

L'évaluation en sociologie : un enjeu au carrefour des dynamiques du champ scientifique mondial

Frédéric Lebaron (Université de Versailles Saint-Quentin-en-Yvelines)

Si l'évaluation occupe une place centrale dans l'activité des sociologues, et si ceux-ci ne cessent d'en débattre de façon formelle ou informelle (voir notre contribution au dossier de *Global Dialogue* en 2014), l'enjeu de l'évaluation n'en a pas moins pris une importance nouvelle depuis quelques années. L'une des raisons évidentes de cette évolution est la conjonction de l'internationalisation de la recherche, avec, entre autres, l'hégémonie accrue de la langue anglaise dans le monde scientifique, et la montée en puissance de l'outil bibliométrique au sein des procédures évaluatives, notamment s'agissant des recrutements et des carrières. Ces évolutions concernent l'ensemble des disciplines académiques et scientifiques.

Elles suscitent en sociologie de fortes résistances, qui prennent des formes variées. Parmi celles-ci, une grande partie semblent relativement proches de celles rencontrées dans de nombreuses autres disciplines des sciences humaines et sociales, particulièrement les disciplines littéraires : les productions particulières de nos disciplines ne sauraient faire l'objet de procédures de standardisation de type industriel, relevant d'une économie des singularités mal traitée par le nouveau modèle d'organisation de la recherche.

Chaque production sociologique nécessite une évaluation interne spécifique, et les critères de celle-ci, multiples et contextuels, ne peuvent être codifiés. Les carrières des sociologues seraient encore plus difficiles à étalonner, du fait de leur caractère encore plus « multidimensionnel », lié à des parcours biographiques plus ou moins complexes, qu'il est pourtant toujours nécessaire de prendre en compte.

Face aux positions de refus, le champ scientifique continue cependant de voir s'affirmer chaque jour la logique de l'évaluation bibliométrique à

divers niveaux, cela en dépit de critiques internes tout autant qu'externes, comme menée par Yves Gingras, utilisateur de l'outil bibliométrique et très critique à l'égard des dérives de son usage.

Faut-il tenter de sortir de l'aporie qui s'est installée, entre le modèle standardisé d'un côté et la résistance au nom de la singularité d'autre part? Ou peut-on faire évoluer le débat en proposant quelques lignes directrices? C'est ce que nous essaierons de faire dans ce qui suit.

Là quantification des activités des sociologues n'a pas attendu le mouvement de standardisation en cours. Le comptage des productions, des heures de travail, et en particulier de cours, la mesure de l'impact des ouvrages par leur diffusion, et bien d'autres formes de mises en nombre, sont présents, à des titres divers, dans le quotidien des sociologues depuis longtemps.

La nouveauté de notre période est l'apparition sur une large échelle de métrologies envahissantes, centrées principalement sur l'activité de production scientifique (l'enseignement et l'administration restant largement à l'écart de ce mouvement de quantification). Elles sont de plus en plus souvent institutionnalisées dans des listes de revues assorties de points, des calculs d'indices ou d'autres procédures formalisées.

L'une des forces de ces pratiques est qu'elles nourrissent la tendance à la création de classements et de palmarès qui peuvent être réalisés à plusieurs échelles : celle des individus, des départements et laboratoires, des universités et des pays. Dans ce système emboîté de mise-en-équivalence, la qualité de science est en théorie facilement évaluable, comparable et cumulable. Le « classement de Shanghai » en est la manifestation la plus visible.

Dans la réalité, l'obsession de la mesure rancontre de nombreux obstacles intellectuels, et la

pratique bibliométrique est beaucoup plus diverse qu'on ne le croit : on peut, schématiquement, décider *a priori* d'attribuer des points à des revues (comme c'est souvent le cas en sciences économiques, en tout cas en France) et calculer la production en conséquence ; on peut utiliser directement des mesures d'impact citationnel à partir de bases de données couvrant un échantillon de l'ensemble des productions (comme avec le *h-index* fondé sur Google Scholar). Dans tous les cas, les biais sont multiples, et les mesures toujours difficiles à interpréter sans une bonne connaissance des domaines scientifiques couverts, et des contextes nationaux de l'activité sociologique.

En particulier, la domination de la langue anglaise est trop souvent considérée comme allant de soi, dans les sciences humaines et sociales, alors qu'une large partie de la production, dans la plupart des pays non-anglophones, continue d'être effectuée dans la langue nationale et reste souvent absente (même si la numérisation fait évoluer cette situation) des classements internationaux pour cause d'invisibilisation. On peut penser que la production dans chaque langue devrait donner lieu, par exemple, à des analyses d'impact spécifiques. Si celles-ci révéleraient évidemment la position asymétrique de la langue anglaise (citée et beaucoup moins « citant »), cela permettrait aussi d'introduire une vision beaucoup plus riche et polycentrique des controverses sociologiques à l'échelle mondiale.

Une des voies qui semblent les plus intéressantes pour réguler la pratique de l'évaluation quantitative des productions consiste en effet à cantonner celle-ci à d'éventuels « éléments d'information » à la disposition des évaluateurs, mais ne pouvant pas conditionner directement la décision évaluative : il s'agit de refuser toute forme d'algorithme qui ferait découler un classement et une décision d'un indice, sans la médiation d'une discussion collégiale.

Une deuxième voie, parallèle, consiste à pluraliser au maximum les outils de mesure, en distinguant clairement ce qui relève de la quantité produite (nombre d'articles, calculé selon diverses

méthodes pour tenir compte de la taille, du nombre d'auteurs, etc.), de l'impact (avec les citations, mais, là encore, calculées dans diverses bases de données, représentatives de sous-espaces et de langues différentes au sein du champ scientifique, et rapportées à divers indicateurs comme l'ancienneté, les conditions de carrière, etc.).

En affirmant la multidimensionalité de l'activité de production scientifique, qui ne relève pas, tant s'en faut, de la seule production d'articles dans les revues pour les pairs, on rend possible une analyse plus solide des dynamiques de recherche, plus ou moins individuelles et collectives et associant des composantes diverses et hétérogènes.

À côté du renouvellement de l'évaluation des productions, les deux autres grands secteurs d'activité professionnelle des sociologues, l'enseignement et l'animation de la communauté, restent mal évalués car souvent peu définis et finalement faiblement constitués en objets légitimes pour la réflexion et l'action de la communauté scientifique. Là encore, on peut penser que l'explicitation de « grilles de critères » complexes et multidimensionnels permettrait au moins de dégager un socle commun de préoccupations. On sait bien sûr que chaque évaluateur a tendance à « pondérer » différemment l'investissement dans des fonctions de direction administrative, l'animation de revues, l'évaluation d'articles, etc. Mais on peut là encore souhaiter que chaque comité d'évaluation consacre à la discussion de ces critères un ensemble de débats préalables, en poussant la formalisation explicite (mais non la quantification) aussi que loin que possible.

Car l'un des enjeux de l'évaluation est bien sûr la légitimité même des décisions prises sur la base des évaluations collégiales. Dans un contexte de tensions budgétaires et de restructuration autoritaire de l'enseignement supérieur et de la recherche, des procédures collégiales ouvertes et pluralistes, fondées sur des critères clairs et appuyant, sans en féliciter aucun, sur un maximum d'éléments objectifs, devraient permettre de renforcer la confiance dans le bon fonctionnement de notre discipline.