

## **Retour sur 35 années d'observation « personnelle » des populations vulnérables**

---

**Benoît RIANDEY**

Institut national d'études démographiques, Paris

À la veille de quitter l'Ined et la vie professionnelle, en juin 2013, je propose aux participants une rétrospective des passionnantes enquêtes cet organisme ou associées à lui relatives aux populations vulnérables ou aux questions sensibles auxquelles j'ai été associé pendant trente-cinq ans. Ce thème pourrait en effet caractériser les enquêtes de l'Ined, comme en témoigne la journée de colloque que j'ai organisée sur ce thème à l'occasion du Cinquantenaire de cette institution en 1995.

L'énumération des thématiques ou sous-populations serait longue et diverse. À travers elles, on jaugera les difficultés de succès de la collecte ; on appréciera son plus ou moins grand succès jusqu'à un éventuel renoncement. L'échantillonnage peut être en cause. Ce peut être aussi le mode d'observation ou encore la méthodologie d'analyse. La communication analysera donc ces enquêtes sous ces trois angles.

Je ne chercherai pas à définir une population vulnérable ou une question sensible car c'est souvent à l'occasion d'une question sensible pour elle qu'une personne révèle sa vulnérabilité, à moins qu'elle ne puisse esquiver la question. Je confondrai abusivement les adjectifs sensibles et vulnérables et, contrairement à tous mes enseignements, j'esquiverai moi aussi la saine posture du démographe ou du sociologue prenant soin de caractériser une population, un groupe social au-delà de ses individus constitutifs.

Tentons donc l'énumération des populations « vulnérables » rencontrées (au sens imprécis précédent) dans un ordre chronologique approché : les mères d'un enfant décédé en bas âge ou d'un enfant handicapé, les femmes qui ont connu l'avortement, les grands blessés de la route et les personnes handicapées, les mères ou pères de familles monoparentales, les femmes et hommes ayant connu le divorce, créanciers ou débiteurs de pensions alimentaires, certaines populations immigrées, notamment les travailleurs logés en foyer, des femmes en situation de polygamie, des immigrés qui sont ou ont été en situation irrégulière, les homosexuels (hommes ou femmes), les toxicomanes, les conjoints infidèles (au sens particulier de leur vulnérabilité), les détenus, les femmes victimes de violence résidant en ménage ordinaire ou en foyer, des femmes excisées, les malades du sida et de toute autre maladie chronique, les handicapés, les sans domicile français ou étrangers, francophones ou non, adultes ou mineurs, les victimes (potentielles ou effectives) de discrimination... Une liste à la Prévert qui trouve sa justification dans l'objet même du colloque.

En premier lieu, on s'intéressera aux difficultés d'échantillonnage, y compris celle d'élaborer un échantillon témoin adéquat. La question du contact et du maintien d'une communication non invasive constitue un second axe de ce retour d'expérience. Le questionnement exige une grande finesse professionnelle. Mais l'évolution des modes de collecte offre des voies nouvelles qui permettent de corriger des insuffisances ou erreurs passées, d'être plus efficace à ce sujet dans le respect de la législation et de la déontologie. Bien sûr, l'aspect réglementaire a marqué toute cette période et cette problématique depuis le vote de la loi Informatique et libertés en 1978 jusqu'aux directives européennes plus récentes.

Il est naturel de commencer par les **thématiques familiales**. L'enquête de 1985 sur le recouvrement des pensions alimentaires concerne pleinement notre propos. D'emblée, l'échantillonnage posait problème : contacter des hommes ou des femmes divorcés est aisé à partir de la base du recensement, mais des hommes ou des femmes ayant divorcé donc pour certains remariés est une tout autre gageure. Guy Desplanques, toujours inventif, nous a suggéré de recourir à l'enquête Famille associée au recensement. Cette excellente suggestion nous a conduits à nous limiter aux femmes créancières, abandonnant le point de vue des hommes payeurs réguliers ou débiteurs, exceptionnellement créanciers. Un dialogue difficile avec certains enquêtés a été évité au prix d'un renoncement par défaut d'échantillonnage. L'expérience a porté réflexion puisque l'Ined a demandé à l'Insee d'introduire un échantillon d'hommes dans l'enquête Famille de 1990, ce qui ne fut réalisé que dans celle de 1999.

Néanmoins, prendre explicitement contact avec une personne en raison de son divorce passé aurait été hautement incorrect, en particulier à l'égard d'une femme remariée depuis son divorce. Les chercheurs ont astucieusement affiché une problématique englobante, rassurante, très mobilisatrice, « Les femmes face au changement familial ». Ce fut la condition du remarquable succès de la collecte avec un taux de réponse de 85 %, alors même que 42 % des femmes de l'échantillon avaient changé de logement, voire même de nom. Cette problématique était si favorable qu'il se révéla possible et profitable de commencer le questionnaire par l'inventaire des unions, contrairement au projet initial qui traitait du changement familial intergénérationnel ; au test, cet abord avait d'emblée été rejeté par les enquêtées par leur réplique « *Allez demander directement à mes parents* ».

La question « *Quelles étaient les raisons qui étaient à l'origine de votre mésentente ?* » n'intervenait qu'en page 12. Des enquêtées ont pleuré pendant l'entretien. L'équipe en a été troublée. Ce trouble a été relativisé suite à une étonnante demande énoncée par lettre et se résumant à ceci « *J'étais présente quand Madame X a été enquêtée pour l'enquête sur le divorce. Moi-même, je suis divorcée. Je voudrais être enquêtée* ». Cette personne – qui n'aurait pas dû assister à l'interview – exprimait le constat que retracer son histoire douloureuse dans un cadre formel en même temps que 3 000 femmes qui avaient vécu la même difficulté pouvait avoir un effet dédramatisant. C'est beaucoup plus rapide et beaucoup moins cher qu'une psychanalyse ; peut-être beaucoup plus efficace ?

Cette enquête m'a convaincu qu'un excellent professionnalisme permettait d'évoquer les vécus les plus sensibles en suscitant l'empathie avec l'enquêté. Une démarche tout à

l'opposé du recueil d'un calendrier biographique sec. L'enquête « Situations défavorisées » de l'Insee avait été un modèle de la sorte en demandant à l'enquêté son appréciation subjective relative à chaque période passée ; de même, le questionnaire « Biographie et entourage » recueillait cette auto-évaluation de l'enquêté.

Cette situation avait prévalu en 1978 pour le volet français de l'enquête mondiale sur la fécondité. Pour beaucoup d'enquêtées, ce fut la première occasion d'évoquer contraception et avortement. L'Insee s'était déchargé sur l'Ined de cette enquête délicate... néanmoins collectée par les enquêtrices de l'Insee. Certaines enquêtées ont évoqué des situations très douloureuses auprès d'enquêtrices très attentives. Celles-ci, initialement inquiètes de l'enquête, sont revenues enthousiastes du terrain et ont communiqué leur satisfaction à leur hiérarchie au point que le directeur de la statistique, Jacques Desabie, citait cette enquête comme l'exemple type d'une enquête supposée impossible qui recevait le meilleur des accueils par les enquêtés.

Le renoncement à échantillonner une sous-population évoqué ci-dessus est intervenu en 1981, mais de façon plus radicale, pour les familles monoparentales lors de l'enquête « Vie familiale-vie professionnelle ». Il s'agit bien souvent de familles vulnérables. L'échantillon, tiré trois ans après le recensement, ne pouvait reposer que sur un échantillon uniforme de logements n'offrant qu'un effectif insuffisant de ces familles pour une analyse spécifique de leur situation. J'ai le remords de n'avoir pas convaincu de l'opportunité de prendre un échantillon complémentaire même partiel, comme les allocataires parents isolés du régime général de la Caisse nationale d'allocations familiales (Cnaf).

Ultérieurement, ces difficultés ont trouvé une solution dans la technique de l'enquête filtre – ou en deux phases – dans le cadre de l'enquête « Situations familiales » relative aux « familles atypiques ». Le tiers sortant de l'enquête « Emploi » fournissait un micro-recensement frais de 20 000 ménages, dont un sous-échantillon de 5 000 ménages surreprésentait toutes les situations précédemment évoquées. C'est une illustration de la stratégie englobante opérée cette fois pour l'échantillonnage.

Cette même stratégie se montre très adaptée aux enquêtes menées auprès des **populations immigrées** ou d'origine immigrée. Ainsi en était-il du premier titre de l'enquête MGIS relative à l'intégration des immigrés et de leurs enfants : « Enquête mobilité sociale et géographique ». L'évaluation du processus d'intégration des immigrés peut être menée par comparaison à l'insertion sociale des autres migrants en métropole que sont les Domiens ou en Île-de-France qu'y sont les provinciaux. Cette stratégie ne satisfait pas la Cnil, qui a réclamé les mots *Intégration* et *Immigrés* dans le titre de présentation de l'enquête. Réagissant à cette rigidité garantissant un échec à l'enquête, l'Ined s'est fixé sur le titre « Mobilité géographique et insertion sociale ». Les tests avaient montré que les immigrés européens refusaient tout à fait cette appellation : « *Les immigrés, ce n'est pas nous, disaient-ils, ce sont les Maghrébins* ». La Cnil s'est satisfaite de « l'effort de l'Ined ». Cette stratégie fut au cœur de la problématique de François Héran pour l'enquête « Identité-Histoire de vie » : l'origine immigrée est l'une des facettes de ces résidents de France et l'échantillon ne fait que

surreprésenter – modérément – les immigrés et leurs enfants. Dans l'enquête « Trajectoire et origine (TéO) », les Domiens et leurs enfants sont effectivement surreprésentés, alors que, tristement, en 1992, l'enquête de l'Insee sur les Domiens en France fut réalisée en même temps et sans coordination avec l'enquête MGIS.

L'échantillonnage des enfants d'immigrés nés en France pose problème après leur émancipation car le recensement ne fait pas part de leur origine. Pour l'enquête MGIS en 1992, Guy Desplanques suggéra de les repérer dans l'« Échantillon démographique permanent » au domicile de leurs parents jusqu'en 1975, puis de les localiser au recensement de 1990. Pour l'enquête Identité, l'enquête « Famille » répondit à la question puisque l'Insee y avait introduit les origines des deux parents à la demande bien antérieure de l'Ined. Cette heureuse situation n'eut qu'un temps, faute d'une enquête Famille associée aux premières vagues du recensement rénové. Guy Desplanques suggéra une procédure assez complexe *via* l'état civil, malgré les incertitudes de cet appariement anonyme.

L'un des sujets abordés était évidemment l'étude des discriminations. Si on en retient la définition d'inégalités résiduelles non expliquées après élimination des inégalités structurelles, se pose avec acuité la question de l'échantillon témoin et celle de sa structure si sa taille est réduite.

La méthodologie de l'échantillon témoin mérite discussion. Si, assez souvent, les sociologues sont focalisés sur leur population d'étude au point d'en oublier de produire un échantillon témoin de référence, cette critique épargne généralement les démographes et toujours les épidémiologistes. Mais cette remarque n'épuise pas le sujet.

Pour les épidémiologistes, la taille de l'échantillon témoin égale celle des cas (les patients) puisqu'on cherche un sosie indemne de chaque cas. Bien plus, cette méthode des sosies assure une grande efficacité statistique, plus recherchée par les épidémiologistes que par les démographes. L'épidémiologie analytique vise en premier lieu à tester les facteurs de risque et les relations causales, tandis que la démographie s'intéresse principalement à produire des estimations. Les épidémiologistes favorisent donc une puissante comparaison toutes choses égales par ailleurs. Revenant à notre propos, l'intégration des immigrés ne prend sens qu'en comparant leurs opinions, comportements ou situations à celles de natifs comparables, s'il en est. Mais peut-on comparer la situation des ouvriers immigrés des grandes entreprises du bâtiment ou des travaux publics à celle d'ouvriers natifs qu'on aurait peine à trouver ? L'enquête MGIS s'est limitée à un petit échantillon témoin de ménages natifs représentatifs de la même plage d'âges, peu enclin à une comparaison terme à terme. À l'opposé, l'enquête de l'Iresco sur les « Stratégies d'intégration des immigrés et des enfants d'immigrés » veillait, en 1987, à sélectionner l'échantillon témoin dans les mêmes communes, et donc dans le même bassin d'emploi. Ainsi, entre ces deux enquêtes, la structure de l'échantillon témoin témoigne respectivement d'une problématique plus descriptive pour la première, plus explicative pour la seconde.

Si l'on s'intéresse plutôt à la perception des discriminations subies, comme dans l'enquête « Identité-Histoire de vie », le problème ne se pose pas dans ces termes. L'objectif

était avant tout de comprendre comment la couche supplémentaire liée à l'origine immigrée compose avec les autres dimensions de l'identité dans une perspective multidimensionnelle. Aussi, l'échantillon devait permettre une simple comparaison entre sous-populations ; les échantillons des immigrés et des enfants d'immigrés étaient seulement légèrement surpondérés mais restaient très minoritaires, contrairement à leur présence massive dans les enquêtes MGIS et Téo.

L'enquête MGIS intervint dans une courte fenêtre politique favorable à ces enquêtes, tandis que l'enquête Téo tomba en pleins feux, aggravés il est vrai par une maladresse dans le projet de questionnaire : il est indiscutable que, pour mesurer les discriminations d'accès au logement des Antillais, on peut difficilement faire table rase du faciès, le département de naissance n'étant pas pertinent pour les békés et les métros... et difficilement s'y référer compte tenu des métissages ; par contre, cette référence posée à un Africain sub-saharien est parfaitement superflue. Il était donc bien inutile de jeter de l'huile sur le feu.

Sur le fond, on était confronté à cette difficulté mal maîtrisée en France selon laquelle les caractéristiques générant les discriminations sont les informations pertinentes pour repérer les discriminations et lutter contre elles. D'où la polémique vouant aux gémonies les statisticiens tentant de mesurer les discriminations comme nous l'imposent les directives européennes. Bien autrefois, la question avait été réglée par une interprétation hyper-rigide de la loi Informatique et libertés : à l'interrogation de la Cnaf sur la perception de leurs droits sociaux par les immigrés, il avait été répondu que, puisque les allocations sont universelles en France, le critère de la nationalité est non pertinent et donc excessif pour la finalité de versement des allocations donc de collecte illégale par cette administration : il n'était pas pertinent d'examiner si l'administration remplissait sa mission. En août 2004, la transcription tardive de la directive du 25 octobre 1995 apurait le problème en validant la stricte finalité statistique.

Doit-on considérer comme appartenant à une population vulnérable toute personne ressentant une discrimination ? Et celle ayant une conduite à risque comme dans l'enquête sur les comportements sexuels en 1992 ? Il s'agissait là du risque de contracter le sida. Sur ce dernier point, notons la discrétion du questionnement téléphonique dont les réponses étaient limitées aux mots sans danger : *oui, non, souvent, parfois, jamais, tout à fait d'accord...*

Laissons donc les multipartenaires et limitons nous aux **pratiques homosexuelles et toxicomanes**. Pour ces dernières, notons à nouveau le recours à une stratégie englobante : à l'ère du partage des seringues, c'était spécifiquement l'injection de drogue qui diffusait l'épidémie de sida. Le caractère illégal de cette consommation conduit à des sous-déclarations de cette pratique. Aussi la question posée a-t-elle été plus large : « *Au cours des 12 derniers mois, avez-vous consommé une drogue dure ou douce* » ? Certes, ce n'est pas la consommation de haschich qui était visée. Mais le repérage des conduites à risque en début de questionnaire était trop délicat et fut englobé dans une question filtre retenant indifféremment pour un « questionnaire long » un échantillon témoin sélectionné sur son jour de naissance

dans le mois et les auteurs de conduites à risque. La procédure très conviviale était astucieuse. Fut-elle un succès ?

Qu'en fut-il de la déclaration de la pratique masculine d'un rapport homosexuel au cours des cinq dernières années ? Elle fut évaluée par cette enquête téléphonique au taux fort bas de 1,4 %, tandis que la proportion de questionnaires longs **sans** conduite à risque déclarée dépassait déjà fâcheusement et largement la taille de l'échantillon témoin<sup>1</sup> attendu **total**. L'inquiétude de l'équipe fut dissipée par la publication par l'équipe anglaise d'un résultat équivalent obtenu par une autre méthode, le recueil d'un questionnaire auto-administré sur papier. Cette inquiétude ne dura que jusqu'à la publication, dans la revue *Population*, de l'article de Rogers prouvant la très forte sous-déclaration des pratiques sensibles, tant oralement par téléphone que par écrit sur papier, par comparaison à l'enregistrement auto-administré sur téléphone ou sur ordinateur portable. En définitive, l'estimation à 4 % par l'enquête sur quotas en Rhône-Alpes était sans doute plus exacte que celle de notre enquête sophistiquée.

En fait, cette sous-déclaration aurait été observable et corrigible dès l'entretien sans l'excès de contraintes imposées par la Cnil, qui avait interdit de collecter la date de naissance complète de l'enquêté pour éviter tout risque d'identification. Précaution excessive, compte tenu de la composition automatisée des numéros de téléphone et de l'absence d'information géographique plus précise que la région et la catégorie de commune. Ainsi, on n'a pas pu enregistrer explicitement l'appartenance de l'enquêté à l'échantillon témoin et, inversement, une dissimulation de conduite à risque. Un rattrapage aurait été possible pendant l'interview, en particulier par un module téléphonique auto-administré. Je regrette que cette procédure n'ait pas été mise en place lors de la réplique de cette enquête en 2006, sous le nom d'enquête CSF.

Aujourd'hui, on chercherait à assurer la sincérité des réponses par une interview par Internet, mais la prise de contact directe par Internet soulèverait de très certains problèmes de représentativité, au même titre que les questionnaires insérés dans la presse spécialisée. Bien sûr, cette sous-estimation des pratiques homosexuelles, ou plus généralement à risque, invalide les projections de l'épidémie, de toutes façons démenties par le succès de la politique de prévention avant l'arrivée des traitements pharmaceutiques. Le potentiel explosif de l'épidémie en a été sous-estimé et donc, également, le succès de la prévention menée.

En définitive, le mode de collecte le plus fiable nous semble bien le recours à la carte filtre avec rattrapage auto-administré, que ce rattrapage soit administré par téléphone, ordinateur portable ou Internet. Mais le cadre médical est toujours le plus rassurant car le plus légitime pour les enquêtés ; ainsi en fut-il de l'enquête « Vespa » en 2003 sur les conditions de vie des malades du sida. Elle fut réalisée à l'occasion des consultations hospitalières

---

<sup>1</sup> Échantillon représentatif de la population, y compris donc des personnes ayant des conduites à risque.

nécessaires pour l'attribution des traitements. Tel fut aussi le cas de l'enquête auprès des femmes victimes de l'excision.

Mais l'enregistrement d'une conduite homosexuelle, éventuellement éphémère, n'induit pas l'auto-reconnaissance identitaire. C'est cette dernière qui permet d'évoquer l'appartenance à une population spécifique. L'énonciation de celle-ci implique bien davantage que la confiance dans la confidentialité du processus d'enquête, mais ce propos dépasserait ma compétence.

Peut-on parler de la population des **femmes victimes de violences conjugales** ? Je ne m'avancerai pas sur ce point. Vulnérables, ces femmes victimes le sont certainement. La lecture des commentaires des enquêtrices de l'enquête de 1985 sur le divorce dessinait cette situation de violence. On y lisait comment des femmes renonçaient à leur pension alimentaire et évitaient tout contact avec l'ancien conjoint de crainte de ses violences. Ainsi s'imposait l'enquête Enveff réalisée en 2000 et relative aux violences subies par les femmes. Ses résultats ont fait date, mais on ne peut que regretter que cette grande enquête ait fait l'objet d'une virulente critique justifiée par l'absence de l'échantillon témoin masculin réclamé par le comité du label du Conseil national de l'information statistique (CNIS) : 7,7 % de harcèlement moral de la part de l'époux n'ont pas la même signification selon que, de la part de l'épouse, il s'élèverait à 1 % ou à près de 7 %. Violence machiste dans le premier cas ; difficultés ou désagréments de la vie de couple de l'autre. Cette faute de méthode est d'autant plus regrettable que, au prix d'un accroissement des intervalles de confiance de guère plus de 10 %, l'échantillon de 7 000 enquêtées permettait, avec 1 000 hommes enquêtés, de fournir une estimation masculine de référence sans compromettre les estimations différentielles féminines. Je ne prône aucunement le partage à égalité de l'échantillon. La gestion intelligente de l'échantillon de l'enquête « Histoire familiale » est éclairante à ce sujet : au moment d'introduire un échantillon masculin dans la collection des enquêtes Famille, l'équipe Insee-Ined a choisi une répartition 3/8, 5/8 entre les sous-échantillons masculin et féminin. Ainsi pouvaient être produites de novatrices estimations de la fécondité masculine, tandis qu'on suivrait avec finesse les évolutions de la fécondité féminine car, dès qu'on veut mesurer des évolutions (ou des différentielles), une taille bien supérieure d'échantillon s'impose.

Mon interprétation – très personnelle – de cette lacune de l'enquête Enveff est que l'évidence des violences subies par les femmes pouvait dispenser d'en établir la preuve. Cette position a affaibli la valeur probante de l'enquête. Heureusement, les enquêtes ultérieures de la Direction de la recherche, des études, de l'évaluation et des statistiques (Drees), de l'Insee avec l'Observatoire national de la délinquance (OND) et la future enquête Virage de l'Ined n'ont pas réitéré ce choix, et nous laissons aux chercheurs impliqués le soin de confirmer les résultats de l'enquête Enveff.

Cette exigence comparative impose-t-elle qu'une sous-population ne peut être enquêtée qu'en présence d'un échantillon couvrant l'ensemble de la population ? Peut-elle toujours l'être dans cette condition ?

La première question appelle bien sûr une réponse négative. Il suffit d'avoir veillé à poser des questions communes aux enquêtes en population générale pour que l'analyse puisse cependant être comparative. À l'opposé, je pense à une publication d'une enquête relative au Rmistes indiquant par exemple leur taux de titulaires du permis de conduire sans aucun élément de comparaison globale ou à structure sociodémographique comparable relative à la population générale ; une publication scientifiquement discutable.

Pour répondre à ces deux questions, Maryse Marpsat m'autorisera à évoquer les enquêtes sur les sans domicile. Si cette population donne lieu à une enquête spécifique, c'est bien parce qu'elle ne peut pas être atteinte par les procédures habituelles d'enquête. Nous n'aborderons pas la question complexe de la définition d'un sans domicile dans tout ce *continuum* de la précarité de l'habitat. Il n'en demeure pas moins que le questionnaire peut comporter des questions spécifiques ne permettant pas de comparer les réponses ni d'évaluer « *Est-ce beaucoup ou pas beaucoup ?* ». Par contre, il est intéressant de se rappeler de la présence, dans une enquête « Logement » de l'INSEE, de questions rétrospectives sur d'éventuelles périodes passées sans logement. Analyse intéressante sur la sortie de la situation de personnes sans domicile, évaluation de l'importance de la population sortie et sur la durée de « séjour » sans domicile, mais biais de mesure sur l'effectif des sans domicile et ignorance des décès dans cette situation, comme dans toute enquête rétrospective sans appariement avec l'état civil.

En définitive, nous avons évoqué de nombreuses situations de personnes vulnérables. Pourrions-nous parler de populations vulnérables ? Les caractériser se révèle souvent délicat et effectué uniquement par un comportement fait ou subi. Telles ne sont pas les situations de personnes ayant divorcé ou des immigrés pour lesquels le critère démographique clair délimite ces populations. Mais avoir passé une ou plusieurs nuits dehors définit-il le SDF ? Il ne s'agit là que d'indices. Le travail sociologique s'impose.