative des économistes, cette idée que les politiques ne s'appuient pas suffisamment sur des preuves, dans plusieurs champs, en particulier l'éducation, l'emploi et l'aide au développement. L'idée de « politique basée sur des données probantes », *evidence-based policy*, c'est que les politiques restent inefficaces parce qu'elles ne prennent pas en compte les preuves scientifiques, et ce à plusieurs niveaux, tant au niveau des pouvoirs publics que des acteurs de terrain. Les tenants de la « politique basée sur des données probantes » en particulier rejettent l'idée que les acteurs de terrain puissent être des artisans qui décident à partir de leur propre expérience, ou sur des questions morales par exemple qui polluent selon eux leur jugement.

Donc il faut sortir de cette logique de l'artisanat pour rentrer dans une logique de la preuve. Alors évidemment la question c'est « qu'est-ce qui fait preuve ? » Et c'est là qu'on retrouve la méthode expérimentale par assignation aléatoire, ce qu'on appelle en anglais le *randomised controlled trial* (RCT), qui serait la seule preuve acceptable.

## LE PROBLÈME AVEC LE RCT

Le RCT c'est potentiellement une méthode extrêmement puissante pour prouver des liens de causalité, parce que c'est une expérimentation qui est conçue pour l'évaluation, ce n'est pas l'inverse, ce qu'on voit habituellement, où on expérimente et ensuite on évalue. Le but d'un RCT ça va être de prouver un lien causal « toutes choses égales par ailleurs », c'est-à-dire en écartant toutes les autres raisons qui pourraient expliquer les résultats.

L'idée derrière, celle que défend par exemple Esther Duflo qui est une figure de l'*evidence-based policy*, c'est qu'avec ces expérimentations on puisse définir des briques « qui marchent », et qu'on reconstitue des politiques publiques en associant les briques qui marchent.

Des RCT il y en a dans les pays anglo-saxons ou dans les pays du sud sur la protection de l'enfance, je n'ai pas d'expertise là-dessus, mais une rapide recherche montre par

exemple des programmes sur le dépistage de la maltraitance, sa déclaration par les enfants, sa détection par des visites à domicile, sur la délinquance juvénile, etc.

Mais le RCT pose des problèmes à plusieurs niveaux, je ne les mets pas dans un ordre en particulier, mais ils sont interconnectés :

- un premier niveau, qui est très important, et qui est un enjeu éthique. Avec le RCT on suit un protocole et on doit le suivre que l'intervention marche ou ne marche pas. Il ne faut pas prendre d'initiatives (car alors on évaluerait la capacité des travailleurs sociaux à gérer des situations, pas l'intervention), et si c'est un échec, c'est un résultat intéressant. C'est très difficile d'accepter cela dans le travail social, peut-être encore plus dans la protection de l'enfance lorsque des enfants sont en danger. Par ailleurs la volonté d'avoir des données peut se mettre en chemin de l'accompagnement lui-même, peut le polluer et poser problème — or dans un RCT la question des données est centrale, on préférera écarter des enfants par exemple si on ne peut pas avoir de données, même si l'action est pertinente pour nous, que de se passer de données. Entre nous ce n'est pas pour rien qu'il y a tant de RCT qui sont faits dans des pays du Sud (souvent par des chercheurs du Nord...), parce qu'on peut contourner ainsi les enjeux éthiques.
- un deuxième souci, il est lié à la méthode, qui est une boîte noire, on a une apparence de scientificité, mais il y a beaucoup de choix, de jugements qui sont faits, ce qui est parfaitement normal dans une évaluation, mais qui n'est pas assumé. Il y a des gens comme Arthur Jatteau qui ont montré ça notamment.
- un troisième souci, ça va être que le RCT teste des relations causales très simples, on dit qu'il faut un « effet tunnel » entre l'intervention et les effets attendus, alors que le travail social il relève du complexe, il dépend beaucoup des relations entre plusieurs acteurs, qui ont chacun leur rationalité, que les effets du travail social se jugent à plusieurs aunes, que l'histoire des jeunes bénéficiaires par exemple, leur expérience des dispositifs va beaucoup influencer les conséquences qu'on peut attendre, etc.
- Enfin, un dernier souci, et qui pour nous est peut-être le plus préoccupant, et qui est l'état d'esprit qui est lié à cette approche, notamment cette idée que la pratique des travailleurs sociaux est un problème plutôt qu'une part de la solution (et qu'il faudrait donc qu'ils appliquent des méthodes « scientifiquement validées » sans discuter), et qu'il faudrait réfléchir au niveau des actions, de chaque sujet, plutôt qu'au niveau global.

Ce dernier point, c'est quelque chose qui est prégnant aujourd'hui sur les questions éducatives et qu'on voit avec la division des classes de CP-CE1 par deux, la distinction faite entre apprentissage de la lecture et parcours scolaire, l'insistance sur les méthodes, etc. C'est peut-être un risque sur le long terme que cette évaluation

---

d'impact, comme les dispositifs de performance, transforme l'action sociale notamment en contrôlant les travailleurs sociaux. Mais j'ai l'impression qu'aujourd'hui c'est plus l'état d'esprit qui va avec, l'idée qu'il y aurait des baguettes magiques et que la « faute » serait aux professionnels qui ne les applique pas, qui me semble un danger.

Je dirais enfin qu'il y a un dernier problème et pas des moindres, c'est l'usage qu'on peut faire de ces évaluations. Elles disent « ça marche » ou « ça ne marche pas », mais elles ne nous disent pas pourquoi : or, c'est le pourquoi qui permet d'améliorer les actions menées.

## DIFFÉRENTES APPROCHES DE L'IMPACT

Alors on peut critiquer le RCT, surtout pour ce qu'il amène comme façon de penser, maintenant la question des impacts, elle reste très intéressante, et tous vos débats de ce matin ont illustré clairement cette question.

Un premier aspect très important c'est de réfléchir un peu plus profondément à la relation causale.

Le RCT se place dans une logique **successionniste et contrefactuelle**, c'est l'idée d'abord que l'effet suit l'intervention, et d'autre part qu'on peut isoler l'effet d'une variable dans un changement qu'on observe, donc par exemple l'effet d'une méthode sur l'apprentissage de la lecture.

Alors sans faire du RCT, on peut quand même rechercher des **régularités**, à travers des études longitudinales, où on va chercher à voir sur le long terme ce qui se passe pour les bénéficiaires de l'intervention, pour les jeunes par exemple. C'est beaucoup ce qui se fait sur la jeunesse et sur l'enfance, on va par exemple voir les trajectoires de jeunes qui ont été en mesure et d'autres qui ne l'ont pas été. C'est intéressant et c'est très utile, ça nous donne des leçons globales, mais c'est aussi très compliqué à faire et votre discussion sur les observatoires montre bien que c'est difficile à faire au niveau local. Comment faire aussi quand les dispositifs changent tout le temps, comment on gère le temps du mandat qui change les orientations, etc.

Bon, en tout cas, cette logique est celle de tester l'effet de la variable sur le changement attendu. Cette façon de faire vous semble de bon sens ?

Mais c'est une vision particulière de l'impact. Une autre vision ce serait de dire qu'un changement dans la trajectoire, qu'un impact n'est jamais obtenu par une action seule, et ça n'a donc pas beaucoup de sens de vouloir isoler une variable. Pour reprendre mon exemple de l'apprentissage de la lecture, c'est de se dire « ce n'est pas la méthode qui m'apprend la lecture, c'est cette méthode peut-être, mais aussi mon enseignant, ma famille, mon entourage, etc. » et dans ce cas ce que l'on va rechercher ça va être de trouver des **configurations de conditions** à cet effet. C'est ce qu'on peut faire avec l'analyse qualitative comparée, qui va tester systématiquement des facteurs qui peuvent expliquer le changement attendu, et on va voir si certains de ces facteurs sont suffisants ou nécessaires, c'est-à-dire : est-ce que chaque fois qu'on a ce facteur, on a le changement attendu ? ou est-ce qu'il faut absolument avoir ce facteur pour avoir le changement attendu ? C'est une autre logique, mais elle est très intéressante, notamment d'un point de vue très pratique, très opérationnel, pour diagnostiquer une situation initiale.

Une autre vision c'est de considérer que ce qui est important ce n'est pas tant d'établir la relation causale, mais d'être capable **d'expliquer comment l'effet est généré**, d'ouvrir la boîte noire, et ça c'est particulièrement utile quand on est dans des situations complexes. Par exemple l'évaluation réaliste, c'est une approche intéressante parce qu'elle prend le contrepied total du RCT. Au lieu de faire une analogie médicale, de considérer que les interventions sont comme un traitement et que les bénéficiaires seraient comme des patients, totalement passifs, on va au contraire considérer que chaque intervention est une opportunité qui est donnée aux personnes de changer de comportement, mais qu'ils vont choisir ou non de s'en saisir. Et ce qui va nous intéresser c'est le pourquoi, c'est les mécanismes qui font qu'on se saisit ou non de l'intervention. Par exemple, si j'évalue ma nouvelle méthode, je vais me demander pourquoi les enseignants s'en saisissent ou pas, comment ils s'en saisissent, comment les élèves réagissent à la nouvelle approche, à ce que disent les familles, aux nouvelles pratiques des enseignants, etc.

Moi je pense que cette approche, l'évaluation réaliste, est particulièrement bien adaptée au travail social, parce qu'elle est vraiment dans la compréhension, et aussi parce qu'elle a une volonté de généraliser, c'est de dire : est-ce qu'on peut identifier les contextes les plus importants, et comment une intervention peut déclencher différents mécanismes qui vont favoriser les effets attendus.

À un niveau plus large, de programmes ou de politiques, on a une approche, c'est celle sur laquelle je travaille le plus, c'est l'analyse de contribution, qui va tester toutes les hypothèses qui peuvent expliquer un changement pour aboutir à une

---

explication, la plus plausible possible, de pourquoi et comment on aboutit ou non aux changements attendus, et on va rechercher des chemins d'impact, c'est-à-dire par quels changements dans les attitudes, les comportements des acteurs on va aboutir ou non aux changements attendus.

## CONCLUSION

Ces différentes approches, elles demandent une certaine technicité, elles sont intéressantes et apportent des éléments pointus. Je ne vous dirai pas qu'il faut forcément des évaluations d'impact aujourd'hui et maintenant.

Ce que je vous propose plutôt en conclusion c'est de réfléchir par ce bout de l'impact, par la question des conséquences que l'on voudrait avoir, aussi par ce que ça signifie la relation de causalité, pour réfléchir à vos dispositifs, et peut-être essayer d'abord de réfléchir à la pertinence de vos actions, à leur cohérence, etc.

Réfléchissez aussi aux données dont vous auriez besoin pour un jour prouver ces liens de causalité. Des données sur la durée, même sans évaluation d'impact, c'est extrêmement précieux, et ça vous sera toujours utile.

Je vous dirais aussi également que ce travail, cette réflexion, vous pouvez l'avoir avec d'autres, avec vos partenaires, ne serait-ce que pour être au clair sur ce qui compte vraiment dans les interventions de protection de l'enfance, pour parler du sens... tout ça sera très utile pour produire des questionnements pertinents et faire en sorte que tout le monde ait envie de s'y impliquer.

Alors seulement, une fois que vous aurez de bonnes questions, que tout le monde sera prêt à se servir des résultats, peut-être lancerez-vous une évaluation d'impact. D'ici là, bonne réflexion à vous tous et à vous toutes !