

Note sur la méthode de l'enquête en population générale incluse dans le rapport de la CIASE

La Commission indépendante sur les abus sexuels dans l'Église catholique (CIASE) a remis son rapport le 5 octobre 2021. Entre autres enquêtes, elle a mené une enquête en population générale confiée aux chercheurs de l'INSERM mobilisés par la commission. Parmi ses résultats les plus spectaculaires figure une estimation du nombre d'adultes ayant subi en France depuis 1950 et avant leur majorité au moins une agression sexuelle perpétrée par une personne en relation avec l'Église catholique : clercs, religieux, religieuses, mais aussi laïcs. Le rapport avance le chiffre de 216 800 personnes abusées par des prêtres ou des religieux, porté à 320 000 si l'on tient compte des abus commis par les laïcs. Chiffres « accablants », révélateurs d'un phénomène « massif », selon les termes de la commission et qui n'ont pas échappé aux médias. Or l'impression produite par les chiffres relatifs, pourtant issus des mêmes données, est tout autre : seuls 0,4 % des 52 642 000 adultes recensés en France au 1^{er} janvier 2020 (Outremer compris) disent avoir été abusés par des prêtres ou des religieux avant leurs dix-huit ans depuis 1950, et 0,6 % si l'on y ajoute les abus des laïcs. Sous cet angle, le phénomène semble ultra-minoritaire.

Ce contraste troublant entre l'effet des chiffres absolus et l'effet des chiffres relatifs est classique en science sociale. On le retrouve par exemple en démographie, en statistique des migrations, en santé publique, en criminologie, en accidentologie. Il pose deux problèmes distincts, qu'on abordera tour à tour : un problème de communication et un problème statistique.

1. Chiffres absolus, chiffres relatifs : un problème classique de communication, bien traité par la CIASE et dans l'ensemble bien compris des médias

Pour les huit membres de l'Académie catholique qui ont publié fin novembre 2021 une critique acerbe des travaux de la CIASE, les choses sont simples : la commission aurait privilégié dans sa communication le nombre absolu de victimes, un nombre si accablant qu'il aurait eu pour effet – voire pour intention – d'étouffer toute velléité de contestation. Les chiffres absolus, accuse l'Académie, ont été « jetés en pâture aux médias et à l'opinion ». À l'en croire, le chiffre de 330 000 victimes aurait été « le seul chiffre retenu par les médias¹ ». Si les Huit ne remettent pas en cause la gravité des actes commis sur les enfants confiés à l'Église, ils dénoncent « l'évaluation disproportionnée de ce fléau », bien faite pour « mettre à bas l'Église-institution² ». Et plus loin : « La fonction, en tout cas l'effet, d'un tel chiffre [les 330 000 victimes] est de clore la discussion.³ »

Un examen attentif des conditions dans lesquelles les résultats ont été présentés aux médias et répercutés par leurs soins inflige un démenti radical à ce diagnostic. Mais, avant d'exposer notre analyse, on rappellera quelques principes de base.

¹ « Analyse du rapport de la Commission indépendante sur les abus sexuels dans l'Église (CIASE) » signée par huit membres de l'Académie catholique, 25 novembre 2021, pp. 2 et 3 (désigné désormais par « le rapport des Huit » ou « les Huit »).

² *Ibid.*, fin de la p. 3.

³ *Ibid.*, début de la p. 4.

Les experts en statistique sociale savent qu'une bonne communication sur l'ampleur d'un phénomène social doit présenter tout à la fois les données absolues (combien de personnes concernées ?), les données relatives (quelle part de la population ?) et les données comparatives (quelle ampleur par rapport à d'autres réalités comparables ?). Devant toute information statistique sur un phénomène jusque-là inconnu ou peu connu, la curiosité bien compréhensible du public et des journalistes est de savoir si c'est « beaucoup » ou « peu », si le phénomène progresse ou recule, s'il est plus important chez nous qu'ailleurs. De ces questions de fait on glisse aisément à une interrogation normative, voire morale : est-ce trop ? est-ce banal ? est-ce bien, est-ce mal ? Le phénomène évolue-t-il dans la bonne ou dans la mauvaise direction ? Etc.

La connaissance statistique des migrations soulève des interrogations analogues. 270 000 primo-arrivants légaux par an en France, c'est beaucoup (la taille d'une capitale de région, disent souvent les journalistes ou les politiques, usant d'un procédé réprouvé par les statisticiens professionnels, puisque la base de calcul varie d'un pays à l'autre), mais c'est « seulement » un accroissement annuel de population de 0,4 %. Il faut être bon connaisseur du domaine pour apprécier un ordre de grandeur à partir des seules proportions, comme le font les démographes ou les géographes de la population quand ils expriment les taux de natalité et de mortalité en « pour mille » : ils savent d'expérience que 46 ‰ est un taux de natalité très élevé (c'est celui du Niger), que 7 ‰ est un taux de natalité très faible (celui de l'Italie), que la France est plutôt bien placée en Europe, avec un taux de 11 ‰, etc. Ces proportions, en revanche, ne disent rien au profane. D'où la préférence des démographes (et des journalistes) pour un indicateur plus évocateur, le nombre moyen d'enfants par femme.

Pour sortir du faux dilemme entre chiffre absolu « massif » et chiffre relatif « ultra-minoritaire », la meilleure méthode est de comparer les phénomènes, et c'est précisément ce qu'a fait la CIASE dans la communication des données. *Le Monde*, par exemple, dans le premier article rendant compte du rapport de la CIASE, écrit fort justement : « Plus encore que la valeur absolue, tragique mais probablement sous-évaluée, c'est la comparaison qui est dévastatrice⁴. » Elle est sous-évaluée, en effet, parce seules les personnes survivantes ont pu évoquer les événements subis dans les années 1950-1970. La journaliste du *Monde* cite le chiffre des 216 000 personnes abusées par un prêtre ou un religieux, mais évoque aussitôt le fait que « la commission (...) a cherché à savoir où se situait l'Eglise catholique, dans cette évaluation, par rapport aux autres sphères de socialisation ». La conclusion de la commission est « formelle », poursuit-elle : « L'Eglise catholique est, hormis les cercles familiaux et amicaux, le milieu où la prévalence des violences sexuelles est la plus élevée », comparé à l'école, au sport, aux colonies de vacances.

Le compte rendu du *Figaro* donne lieu à la même observation. Le dossier contient une interview du président de la commission, qui met en avant le nombre des 216 000 victimes mais le replace aussitôt dans l'ensemble des victimes d'abus sexuels toutes institutions réunies, y compris la famille : 4 % de l'ensemble des abus. Après quoi le président compare les taux de prévalence selon les milieux de socialisation. Le dossier du *Figaro* consacre un second article à l'enquête de l'INSERM et restitue en détail la comparaison des taux d'incidence⁵.

⁴ Cécile Chambraud, « Abus sexuels dans l'Eglise catholique : l'état des lieux accablant de la commission Sauvé », *Le Monde*, 5 octobre 2021. La sous-évaluation tient au fait que l'enquête, par définition, enregistre les déclarations des adultes encore en vie lors de l'enquête.

⁵ Agnès Leclair, « Violences sexuelles : famille, Église, école... la part des abus sur les enfants dévoilée par l'INSERM », *Le Figaro*, 5 octobre 2021.

France-Inter a suivi la même stratégie de présentation, mais avec l'apport d'une infographie soignée⁶. Un premier graphique montre ce que représentent les 216 000 et 330 000 victimes de l'Église par rapport au nombre total de personnes abusées en population générale, tous contextes réunis. Ce mode de présentation est clairement à l'opposé d'un choix sensationnaliste. Le graphique suivant compare les taux de prévalence selon les milieux de socialisation ; il révèle que l'Église se place loin derrière le milieu familial mais nettement au-dessus des autres milieux. Au total, le site de France-Inter présente tour à tour les chiffres absolus, les proportions et les données comparatives.

Ces trois exemples – *Le Monde*, *Le Figaro* et France-Inter – montrent que la commission, contrairement à la formule vindicative employée par le rapport de l'Académie catholique, n'a aucunement décidé de « jeter en pâture » les chiffres absolus dans l'arène médiatique. Prétendre que le nombre absolu des victimes aurait été « le seul chiffre retenu par les médias » est contraire à la vérité. La commission a pris soin de compléter les chiffres relatifs par des chiffres proportionnels ; elle s'est employée à présenter une comparaison des taux de prévalence situant clairement l'Église par rapport aux autres milieux. La plupart des médias ont bien saisi le message.

D'autres médias, il est vrai, ont suivi une autre voie. Le quotidien *Libération*, par exemple, a titré sur le nombre de victimes, mais le corps de l'article n'en contient pas moins une comparaison détaillée des taux de prévalence⁷. De son côté, le quotidien catholique *La Croix* a mobilisé une dizaine de journalistes pour rendre compte du rapport de la commission et recueillir de nombreuses réactions de fidèles. L'article central du dossier est précédé d'un chapeau soulignant le caractère « massif » et « systémique » des abus commis dans l'Église, mais sans donner de chiffres⁸. Ceux-ci apparaissent dans le corps de l'article :

« 216 000 victimes estimées. Le chiffre est accablant, bien supérieur aux prévisions les plus sombres qui pouvaient être envisagées, et c'est d'abord celui-ci qui ressort des 485 pages du rapport (assorties de 2 500 pages d'annexes), fruit de trois ans de travaux. »

Ces données sont ensuite détaillées, et suivies de ce commentaire :

« Ces chiffres sont considérables – et ne concernent que les personnes vivantes –, mais ils sont aussi à resituer dans le contexte plus général des violences sexuelles en France. »

Au total, le quotidien catholique a pris au sérieux le rapport de la CIASE, dont il a détaillé les nombreuses recommandations des semaines durant. Il s'est clairement rangé du côté des défenseurs du rapport.

On ne poussera pas plus loin l'analyse de l'écho rencontré par le rapport de la CIASE dans les médias. Les exemples évoqués, qui couvrent un large spectre de la presse quotidienne, complété par le site de France-Inter, confirment que la CIASE n'a pas éludé le dilemme de communication inhérent aux phénomènes qui paraissent à la fois massifs en nombre absolu et faibles en proportion. Si la commission, à commencer par son président, a été légitimement heurtée par l'ampleur des abus sexuels dans l'Église, elle a pris soin de les situer dans un cadre comparatif que les médias ont correctement restitués. Il était inévitable qu'un rapport de

⁶ Juliette Geay, « Rapport Sauvé : cinq infographies pour comprendre l'ampleur des abus sexuels dans l'Église catholique », www.franceinter.fr/societe/rapport-sauve-cinq-infographies-pour-comprendre-l-ampleur-des-abus-sexuels-dans-l-eglise-catholique, 5 octobre 2021.

⁷ Bernadette Sauvaget, « Pédocriminalité dans l'Église : 216 000 mineurs victimes de prêtres et religieux catholiques », *Libération* 5 octobre 2021.

⁸ Céline Hoyeau et Christophe Henning, « Abus sexuels dans l'Église : ce qu'il faut retenir du rapport Sauvé », *La Croix*, 5 octobre 2021.

3000 pages subisse une perte en ligne dans la communication aux médias et au grand public. Mais cette perte s'est avérée très limitée au regard de la complexité du sujet. La commission n'a pas sacrifié les données de base au profit d'une quête du sensationnel, encore moins cherché à nourrir une hostilité de principe envers l'Église. De leur côté, les médias n'ont pas dénaturé les résultats présentés par la commission. Au total, la *communication proprement statistique*, c'est-à-dire la capacité à transmettre aux médias des résultats chiffrés sous une forme lisible et digeste, tout en gérant l'épreuve douloureuse de la révélation pour l'Église de France et ses fidèles, fut une pleine réussite. Il était difficile de faire mieux.

2. L'enquête INSERM-CIASE : une méthode débattue mais suffisamment contrôlée pour fournir des ordres de grandeur fiables

Indépendamment des problèmes de communication, reste à évaluer la qualité statistique des données chiffrées présentées par la CIASE. Le rapport critique des Huit n'hésite pas à jeter le discrédit sur la valeur scientifique de la méthode : « En raison de la faiblesse du chiffre de départ et des biais inévitables de l'enquête, il n'est pas possible d'extrapoler pour transposer à l'échelle de la population française adulte (47 millions de personnes)⁹ ».

Sur le plan technique, la question posée est de savoir dans quelle mesure on peut cerner les contours statistiques d'une minorité dont les effectifs absolus sont impressionnants mais qui représentent moins de 1 % de la population générale. Les enquêtes représentatives n'y parviennent qu'en s'appuyant sur des échantillons de grande taille, sans commune mesure avec ceux qui alimentent les sondages politiques, souvent proches du millier d'électeurs.

Pour estimer le nombre de victimes des abus sexuels dans l'Église, le rapport s'appuie sur une enquête de l'IFOP (Institut français d'opinion publique) menée auprès d'un échantillon de 28 010 personnes. Sur ce nombre, 171 répondants ont déclaré avoir subi avant leurs 18 ans un abus sexuel commis par une personne liée à l'Église catholique, un nombre ramené à 118 si l'on écarte les abus des laïcs. C'est en extrapolant ces données d'enquête que les chercheurs de l'INSERM mobilisés par la CIASE ont avancé respectivement pour la France entière les chiffres de 320 000 et 216 800 victimes. Les 118 victimes de prêtres ou religieux extrapolées à la France entière représentent 216 800 personnes sur une population de 52 642 000, soit 0,4 % des adultes vivant en France. Les 53 victimes de laïcs (171 moins 118) représentent 103 200 adultes sur 56 642 000, soit 0,2%. Les deux groupes réunis atteignent 0,6 % de la population adulte.

Le problème de l'extrapolation

Pour le profane, un effectif de cent à deux cents réponses peut paraître bien maigre pour étayer une extrapolation sur plusieurs centaines de milliers de réponses. Or une déduction de cette ampleur est légitime en statistique sociale pour peu que soient remplies certaines conditions :

1. La taille de l'échantillon doit être suffisante (en l'occurrence, elle dépasse 28 000 personnes) ;
2. L'échantillon doit être représentatif de la population-cible (un objectif visé par l'IFOP au moyen de la méthode des quotas et par des redressements *a posteriori* – un point sur lequel on reviendra en détail) ;

⁹ « Analyse du rapport de la CIASE », haut de la p. 3.

3. Les résultats produits doivent être assortis d'une marge d'erreur (consistant à signaler, par exemple, qu'il y a 95 % de chances qu'ils tombent dans un intervalle donné).

Ces points sont contestés par les Huit.

S'agissant du premier, leur commentaire est pour le moins étrange :

« Le chiffre de 118 personnes déclarant avoir été abusées par un prêtre donne un pourcentage de 0,42 % et celui de 53 pour les personnes se déclarant abusées par un laïc un pourcentage de 0,19 % soit des chiffres nettement en deçà de ceux qui permettent une interprétation statistique (le chiffre de valeur absolue retenu permettant une interprétation doit être supérieur ou égal à 15, ce qui écarte toutes les enquêtes relatives à l'Église catholique) et la marge d'erreur qui subsiste dans ce type de sondage (liée à la représentativité, au choix de l'échantillon, à la formulation et à la séquence des questions) est supérieure aux résultats obtenus. »

Ce passage ne fait sens pour aucun statisticien connaissant un tant soit peu la question. Les seuils numériques ainsi fixés à l'interprétation des données de sondage sont totalement inconnus de la littérature méthodologique. On aimerait savoir sur quelles références internationalement reconnues l'académie s'appuie.

À moins d'une coquille dans le texte, on ne voit pas d'où vient le seuil minimal de 15 unités, ni en quoi il invaliderait les résultats de l'enquête INSERM, puisque ces derniers dépassent largement ce seuil. Plus généralement, nombre d'enquêtes de la statistique publique ont des échantillons avoisinant 20 000 personnes (la cohorte Elfe¹⁰, les deux enquêtes Trajectoires et Origines, le volet de l'Enquête Santé européenne, etc.), qui permettent de calculer des prévalences inférieures à 1 % (10 pour mille en démographie) pour des sous-échantillons limités (tels que des tranches d'âges)¹¹.

Une mise en cause trop générale des « quotas sur panel »

Les Huit opposent à la représentativité de l'échantillon INSERM des objections plus sérieuses¹². L'enquête INSERM, relèvent-ils, a été réalisée sur Internet en sollicitant un panel de volontaires rompus au remplissage de questionnaires de toute sorte et, de surcroît, libres de déclarer un profil imaginaire. Elle souffrirait d'un « biais générationnel et culturel » engendré par la fracture numérique qui tient les personnes âgées et les personnes vulnérables à l'écart d'Internet. Autant d'« inconvénients indéniables » qui priveraient l'enquête de toute rigueur scientifique. Les Huit citent incidemment à l'appui de leurs remarques le Rapport d'information produit en 2009 par le Sénat sur la surveillance des sondages politiques par la Commission des

¹⁰ Enquête longitudinale française depuis l'enfance, portée par l'Ined, l'INSERM et l'Établissement français du sang (18 000 enfants suivis depuis leur naissance en 2011). L'enquête britannique équivalente, la Millenium Cohort, a la même taille.

¹¹ Pour mémoire, la taille minimale d'un échantillon représentatif pour obtenir des résultats significatifs est déterminée par la formule $n = t^2 \times p \times (1 - p) / m^2$, dans laquelle
t est le niveau de confiance (1,96 pour un niveau à 95 %)
p la proportion estimée de la population qui possède la caractéristique recherchée
m la marge d'erreur (généralement fixée à 5 %).

Si p était une valeur proche de 50 % (comme dans un sondage électoral visant à séparer significativement les oui des non dans un scrutin indécis), l'échantillon devrait comporter un minimum de 384 personnes pour produire un résultat significatif à 95 % avec une marge d'erreur fixée à 5 %. Si la marge d'erreur était limitée à 2 %, l'échantillon devrait être porté à 4144 personnes. L'enquête INSERM-CIASE, avec ses 28 000 répondants et malgré l'option pour la méthode des quotas sur panel, se situe très au-delà de ces valeurs.

¹² « Analyse du rapport de la CIASE », p. 2.

sondages¹³. Or ce rapport était co-signé par le sénateur Hugues Portelli, qui se trouve être le président actuel de l'Académie catholique et le premier signataire du rapport des Huit. Le lecteur est tenté d'en conclure que les Huit ont toute compétence pour discréditer l'équipe de l'INSERM, sans le moindre égard pour l'expérience acquise de longue date par cette dernière dans le domaine des enquêtes nationales sur les orientations sexuelles, les violences sexuelles, les infections à VIH et les discriminations.

Le biais d'Internet : une fracture numérique considérablement réduite en dix ans

Or, depuis le rapport du Sénat, la fracture numérique s'est considérablement réduite en France¹⁴. Selon l'enquête SRCV de l'INSEE (Statistiques sur les revenus et les conditions de vie), la part des ménages connectés à Internet est passée de 61 % à 86 % de 2009 à 2019, celle des possesseurs d'ordinateur ou de tablettes de 67 % à 83 %, tandis que la possession de téléphone portable atteint désormais 95 %.

Certes, comme le souligne l'INSEE, des inégalités persistent entre les générations et les milieux sociaux, et l'aisance vis-à-vis de l'outil reste inégale. Mais la situation des personnes âgées ne cesse de progresser : de 2009 à 2019, l'équipement des ménages en ordinateur selon l'âge de la personne de référence est passé de 50,4 % à 81,5 % pour les 60-74 ans et de 14,2 % à 46,8 % pour les 75 ans ou plus. Parmi les 10 % de ménages les plus modestes (sous le 1^{er} décile, comme disent les statisticiens), 75 % disposent désormais d'un accès à Internet.

Dans ces conditions, un coefficient de redressement qui ajuste la structure sociodémographique des répondants à celle de la population générale peut assurer aux ménages les plus modestes et les plus âgés une représentation tout à fait correcte, bien meilleure qu'à l'époque du rapport sénatorial. L'objection des Huit serait pertinente s'il existait une corrélation perceptible entre les facteurs actuels de connexion à Internet (niveau d'éducation, initiation par les enfants, programme de diffusion par les associations ou les municipalités, extension géographique des réseaux, etc.) et l'exposition au risque d'abus sexuel d'origine ecclésiastique subi il y a plusieurs décennies avant l'âge de 18 ans. On ne voit pas quel pourrait être le mécanisme d'une telle corrélation.

Du reste, le raisonnement tenu par les Huit sur le biais générationnel de l'enquête INSERM ne convainc pas. À supposer que le vivier initial des panélistes sous-représente encore fortement les personnes les plus âgées et les plus modestes (et, ajoutent les Huit, celles aux comportements les plus « traditionnels »), en quoi ce biais pourrait-il *accroître* l'ampleur des abus sexuels déclarés dans l'enquête ? On pourrait penser, tout au contraire, qu'il le *réduit* : ces segments de la population étaient davantage en contact avec les institutions de l'Église dans les années 1950 à 1970 et, partant, davantage exposés au risque d'abus sexuel.

Une enquête sur vivier d'abonnés dûment contrôlée en amont et en aval

Le rapport critique des Huit soutient qu'un échantillon par quotas formé à partir d'un panel de volontaires ne saurait être représentatif. Mais la question est bien plus complexe.

Comme le rappelle bien le rapport de la CIASE, les spécialistes des sondages distinguent deux grandes familles d'échantillons, aléatoires et non aléatoires¹⁵.

¹³ « Sondages et démocratie : pour une législation plus respectueuse de la sincérité du débat politique. Rapport d'information n° 54 » de MM. Hugues Portelli et Jean-Pierre Sueur, fait au nom de la commission des lois du Sénat, déposé le 20 octobre 2010. »

¹⁴ *Insee-Focus* n° 226, 23 février 2021.

¹⁵ On oppose, en anglais, les *probability samples* aux *nonprobability samples*.

Les **échantillons aléatoires (ou probabilistes)** sont issus d'un tirage au hasard dans la liste exhaustive des personnes composant la population-cible à représenter (comme la liste des abonnés sur téléphone fixe ou mobile fournie par l'ARCEP, le recensement de la population quand il était encore exhaustif ou le fichier des foyers fiscaux, y compris ceux qui ne paient pas l'impôt sur le revenu). Les probabilités de tirage peuvent être différenciées selon les catégories de personnes, à condition que ces probabilités soient connues. On peut ensuite les compenser *ex post* par des coefficients de redressement. En toute rigueur (mais on verra que ce point ne fait pas l'unanimité chez les spécialistes), seuls les tirages aléatoires permettent de calculer des intervalles de confiance sur les résultats observés.

Les **échantillons non aléatoires (non probabilistes)** sont de divers types : interview au petit bonheur (« micro-trottoir »), sélection *a priori* de personnes jugées intéressantes par un expert (recrutement des membres d'un *focus group*, par exemple), échantillon « boule de neige » (chaque répondant mettant sur la piste d'autres répondants), *crowdsourcing* (appel massif à la collaboration du public, utilisé par Statistics Canada pour son enquête sur les effets du Covid-19).

Mais la méthode non probabiliste la plus utilisée et de loin la plus éprouvée est la **méthode des quotas**, qui structure *a priori* l'échantillon selon quelques variables sociodémographiques (sexe, âge, région, taille d'agglomération, catégorie socioprofessionnelle, tranche de revenus, niveau de diplôme...) en vue de reproduire au mieux leur distribution en population générale telle qu'elle est connue par les données exhaustives ou aléatoires des instituts nationaux de statistique. La théorie des sondages peine à justifier en théorie l'efficacité constatée des enquêtes par quotas quand on compare leur capacité respective à estimer la valeur de la même donnée. On reviendra sur ce retard de la théorie sur l'empirie.

Depuis une vingtaine d'années, s'est développée, en amont de l'échantillon par quotas, une méthode non aléatoire consistant à solliciter en ligne un **panel d'abonnés** de grande dimension (plusieurs centaines de milliers de personnes en France, plusieurs millions aux États-Unis), dont on s'est assuré au préalable qu'ils sont disposés à répondre aux enquêtes en ligne et qui touchent en contrepartie une maigre indemnité. Ces viviers de répondants potentiels (*web panels* ou *access panels* dans le jargon des praticiens) sont recrutés et entretenus par des sociétés spécialisées auxquelles peuvent s'adresser les instituts de sondage moyennant finance. Certains viviers appartiennent à des instituts de sondage (*proprietary panels*), mais ces instituts ne s'en contentent pas car ils ont intérêt à élargir le vivier des répondants en nouant des accords de partage de viviers avec d'autres panels. C'est le cas de l'IFOP, qui apportait les précisions suivantes dans une profession de foi méthodologique publiée sur son site en 2018 :

« Le recrutement des panélistes de la principale base sollicitée par l'IFOP est assuré via une centaine de sites partenaires issus de secteurs d'activité très variés (voyages, locations, hôtellerie, beauté, nouvelles technologies, textile, sites d'information, etc.) et via des liens sponsorisés, des parrainages ou encore des échanges de bases de données¹⁶ ».

Si les abonnés d'un panel sont recrutés sur la base du volontariat, c'est en fonction de l'appartenance à certaines catégories de la population et par des voies aussi diverses que possible. Un institut de sondage qui souhaite lancer une nouvelle enquête puise à moindre coût dans ce vivier de répondants potentiels, mais il le fait selon certaines procédures, qui varient

¹⁶ Yves-Marie Cann [directeur d'études au département Opinion et stratégies d'entreprise de l'IFOP], « Sondages en ligne : une méthodologie éprouvée », https://www.ifop.com/wp-content/uploads/2018/03/236-1-document_file.pdf? Ce document mentionne des données des années 2010. Une nouvelle version est accessible en décembre 2021 sous la signature de Jérôme Fourquet.

d'une société à l'autre ou d'une commande à l'autre¹⁷ : seule une partie des membres du vivier acceptent de répondre à une nouvelle enquête, mais ils doivent impérativement en ignorer le thème. Cette condition d'ignorance est capitale, pour éviter le « biais d'intérêt » ; elle est respectée par l'enquête de l'INSERM.

Certaines sociétés gestionnaires de panels utilisent à cet effet un algorithme semi-aléatoire, qui applique aux répondants une série de quotas ou une combinaison de quotas. L'algorithme intègre également certains contrôles : il peut plafonner le nombre d'enquêtes par personne pendant une certaine durée (afin de limiter l'émergence de « répondants professionnels »), veiller à ce que la même personne ne remplisse pas plusieurs questionnaires d'une même enquête, récupérer les répondants excédant les quotas d'une autre enquête, voire d'un autre institut de sondage (technique du déroutage), éliminer les questionnaires remplis à une vitesse excessive ou à l'aveugle (comme la pratique consistant à cocher systématiquement la même case sur une colonne d'items), déceler l'intrusion massive d'un groupe de répondants cherchant à influencer les résultats, empêcher le remplissage sériel par des robots, vérifier la cohérence des réponses avec le profil déclaré de la personne, etc.

Les enquêtes par quotas sur panels sont parfois décrites – et décriées – par des journalistes d'investigation, qui racontent avec quelle facilité ils ont pu adhérer à plusieurs panels sous des identités diverses¹⁸. Ces examens critiques sont évidemment salutaires s'ils poussent les instituts à améliorer la gestion des panels, mais il faut savoir que les problèmes ainsi mis au jour ne sont pas tous propres aux panels en question ; ils se posent pour toute interaction à distance entre une institution et des adhérents ou des administrés. L'institution doit savoir miser *a priori* sur la relation de confiance, tout en prenant les précautions nécessaires pour prévenir l'abus de confiance. Le deuxième objectif ne doit pas tuer le premier. La statistique publique elle-même ne pourrait pas décrire correctement les populations si elle passait son temps à les soupçonner de fraude *a priori*.

De fait, l'objection des profils imaginaires incontrôlés, adressée par les Huit à l'enquête INSERM-CIASE, n'est pas propre aux viviers de répondants en ligne. Elle vaut aussi pour les répondants tirés de façon aléatoire sans interaction visuelle avec des enquêteurs, par exemple dans les sondages tirés dans la base nationale des numéros de téléphone fixes et mobiles tenue par l'ARCEP, ou à partir du recensement de la population¹⁹.

L'enquête par quotas sur panel est devenue la méthode dominante dans le monde des sondages. À mesure que se propage la connexion à Internet au sein de la population, la constitution d'échantillons sur Internet se généralise, au point de supplanter le tirage aléatoire des numéros de téléphones, qui avait lui-même supplanté l'interrogation à domicile. Chacune de ces innovations suit l'évolution des comportements et doit son succès à sa capacité à réduire fortement les coûts et les délais par rapport à la technique précédente. Dans ce contexte, une position de stricte orthodoxie consistant à écarter d'office les échantillons non probabilistes aboutirait *de facto* à réserver le monopole des enquêtes et des sondages représentatifs à l'INSEE

¹⁷ La littérature de langue française est discrète à ce sujet. Pour une description plus circonstanciée des méthodes, voir la référence de la note 26.

¹⁸ Luc Bronner, « Dans la fabrique opaque des sondages », *Le Monde*, 4 novembre 2020.

¹⁹ Rappelons que le recensement français repose sur la méthode du dépôt-retrait des bulletins dans les boîtes aux lettres physiques ou électroniques des habitants et qu'il implique rarement une rencontre physique avec l'agent recenseur (sauf à la demande, dans le cas des personnes âgées ou handicapées, par exemple). L'agent recenseur n'est pas un enquêteur ; il n'est pas tenu de contrôler la véracité des déclarations.

tous sujets confondus, – ce qui n’est pas souhaitable et que l’INSEE lui-même, du reste, ne souhaite pas.

Les chercheurs de l’INSERM mobilisés par la CIASE exposent correctement les deux méthodes disponibles de tirage d’échantillon, probabiliste et non probabiliste (pp. 375-377). Ils assument le choix de la seconde méthode proposé par l’IFOP, qui a exploité son vivier de répondants, avec des extensions sur les viviers partenaires. L’enquête INSERM-CIASE a sollicité en définitive 243 601 panélistes, *sans leur annoncer à l’avance le thème de l’enquête*. Cette précaution est fondamentale si l’on veut éviter le « biais d’intérêt » consistant à sélectionner en priorité les personnes qui s’intéressent à l’enquête. C’est seulement une fois obtenu l’accord du panéliste que le thème de l’enquête a été dévoilé. Il l’est, certes, au début du questionnaire, comme le relève l’académie en croyant dénoncer là une faiblesse de l’enquête, mais cette annonce fait partie de l’exigence déontologique de loyauté que la CNIL veille à faire respecter : tout questionnaire de la statistique publique ou de la recherche publique doit annoncer clairement le thème du questionnaire et s’y tenir. Mais, une fois le thème dévoilé, les non-réponses ont été minimales : aucun biais majeur ne peut être décelé à ce stade.

L’algorithme mis en œuvre par le prestataire du vivier à la demande de l’IFOP a dispersé les panélistes entre les enquêtes du moment, pré-structuré l’échantillon par le biais des quotas sociodémographiques classiques et appliqué *in fine* une post-stratification de l’échantillon final des répondants qui attribue à chacun un coefficient de redressement destiné à améliorer la représentativité.

Les chercheurs de l’INSERM précisent bien (au bas de la p. 376) que l’application de ces procédures non aléatoires à un panel d’abonnés ne permet pas, en théorie, de calculer des intervalles de confiance. Ils notent cependant que la question est discutée dans la littérature scientifique, ce qui est vrai²⁰. C’est pourquoi ils annoncent avoir pris la décision de calculer des intervalles de confiance qui encadrent les valeurs estimées à 95 %.

Mises en œuvre sérieusement, les enquêtes par quotas résistent à la critique

La diffusion de ces nouvelles pratiques met-elle en cause radicalement la qualité de l’enquête ? Les spécialistes du sondage aléatoire sont eux-mêmes divisés à ce sujet. Sur ce point, on peut conseiller à M. Hugues Portelli, président de l’Académie catholique, de relire attentivement l’annexe statistique du rapport qu’il avait préparé en octobre 2010 avec son collègue Jean-Pierre Sueur, rapport cité à la note 5 du rapport de l’Académie. Le spécialiste interrogé, Pascal Ardilly, était alors administrateur de l’INSEE, membre de l’Association française de statistique, auteur de plusieurs manuels de référence sur la théorie des sondages. Il est aujourd’hui inspecteur général de l’INSEE. Il a beau représenter l’institution nationale qui détient de fait le monopole des sondages probabilistes en France, il ne condamne pas pour autant les techniques non aléatoires :

« L’auteur de cette note est convaincu que la méthode des quotas est acceptable si et seulement si elle est mise en œuvre par un organisme professionnel, expérimenté, offrant des garanties quant au processus de collecte et un minimum de transparence pour ce qui concerne sa pratique et ses calculs. Certes il est toujours mathématiquement préférable, à taille d’échantillon donnée, de pratiquer de manière aléatoire, mais en cas d’impossibilité (base de sondage inexistante, coût trop

²⁰ Question développée à propos des enquêtes de santé sur les comportements à risque par Carol A. Gotway Crawford, “Comment”, *Journal of Survey Statistics and Methodology*, 2013, 1(2), p. 118.

élevé, impératif de délais...), il n'y a pas non plus lieu de stigmatiser outre mesure l'approche empirique dès lors qu'elle est sérieusement mise en œuvre²¹. »

Or, s'il est un organisme professionnel expérimenté dans le monde des sondages, c'est bien l'IFOP, fondé en 1938 par Jean Stoetzel, comme le rappelait opportunément M. Portelli dans la discussion du Sénat.

Le statisticien de l'Insee auditionné en 2010 par le sénateur Portelli jugeait même possible de calculer des marges d'erreur dans des enquêtes par quotas. Il terminait sur cette précision, qu'on peut étendre des sondages politiques à tous types de sondage :

« Il est aussi nécessaire de relativiser l'impact de l'échantillonnage : d'une part il ne faut pas oublier qu'il existe bien d'autres erreurs que l'erreur d'échantillonnage, probablement au moins aussi perturbatrices en matière de mesure des opinions politiques : non-réponse, erreurs de réponse (volontaires ou non) sur ce sujet sensible, volatilité de l'opinion dans le temps ; d'autre part, l'échantillonnage ne préjuge pas de tous les traitements qui sont effectués ultérieurement, en phase d'estimation, et qui ont évidemment des conséquences déterminantes sur les résultats diffusés : l'échantillonnage probabiliste ne protégerait l'utilisateur contre aucun de ces risques ! »

Une difficulté majeure des sondages probabilistes est l'ampleur croissante des taux de non-réponses, contrée tant que bien que mal par des procédures de remplacement ou d'imputation, elles-mêmes sujettes à caution.

Invitée à commenter l'une des premières études comparant les sondages probabilistes et les sondages non probabilistes²², la directrice de la Surveillance des comportements à risques au CDC d'Atlanta a émis un jugement équilibré sur les biais respectifs des deux méthodes : « Dans la pratique, je ne vois guère de différence entre l'engagement des volontaires et le désengagement des non-répondants »²³.

L'analyse comparée des méthodes d'enquêtes : de nombreuses expériences en cours

De fait, l'évaluation comparée des enquêtes aléatoires et des enquêtes par quotas en ligne a donné lieu à une vaste littérature méthodologique, que les Huit semblent méconnaître.

Un premier courant, celui des « probabilistes » officiant généralement dans les instituts statistiques nationaux, reste attaché à la méthode canonique des tirages aléatoires, alors qu'un autre courant, non probabiliste et pratiquant de longue date les *web surveys*, rappelle cette réalité historique récurrente : la pratique empirique des enquêtes a prouvé le mouvement en marchant, sans attendre que la théorie des sondages parvienne à la justifier scientifiquement. Ce fut déjà le cas pour les enquêtes par quotas, inaugurées dès 1905 par Anders Kiaer, le fondateur de Statistics-Norway, mais validées seulement en 1934 – et de façon partielle – par Jerzy Neyman dans un article fondateur²⁴. La discussion se poursuit de nos jours aujourd'hui. La *web bibliography* ne cesse de progresser et elle évolue très rapidement. On se contentera d'évoquer quelques éléments du débat. Les publications les plus intéressantes sont celles qui mettent sur

²¹ « Sondages et démocratie : pour une législation plus respectueuse de la sincérité du débat politique. Rapport d'information n° 54 » de MM. Hugues Portelli et Jean-Pierre Sueur, fait au nom de la commission des lois du Sénat, déposé le 20 octobre 2010, « Annexe 2. Éléments sur la précision des sondages : contribution de M. Pascal Ardilly (Société française de statistiques). »

²² R. Baker et al., "Summary report of the AAPOR Task Force on non-probability sampling", *Journal of Survey Statistics and Methodology*, 2013, 1(2), pp. 90-143.

²³ "I see little practical distinction between 'opting in' by volunteering and 'opting out' by not responding" (Carol A. Gotway Crawford, "Comment", *ibid.*, p. 118).

²⁴ J. Neyman, "On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection." *Journal of the Royal Statistical Society* 97, 1934.

piéd des expériences comparant les performances des divers systèmes, à savoir leur capacité à se rapprocher des estimations réalisées par les méthodes classiques de la statistique publique.

Statistics Netherland, l'Office statistique des Pays-Bas, a publié en 2016 des expériences comparatives sur des enquêtes des années 2010. Elles débouchent sur un jugement défavorable à l'égard des enquêtes sur panel²⁵. Les tests effectués sur 18 panels de volontaires sur Internet (*opt-in online panels*) font apparaître des performances inférieures à celles des sondages aléatoires, y compris des sondages menés sur Internet auprès de non-volontaires. Cela prouve au passage que les biais statistiques n'étaient pas liés à l'usage d'Internet, comme le suggère le rapport des Huit, mais au mode de sélection des répondants.

La comparaison la plus étendue et la plus fouillée est venue d'une équipe du département des études de communication de Stanford²⁶. Les valeurs qu'il s'agissait de prédire sont classées en variables démographiques de base (utilisées pour les quotas), variables démographiques secondaires et variables non démographiques. Les écarts aux valeurs données par les enquêtes aléatoires de la statistique publique sont mesurés pour six entreprises de panels différentes. Les auteurs concluent en faveur du sondage aléatoire, mais la synthèse des résultats montre que trois des six panels non aléatoires ont des performances assez proches des panels mixtes.

Plus récemment, l'Université Pompeu Fabra a comparé l'efficacité de la 8^e vague de l'Enquête sociale européenne avec celle du panel CRONOS (Cross-national Online Survey), financé également par l'Union européenne, mais sur les fonds du programme d'infrastructure SERISS²⁷. Les auteurs montrent qu'un panel en ligne recruté avec de solides garde-fous et bénéficiant d'un suivi attentif offre des performances en matière d'estimation statistique très proches de celles du sondage aléatoire, avec des résultats largement interchangeables.

En France, enfin, le CREDOC (Centre de recherche pour l'étude et l'observation des conditions de vie) a mené en 2015 une expérience consistant à comparer les résultats de l'enquête Conditions de vie selon deux méthodes de collecte : l'interrogation face à face de personnes sélectionnées selon un échantillon par quotas classique et l'interrogation en ligne d'un échantillon de répondants puisés par quotas dans un *access panel*. L'expérience a été présentée à Sciences Po en octobre 2021 au séminaire commun du CERI et du CEE à Sciences Po, par Mme Sandra Hoibian, directrice du pôle Société du CREDOC. Le compte rendu détaillé de la séance est consultable en ligne²⁸. Il en ressort une convergence totale des résultats sur un grand nombre de variables, mais aussi quelques différences notables, liées à l'effet de « désirabilité sociale » : n'ayant pas à se justifier devant un enquêteur physiquement présent, les panélistes fidélisés s'avèrent plus conservateurs, moins compatissants envers les personnes vulnérables, politiquement plus à droite que les répondants en face-à-face – soit l'inverse du biais suggéré par l'Académie. De ce fait, ils sont sans doute moins enclins à critiquer une autorité établie telle que l'Eglise. Là encore, on ne voit pas pourquoi les biais de recrutement des viviers panélisés aboutiraient à grossir fortement les déclarations d'abus sexuels.

²⁵ E. Bruggen, J. van den Brakel, J. A. Krosnick, "Establishing the accuracy of online panels for survey research", Statistics Netherland, *Web Survey Methodology*, 2016.

²⁶ B. MacInnis, J. A. Krosnick, A. S. Ho & M.-J. Cho, "Accuracy or measurements with probability and nonprobability survey samples: replication and extension", *Public Opinion Quarterly*, 82(4), 2018.

²⁷ A. Cernat & M. Revilla, "Moving from Face-to-Face to a Web Panel: Impacts on Measurement Quality", *Journal of Survey Statistics and Methodology* 9(4), 2020.

²⁸ Séminaire CERI/CEE, « Les sciences sociales en question : grandes controverses épistémologiques et méthodologiques », compte rendu de la 56^e séance : « Peut-on se fier aux enquêtes en ligne ? », 21 oct. 2021, https://www.sciencespo.fr/ceri/sites/sciencespo.fr.ceri/files/cr_04102021.pdf.

Dans la dernière décennie, les promoteurs des enquêtes de panels en ligne ont publié des manuels de référence qui ne dissimulent pas les difficultés rencontrées par les sondages non aléatoires. Ils soulignent néanmoins qu'un *monitoring* plus serré des panels dans les trois phases de développement (recrutement, imposition de quotas, pondération *ex post*) peut améliorer les performances des méthodes d'enquête non aléatoires et les rapprocher des méthodes aléatoires classiques, tout en réduisant les coûts et les délais²⁹. Ils insistent sur un point absolument décisif : les probabilités d'entrer dans le vivier de départ et, ultérieurement, de figurer dans l'échantillon des répondants doivent être indépendantes des thématiques de l'enquête. Cette condition est bien remplie dans le cas de l'enquête de l'INSERM.

Les trois sources d'information de la CIASE : des discordances aisément explicables

Les Huit ont tenté de jeter le discrédit sur la méthode de l'INSERM en évoquant le « gouffre » qui séparerait la quantification des abus opérée par voie d'enquête et celle qu'on peut tirer d'autres sources : les archives ecclésiastiques et judiciaires fouillées par l'équipe de l'EPHE (4832 cas) et les victimes ayant témoigné devant la Commission (2738 cas).

Le rapport de la commission est pourtant clair à ce sujet, et la tribune publiée dans *Le Monde* du 14 décembre 2021 en réaction aux critiques de l'Académie, n'a fait que le rappeler : seuls 4 % des répondants de l'enquête générale ont déposé une plainte, ce qui atteste l'ampleur du chiffre noir des abus sexuels commis par l'Église catholique. Par ailleurs, seule une minorité de plaintes ont débouché sur la constitution de dossiers conservés dans les archives des institutions et restés accessibles à l'heure actuelle. Le rapport de la commission, quant à lui, ne parle pas de « gouffre » mais d'un « écart massif » qui tient objectivement à la nature des sources et des procédures et non pas à une sombre volonté de partialité. Si l'on peut parler de « gouffre », c'est à propos du vide où sont tombées tant de plaintes réelles, sans compter les plaintes rentrées, celles qui n'ont jamais pu s'exprimer.

Un solide argument avancé par la CIASE (§ 565) est la congruence du chiffrage des agressions sexuelles imputables à l'Église avec celui des abus sexuels perpétrés dans l'ensemble de la société française, plus spécialement, dans d'autres institutions de socialisation de l'enfance ou de la jeunesse, à savoir l'école, les colonies de vacances et les activités sportives, où le nombre de cas, selon la même enquête, s'élève respectivement à 141 000, 103 000 et, derechef, 103 000. Comparés au bilan de chacune de ces trois institutions, les abus imputables à l'Église sont environ deux fois plus nombreux, mais on reste bien dans le même ordre de grandeur. Les agressions sexuelles commises au sein des familles sont si nombreuses qu'au total, quand on considère l'ensemble des cas déclarés toutes institutions réunies, seuls 6,1 % d'entre eux (environ 330 000 sur 5,4 millions), sont imputables à l'Église catholique. Dans ces conditions rien ne permet de soupçonner la commission d'avoir ciblé spécialement l'Église, que ce soit en raison d'une hostilité a priori ou, comme disent les psychologues, d'un « biais implicite » qui aurait faussé son jugement.

²⁹ R. Tourangeau, F. G. Conrad & M. P. Coupe, *The Science of Web Surveys*, New York: Oxford University Press, 2013. Voir aussi le manuel de M. Callegaro, K. Lozar Manfreda & V. Vehovar, *Web survey methodology*, London: Sage, 2015. – C'était déjà la conclusion avancée en 2003 dans une excellente revue de la question : Ronan Divard, « La représentativité des échantillons issus d'*access panels* en ligne : une question majeure pour l'avenir des études de marché », 2009, édité en ligne en 2013 <hal-00819324>. Après un exposé lumineux des problèmes techniques posés par ce type d'enquête, l'auteur concluait qu'il fallait les relativiser sans les minorer ; il détaillait les mesures de contrôle nécessaires pour améliorer leur qualité.

Expérience et compétences de l'équipe de l'INSERM

L'art du statisticien ou du sondeur est de savoir travailler sur des données imparfaites. Il ne peut éliminer totalement les biais de sélection, les non-réponses ou les défaillances de la mémoire. Mais il dispose d'un arsenal de moyens adéquats pour estimer l'ampleur et la direction des biais et en limiter raisonnablement les effets : quotas a priori, coefficients de redressement, données comparatives... Ces moyens ont été utilisés à bon escient par l'équipe de l'INSERM, qui ne mérite aucunement le discrédit qu'on a voulu jeter sur ses méthodes. L'équipe a déclaré en toute honnêteté les méthodes d'échantillonnage, de redressement et d'exploitation qu'elle a dû adopter dans les limites de temps et de coût qui lui étaient imparties. Ces méthodes – y compris l'appel au système de vivier de répondants potentiels dans lequel a été tiré l'échantillon final – ont fait leurs preuves dans d'autres enquêtes.

Après lecture attentive du rapport, y compris de ses notes et annexes méthodologiques, force est de conclure que la CIASE ne pouvait choisir équipe plus expérimentée que celle de l'INSERM pour mener au mieux l'enquête générale de victimation. Pour assurer la qualité scientifique de l'ensemble, y compris la rédaction finale, elle a fait appel à des conseillers scientifiques de premier ordre, Henri Leridon pour la démographie, Alfred Spira pour l'épidémiologie et la santé publique, rompus l'un et l'autre à l'organisation et à l'exploitation de grandes enquêtes sur des sujets sensibles. Il n'existe pas à l'heure actuelle, en France, d'expertise plus solide en la matière. La présidence de la CIASE a donc fait le meilleur choix possible en recrutant de tels collaborateurs. Le travail accompli dans ce rapport ne laisse transparaître aucune hostilité de principe envers l'Église catholique ; il témoigne d'une quête indépendante de vérité et de justice, qui illustre les valeurs profondes de la statistique sociale et des études de santé publique.

François Héran

Professeur au Collège de France, chaire « Migrations et sociétés »

Directeur de l'Institut Convergences Migrations (coordonné par le CNRS) depuis 2017

Chargé du cours de « statistique appliquée » à l'École nationale de la statistique et de l'administration économique de 1982 à 1995

Chef de la division des « Enquêtes et études démographiques » de l'INSEE de 1993 à 1998

Directeur de l'Institut national d'études démographiques de 1999 à 2009

Président de la *European Association for Population Studies* de 2012 à 2016