



**HAL**  
open science

## Nouveaux arguments en faveur des enquêtes qualitatives

Nicolas Bouleau

► **To cite this version:**

| Nicolas Bouleau. Nouveaux arguments en faveur des enquêtes qualitatives. 2013. halshs-00846657v2

**HAL Id: halshs-00846657**

**<https://shs.hal.science/halshs-00846657v2>**

Preprint submitted on 30 Apr 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Nouveaux arguments en faveur des enquêtes qualitatives

Nicolas Bouleau  
Cired, été 3013

**Résumé.** Nous abordons à nouveau la question déjà très travaillée des vertus respectives de l'enquête quantitative et de l'enquête qualitative, par une approche originale qui consiste à raisonner sur un modèle conceptuel simplifié afin de dégager des interrogations philosophiques sur les cas réels. Cette conceptualisation est décrite au détail et débouche sur une vision tranchée des deux démarches quantitative et qualitative. Cela permet de montrer une véritable contradiction des images fournies par les deux approches dans certaines situations qui sont illustrées par quelques exemples, dont celui du risque sanitaire devant la présence de pesticides dans l'alimentation.

Il est un peu surprenant qu'en interrogeant deux mille électeurs, on puisse savoir qui de Nicolas Sarkozy ou de François Hollande est finalement élu. A quoi a réellement servi le vote de tous ceux dont les instituts de sondage n'ont pas recueilli l'avis ? Il y a là un côté magique qui peut étonner. En l'occurrence on n'a besoin ici que de *la moyenne générale* et les lois du hasard et des grands nombres fournissent bien ce genre d'information, avec une précision dont cependant la signification est subtile et relève de l'inférence statistique. En revanche sur ce que pensent des groupes spécifiés de la population de tels sondages sont insuffisants et pour connaître l'opinion, ou, plus ambitieux encore, le mode de vie et les raisons du comportement de telle fraction des habitants, des enquêtes d'un type différent doivent être menées.

Les enquêtes qualitatives sont l'outil privilégié des sociologues. Elles font l'hypothèse qu'une série de questions-réponses à quelques individus, même en petit nombre, sur un thème de préoccupation qui les concerne peut donner des compréhensions nouvelles, apporter de la lumière à la problématique même que l'enquête est sensée servir.

Il y a donc deux manières de faire et le débat entre enquêtes quantitatives et enquêtes qualitatives a déjà une longue histoire marquée notamment par les avancées du département de sociologie de l'université de Chicago où naissent les concepts d'*observation participante* et d'*histoires de vie* et le renouveau des études quantitatives après 1935 aux Etats-Unis. Ce débat a même depuis 1967 sa propre revue, signe que cette dualité est maintenant en elle-même légitime.<sup>1</sup> Mais des opposants farouches s'affrontent encore comme il fut toujours de tradition dans les lettres, les arts et les sciences depuis la querelle des Anciens et des Modernes jusqu'à celle des logicistes et des intuitionnistes, etc. La critique majeure faite aux enquêtes qualitatives porte sur la généralisabilité des savoirs recueillis. Si l'on accompagne durant leurs déplacements quotidiens successivement une vingtaine de banlieusards pour les questionner sur leurs choix et leurs contraintes, on peut tomber sur vingt personnes atypiques ou en situation tout à fait transitoire dans leur propre scénario de vie ou de santé. Et *a contrario* le sociologue se défie des enquêtes quantitatives par le simple fait qu'elles sont nécessairement fondées sur des catégories conceptuelles *a priori*, des grilles, dont la signification n'est pas réellement remise en cause, alors que précisément c'est le contour social de ce qui fait l'émergence des significations qui intéresse le sociologue.

Comme souvent, au delà des chapelles, la tendance *main stream* prône maintenant une paix pluraliste fondée sur ce que l'on appelle *les méthodes mixtes*, où l'on récuse l'idée qu'une quelconque des deux démarches pourrait prétendre être la seule scientifiquement pertinente.

Le débat est toujours finalement pensé de la façon suivante : la réduction quantitative produit seule les garanties de précision et de vérifiabilité, mais elle trahit la réalité en lui substituant des schémas, opposition entre une épistémologie fondée sur *des modèles* ou des lois dans le sillage positiviste (Comte, Durkheim) et une recherche de compréhension, par des récits et des

---

<sup>1</sup> Cf. *Quality & Quantity*, Springer.

interprétations davantage nourrie de la philosophie allemande (Max Weber, Husserl) et de l'herméneutique (H.-G. Gadamer).

Le but de notre essai est d'aller plus loin dans l'étude des propriétés cognitives de ces deux approches et de les différencier plus nettement encore afin d'être à même de fournir des exemples explicites concrets où elles donnent des résultats *en contradiction* l'une et l'autre. Ceci fera apparaître des faiblesses des méthodes quantitatives insuffisamment relevées jusqu'à présent.

## 1. Démarche philosophique adoptée

Pour aborder une telle problématique, il serait périlleux — voir naïf — de se placer dans une vision épistémique prédéfinie où l'on prétendrait avancer avec la science derrière soi, en avant-garde de la grande armée des connaissances vérifiées. Notre démarche est au contraire purement analogique. Nous proposons un modèle-réduit du débat, une maquette comme Brunelleschi en fit pour le dôme de Florence, dans le domaine des idées et non en bois, mais porteur néanmoins d'une phénoménologie dans laquelle on peut trouver suffisamment de similitude avec notre débat pour pouvoir raisonner.

Notre modèle est mathématique. Alors tout de suite je vois les sociologues hausser les épaules "encore des propos à côté du problème !". N'allons pas trop vite, je demande au lecteur quelque attention. Les mathématiques ne sont pas seulement le lieu du quantitatif, erreur que font souvent les gens de lettres, elles sont aujourd'hui largement une pensée — particulière certes — qui concerne le qualitatif. Tout l'*analysis situs* (Gauss, etc.), la géométrie différentielle (Riemann, Levi-Civita, Betti, etc.) et surtout le formidable développement au 20<sup>ème</sup> siècle des espaces fonctionnels (espaces de Hilbert, espaces de Banach, espaces vectoriels topologiques, etc.) qui furent et restent d'une fécondité remarquable sont essentiellement le maniement *de concepts* où l'on est fort loin du quantitatif, des dénombrements, et des mesures avec des appareillages gradués.

De plus le modèle à l'aide duquel je prétends raisonner n'a pas vocation de tirer de la réalité des quantités mais au contraire, dans l'autre sens, de suggérer des questions méthodologiques par l'apparition de limitations déjà dans une situation simplifiée.

Philosophiquement je me référerais volontiers, comme exemple d'une telle démarche, aux travaux de W. V. O. Quine lorsque celui-ci part de propriétés obtenues en logique mathématique sur les systèmes formels axiomatisés pour en déduire des limitations des langages ordinaires (avec ses fameuses thèses de l'indétermination de la traduction, de la relativité de l'ontologie, et de la sous-détermination des théories par l'expérience) selon le raisonnement assez convaincant que la situation des langages réels est certes plus complexe, mais qu'on voit mal justement pourquoi ils seraient épargnés par les limitations qui existent déjà dans les langages formels.

Remarquons que cet usage des mathématiques à fins philosophiques est un peu passé de mode par la sophistication des techniques sur lesquelles travaillent les chercheurs en pure monodisciplinarité aujourd'hui. Mais qu'elle était courante dans le passé. Comme en attestent les discussions sur l'infini actuel ou potentiel, et les visions des mathématiques qui s'opposaient et s'opposent encore (platonisme, formalisme, intuitionnisme, esthétisme, ancillarité, faillibilisme, etc.) et qui nourrissent les œuvres de philosophes tels que Jean Cavaillès et Albert Lautman.

Nous allons décrire le modèle sur lequel nous voulons nous appuyer. Et pour aider le lecteur non mathématicien à se familiariser avec ses propriétés nous allons dans le paragraphe prochain le manier de façon purement didactique.

## 2. L'outil conceptuel

Nous allons raisonner dans un espace de Hilbert, ce sera notre modèle. Cet objet est tellement fondamental en mathématiques aujourd'hui, si fécond de tant d'applications, si simple conceptuellement qu'il doit être accueilli comme catégorie fondamentale de la pensée au dessus de toute référence disciplinaire.

Il y a beaucoup d'espaces de Hilbert en mathématiques mais en même temps il n'y en a qu'un, ils sont tous le même, ce qui n'empêche pas d'avoir besoin de plusieurs, comme l'enfant a besoin de plusieurs cubes pour jouer.

Au préalable il faut parler de l'espace euclidien à  $n$  dimensions. Là, rien de fondamentalement différent par rapport à notre espace métrique usuel à trois dimensions. Entre deux points il y a une distance qui s'exprime par la formule habituelle avec  $n$  coordonnées au lieu de trois. Si  $u$  et  $v$  sont des vecteurs (de l'origine vers un point) on peut définir le produit scalaire  $\langle u, v \rangle$  comme la somme des produits deux à deux des coordonnées de  $u$  et de  $v$ . C'est une opération qui fournit un nombre réel et qui est linéaire par rapport à  $u$  et par rapport à  $v$ .

Dès le milieu du 19<sup>ème</sup> siècle les mathématiciens, et en particulier Joseph Liouville, s'étaient rendu compte qu'on pouvait former des produits scalaires dans d'autres contextes que l'espace euclidien  $n$ -dimensionnel. En particulier si  $f$  et  $g$  sont deux fonctions réelles définies sur l'intervalle  $[0,1]$ , l'expression  $\int_0^1 f(x)g(x)dx$  est de cette même sorte. Il faut alors penser la longueur d'un vecteur  $f$  comme donnée par l'expression  $\sqrt{\int_0^1 f^2(x)dx}$ , et écrire l'orthogonalité de  $f$  et  $g$  par la condition  $\int_0^1 f(x)g(x)dx = 0$ .

La notion d'espace de Hilbert axiomatise un espace de dimension infinie doté d'un produit scalaire aux propriétés habituelles et donc d'une distance, avec en outre une hypothèse de complétude — c'est-à-dire que toute suite de Cauchy est convergente — fondamentale en mathématique mais dont nous n'aurons pas à nous servir dans la suite.

Ainsi l'espace de Hilbert est la généralisation la plus naturelle de l'espace usuel (l'espace kantien) au cas de la dimension infinie. L'ensemble des suites de nombres réels  $(a_n)$  telles que  $\sum_n a_n^2 < \infty$  est un espace de Hilbert. L'espace des fonctions réelles sur  $[0,1]$  telles que  $\int_0^1 f^2(x)dx < \infty$  est un espace de Hilbert noté  $L^2[0,1]$ . Les fonctions  $(1, \sqrt{2}\sin 2\pi kx, \sqrt{2}\cos 2\pi kx, k \geq 1)$  en sont une base orthonormée.

Les polynômes trigonométriques

$$P(x) = a_0 + \sum_{n=1}^N a_n \cos 2\pi nx + b_n \sin 2\pi nx$$

sont des fonctions avec un nombre fini de coordonnées non nulles sur cette base. Toute fonction sur  $[0,1]$  telle que  $\int_0^1 f^2(x)dx < \infty$  se représente comme une série :

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos 2\pi nx + b_n \sin 2\pi nx$$

vérifiant  $\sum_{n=1}^{\infty} a_n^2 + b_n^2 < \infty$ .

Notre espace de Hilbert  $L^2[0,1]$  muni de son produit scalaire et de sa distance ressemble donc comme un frère à l'espace euclidien usuel, la notion d'orthogonalité, celle de projection orthogonale y ont un sens tout à fait similaire.

Qu'il soit de dimension infinie fait apparaître, cependant, un phénomène nouveau : il y a sur cet espace plusieurs topologies non équivalentes. Cela veut dire qu'on peut donner plusieurs sens différents à la convergence d'une suite de points, en particulier la convergence faible et la convergence forte.

La convergence forte est la plus simple à comprendre, elle signifie que *la distance* des points de la suite avec le point limite tend vers zéro.

La convergence faible s'énonce ainsi : la suite  $f_n$  converge faiblement vers  $f$  si pour toute fonction  $g \in L^2[0,1]$  le produit scalaire  $\langle f_n, g \rangle = \int_0^1 f_n(x)g(x)dx$  converge vers  $\langle f, g \rangle = \int_0^1 f(x)g(x)dx$ .

Il y a des suites qui convergent faiblement et qui ne convergent pas fortement. D'où ce vocabulaire. Par exemple les vecteurs de la base orthogonale  $\sqrt{2}\sin 2\pi nx$  convergent faiblement vers zéro, mais leur distance au vecteur nul reste toujours égale à 1.

Donc notre "modèle" consiste à considérer que les humains sont des points de l'espace de Hilbert  $L^2[0,1]$ , c'est-à-dire des fonctions réelles sur l'intervalle  $[0,1]$  vérifiant  $\int_0^1 f^2(x)dx < \infty$ .

Philosophiquement ce n'est pas si stupide qu'il pourrait sembler *a priori*. Considérer qu'il faut une suite infinie de grandeurs pour "décrire" un individu, est compatible avec une vision

humaniste où, de fait, il n'est possible de connaître qu'un nombre fini de caractéristiques et où donc chaque personne garde en réserve, toujours, une infinité de propriétés inconnues.

Je trouve au contraire plaisant qu'on puisse faire une place dans la "rationalité formelle" à l'idée que les individus soient infini-dimensionnels, si on garde à l'esprit que le modèle est volontairement extrêmement simple.

### 3. Comment modéliser une question et sa réponse ?

Poser une question et obtenir une réponse fournit un renseignement — j'évite le plus possible le terme d'information trop polysémique<sup>2</sup> — sur la personne interrogée qui dépend aussi de l'interrogateur. Nous représenterons une question par une fonction et le renseignement obtenu par la réponse de l'individu  $f$  à la question  $g$  par le produit scalaire  $\langle f, g \rangle = \int_0^1 f(x)g(x)dx$ .

La fonction  $g$  peut être très simple, une des fonctions de base ( $1, \sqrt{2}\sin 2\pi kx, \sqrt{2}\cos 2\pi kx, k \geq 1$ ) ou plus compliquée.

Les vecteurs  $g$  avec lesquels on compose peuvent être ceux que le questionnant possède *a priori* dans sa besace. Mais ils peuvent aussi bien être engendrés par un processus réactif des réponses aux questions, le questionnant disposant de l'appareillage, ou du talent, lui permettant d'imaginer une nouvelle question à partir des réponses aux précédentes. On peut aussi supposer que la proposition d'une nouvelle question fait intervenir le contexte, l'extérieur, l'historicité, etc.

Nous voyons immédiatement qu'avec beaucoup de questions, on ne connaît pas forcément complètement  $f$  mais seulement *la projection* de  $f$  sur l'espace engendré par les questions. Cela est compatible avec les soucis des sociologues qui savent bien qu'en réalité la question peut être mal comprise et qu'elle peut correspondre dans l'esprit de la personne interrogée à un champ métaphorique très vaste dont on n'obtient qu'un indicateur.

Avec le produit scalaire, la réponse est la "projection" au sens mathématique sur la question. Cela rend bien compte également de la remarque faite dans le sillage marxiste ou pragmatiste que l'indication verbale de la réponse ne renseigne pas sur les capacités d'agir du sujet. Ici le rôle que peut jouer la fonction  $f$  inconnue au sein d'une structure reste hors de portée.

Cette représentation de la question-réponse est également conforme à ce qu'attendent les quantitativistes : si  $(f_n)$  sont les personnes interrogées sur la question  $g$ , le résultat du sondage s'exprimera par

$$\frac{1}{N} \sum_{n=1}^N \langle f_n, g \rangle.$$

Est-ce que cette conceptualisation est susceptible d'apporter quelque lumière à la discussion quantitatif-qualitatif ? Elle a la vertu déjà de bien séparer les approches.

Dans l'approche quantitative on s'intéresse à un nombre fini, restreint, de caractères et on interroge beaucoup de personnes, pour faire des moyennes, voire des traitements statistiques plus élaborées comme des régressions, des analyses en composantes principales, etc.

Dans la démarche qualitative on s'intéresse à un groupe restreint d'individus et c'est le questionnement qui est poursuivi, approfondi, parce qu'on suspecte une complexité intime qu'on ignore et en relation avec un contexte complexe lui aussi. C'est typiquement le souci de l'anthropologue tel que Claude Lévi-Strauss le situe : "Comme l'a remarqué von Neumann (von Neumann et Morgenstern, 1944) *Il est infiniment plus simple d'élaborer la théorie presque exacte d'un gaz contenant environ  $10^{25}$  particules libres que celle du système solaire qui comprend seulement 9 grands corps*. Or, l'anthropologue en quête de modèles se trouve devant un cas intermédiaire : les objets dont nous nous occupons — rôles sociaux et individus intégrés dans une société déterminée — sont beaucoup plus nombreux que ceux de la mécanique newtonienne, tout en ne l'étant pas assez pour relever de la statistique et du calcul des probabilités. Nous sommes donc placés sur un terrain hybride et équivoque"<sup>3</sup>.

<sup>2</sup> Cf. N. Bouleau "Dommages et intérêt de la spéculation" halshs-00823520.

<sup>3</sup> *Anthropologie structurale* Plon 1958, pp349-50.

Dès lors que la grille de questions s'applique à un grand nombre de personnes interrogées, elle traduit une quête de connaissance *inconditionnelle* alors que le sociologue s'intéresse, toujours, aux conditions de production de l'information. Ainsi dans le cas d'un tableau statistique avec comme entrée d'un côté des tranches d'âge, et de l'autre des propriétés étudiées, le statisticien perfectionne son enquête en augmentant la taille et la représentativité de son échantillon, alors que le sociologue souhaitera poursuivre avec d'autres questions pour savoir si la variable "âge" exprime par exemple *un effet de vieillissement* ou *un effet de génération*<sup>4</sup>.

Notre modèle permet de bien séparer les approches, mais comme nous allons le voir, il permet aussi de faire apparaître des phénomènes nouveaux. C'est-à-dire des propriétés bien visibles sur le modèle qui sollicitent en retour un questionnement du monde réel.

Mais auparavant nous devons clarifier certaines questions relatives au fini et à l'infini, au "small  $n$ " et "large  $N$ " comme on dit dans la littérature de langue anglaise.

#### 4. Fini vs infini

Nous allons raisonner comme si la population totale ('les humains, ou "la France", ou 'les entreprises de moins de dix salariés', etc.) était *finie mais très grande*.

Pourquoi ne pas la penser infinie ? Au point où nous en sommes ! Pour la raison technique mais incontournable que sur une population infinie, il n'y a pas moyen de tirer un échantillon fini "au hasard". Il n'y a pas de loi de probabilité uniforme sur l'ensemble des entiers. Or nous voulons pouvoir faire des échantillons comme en font les instituts de sondage.

Donc la population est un ensemble fini très grand de fonctions  $f_i \in L^2[0,1]$ .

Une enquête quantitative consiste à considérer un groupe de questions  $q_1, \dots, q_K$  et d'étudier les réponses

$$(\langle f_1, q_1 \rangle, \dots, \langle f_1, q_K \rangle), \dots, (\langle f_m, q_1 \rangle, \dots, \langle f_m, q_K \rangle), \dots, (\langle f_n, q_1 \rangle, \dots, \langle f_n, q_K \rangle)$$

pour un échantillon

$$(f_1, \dots, f_m, \dots, f_n)$$

de la population, en le choisissant et en choisissant  $n$  assez grand pour obtenir une précision souhaitée lorsqu'on fait la moyenne

$$\frac{1}{n} \sum_{m=1}^n \langle f_m, q_1 \rangle$$

et de même pour les autres questions, ou d'autres opérations statistiques.

Une enquête qualitative consiste à considérer un groupe  $(f_1, \dots, f_n)$  fixé, des réponses aux questions  $(q_1, \dots, q_K)$  soit

$$(\langle f_1, q_1 \rangle, \dots, \langle f_1, q_K \rangle), \dots, (\langle f_m, q_1 \rangle, \dots, \langle f_m, q_K \rangle), \dots, (\langle f_n, q_1 \rangle, \dots, \langle f_n, q_K \rangle)$$

où cette fois-ci c'est le nombre  $K$  de questions qui augmente,  $n$  restant fixé.

#### 5. Convergence faible et convergence forte

Nous pouvons aisément nous convaincre que ce dont il s'agit lorsque le nombre de questions augmente est une convergence forte.

Pour chaque personne interrogée fixée, lorsque l'on questionne davantage, la collection des réponses que l'on obtient converge, au sens fort. Cela ne converge pas forcément vers "la personne interrogée" mais cela converge, au sens fort, vers la projection de la personne interrogée sur "l'espace des questions".

Prenons un exemple explicite. Si une personne  $f \in L^2[0,1]$ , est interrogée avec les questions  $q_0=1, q_1=\sqrt{2}\cos 2\pi x, \dots, q_K=\sqrt{2}\cos 2\pi Kx$ , la suite des vecteurs

$$\langle f, q_0 \rangle q_0 + \langle f, q_1 \rangle q_1 + \dots + \langle f, q_K \rangle q_K$$

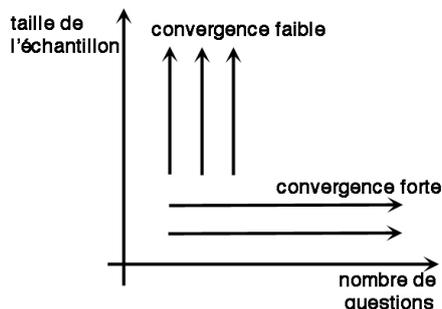
converge fortement non pas vers  $f$ , mais vers la projection orthogonale de  $f$  sur l'espace des fonctions dont le graphe est symétrique par rapport à la droite d'abscisse  $\frac{1}{2}$ , espace engendré par les fonctions  $\cos 2\pi kx$ . Cette fonction limite a quelque chose de  $f$  mais n'est pas  $f$ . Elle est l'image qu'on a de  $f$  à travers l'enquête qualitative ainsi conduite.

<sup>4</sup> Cf. Cl. Passeron, *Le raisonnement sociologique*, Nathan, 1991.

Au contraire lorsqu'on fait des enquêtes quantitatives, ce dont il s'agit est une convergence faible.

Les sondages d'opinion s'attachent en premier lieu à choisir l'échantillon et sa taille, le groupe de question étant fixé, pour obtenir la précision souhaitée. On a ainsi une image de la population totale ou plutôt des propriétés moyennes de la population totale sur ce groupe de questions fixées. On peut ensuite procéder à d'autres sondages sur d'autres questions.

Nous pouvons schématiser la situation ainsi:



Notons en passant, pour le lecteur spécialiste, que les sondages quantitatifs sont faits généralement sur une seule question ou un petit nombre de caractéristiques, à cause du fait qu'une bonne précision lorsque la dimension de l'espace (le nombre de questions) croît est de plus en plus difficile à obtenir. C'est le phénomène connu dans la littérature sous le nom de "malédiction de la dimension" ou "curse of dimensionality"<sup>5</sup>.

Revenons au problème de fond de la valeur des deux types d'enquête ainsi représentés schématiquement par ces deux types de convergence. Si nous fixons les personnes interrogées et augmentons le nombre de questions, nous voyons bien que nous obtenons toujours plus de renseignements sur ces personnes (ou du moins sur ce qu'elles disent) mais nous n'en déduisons des propriétés de la populations que si on peut tabler sur une certaine ressemblance de notre groupe avec la population générale dans le registre d'idées qui nous intéresse. Par exemple en tentant, à partir des réponses, de dégager des "faits sociaux" à la manière d'Emile Durkheim ou en attachant de l'importance à ce qui peut relever de phénomènes d'imitation comme préconise Gabriel Tarde<sup>6</sup>.

Quant à l'enquête quantitative, le problème intéressant n'est pas tant de savoir si sur une question on a choisi correctement l'échantillon et sa taille, à ce sujet les manuels donnent toutes les explications que l'on peut souhaiter, mais bien de savoir si en multipliant les enquêtes bien calibrées sur des questions différentes on aboutit à une "vision" satisfaisante de la population générale.

Déjà les statistiques en dimension finies nous recommandent une extrême prudence à cet égard, car on sait — contrairement à ce que pensait Adolphe Quételet — que l'idée d'homme "moyen" est contradictoire, dès lors que les qualités mesurées sont liées par des relations non linéaires, comme la hauteur, le volume etc.

Mais, au delà de ces questions de similitude dimensionnelle, il existe *un véritable phénomène d'effacement* dans la procédure de l'enquête quantitative qui est bien mis en lumière par notre modèle.

<sup>5</sup> Cf. le bel article de U. Gather et C. Becker "The curse of Dimensionality — A Challenge for Mathematical Statistics" *Jber. D. Dt. Math.-Verein.*, 103 (2001) 19-36.

<sup>6</sup> "Mais, qu'il s'agisse de contrats, de services ou de contraintes, il s'agit toujours de faits d'imitation. Que l'homme parle, prie, combatte, travaille, sculpte, peigne, versifie, il ne fait rien que tirer des exemplaires nouveaux de signes verbaux, de rites, de coups d'épée ou de fusil, de procédés industriels ou artistiques, de formes poétiques, de modèles en un mot, objets de son imitation spontanée ou obligatoire, consciente ou inconsciente, volontaire ou involontaire, intelligente ou moutonnaire, sympathique ou haineuse, admirative ou envieuse, mais de son imitation toujours" G. Tarde *Logique sociale* 1893.

## 6. Gommage des particularités individuelles par les enquêtes quantitatives

Nous allons voir que les propriétés (bien connues en mathématiques<sup>7</sup>) qui différencient la convergence faible et la convergence forte se traduisent par des contradictions entre les images fournies par les enquêtes quantitatives et qualitatives. Nous prendrons simplement des exemples.

a) **Lecture de romans.** Nous procédons à un sondage quantitatif en France sur la question  $q_1$  suivante "Avez-vous lu complètement un roman de Victor Hugo ?" et nous supposons — ce qui est une hypothèse plausible — qu'au seuil de vraisemblance utilisé la proportion de réponses positives dans la population générale est zéro.

Nous recommençons cette fois avec Alexandre Dumas père, question  $q_2$ , nous obtenons encore zéro. Puis Marcel Proust, Anatole France, Romain Gary, William Faulkner, Ernest Hemingway, etc.,

A chaque enquête, sur chaque auteur particulier supposons que nous obtenions zéro. Et, en poursuivant les sondages sur des romanciers de moins en moins connus nous obtenons toujours zéro.

Il serait évidemment absurde d'en déduire que les Français n'ont lu aucun roman. Car si au contraire nous procédons à une enquête qualitative, sur chaque personne interrogée nous pouvons obtenir qu'elle a lu un roman, au moins un.

b) **Facebook.** Même scénario sur la question "Avez-vous François Hollande parmi ceux qui se sont déclarés vos 'amis' ?". Puis en descendant des célébrités vers des gens moins connus nous obtenons une suite en contradiction apparente avec le fait que chaque membre de Facebook a tout de même quelques 'amis'.<sup>8</sup>

c) **Toxicité des produits chimiques dans l'alimentation.** Ce phénomène est plus préoccupant si l'on pense que notre société et nos organisations internationales ont cru pouvoir s'appuyer sur une scientificité dure comme fer en instaurant la notion de "Dose Journalière Admissible" ou DJA.

Rappelons brièvement la situation. On dénombre environ 150000 produits chimiques pour lesquels on tente d'identifier un effet toxique éventuel et de définir la relation effet/dose : Quelle fréquence ? A quelle dose ? Quels risques ? Il s'agit de produits phytosanitaires, de médicaments ingérés par les animaux que l'on consomme et de produits chimiques dans les additifs de l'industrie agroalimentaire. Le problème concret est double. D'une part l'insuffisance des données toxicologiques, d'autre part l'impossibilité de tester 150000 produits sur les animaux.

Plusieurs études soulèvent également le problème moins classique *des faibles doses*, certaines molécules pouvant avoir une fonction combinatoire de déclenchement de processus endocriniens complexes visibles uniquement sur les rejets.

Notre conceptualisation ne vise pas à résoudre ce problème sanitaire dans son intégralité mais juste de faire comprendre qu'étant donné qu'il y a des milliers de produits chimiques artificiels dans notre alimentation il est parfaitement possible qu'à l'examen de chaque produit particulier les autorités les plus sérieuses obtiennent par des études quantitatives qu'il est inoffensif au seuil de concentration donné. Et que néanmoins on trouve des désordres de santé pour une forte proportion de la population dus à des sensibilités et des usages particuliers.

Non seulement il conviendrait de tester les combinaisons de produits chimiques absorbés pour étudier les effets conjoints, mais le principe même d'un seuil obtenu statistiquement pour chaque produit ou groupe de produits donne une idée fautive de la réalité qui se trouve tronquée par ce que nous avons appelé la convergence faible.

On obtient une connaissance moins réductrice et parfaitement scientifiquement légitime en étant à l'écoute des désordres existant sur chaque patient. N'oublions pas la vertigineuse variété des

---

<sup>7</sup> Notamment le phénomène de bosse glissante ou "gliding hump".

<sup>8</sup> Un phénomène de même nature est pointé par les chercheurs des sciences du territoire, car le lieu économique et culturel est une propriété complexe. Comme dit L. Davezies "Avec un échantillon de quelques milliers de Français, on apprend beaucoup de choses sur ce que pensent ou font 60 millions de concitoyens. Ce n'est pas vrai pour les territoires du fait de leur singularités." (*La crise qui vient*, Seuil 20013).

humains les uns vis à vis des autres par le système de la procréation par partage binaire au hasard du patrimoine génétique et l'immense variété des fréquentations de produits chimiques que chacun a rencontrée dans sa vie.

Cette démarche duale, qualitative, surtout poussée par les associations de malades ou de parents (des *groupes concernés* comme les appelle Michel Callon) est une approche fondamentale pour faire émerger des connaissances qui seraient complètement gommées par l'usage exclusif d'enquêtes quantitatives.

d) **Médicaments.** Le problème est similaire d'un point de vue épistémologique. Se surajoute la difficulté que dans le processus décisionnel actuel d'autorisation (avec ses phases I à IV) une grande part des études sont financées par l'industrie pharmaceutique, ce qui est propice aux dissimulations, voire aux ambiguïtés de comportements d'experts. Nous n'insistons pas davantage ici.

===

La représentation d'une population par des points d'un espace de dimension infinie, permet d'envisager facilement des situations où les individus sont très différents, tellement différents que les enquêtes quantitatives et les procédés statistiques donneront toujours une image très simplifiée de cette réalité complexe.

Pour donner encore une autre image qui parlera peut-être davantage à certains, ce n'est pas en faisant des statistiques sur les décimales de nombres réels tirés au hasard entre 0 et 1, que l'on découvrira qu'il y a des nombres irrationnels ni des nombres transcendants.

## Bibliographie

- Abbott A., *Methods of Discovery, Heuristics for the Social Sciences* Jeffrey C. Alexander 2004
- Adorno T., Popper K., *De Vienne à Frankfort, la querelle allemande des sciences sociales*, Ed. Complexe 1979
- Armatte M., "La notion de modèle dans les sciences sociales: anciennes et nouvelles significations" *Math & Sc. Hum.* 2005
- Bouleau N., *Risk and Meaning, Adversaries in Art, Science and Philosophy*, Springer 2011
- Bouleau N., "Can there be Excessive Mathematization of the World ?" in *Seminar on Stochastic Analysis, Random Fields and Applications: VII: Centro Stefano Franscini, Ascona, May 2011 R.C.*
- Bryman A., "Integrating quantitative and qualitative research: how is it done?" *Qualitative Research* vol. 6(1) 97–113.
- Dalang, M. Dozzi et F. Russo, Springer, Progress in Probability, 2013
- Charolles V., *Et si les chiffres ne disaient pas toute la vérité* Fayard 2008
- Desrosières A., "Est-il bon, est-il méchant ? Le rôle du nombre dans le gouvernement de la cité néolibérale" *Rev. delle Politiche Sociali* 2010
- Gadamer H.-G., *Esquisses herméneutiques*, Vrin 2004
- Moreau Y., *Catastrophes et mondes*, Thèse EHESS, 2013.
- Passeron Cl., *Le raisonnement sociologique*, Nathan, 1991
- Sale J. E. M., Lohfeld L. H., Brazil K., "Revisiting the Quantitative-Qualitative Debate: Implications for Mixed-Methods Research" *Quality & Quantity* **36**: 43–53, 2002.