

La preuve sociologique

par Emmanuel Didier

Howard Becker, l'une des plus grandes figures de l'École de Chicago, s'interroge sur les conditions qui rendent possible la formulation d'un énoncé sociologique solide et robuste. Son dernier ouvrage est en effet consacré à la production des données, par-delà le clivage entre quantité et qualité.

Recensé : Howard Saul Becker, *Evidence*, Chicago ; Londres: The University of Chicago Press, 2017.

Faut-il encore présenter Howard Becker, ce monstre sacré de la sociologie ? Né en 1928, deux de ses ouvrages, *Outsiders* (1963) et *Les Mondes de l'art* (1982) ont particulièrement fait date. Becker a aussi eu une très importante production méthodologique. On lui attribue la « labelling theory » et il est aujourd'hui l'un des principaux représentants du symbolisme interactionniste qui insiste sur l'ethnographie, les données qualitatives, l'observation participante et les entretiens.

Evidence, l'ouvrage qu'il nous livre aujourd'hui (à l'âge canonique de 90 ans !) prend place dans la série de ses travaux méthodologiques. La question à laquelle il propose de répondre est fort directe : comment établit-on des propositions vraies, ou du moins convaincantes, en sociologie ?

Le fait qu'il s'agisse d'un livre méthodologique ne veut pas dire qu'il soit aride, bien au contraire. Becker ne rechigne pas à parler de ses souvenirs et de ses expériences de sociologue – par exemple avec Everett Hughes, son directeur de thèse (p. 167-170). En effet, une des caractéristiques frappantes de la plume de Becker est qu'il est *cool*, très MidWest. Lire son livre, c'est un peu comme boire un verre avec lui. Mais cool ne veut pas dire mou. Son livre, solidement charpenté, est constitué de deux parties. La première présente son modèle général de scientificité ; la seconde examine différentes méthodes de production des données.

Transformer les données en idées

Pour comprendre son épistémologie, influencée par le pragmatisme américain de James et de Dewey, il faut partir d'un axiome répété plusieurs fois dans le livre : le monde est toujours plus compliqué qu'on ne le croit, et quel que soit le plan que l'on mette en œuvre pour le connaître, des imprévus surviendront qui fragiliseront les résultats. L'imprévu, la surprise, la circonstance tiennent donc une place fondamentale dans ce texte. Mais pourtant, il s'agit bien de méthodologie. Alors, comment fait-on pour produire des connaissances dans des conditions si instables ?

Becker, dès les premières pages, pose un modèle. Pour convaincre, le sociologue, premièrement, combine des données, c'est-à-dire les traces de faits qu'il a observées et relevées dans le monde ; deuxièmement il transforme ces données en preuves (*evidence*, le titre du livre) lorsqu'il les articule dans un argumentaire ; et troisièmement, les preuves supportent une idée générale qu'il veut transmettre à d'autres. La science ressemble donc d'abord, à ce stade, à un processus de généralisation qui part des faits pour produire des idées.

Mais, point capital, l'organisation de ces trois éléments n'est pas systématique : « l'utilité de chacune de ces trois composantes dépend de la façon dont elles sont connectées aux deux autres. » (p. 5). L'idée n'est pas toujours l'aboutissement de l'enquête ; elle peut aussi être l'aiguillon poussant à réunir des données. Il arrive aussi que les données réunies ne permettent pas, *in fine*, de soutenir l'idée qui avait pourtant incité à les réunir, auquel cas le chercheur est incité à reformuler son idée initiale. Bref, la science est un processus par lequel on articule par tâtonnements et par reformulations, ces trois composantes : données, preuves, idées.

L'erreur est ce qui résulte d'un triplet mal ajusté : par exemple telles données ne soutiennent pas telle idée. Or « ce qui s'est produit par le passé peut se reproduire » (p. 10). En particulier, certaines erreurs se répètent. Le livre sert ainsi à pointer les erreurs les plus fréquentes, afin d'éviter aux sociologues de les commettre à nouveau. « La plus grande partie de ce livre reprend certaines erreurs typiques dans les données sociologiques et les rapporte à des méthodes de travail communément acceptées qui expliquent qu'elles perdurent » (p. 11).

Becker propose alors une analyse sociologique du travail de production des données. Il propose d'identifier le hiatus fréquent entre les intérêts de ceux qui conçoivent ou analysent les enquêtes, et ceux des « glaneurs de données » (*data gatherers*), qui ne peuvent qu'être différents puisqu'ils ne sont pas dans la même situation sociale, car il juge ce hiatus à la source de bien des erreurs. Ce livre de méthodologie se transforme ainsi en une sociologie de la production des données sociologiques.

Le chapitre deux opère un retour historique sur l'apparition de la dichotomie méthode quantitative / qualitative et montre qu'elle n'a rien de central dans l'argument. Au contraire, Becker rappelle le fait beaucoup trop méconnu, aux États-Unis comme en France, que cette

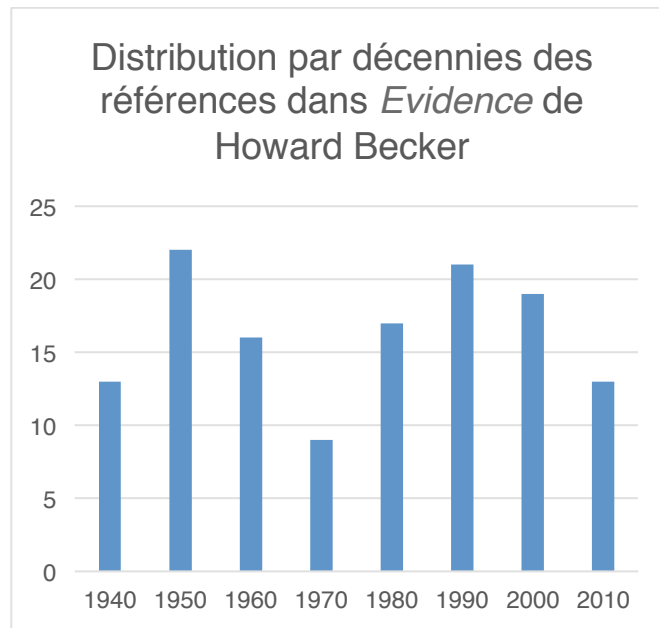
dichotomie n'a jamais été vraiment opérante à l'université de Chicago, et qu'il s'est lui-même toujours tenu à des vues parfaitement « œcuméniques » (22) sur ces différentes méthodes sociologiques. L'une comme l'autre permettent d'articuler données / preuves / idée de façon valide comme de façon erronée.

Produire des données sociologiques

La seconde partie du livre est construite autour de la taille des équipes qui produisent des données, car différentes tailles vont de pair avec différentes contraintes situationnelles. Cinq méthodes de production des données sont passées en revue : premièrement le recensement, deuxièmement les données officielles produites par les administrations publiques, troisièmement les données produites par les instituts de sondage, quatrièmement les données produites par les petites équipes regroupées autour d'un « chief investigator » (un professeur et ses élèves par exemple) et enfin, cinquièmement, les données produites par une personne seule. Chaque organisation sociale engendre des contraintes sur l'articulation des données et des idées et ainsi des risques d'erreurs classiques que l'auteur propose d'identifier.

Becker, contrairement à son habitude de sociologue, ne présente pas ici de terrain d'enquête. Il le remplace par une bibliographie qui, sans être gigantesque, a deux particularités. La première est que sur 145 références, 16 soit 11% sont écrites par un de nos compatriotes. La pensée française continue clairement d'être influente aux États-Unis (on ne trouvera aucun auteur allemand ou italien). Becker utilise en particulier notre tradition de socio-histoire de la statistique, représentée par les travaux d'Alain Desrosières qui occupe une place de premier ordre mais aussi, entre autres, par un dossier organisé autour d'un excellent article de Jean Péneff portant sur les enquêteurs de l'INSEE, qui est beaucoup trop peu (re)connu de notre côté de l'Atlantique. Becker lui accorde ici l'importance qu'il mérite.

L'autre caractéristique de la bibliographie de Becker est sa profondeur chronologique.



Il cite en moyenne un peu plus de 16 références par décennie, sans grande variation (avec l'exception intéressante des années 1970 qui lui ont laissé moins de souvenirs que les autres) sur toute la période durant laquelle il a travaillé. L'effet très plaisant de cette bibliographie est qu'elle plonge très en deçà des modes. La sociologie des données de Becker est clairement une socio-histoire.

Les impensés de la statistique

Concernant le recensement, Becker réfléchit sur le fait que le questionnaire contient toujours trop peu d'informations ou une information trop imprécise pour les sociologues. Par exemple, il est très difficile pour le recensement de compter les sans-abri. Pourquoi ? Car le processus d'énumération des individus repose sur une cartographie préalable des lieux d'habitation. Ceux qui n'ont pas de toit passent ainsi au travers des mailles du filet. Dans ce cas, ou bien on se satisfait de n'avoir un décompte qu'imprécis de la population totale – mais alors il faut en tenir compte lorsqu'on met en place des politiques sociales. Ou bien on met en place des procédures alternatives qui permettent de compter les sans-abri. Becker cite l'exemple d'une enquête conçue par Peter Rossi qui organisa des enquêtes nocturnes où l'enquêteur, étant donnée la violence endémique, était accompagné de deux policiers ! (p. 80-84). Becker en conclut que si l'on veut avoir des données vraiment précises, il faut alors y mettre le prix et les efforts, ce que, déplore-t-il, le gouvernement n'est pas prêt à faire.

Becker insiste ensuite sur les données des autres administrations publiques, dont les producteurs ont souvent des tâches à accomplir jugées prioritaires (par exemple, dans le cas de

la police, faire régner l'ordre au quotidien). De sorte que leurs données sont le résultat de comportements visant à « faire en sorte que leur vie professionnelle se déroule sans heurts » (*go smoothly*, p. 116). Il expose alors de nombreux exemples où cet impératif engendre des biais et insiste longuement sur celui des statistiques de délinquance qui informent autant sur l'activité policière (qui décide de faire de la lutte contre ces trafics une priorité ou non, et qui donc constate, ou non, des faits de délinquance) que sur l'activité délinquante (p. 121-128). Becker a bien sûr entièrement raison, mais le lecteur reste légèrement sur sa faim. En particulier, s'il montre le caractère socialement construit de la statistique, il ne prolonge pas son analyse sur les effets sociaux propres à ces statistiques une fois construites – par exemple comment elles transforment le management, engendrent différentes techniques de triche ou provoquent la production de statistiques alternatives.

Ainsi, il reste peut-être chez Becker un noyau de défiance concernant les données administratives. Ceci s'explique par l'histoire de la discipline. Aux États-Unis, l'opposition à la recherche quantitative par les sociologues « radicaux » a atteint son acmé pendant les années 1970, celles de la dénonciation libertaire de toutes les institutions et en particulier de l'État. Les radicaux considéraient que travailler avec les statistiques de l'administration revenait déjà pactiser avec elle. Becker, un des principaux représentants de ces derniers (Jack Katz, en entretien, rapporte que Becker portait un tee-shirt qui disait « Hey Kids, let's fuck the State ! »), professait déjà la position œcuméniste développée ici. Mais à l'époque, il concluait en invitant plutôt à « voyager léger » (*travel light*), c'est-à-dire de se passer de ces données. Ceci explique d'ailleurs aussi le peu de citations prises à cette décennie.

Pour les enquêtes ethnographiques légères, Becker insiste sur les avantages qu'il y a à produire *aussi* des données quantitatives. Il rappelle de nombreux exemples où des ethnographes ont effectué eux-mêmes des dénombrements simples.

Au final, il recommande, pour utiliser une enquête quantitative, de tirer les leçons d'une ethnographie des enquêteurs, car celle-ci fait apparaître un sens spécifique souvent inaperçu aux données. Mieux, elle permet souvent de produire des connaissances inattendues, surprenantes, sur la société. Inversement, une bonne ethnographie n'a aucune raison de se passer de données quantitatives que l'ethnographe peut souvent produire seul. Une bonne enquête sociologique articule données qualitatives et quantitatives. C'est dans l'espoir de confirmer cet argument que, dans cette recension même, nous avons procédé à quelques décomptes dans le texte de Becker.

Politique des données

En refermant ce livre, on peut exprimer deux regrets. D'abord, l'absence presque complète de discussion avec les ethnométhodologues. Ces derniers ont pourtant été pionniers

dans l'analyse sociologique de la quantification, en particulier les enquêtes de Garfinkel sur les usages des nombres dans l'hôpital, celles de Kitsuze et Cicourel avec leur argument de la « mesure par décret » (*measurement by fiat*) selon lequel les administrations imposent, comme par décret, leur définition des catégories sociales à la société, ou celles de Maynard sur la situation de passation de questionnaires. Ce silence reste d'autant plus étonnant qu'ils se connaissaient et se citaient auparavant.

D'autre part, la sociologie des données opérée par Becker rabat toujours ses observations vers le critère de la précision des données. Il qualifie ces dernières en termes de « accuracy », « precision » « good » « errors », « mistakes » (p. 69-74). Mais il ne s'empare pas réflexivement du critère de la précision. L'exemple archétypique est celui du recensement dont, nous l'avons vu, Becker regrette qu'il ne soit pas plus précis. Mais il n'entre pas dans une analyse socio-politique, au sens de l'école française de socio-histoire de la statistique, qui observerait comment les différents acteurs jugent et évaluent ces données.

Bien sûr, Becker parle de politique. Les mots dont la racine est « politic- » apparaissent 61 fois dans son texte, soit en moyenne près d'une fois toutes les trois pages, bien que la moitié des occurrences se trouve au chapitre 4 sur le recensement. Mais pour lui, la politique est systématiquement « manipulation » (p. 86). Or la politique est bien plutôt la condition de possibilité des données. C'est politiquement qu'un accord se fait sur la nécessité d'investiguer un problème, sur les sommes qu'il faut allouer, sur les meilleurs usages qui en seront faits. Le sociologue ne peut se contenter de déplorer l'attitude du gouvernement. Il doit comprendre sociologiquement comment et pourquoi telle solution politique s'est imposée.

Au final, nous recommandons bien sûr vivement la lecture de ce livre. Se présentant comme un livre de méthodologie, il propose en fait une socio-histoire des données sociologiques. Les propositions œcuméniques qui y sont exposées et son style détendu en font une source inépuisable et fort plaisante d'exemples qui enrichiront les cours de méthode et la réflexion épistémologique de tous les chercheurs en sciences sociales et humaines.

Publié dans lavidessidees.fr, le 30 mai 2018.