



**Business School**

**WORKING PAPER SERIES**

**Working Paper**

2015-618

**La place du raisonnement statistique  
dans les sciences sociales et dans  
l'analyse des comportements humains**

Frédéric Teulon

<http://www.ipag.fr/fr/accueil/la-recherche/publications-WP.html>

IPAG Business School  
184, Boulevard Saint-Germain  
75006 Paris  
France

IPAG working papers are circulated for discussion and comments only. They have not been peer-reviewed and may not be reproduced without permission of the authors.

# La place du raisonnement statistique dans les sciences sociales et dans l'analyse des comportements humains

---

Frédéric Teulon

IPAG Business School

*« Les chiffres sur la déviance sont obtenus par des statistiques établies dans des organisations particulières et ils peuvent être expliqués par les activités de ces organisations et leur rôle dans la désignation de la déviance. »*

Aaron CICOUREL, John KITSUSE (1963)

*« Si les sciences sociales doivent devenir véritablement des sciences et, pour parler court, si elles doivent continuer d'exister d'ici vingt ans, il est indispensable qu'une réforme soit opérée de toute urgence. On peut, dès aujourd'hui, être certain que les jeunes spécialistes de sciences sociales devront désormais posséder une solide et moderne formation mathématique, sans quoi ils seront balayés de la scène. »*

Claude LEVI-STRAUSS (1954)

## Résumé

Les enjeux liés à l'utilisation des statistiques en sociologie sont présentés au travers d'une revue de la littérature. La sociologie positiviste et empirique a conféré aux enquêtes et aux chiffres une place centrale. Cette place est inscrite dans les différentes étapes du raisonnement sociologique. L'utilisation démultipliée des statistiques fait de cette pratique un fait social et révèle le fonctionnement d'une société basée sur la production et la manipulation de chiffres.

Mots clés : sciences sociales, objectivité, statistiques, suicide.

## Introduction

Le raisonnement statistique repose sur des méthodes mathématiques de traitement de données chiffrées. Il est un outil puissant d'analyse permettant d'évaluer de façon quantitative les phénomènes sociaux. Il s'agit de recueillir et d'élaborer des données chiffrées au travers de sondages, d'enquêtes d'opinion, de données d'état civil, de recensements de population... Les données numériques permettent d'effectuer des classements, dénombrements, inventaires ou répertoires servant à classer et à organiser les relations sociales. Elles permettent aussi de valider ou d'infirmier des modèles mathématiques.

On le sait, les mathématiques sont devenues une discipline intellectuelle sous l'Antiquité avec Archimède, Euclide ou Pythagore. Elles sont un ensemble de connaissances abstraites provenant de raisonnements logiques et de modélisations, établis en utilisant des chiffres, des symboles, des équations ou des graphiques.

A l'évidence, les mathématiques et les statistiques ne peuvent pas jouer le même rôle dans les sciences humaines que dans les sciences dures. Les analyses sociologiques ou historiques portent sur des sociétés, donc sur des entités dynamiques et mouvantes qui changent en permanence.

La question du statut des statistiques s'est posée dès les origines de la sociologie. Cette discipline est née à la fin du XIX<sup>e</sup> siècle - époque qui bénéficie d'un essor des mathématiques (voir les travaux des mathématiciens Gauss, Laplace, Cauchy, Poisson...) - avec Emile Durkheim qui affirme que la sociologie ne peut être bâtie que sur l'étude des faits. Allant à l'encontre de ceux qui reprochent à la statistique d'atomiser le social, Franklin Giddings - le pionnier de la sociologie inductive et quantitative aux Etats-Unis - insiste dans son livre de 1911 sur le fait que le recours aux statistiques amène les sociologues à étudier le cœur de leur sujet : les phénomènes sociaux à un niveau agrégé. De son côté, Ferdinand Tönnies en appelle à la création d'une science empirique du social : la sociographie, afin « *d'affronter le réel dans sa vérité et sa totalité* »... Comment le raisonnement statistique s'articule aux théories ou à aux problématiques de recherche en sociologie ? Quelle est la nature de ce débat ?

Nous verrons dans une première section que les statistiques occupent une place centrale dans les sciences humaines. La section 2 rattache l'utilisation des statistiques aux deux étapes de l'analyse en sciences sociales. Enfin la section 3 montre que cette utilisation des statistiques pour étudier les comportements humains ne peut pas être exclusive et qu'elle est contestée.

## I. Une place centrale ?

### Une posture méthodologique

Les statistiques portent sur des faits. Leur utilisation est donc supposée être une posture méthodologique permettant d'atteindre l'objectivité. Dans les sciences sociales, les statistiques permettent l'objectivation ; elles s'efforcent d'organiser le réel pour le rendre plus intelligible. Elles permettent d'installer une distance entre l'observateur et les sujets de son étude (distanciation du sujet).

Cette posture est à la base des travaux de Durkheim :

- lorsqu'il affirme qu'il faut traiter les faits sociaux comme des choses ;
- lorsqu'il effectue sa fameuse distinction entre le normal et le pathologique (le pathologique n'est pas défini comme le mauvais état de santé des individus, mais comme un état du corps social qui s'écarte de la norme statistique habituellement observée).

C'était déjà l'ambition de Condorcet lorsqu'il voulait construire une science de l'homme basée sur l'utilisation des mathématiques. Ambition également d'Adolphe Quetelet (1835) - astronome, statisticien et démographe belge - qui souhaitait créer une "physique sociale" où le calcul des probabilités jouerait un rôle central. Quetelet espérait qu'il serait un jour possible d'établir en démographie, des lois aussi rigoureuses que celles qui commandent le mouvement des astres.

Dès le début du XIX<sup>e</sup> siècle, les précurseurs de la sociologie ont affiché leur désir d'utiliser des statistiques, volonté qui n'a été freinée que par le manque de données ou le manque de fiabilité des chiffres à leur disposition. Au travers des statistiques, leur ambition était de dépasser les situations individuelles pour faire émerger des tendances collectives. Comme le rappelle Maurice Halbwachs (1935) : « *La statistique permet d'atteindre les caractères d'un groupe, qui ont une réalité pour le groupe tout entier, mais qu'on ne découvre dans aucun membre de ce groupe pris à part isolément : par exemple, la durée de vie moyenne de vie dans un groupe d'hommes, qui ne sera, en général, la durée de vie exacte d'aucun de ces hommes pris au hasard, qui est cependant une réalité, puisqu'on la retrouve identique, pour le même groupe à plusieurs époques successives, et, à la même époque, pour plusieurs groupes composés d'individus répartis de la même manière quant à l'âge.* »

L'étude de Durkheim sur le suicide est emblématique. Au XIX<sup>e</sup> siècle, le suicide était un phénomène que l'on tentait d'expliquer par la psychologie, l'astrologie ou la biologie. On considérait que l'acte de se suicider était une décision personnelle par excellence et qu'à ce titre la seule explication possible se trouvait dans l'individu lui-même. Pourtant Durkheim montre que quelle que soit la période étudiée le taux social de suicide reste constant dans le temps. Cette régularité statistique ouvre la voie à une analyse proprement scientifique de cette question.

L'après Seconde Guerre mondiale s'est traduite par un regain d'intérêt pour l'utilisation des mathématiques dans les sciences humaines : « l'idée d'une « sociologie mathématique » c'est-à-dire d'une science du social reposant directement sur les théories et les instruments mathématiques est prise au sérieux, notamment outre-Atlantique ; et plusieurs sociologues qui deviendront des figures de la discipline recourent aux mathématiques (Paul Lazarsfeld,

Harrison White, James Coleman...). Ils vont faire feu de tout bois pour donner à la sociologie des fondements mathématiques ou, plus simplement, pour étayer leurs analyses sociologiques par des outils mathématiques » (Martin, 2002).

## Une pensée d'Etat

La production de statistiques peut obéir à des objectifs de gestion administrative, cela été le cas par exemple de la statistique de la justice criminelle, qui s'est développée en France dès les années 1820 et a évolué par la suite parallèlement avec la politique pénitentiaire, la criminologie puis la sociologie de la délinquance.

Une rupture s'est opérée avec la création des Instituts de sondage et des Instituts spécialisés dans la production de statistiques. Dès 1850, en Allemagne, Ernst Engel crée l'Office statistique saxon. Les statistiques deviennent alors un moyen pour l'Etat d'avoir une emprise plus forte sur la société. Elles permettent d'administrer et de contrôler.

Les statistiques cristallisent des débats dans lesquels l'Etat est partie prenante. C'est le cas aujourd'hui en France pour la question de l'ethnicisation des statistiques (Héran, 2014) ; en interdisant ce type de statistiques, l'Etat français a fait le choix de l'ignorance pour défendre le projet d'une société *color blind*. L'objectif poursuivi ici est de limiter les discriminations : « *En codifiant a priori des groupes, elles fournissent simultanément les moyens de l'analyse scientifique quantitative et les labels qui serviront à qualifier les groupes qu'elles représentent, et éventuellement à les discréditer* » (Simon, 2008).

Dans la tradition interactionniste (Becker, 1963), la déviance n'est pas une propriété inhérente à certains comportements, mais elle résulte plutôt d'une construction lors d'interactions entre plusieurs types de personnes : ceux qui sont qualifiés de "déviant", ceux qui font respecter les normes et ceux qui cherchent à en imposer de nouvelles. La déviance n'est une qualité de l'acte commis par une personne, mais la conséquence de l'application par les autres, de normes, de catégories statistiques et de sanctions.

## Une mesure de l'opinion

Aux Etats-Unis, George Gallup fonde un Institut de sondage en 1935 et développe une méthode d'enquête par questionnaires individuels qui reçoit une consécration éclatante en 1936, lorsqu'elle lui permet de prédire la réélection de F. D. Roosevelt lors de l'élection présidentielle, avec une erreur très réduite. En France, c'est Jean Stoetzel fonde et développe l'IFOP (1939) à partir de plusieurs idées fortes :

- la recherche en sciences sociales ne doit pas relever de méthodes artisanales. Elle nécessite une organisation de type industriel ;
- l'analyse des enquêtes d'opinion doit révéler les lois qui régissent les attitudes et les comportements ;
- les résultats des enquêtes passées doivent rester disponibles.

« Le vulgaire pointage des opinions individuelles qu'effectuent de manière régulière les sondages s'est imposé de manière consensuelle dans le monde entier comme la définition de base de l'opinion publique » (Converse, 1987). Les sondages sont à la fois un outil de construction et d'expression de l'opinion publique. Leur succès s'explique par la fascination pour les chiffres propres à nos sociétés contemporaines. Comme le dit avec justesse Loïc Blondiaux (1997) : « *Par la magie de l'agrégation statistique, la masse informe des désirs et des passions du public se transforme en chiffres et en données propres au raisonnement arithmétique et à l'analyse politique* ».

La sociologie peut difficilement se passer de statistiques, il ne s'agit pas uniquement d'un lien pratique, mais d'une dépendance théorique (Héran, 1984). Il convient de prendre au sérieux le calcul statistique et de la considérer comme un fait social et pas seulement comme un simple outil scientifique. Les chiffres ont un triple visage :

- rationalité pratique et théorique qui renforce la rationalité épistémique des sociologues ;
- instrument de lutte pour la détention de l'autorité intellectuelle dans l'espace académique ;
- méthode performative qui instaure le primat du calcul.

## **II. Les étapes de l'analyse en sciences sociales**

Une démarche de recherche repose sur la définition d'un problème, la formulation des hypothèses, la définition d'un cadre méthodologique, la production ou collecte des données, le traitement qualitatif et/ou quantitatif des données, l'analyse des résultats et administration de la preuve. Les statistiques sont utilisées dans les deux grandes étapes de la démarche en sciences sociales.

Lazarsfeld (1954) s'est efforcé de codifier les procédures méthodologiques des enquêtes, d'élaborer des techniques d'analyse des données (analyse multivariée, panel). Il a insisté sur l'ambiguïté des mots issus du langage courant et propose de spécifier précisément chaque concept de façon à pouvoir le définir par des relations mathématiques.

### **1<sup>ère</sup> étape : Sélection, production et classement des données**

Pour faire œuvre de sociologie empirique, il faut disposer de données. Celles-ci peuvent être disponibles, il suffit alors de les extraire. C'est ce que fait Durkheim lorsqu'il utilise les statistiques de l'état civil dans son étude sur le suicide. Dans le cas contraire ces données doivent être produites par des enquêtes ou des sondages.

Ensuite, le classement de ces données suppose la construction des catégories statistiques (des nomenclatures). Il est plus simple pour le sociologue de reprendre les catégories déjà socialement validées par l'usage et utilisées par les appareils statistiques officiels et les administrations publiques. Ces catégories sont issues d'un lent travail de maturation et de rationalisation administrative. Elles ont souvent une dimension juridique (à l'exemple des données de l'état civil).

La répartition professionnelle des individus dans différents groupes pose problème. Chenu (1997) insiste sur l'identité incertaine de certains groupes socioprofessionnels. Il montre que les catégories sociales en déclin numérique (agriculteurs, ouvriers) sont celles dont les frontières sont les plus nettes et la segmentation interne la plus réduite, alors que les catégories en expansion présentent les caractéristiques inverses : des contours flous et une segmentation interne élevée, selon les métiers, selon le sexe et selon le statut (salariés du public versus salariés du privé, notamment).

L'explication en sociologie repose souvent sur l'exploitation de données issues d'une enquête qui sont présentées dans un tableau croisé. Ainsi pour étudier le comportement électoral, les sociologues/politologues recourent à des enquêtes par questionnaires ou à des sondages afin : 1/ de mettre en évidence des caractéristiques sociodémographique (sexe, âge...), socio-économique (niveau de revenu...) et socioculturelle (niveau d'études...) des électeurs ; 2/ d'explicitier les mécanismes et les modalités des décisions ; 3/ d'analyser la signification du vote (l'ensemble des représentations, sentiments, opinions, attitudes associées au comportement électoral).

## **2<sup>ème</sup> étape : Explication, mise en évidence de régularités, validation théorique et administration de la preuve**

Les chiffres ne suffisent pas, ils ne sont qu'un support, la compréhension des phénomènes sociaux repose sur la précision des concepts, la rigueur et l'intelligence de l'observateur.

Les statistiques sont souvent utilisées pour affirmer le primat du collectif et la force que la société exerce sur les individus. Ainsi dans une étude effectuée en 1831, Quetelet analyse l'influence de facteurs tels que le sexe, l'éducation, l'âge, le climat et les saisons sur le taux de criminalité. Il affirme qu'il est impossible de connaître qui est susceptible de commettre tel crime particulier, mais, selon lui, la régularité des statistiques permettent de « prévoir le nombre d'individus qui souilleraient leurs mains du sang de leurs semblables, le nombre de faussaires ou le nombre d'empoisonneurs. » La découverte de ces régularités l'amène à une conclusion radicale : « C'est d'une certaine manière la société qui prépare tous ces crimes, dont les criminels ne sont que les exécutants. » Durkheim se situe dans la même perspective lorsqu'il dit à propos du suicide : « Chaque année nous apporte son contingent de morts volontaires. »

L'utilisation des statistiques permet de valider ou non des hypothèses ou des questions de recherche. Elle permet de décrire des évolutions. Alors qu'Halbwachs (1913) insistait sur la ligne de démarcation entre le groupe des ouvriers et celui des employés (avec chacun leur logique propre), Alain Chenu (1990) montre que le monde des employés est un monde stratifié, avec une aristocratie des emplois administratifs et un prolétariat des services qui est proche de la classe ouvrière.

Associées à un calcul de probabilités les statistiques peuvent être utilisées pour valider des hypothèses théoriques. C'est ce que fait Casey Mulligan (2003) lorsqu'il teste l'hypothèse d'un calcul conduisant les électeurs à ne pas se déplacer pour aller voter. En étudiant les résultats de plusieurs dizaines de milliers d'élections au Congrès et dans les assemblées des Etats américains, Mulligan montre que la probabilité que la voix d'un électeur particulier influence une élection est rarissime. Il est extrêmement peu probable que le résultat se joue à une voix. Le cas s'est présenté à sept reprises en 40 000 élections aux assemblées d'Etat, et une fois seulement lors des 16 000 élections au Congrès, en 1910 à Buffalo ! En conséquence les statistiques confirment le calcul de l'électeur rationnel.

L'utilisation de statistiques fait de la sociologie une discipline d'observation de la société dégagée de toutes perspectives métaphysiques (Comte, 1830). Cette utilisation ouvre la voie à l'élaboration de lois. Elle est au cœur de nombreuses analyses sociologiques à caractère déterministes :

- le positivisme. Ainsi Quetelet voit dans les statistiques de la nuptialité une première confirmation de ses hypothèses : « Le mariage suit une marche si régulière que le nombre annuel de mariages se reproduit avec une constance plus grande que les phénomènes naturels. » Par ailleurs, Quetelet propose de définir les individus à partir de distributions statistiques, il construit une théorie de "l'homme moyen", considérant que tout individu est une variable aléatoire distribuée autour d'un type moyen. Toutes les caractéristiques de cet homme moyen (mensuration, propension à l'alcoolisme...) seraient les moyennes des caractéristiques des hommes réels. Le positivisme consiste à appliquer les méthodes utilisées en mathématiques et dans les sciences expérimentales aux phénomènes sociaux et politiques afin de dégager les lois qui régissent la structure et le développement des sociétés.

- la sociologie durkheimienne et ses prolongements. Ainsi Simiand (1908) se démarque d'une vision formelle ou hypothético-déductive de l'économie. Il s'oppose à « l'économie abstraite », non par ce qu'elle fait usage de l'abstraction, mais parce qu'il estime qu'une théorie ne doit pas tirer sa validité de sa seule cohérence logique ; les propositions déduites des hypothèses doivent être vérifiées. Il dénonce la méthode des partisans de l'homo œconomicus (ou homo sociologicus) qui ne disent pas que les choses se passent ainsi parce que l'expérience l'a établi, mais qui soutiennent plutôt qu'elles doivent se passer ainsi parce qu'elles seraient absurdes autrement.

- la théorie des « champs ». Pierre Bourdieu (1984) met en évidence le « système de positions » des Universitaires, réseaux de relations objectives entre les acteurs, en utilisant des analyses factorielles de correspondance (qui permet de déterminer et de hiérarchiser les relations de dépendance entre les lignes et les colonnes d'un tableau statistique).

### **III. Une place contestée**

L'analyse critique des traitements statistiques envoie à des problèmes méthodologiques de construction des enquêtes, de fiabilité des chiffres utilisés et d'interprétation.

#### **Problèmes de méthode**

La fiabilité des chiffres utilisés en sciences humaines pose problème.

Les résultats des enquêtes dépendent étroitement de la manière dont les questions sont posées et de la nature de l'interaction entre l'observateur et les personnes observées. Peneff (1984) s'intéresse à l'utilisation des données chiffrées par les sociologues en prenant comme exemple la statistique du métier du père des étudiants de première année de l'université de Nantes. Il montre que le contexte dans lequel les questions sont posées est important (selon le type



d'interaction avec l'enquêteur, les étudiants vont percevoir le questionnaire de manière différente). Les statistiques officielles sont biaisées car on demande aux étudiants des réponses trop courtes et parce que les codeurs ne sont pas dans de bonnes conditions de travail.

Les problématiques que fabriquent les instituts de sondages d'opinion sont subordonnées aux demandes particulières des commanditaires. Les données statistiques peuvent être présentées de façon à obtenir telle ou telle réponse souhaitée.

Les sondages comportent une marge d'erreur. Aucune enquête même si elle est réalisée à grande échelle ne peut prétendre échapper à l'approximation. Par ailleurs, on sait que l'activité des laboratoires de recherche est avant toute chose orientée vers la publication d'articles scientifiques plus que vers la mise en évidence de la vérité (Latour & Woolgar, 1979).

Dans les enquêtes les « non réponses » posent des problèmes d'interprétation. Il est difficile de savoir si les opinions émises par les individus correspondent exactement à ce qu'ils pensent. Elisabeth Noelle-Neumann (1993) montre que les individus peuvent renoncer à leur propre jugement (ils préfèrent se taire plutôt que de se sentir rejetés). C'est là une condition nécessaire au fonctionnement de toute société humaine, s'il en allait autrement l'intégration serait impossible. En conséquence, Neumann définit l'opinion publique comme « celle qui peut être exprimée en public sans risque de sanctions. »

Un autre problème est plus rarement évoqué : celui des personnes qui n'apparaissent jamais dans les statistiques. Stéphane Beaud (2006) montre qu'il existe des populations qui, malgré leur nombre, sont masquées (en dehors des statistiques), volontairement ou non, par les chiffres, le droit, le discours politique, les représentations médiatiques, les politiques publiques, les études sociologiques ou les catégorisations dépassées qui occultent leurs conditions d'existence. Elle conduit à des conclusions parfois surprenantes : la proportion de précaires est plus élevée dans le public que dans le privé, de plus en plus de personnes ne demandent pas les prestations sociales auxquelles elles ont droit, la plupart des SDF ont une adresse, la moitié des adolescents qui se suicident sont homosexuels, les licenciés qui retrouvent un emploi connaissent presque systématiquement une perte de revenu...

Les sociologues insistent sur le fait que les statistiques sont socialement construites (elles sont élaborées à partir de choix qui ont un impact sur les résultats obtenus). Alain Desrosières (2000 et 2008) montre que la production de chiffres et la construction de catégorie traduisent un processus social, émanant non seulement des organismes scientifiques officiels et labellisés (Instituts publics de recherche, ministères...), mais aussi d'autres sources (syndicats, groupes d'intérêt...) et de ce fait objet de conflit entre acteurs prenant part au débat social.

Par ailleurs, les statistiques tendent à devenir un outil de gouvernance avec le risque que cette activité d'expertise se substitue à la délibération publique, instituant une coupure entre ceux qui manipulent les chiffres et le reste de la population.

## **A propos des statistiques du suicide**

Par nature les données statistiques sont imparfaites et il y a des doutes sur les usages que l'on peut faire des statistiques officielles. Ainsi la pertinence des données statistiques utilisées par Durkheim dans son étude sur le suicide a été par la suite fortement contestée (Pescosolido et Mendelsohn, 1986).

Halbwachs (1930) insiste sur la fragilité de certaines démonstrations de Durkheim dans son analyse du suicide (il nuance les analyses sur les facteurs de l'intégration sociale et sur les causes des suicides). Les données statistiques supplémentaires dont il dispose lui permettent de faire apparaître la fragilité de certaines démonstrations de Durkheim. Par exemple, l'opposition entre catholiques et protestants en Allemagne n'est pas forcément pertinente car les protestants ne sont pas seulement protestants, ils sont aussi plus urbains que les catholiques, en outre leur distribution dans l'espace socioprofessionnel est différente (effets de structure). Halbwachs souligne l'importance qu'il y a à considérer non seulement les suicides "réussis", mais aussi les tentatives de suicide ; les uns et les autres apparaissent comme étant différemment distribués et liés à des variables telles que le sexe ou l'âge. N'oublions pas que dans la société française du XIX<sup>e</sup> siècle la religion occupait une place prépondérante, les familles et l'entourage pouvaient donc être tentées de transformer les suicides de leurs proches en accidents lorsque cela était possible.

Selon Jack Douglas (1967), le sens social du suicide diffère d'une société à une autre. Plus une société tend à intégrer ses membres dans un projet collectif, plus le suicide est socialement réprouvé comme contraire à la volonté de vivre ensemble, plus les suicides sont occultés et plus les statistiques sont trompeuses. Ainsi les groupes qui ont une forte cohésion sont plus enclins que les autres à transformer des suicides en morts accidentelles. Douglas estime que l'étude de la mort volontaire ne doit pas se faire à partir de statistiques, mais être basée sur l'étude de cas individuels.

Dans la même perspective, Maxwell Atkinson (1968, 1978) s'interroge sur les critères de classification qui amènent à trancher en faveur d'une mort naturelle ou d'un suicide. Il montre comment les statistiques du suicide sont construites socialement par les interprétations des « coroners » (officiers royaux qui sont chargés en Grande Bretagne d'élucider les causes des morts suspectes). Les suicides ne sont pas des « faits », ils résultent d'interprétation de situations et ils sont affectés par les conditions de leur collecte.

Le suicide peut être étudié à partir de matériaux non chiffrés divers : rapports de police, les études ethnologiques de Malinowski, le roman de Truman Capote *De Sang Froid*, des lettres de suicidés... On peut prendre le contre-pied de Durkheim et réfuter l'idée de « courants suicidogènes » ayant uniquement un caractère collectif et social.

## **Un test pour la statistique : la mesure de la délinquance**

La déviance n'est pas en soi le fait de transgresser des normes, mais le fait d'être étiqueté comme déviant par la société. Le caractère déviant ou non de l'acte dépend de la manière dont l'entourage, le voisinage ou l'opinion réagissent. Il est donc important de distinguer le fait même de la déviance, de sa perception.

Cicourel et Kitsuse (1963) montrent que les enfants de milieux populaires sont surreprésentés dans les statistiques de la délinquance juvénile car leurs parents n'ont pas le bagage culturel qui leur permettrait de négocier un arrangement avec la police comme cela est le cas dans les classes moyennes. Ils considèrent que les délinquants sont produits par les agences de contrôle social. Certains individus sont labélisés déviants au cours d'un processus d'interaction.

Cicourel (1968) étudie deux villes de Californie de taille comparable ayant des caractéristiques socio-économiques similaires. Pourtant le taux de délinquance est différent. Pour Cicourel et Kitsuse (1963), cette situation s'explique par des différences dans la taille, l'organisation et les politiques menées par les polices locales : « *Les chiffres sur la déviance*

*sont obtenus par des statistiques établies dans des organisations particulières et ils peuvent être expliqués par les activités de ces organisations et leur rôle dans la désignation de la déviance. Ainsi, ces statistiques doivent être considérées comme les indicateurs de processus institutionnels plutôt que comme les indicateurs d'incidence de certaines formes de comportements: les variations détectées de comportements déviants (par exemple parmi les Noirs) telles qu'elles apparaissent dans les statistiques de différentes organisations peuvent être le produit de définitions différentes du comportement déviant utilisées par ces organisations, différences résultant des conditions idéologiques, politiques et organisationnelles affectant le processus de production statistique. »*

## **L'individualisation statistique**

La statistique sépare l'individu de son groupe d'appartenance en considérant qu'il est une unité autonome dont le poids identique à celui de n'importe quelle autre personne vient s'ajouter à l'ensemble.

De nombreux auteurs, comme Herbert Blumer (1948), Pierre Bourdieu (1973) ou Philip Converse (1987), se sont montrés très critiques vis-à-vis des sondages d'opinion. Pour eux, les opinions recueillies dans les sondages politiques sont instables, superficielles et incohérentes, comparables à des pseudo-attitudes ou « non attitudes ».

Bourdieu conteste l'utilisation des sondages d'opinion considérant qu'ils conduisent à des distorsions qui s'observent lors même que toutes les conditions de la rigueur méthodologique sont remplies dans la collecte et l'analyse des données : « Toute enquête d'opinion suppose que tout le monde peut avoir une opinion ; ou, autrement dit, que la production d'une opinion est à la portée de tous. Quitte à heurter un sentiment naïvement démocratique, je conteste ce premier postulat. Deuxième postulat : on suppose que toutes les opinions se valent. Je pense que l'on peut démontrer qu'il n'en est rien et que le fait de cumuler des opinions qui n'ont pas du tout la même force réelle conduit à produire des artefacts dépourvus de sens. Troisième postulat implicite : dans le simple fait de poser la même question à tout le monde se trouve impliquée l'hypothèse qu'il y a un consensus sur les problèmes, autrement dit qu'il y a un accord sur les questions qui méritent d'être posées.»

Lazarsfeld parle des deux étapes du flux de la communication » (*the two step flow of communication*). Non seulement les individus ne sont pas isolés dans l'exposition aux messages parce qu'ils les reçoivent à l'intérieur d'un groupe social d'appartenance (d'où une réinterprétation), mais ils ne les reçoivent pas directement : les messages sont réinterprétés par des « leaders d'opinion » qui peuvent les relayer ou au contraire les neutraliser (Lazarsfeld et Katz, 1955).

## **Causalité et hasard**

Les neurosciences et les sciences comportementales montrent que les comportements individuels sont sujets à des biais (Shiller, 2000). Selon Paul Bloom (2014), les êtres humains ont un « penchant fataliste ». Ils considèrent que ce qui peut leur arriver se produit nécessairement pour de bonnes raisons (ils ont tendance à sur-interpréter les événements et ne peuvent se résigner au hasard).

Ce problème se retrouve au niveau des études scientifiques. Le “préjugé nomologique” est celui qui consiste à vouloir utiliser les statistiques pour établir des lois. Selon Boudon (1984), il existe une tendance naturelle chez les sociologues à sous-estimer la place du hasard et à attribuer une validité générale à des réalités macroscopiques. Au-delà de l'influence du positivisme, la sous-estimation du hasard et la vision déterministe du monde souvent développée par les sciences sociales et par l'histoire trouve son origine dans les processus naturels de la pensée humaine. Pour Boudon, les sciences sociales ne peuvent qu'échouer à découvrir un ordre dans le désordre qu'elles étudient sous la dénomination de changement politique, social ou économique. Le programme qui est à la base de cette démarche ne saurait s'appliquer à la société, car le changement prend mille formes imprévisibles dont on ne peut démontrer ni la tendance, ni la loi, ni la cause. Ce qui est en jeu ce n'est pas l'étude du changement social qui est indispensable, c'est la volonté de généraliser des relations qui sont locales. Dans la même perspective, lorsque Christopher Jencks (1972) étudie les modes d'acquisition du statut social de manière probabiliste, il affirme que: « Bien que des origines sociales élevées rendent plus probable un niveau scolaire élevé et qu'un niveau scolaire élevé rende plus probable un statut et un revenus élevés, il existe une importante mobilité d'une génération à la suivante, de sorte que l'on trouve à peu près autant d'inégalité économique entre des frères élevés ensemble qu'entre des individus quelconques. »

On peut se poser la question du statut de la « réfutabilité » (Karl Popper) en sociologie. La collecte et le traitement de statistiques peut créer une illusion de scientificité (Passeron, 1991). La sociologie appartient au même cadre épistémologique que l'histoire. Elle ne peut avoir véritablement le statut de science du fait de la conjonction de plusieurs problèmes :

- il est difficile de mener des expériences (la base empirique fait défaut) ;
- les conditions de l'observation historique excluent les plus souvent la généralisation (existence de situations non reproductibles). Les faits observés dépendent du contexte spatio-temporel ;
- le statut des mots utilisés par les sociologues n'est pas suffisamment stabilisé pour que le savoir puisse devenir cumulatif ;
- les conditions sociales dans lesquelles l'énonciation sociologique se fait ont impact sur ce qui est dit.

### **La critique de Wilhelm Dilthey**

Wilhelm Dilthey s'est efforcé de garantir l'autonomie et la spécificité des “sciences de l'esprit” qui, du fait de leur objet, ne peuvent se développer sur le modèle des sciences de la nature. Il entend fonder la psychologie, la métaphysique et les conceptions du monde sur l'histoire (en reconstruisant le contexte historique et culturel dans lequel tel ou tel phénomène s'inscrit). Dilthey s'interroge sur l'objectivité des sciences historiques. Il veut répondre à la crise de l'historicisme (analyses de Schmoller selon lesquelles le savoir historique ne peut être construit que sur une accumulation considérable de faits) par une “critique de la raison historique”. Il est le fondateur de la tradition allemande de la “sociologie compréhensive” et à ce titre il exercera une influence déterminante sur Max Weber. Les sciences de l'homme ne peuvent déboucher sur la mise en évidence de lois générales car l'homme est un être libre qui apporte des significations à ses actes, il y a donc nécessité d'interpréter les actions humaines et les mathématiques en sont incapables (Levi-Strauss qualifie la conscience humaine « d'ennemie secrète des sciences de l'homme »). Dilthey a écrit son œuvre principale en 1910: *Der Aufbau der geschichtlichen Welt in den Geisteswissenschaften (L'Édification du monde historique dans les sciences de l'esprit)*. Il est également l'auteur d'une *Introduction*

*aux sciences de l'esprit* (1887). Dilthey s'insurge contre le positivisme de la sociologie d'Auguste Comte qui s'en tient à une conception de la science limitée aux sciences de la nature. Dilthey considère que dans le domaine des sciences sociales, la connaissance ne peut reposer sur une reconstitution du réel (à l'image de l'historien qui reconstitue et analyse le passé en fonction de valeurs auxquelles il adhère). Cependant, les sciences sociales ne peuvent pas renoncer à toute forme d'objectivité. Il est nécessaire de construire des méthodes adaptées à l'objet étudié. Dilthey centre son interprétation sur l'histoire de l'esprit humain, qui se manifeste dans des cultures différentes et dans les différentes circonstances selon des "visions du monde" spécifiques (*Weltanschauung*). Puisque la sociologie s'occupe de faits signifiants, ils doivent donc donner lieu à une compréhension qui en dégage le sens. Cette compréhension est présentée comme un effort que doit faire le chercheur pour retrouver la signification que les acteurs sociaux donnent à leurs actions. Il faut revivre ce que d'autres ont vécu sans pour autant tomber dans la psychologie (il s'agit d'éviter que la compréhension s'inscrive dans une perspective uniquement subjective). Selon Dilthey, il est nécessaire de replacer les phénomènes étudiés dans des ensembles plus vastes où ils trouvent leur signification.

## Conclusion

Les mathématiques et les statistiques occupent une place prépondérante dans l'analyse sociologique du fait d'une volonté d'objectivité rendue possible par la distanciation.

L'objectif initial des sociologues a été de faire de leur discipline une science, ils ont donc tenté autant que possible d'utiliser les méthodes se rapprochant de celles des scientifiques. Cependant, dans la mesure où la sociologie s'attache aux comportements humains subjectifs et où le sociologue a son propre système de valeurs, l'analyse sociologique ne peut être neutre et les sociologues ont une démarche interprétative (et parfois militante). Malgré cette apparente contradiction, les sociologues continuent officiellement d'essayer d'objectiver leurs analyses.

Le raisonnement statistique est la base de la plupart des analyses sociologiques. Il intervient aux différentes étapes de l'analyse : production de données, observation, inférence sur la base des régularités observées, explication, administration de la preuve. Il est indispensable, sauf si l'on s'en tient aux monographies et des approches interactionnistes.

Les statistiques permettent d'objectiver des résultats et de les présenter comme des lois générales à caractère scientifique. Elles permettent de mieux comprendre la société, mais le sociologue n'a-t-il pas aussi une responsabilité, celle d'aider la société à changer (Lynd, 1939) ? Ceci nous renvoie au débat sur la neutralité l'observateur et au rapport aux valeurs. Les statistiques ouvrent la voie à une sociologie quantitative, mais l'usage croissant des modèles mathématiques dans l'analyse des faits sociaux n'est-elle pas une intrusion épistémologique injustifiée du fait de la nature de ce qui est observé ? Cela nous renvoie au débat sur la légitimité de l'extension du calcul économique à toutes les sphères de la société.

Il convient d'ouvrir le débat sur le statut et les usages de la statistique et de mettre en avant un plaidoyer en faveur d'une science empirique du social que Tönnies (1929) nommait la sociographie. Mais il convient de dénoncer la prétention de la statistique à satisfaire la totalité des exigences d'une analyse empirique du social. Selon nous, la sociologie ne peut pas se passer d'une articulation entre les méthodes quantitative et qualitative.

## Références

- ATKINSON Maxwell (1968), "On the Sociology of Suicide", *The Sociological Review*, vol. 16 (1), pp. 83-92.
- \_\_\_\_\_ (1978), *Discovering Suicide. Studies in the social Organisation of sudden death*, London, Macmillan.
- BEAUD Stéphane, éd., (2006), *La France invisible*, La Découverte.
- BECKER Howard (1963), *Outsiders: Studies in the Sociology of Deviance*, New York, The Free Press.
- BEIRNE Piers. 1987. "Adolphe Quetelet and the Origins of Positivist Criminology", *American Journal of Sociology*, vol. 92, 1140-1169.
- BESNARD Philippe (1976), "Anti- ou anté-durkheimisme ? Contribution au débat sur les statistiques officielles du suicide", *Revue française de sociologie*, XVII, n° 2, 313-341.
- BLONDIAUX Loïc (1997), "Ce que les sondages font à l'opinion", *Politix*, n° 37, 117-136.
- \_\_\_\_\_ (1998), *La fabrique de l'opinion. Une histoire sociale des sondages*, Paris, Le Seuil.
- BLOOM Paul (2014), "The war on reason", *The Atlantic Journal*.
- BOUDON Raymond (1967), *L'analyse mathématique des faits sociaux*, Paris, Plon.
- \_\_\_\_\_ (1984), *La place du désordre*, Paris, Presses Universitaires de France.
- BOURDIEU Pierre (1973), "L'opinion publique n'existe pas", *Les Temps modernes*, 318, 1292-1309.
- \_\_\_\_\_ (1984), *Homo Academicus*, Paris, éd. De Minuit.
- BLUMER Herbert (1948), "Public opinion et public opinion polling", *American Sociological Review*, vol. 13, 542-554.
- CHENU Alain (1990), *L'archipel des employés*, Paris, éd. de l'Insee.
- \_\_\_\_\_ (1997), "La descriptibilité statistique des professions", *Sociétés contemporaines*, n°26, 109-136.

- CICOUREL Aaron, KITSUSE John (1963), "A Note on the utilization of statistics", *Social Problems*, vol.11, 131-139.
- CICOUREL Aaron (1968), *The Social Organization of Juvenile Justice*, New York, John Wiley & Sons.
- COMTE Auguste (1830), *Cours de philosophie positive*, Paris, Alcan.
- CONVERSE Philip (1987), "Changing conceptions of public opinion in the political process", *Public Opinion Quarterly*, vol. 51 (2), 12-24.
- DESROSIERE Alain (2000), *La politique des grands nombres : histoire de la raison statistique*, Paris, La Découverte.
- \_\_\_\_\_ (2008), *Gouverner par les nombres*, Paris, Presse de l'Ecole des Mines.
- DILTHEY Wilhelm (1910), *Der Aufbau der geschichtlichen Welt in den Geisteswissenschaften*, Berlin.
- DOUGLAS Jack (1967), *The Social Meaning of Suicide*, Princeton University Press.
- DURKHEIM Emile (1895), *Les règles de la méthode sociologique*, Paris, Presses Universitaires de France.
- GIDDINGS Franklin (1911), *The Principle of Sociology*, MacMillan.
- HALBWACHS Maurice (1913), *La classe ouvrière et les niveaux de vie*, Paris, Alcan.
- \_\_\_\_\_ (1930), *Les causes du suicide*, Paris, Alcan.
- \_\_\_\_\_ (1935), "La statistique en sociologie" in *La statistique, ses applications, les problèmes qu'elle soulève*, Paris, Presses Universitaires de France, 7<sup>ème</sup> semaine de synthèse, 113-134.
- HERAN François (1984), "L'assise statistique de la sociologie", *Economie et Statistique*, vol. 168, 23-35.
- \_\_\_\_\_ (2014), "Cinq idées reçues sur l'immigration", *Population et Société*, n°397.
- JENCKS Christopher (1972), *Inequality: a Reassessment of the Effect of Family and Schooling in America*, Basic Books.
- LATOUR Bruno, WOOLGAR Steve (1979), *Laboratory Life: the Social Construction of Scientific Facts*, London, Sage Publications.
- LAZARSFELD Paul (1954), *Mathematical Thinking in the Social Sciences*, Glencoe, The Free Press.
- LAZARSFELD Paul, KATZ Elihu (1955), *Personal Influence*, Glencoe, The Free Press.

- LAZARSFELD Paul (1961), "Notes on the History of Quantification in Sociology, Trends, Sources and Problems", *Isis* vol. 52, 277-333.
- LEVI-STRAUSS Claude (1949), *Les structures élémentaires de la parenté*, Paris, Mouton & Co.
- \_\_\_\_\_ (1954), "Les mathématiques de l'homme", *Bulletin International des Sciences Sociales*, vol. 6 (4), 643-653.
- LYND Robert (1939), *Knowledge for what ? The Place of the Social Science in the American Culture*, Princeton University Press.
- MARTIN Olivier (2002), "Mathématiques et sciences sociales au XX<sup>e</sup> siècle", *Revue d'Histoire des Sciences Humaines*, vol. 1 (6), 3-13.
- MULLIGAN Casey (2003), "The Empirical Frequency of Pivotal Vote", *Public Choice*, vol.116, 31-54.
- NOELLE-NEUMANN Elisabeth (1993), *The Spiral of Silence*, The University of Chicago Press.
- OTIS Duncan (1984), *Notes on Social Measurement, Historical and Critical*, Hardcover.
- PASSERON Jean-Claude (1991), *Le raisonnement sociologique. L'espace non-poppérien du raisonnement naturel*, Paris, Nathan.
- PENEFF Jean (1984), "Le travail statistique ou le métier du père", *Sociologie du Travail*, n°2, 195-211.
- PESCOSOLIDO Bernice, MENDELSON Robert (1986), "Social Causation or Social Construction? An Investigation into the Social Organization of Suicide Rates", *American Sociological Review*, vol. 51 (1), 80-101.
- QUETELET Adolphe (1831), *Recherches sur le penchant au crime aux différents âges*, Bruxelles.
- \_\_\_\_\_ (1835), *Sur l'homme et le développement de ses facultés, Essai de physique sociale*, Bruxelles.
- SHILLER Robert (2000), *Irrational Exuberance*, Princeton University Press.
- SIMIAND François (1908), "La Méthode positive en science économique", *Revue de Métaphysique et de Morale* (reproduit dans *Méthode historique et sciences sociales*, 1912).
- SIMON Patrick (2008), "Les statistiques, les sciences sociales françaises et les rapports sociaux ethniques et de « race »", *Revue Française de Sociologie*, vol. 49, 153-162.
- TÖNNIES Ferdinand (1929), "Statistik und Soziographie", *Allgemeines Statistisches Archiv*.