

COPAS

A l'aise dans l'évaluation



Partons d'un constat général, que je ne peux pas illustrer longuement ici, mais dans lequel, je pense, chacun se reconnaîtra pour sa part : il existe un décalage extrême entre les discours sur l'évaluation (tout le monde s'accorde à dire qu'il faut en faire et que c'est important) et la réalité des pratiques (il s'en fait finalement assez peu et, surtout, avec une utilité réelle très limitée). Je voudrais tenter d'expliquer ce « malaise dans l'évaluation » et d'explorer quelques chemins qui pourraient nous permettre d'en sortir.

Définissons rapidement l'évaluation comme l'activité qui conduit à la production d'un jugement de valeur sur un objet donné (ici, nous nous limiterons à l'évaluation des actions à finalité « sociale »). J'évalue lorsque je dis si telle action me paraît « bonne » ou pas.

L'évaluation est une activité cognitive naturelle, parmi les plus automatiques : placé face à un objet nouveau, notre première réaction sera inévitablement de formuler un jugement de valeur (bien avant de se poser toutes les autres questions possibles) : j'évalue la nourriture qu'on me propose, les spectacles auxquels j'assiste, les gens que je rencontre, etc. Mais, bien sûr, nous ne nous intéressons pas ici aux évaluations purement privées, celles que chacun accomplit sans cesse pour mettre de l'ordre dans son univers, mais à celles qui vont prendre place dans des systèmes collectifs, dans des institutions : les évaluations qui constituent un enjeu pour l'action collective.

Puisque nous sommes partis du discours affirmant fortement la nécessité de l'évaluation des actions, nous devons en particulier interroger cet enjeu en termes d'utilité : à quoi sert l'évaluation ? En tant que telle, elle n'a aucune utilité sociale particulière. Elle n'acquiert cette utilité qu'en fonction de la place qu'elle va venir prendre dans les systèmes d'action. Or, de ce point de vue, sa position est assez ambiguë.

Evaluation, régulation et recherche

D'une part, l'évaluation appartient à l'ensemble des fonctions de régulation, c'est-à-dire des mécanismes qui permettent la transmission et le traitement des informations qui viendront alimenter les décisions : l'évaluation va déboucher sur le choix d'arrêter ou de développer une action, de conclure ou pas une négociation, de confirmer ou remettre en cause une convention, etc. Bien entendu, l'évaluation n'est qu'une activité régulatrice parmi d'autres : elle côtoie par exemple la concertation, le suivi, le contrôle, etc.

Mais d'autre part, l'évaluation appartient à l'ensemble des fonctions de recherche¹, c'est-à-dire de production de connaissances. Bien entendu, à nouveau, elle n'est qu'une forme de recherche parmi d'autres, à côté de la formalisation, de la recherche-développement,

¹ J'emploie ici le mot recherche dans son sens large : toute activité qui vise à produire des connaissances valides. Ceci ne se limite pas à la recherche formelle, scientifique : tout acteur peut travailler à produire des connaissances sur l'action à laquelle il est mêlé. De même, quand je parlerai plus loin de chercheurs, cela ne concerne pas seulement (ni même principalement) les professionnels de la recherche mais tous ceux qui ... cherchent !



de la modélisation, de la validation d'hypothèses, de la théorisation, etc. L'utilité sociale de la recherche est moins immédiate que celle de la régulation, mais elle fait l'objet d'un consensus culturel encore assez fort pour qu'il soit inutile de l'interroger ici : mieux on connaît un système social, mieux on pourra l'améliorer.

Illusion et frustration

Je peux maintenant en venir à la formulation des deux hypothèses principales à propos dudit malaise.

- Lorsqu'on parle d'évaluation et qu'on en affirme la nécessité, on ne pense pas à toute l'évaluation, mais à une forme particulière, celle qui se trouve à l'intersection des deux univers que je viens de définir.

- Or, les évaluations qui sont satisfaisantes à la fois sur le plan de la connaissance (évaluations justes) et sur le plan de la régulation (évaluations efficaces pour déclencher la bonne décision) sont très rares, sinon inexistantes.

Le premier point s'explique par un pente bien connue de notre rapport à l'action et au pouvoir : puisque nous devons prendre des décisions, nous aimerions pouvoir les présenter (y compris à nous-mêmes), non comme un véritable choix de notre part, dont nous devrions assumer les incertitudes et la part d'arbitraire, mais comme une conséquence des faits eux-mêmes. D'où l'espoir d'ancrer l'évaluation- régulation dans l'évaluation-recherche : la première viendrait nous dire comment les choses sont, et il n'y aurait plus qu'à en déduire ce qu'elles doivent être et donc les décisions qui doivent être prises.

Ce projet, surtout porté par les acteurs de la décision, peut rencontrer l'aval des acteurs de la recherche, qui y voient une manière d'accéder à un pouvoir social ou à des rétributions qui, en tant que chercheurs « purs », leur échappent. Or ce projet est illusoire, d'abord parce que les règles du jeu des deux univers sont différentes, voire contradictoires. Le propre du choix pratique ou politique est d'être un choix nécessaire (on ne peut pas ne pas choisir, même si l'un des choix possibles reste toujours de ne rien changer) en l'absence de référent rationnel externe (on n'a aucun moyen de savoir à coup sûr quel est le bon choix).

Or les règles de la démarche de recherche sont opposées : elles commandent de suspendre le jugement, aussi longtemps qu'il le faudra, tant qu'on n'a pas atteint un haut degré de certitude.

Autre différence de règle : le jugement pratique est nécessairement global, il doit prendre en compte tous les éléments disponibles, aussi hétérogènes et informels soient-ils. Au contraire, le jugement de recherche sera sélectif : on choisira de ne prendre en compte que quelques paramètres qu'on aura sélectionnés et qu'on traitera à partir d'une problématique particulière.

Une évaluation-décision doit être totale et elle est souvent irréversible ; une évaluation-recherche ne peut être que partielle et provisoire.

Mais, surtout, ce projet se heurte aux stratégies mêmes des acteurs. Ceux qui doivent prendre la décision affirment vouloir la soumettre au jugement issu de la recherche, mais, le moment venu, ils n'abdiqueront pas si facilement leur liberté de choix, soit parce qu'ils doivent, pour préserver leur pouvoir, protéger leur marge de manœuvre, soit parce que, souvent à juste titre, ils resteront convaincus jusqu'au bout que leur propre jugement est finalement mieux fondé que celui des chercheurs. Qui n'a déjà observé cette curieuse situation : un responsable demande à un spécialiste une coûteuse évaluation puis en néglige les résultats parce qu'ils ne viennent pas confirmer ses propres intuitions ? Symétriquement, le chercheur qui prétendait accéder au pouvoir se dérobera à la dernière



minute en produisant des jugements évaluatifs tellement prudents et nuancés qu'ils ne pourront fonder aucune décision.

Personne n'est à blâmer, et celui qui se dit déçu n'a qu'à se reprocher à lui-même son espoir illusoire : celui de réduire l'art complexe de la décision politique à de la méthodologie pseudo-scientifique, ou les subtilités infinies de la connaissance à la production d'arguments pratiques.

Que faire, alors ?

Quelle attitude adopter face à la demande sociale d'évaluation, qui ne semble pas devoir se réduire, et à juste raison, puisque la question de savoir ce que valent nos actions est bel et bien la plus sérieuse des questions ? Généralement, nous sommes conduits à explorer l'une ou l'autre des pistes suivantes. Tout d'abord, donner toute leur place aux autres pratiques utiles dans les deux champs considérés. Car, comme je l'ai déjà dit, l'évaluation n'est pas la seule pratique, et les autres sont parfois bien plus intéressantes.

◆ Ne pas négliger les autres pratiques utiles

Combien de commanditaires demandent une évaluation des programmes dont ils ont la charge, sans même avoir mis en place les conditions nécessaires au simple suivi des actions (c'est-à-dire les moyens d'être informés de ce que sont ces actions et leurs conditions de déroulement) ?

Combien de responsables de structures décident une évaluation de leurs services, alors qu'il serait bien plus urgent de créer des lieux de concertation (c'est à dire de maturation des décisions par le dialogue) ?

Même chose dans le champ de la connaissance : à quoi bon évaluer une action si l'on n'a pas encore commencé un travail minimum de formalisation (description structurée des pratiques et des principes qui les orientent) ou de modélisation (exploration des variantes de l'action, par la méthode des types idéaux) ? Et si l'on veut aller plus loin, n'est-il pas souvent plus utile, plutôt que prétendre fonder des jugements de valeur, de créer des moyens d'action nouveaux (recherche-développement) ou de mettre à l'épreuve les hypothèses théoriques ou techniques qu'on est parvenu à formuler ?

Mais on ne peut pas toujours esquiver l'évaluation, parce que celle-ci, en tant qu'activité cognitive élémentaire, est en fait omniprésente : le suivi comme la concertation, la formalisation comme la modélisation ne sont jamais indifférentes et sont au contraire généralement saturées de jugements évaluatifs. Autant le reconnaître et affronter les formes élargies d'évaluation, qui certes ne seront pas conformes au modèle idéal de l'évaluation-recherche-décision, mais qui n'en sont pas moins intéressantes parce qu'elles sont vraiment utiles.

Du côté de la régulation, cela revient à risquer l'exploration du monde de l'évaluation informelle, c'est-à-dire de ces jugements de valeurs qui ne sont pas fondés sur une démarche rigoureuse de recherche, mais qui n'en sont pas moins à l'origine réelle des décisions.

◆ Réguler les évaluations informelles

Du côté de la régulation, cela revient à risquer l'exploration du monde de l'évaluation informelle, c'est-à-dire de ces jugements de valeurs qui ne sont pas fondés sur une démarche rigoureuse de recherche, mais qui n'en sont pas moins à l'origine réelle des décisions. On aurait plutôt envie de ne pas voir cette évaluation-là, ou de ne la considérer que comme une dangereuse faiblesse à remplacer rapidement par quelque chose de plus sérieux, d'une part parce qu'elle semble contradictoire avec le souci d'« objectivité »,



voire de scientificité, d'autre part et surtout parce qu'elle risque de déboucher, si on accepte de la voir, sur des relations conflictuelles en cas de désaccord.

Mais de toutes façons ces jugements existent, et ils pèsent lourd dans les décisions effectives : il vaut donc mieux les prendre en compte et les réguler plutôt que de les laisser jouer de manière sauvage.

Il y a en effet des risques réels, car les évaluations implicites peuvent être soumises à des biais importants : pour émettre ces jugements, les acteurs impliqués n'activent généralement pas toutes leurs capacités cognitives (d'observation, d'interprétation, etc.), parce qu'ils ont une vision avant tout affective de leur action et des relations avec les autres acteurs ; d'autre part, il arrive souvent que leurs jugements soient déformés par un souci de rationalisation des engagements (par exemple : je trouve que c'est bien, puisque je l'ai fait ; ou je trouve que tel acteur n'est pas bon, puisque j'ai jusqu'à présent refusé de travailler avec lui) ou des normes (je trouve que c'est bien, puisque tout le monde le dit) ; enfin, les jugements informels donnent souvent lieu à des biais de confirmation (à partir du moment où mon jugement est ébauché, je ne cherche plus et je ne retiens plus que les observations qui tendent à le confirmer) et d'auto-confirmation (le jugement que je porte tend à agir sur la réalité dans le sens de mon appréciation : puisque je juge une action négativement, je tends à la faire échouer).

Mais, encore une fois, on ne peut prétendre se débarrasser de ces risques en les ignorant : ce qu'il faut, c'est prendre les moyens de travailler ces jugements informels de manière à les rendre plus fiables, par un triple processus : d'explicitation (dire les jugements qu'on garde habituellement pour soi) ; d'argumentation (au delà du jugement brut, apporter des informations qui l'éclairent) ; de confrontation (accepter des arguments contradictoires et engager le dialogue). Plus question ici de singer une démarche de recherche scientifique : nous sommes dans une autre vision de la production du juste : non plus par l'expérimentation mais par la discussion.

L'évaluation éclairante

Dans l'univers de la recherche, là où ce qui compte n'est pas de décider ce qui est juste, mais de décrire et comprendre ce qui est, il y a également place pour une pratique élargie de l'évaluation. En effet, une recherche d'évaluation n'a pas besoin de déboucher sur des décisions qui en découleraient nécessairement pour être intéressante : on peut se contenter de lui demander de nous éclairer, et on aura d'autant plus de chance de l'être effectivement qu'on ne lui en demandera pas plus.

Si l'on s'engage sérieusement dans le chemin complexe, plein d'embûches, mais ô combien passionnant d'une telle recherche (délimitation précise de l'objet à évaluer, construction du référent, définition et production des informations pertinentes, rapprochements et discussion des résultats) on apprendra beaucoup. Et pourtant, tout ce qui fera la qualité de cette recherche (réduction de l'objet et du référent, délai de collecte des données, prudence dans l'interprétation) la rendra peu intéressante dans une perspective de régulation.

Ainsi, si l'on veut faire de l'évaluation de qualité, il nous faudra souvent choisir : soit se centrer sur l'utilité régulatrice, et tourner le dos aux procédures de recherche pour se concentrer sur la formation et l'ajustement des jugements informels ; soit se centrer sur la fonction cognitive et accepter que les voies de la connaissance ne reconduisent que très indirectement vers celles du pouvoir.

Les évaluations paradoxales

Si, toutefois, on continue à vouloir faire les deux, à construire une recherche méthodique qui serait directement liée aux décisions, il faudra trouver les moyens pour que les processus de connaissance d'une part, et de pouvoir, d'autre part, ne soient pas trop mutilés par leur rapprochement forcé. Pour cela, il faudra paradoxalement y introduire (ou accepter d'y laisser) des imperfections, un relâchement suffisant : ce n'est qu'à cette condition que l'esprit de recherche pourra ne pas être paralysé par les responsabilités attachées à la décision, ou que les décideurs cesseront de craindre de devoir se soumettre à des résultats incontrôlables.

C'est ainsi qu'il nous arrive de pratiquer des évaluations-recherches d'autant plus utiles qu'elles sont incomplètes (par exemple, on fait le travail de conception du référent et on imagine les procédures de collecte de données pertinentes, mais on s'arrête là : on a eu au moins l'occasion de s'interroger et d'échanger sur ce qui fait la valeur d'une action et les indices qui permettraient de la détecter) ou inversées (on part d'un jugement a priori, et on remonte la chaîne du raisonnement pour mieux le comprendre).

Je pense qu'en ouvrant ces trois perspectives, il est possible de nouer avec l'évaluation une relation bien moins raide et bien moins frustrante que ce que l'on observe le plus souvent. Mais, pour réussir à se mettre ainsi à l'aise, il faut accepter un triple renoncement (et ça, c'est toujours une épreuve) : renoncement à l'exclusivité (il y a bien d'autres choses utiles à faire, et parfois plus prioritaires), renoncement au modèle unique de l'évaluation-recherche-décision (l'évaluation est morte, vive les évaluations), et, le plus difficile, renoncement à la perfection (on peut faire des évaluations « anormales » et pourtant très utiles).

Claude Coquelle COPAS N° 17 - Mars 1998