

# Étude de l'histoire familiale : L'édition 1999 de l'enquête Famille

---

Francine CASSAN, François HÉRAN,  
Laurent TOULEMON, Cécile LEFÈVRE<sup>(1)</sup>

À l'occasion du recensement de la population de mars 1999, 380 000 hommes et femmes vivant en domicile ordinaire ont rempli un bulletin complémentaire sur le thème de leur « histoire familiale », incluant des questions sur leurs origines, leurs enfants, leurs périodes de vie en couple et leur parcours social, ainsi que sur les langues d'usage au sein de leur famille, nationales et régionales ; ces mêmes informations ont été recueillies auprès de 6 600 personnes vivant en communauté, dont 1 700 détenus de sexe masculin. Profondément rénové, le questionnaire démographique de cette édition 1999 de l'enquête Famille, la première à s'adresser également à des hommes, permet de nombreuses études, couvrant les histoires familiales dans toute leur complexité, comme en témoigne le présent ouvrage. L'exploitation des questions complémentaires sur les langues et parlars apporte par ailleurs une information très attendue sur la diversité linguistique de notre pays et la dynamique de l'intégration nationale tout au long de ce siècle. Les tout premiers résultats, portant sur les détenus, ont été publiés en avril 2000. Un groupe d'exploitation de l'enquête a été créé en septembre 2000, qui a pu dès lors travailler sur un fichier provisoire de l'enquête. Le fichier définitif a été mis à la disposition du groupe d'exploitation en juillet 2002, et à la disposition du public au cours de l'été 2003. Pour ne pas alourdir la présentation, il sera ici essentiellement question, s'agissant des aspects méthodologiques, de l'enquête auprès des ménages ordinaires.

Le principe de réalisation est très simple : dans une fraction des secteurs d'agent recenseur constitués pour la collecte du recensement (un sur cinquante depuis 1962)<sup>(2)</sup>, les personnes entrant dans le champ de l'enquête Famille reçoivent, en plus de leur bulletin individuel (BI), un questionnaire spécifique qu'elles sont invitées à remplir elles-mêmes et à remettre à l'agent

---

<sup>(1)</sup> Une première version de ce texte a déjà été publiée dans *Le Courrier des statistiques*, n° 93, Insee, mars 2000. Cet article de François Héran, Laurent Toulemon, et Francine Cassan, est ici actualisé par Cécile Lefèvre.

<sup>(2)</sup> La taille de l'échantillon enquêté s'explique par la nécessité de ventiler la plupart des données selon de nombreux critères, dont l'âge et la génération.

recenseur avec les autres documents du recensement (voir les questionnaires en annexe 2. Pour l'enquête Famille à strictement parler : « Étude de l'histoire familiale »). Pour le recensement : « Bulletin individuel » et « Feuille de logement ». Entre autres avantages, cette organisation permet de décharger le BI d'un certain nombre de questions démographiques qui figurent dans les recensements étrangers ou sont recommandées par l'Onu, mais qu'il serait coûteux de poser à l'ensemble de la population (nombre d'enfants que l'on a déjà eus, état matrimonial détaillé, dates de décès du conjoint ou des enfants...). Réciproquement, l'enquête Famille utilise les informations collectées dans le recensement.

### ***1) Une rénovation importante, appuyée sur les travaux de l'Ined***

L'enquête de 1999 a été l'occasion d'une rénovation importante<sup>(3)</sup>, inspirée en grande partie par les travaux de l'Ined sur les situations et histoires familiales (encadré 1).

Les questions sur les situations et les histoires conjugales et familiales (volet démographique) ont été sensiblement remaniées, en prenant soin de distinguer les situations de fait des situations légales et de repérer le cas des personnes ayant vécu plusieurs unions successives ; y ont été ajoutées des questions sur les enfants adoptés et les enfants du conjoint, ainsi que sur les petits-enfants et arrière-petits-enfants.

D'autre part, le champ enquêté a été élargi, dans deux directions : les hommes et les personnes âgées de plus de 64 ans. Il a également été décidé, s'agissant des communautés, d'étendre le champ de l'enquête au cas des personnes vivant en maison de retraite ou détenues en prison (encadré 2)<sup>(4)</sup>.

### ***2) Hommes et femmes***

Jusqu'à l'édition de 1990, l'enquête Famille était réservée aux femmes, selon une habitude chère aux démographes. À l'origine, elle portait sur les femmes mariées ayant achevé de constituer leur descendance (entre 45 et 54 ans). Puis le champ s'est élargi aux femmes non célibataires de moins de 65 ans, et par la suite aux femmes non mariées : ainsi, l'enquête de 1990 a porté sur 340 000 femmes âgées de 18 à 64 ans.

Prolongeant cette logique d'élargissement du champ de l'enquête, l'extension aux hommes permet d'apporter des informations fiables sur la paternité, sujet mal connu<sup>(5)</sup>. Les contrastes sociaux ne sont pas identiques pour les hommes et pour les femmes, le célibat étant croissant pour les femmes

<sup>(3)</sup> Pour un historique détaillé de l'enquête Famille, voir l'article de Guy Desplanques.

<sup>(4)</sup> L'enquête Famille était déjà étendue depuis 1975 aux femmes vivant en cité universitaire ou dans un foyer de travailleuses. Seuls sont ainsi restés exclus de l'enquête de 1999 les hôpitaux, les casernes et les communautés religieuses.

<sup>(5)</sup> Voir sur ce thème l'article, dans cet ouvrage, de Laurent Toulemon, *Enfants et beaux-enfants des hommes et des femmes* (chap. 3).

### Encadré 1

#### Les travaux de l'Ined sur les situations et histoires familiales

Les situations légales recueillies par le recensement et l'état civil ne tiennent pas compte des histoires familiales. Ainsi, le recensement ne fait pas de différence entre personnes mariées et personnes remariées, il ne repère précisément ni les couples non mariés ni les relations entre adultes et enfants d'un même ménage. De son côté, l'état civil distingue entre enfants « légitimes » et « naturels », mais n'informe pas sur le type de relation qui unit les parents des enfants nés hors mariage (corésidence éventuelle, durée de la cohabitation).

Pour prendre la mesure de ces phénomènes, l'Ined a mené en 1986 une enquête spécifique sur les situations familiales, à partir d'un échantillon de 5 000 personnes issu de l'enquête Emploi de l'Insee, dans lequel étaient surreprésentés les ménages comprenant des personnes en situation « atypique » : adulte non célibataire vivant seul, adultes vivant en couple sans être mariés, enfant dont la personne de référence du ménage n'est pas le parent, enfant dont un seul parent réside dans le ménage. Le questionnaire comprenait des questions rétrospectives sur l'ensemble des périodes de vie en couple, avec ou sans mariage, sur les dates de naissance et de changement de résidence de l'ensemble des enfants de la personne, résidant dans le ménage ou non, ainsi que sur l'ensemble des enfants vivant dans le ménage sans être les enfants de la personne interrogée.

Cette enquête a permis de décrire les situations familiales de fait des adultes et des enfants. Elle a montré que les « histoires familiales », même différentes de leur sanction légale par le mariage, n'étaient pas des « secrets de famille » indicibles, mais à l'inverse que les personnes interrogées répondaient volontiers à des questions sur leur situation et histoire familiales, questions qu'elles jugeaient concrètes et bien adaptées à leur situation. Toutefois, elle souffre de deux limites : malgré l'efficacité du plan de sondage, la taille réduite de l'échantillon ne permet pas de produire une information précise sur des situations relativement rares ; en outre, les personnes de plus de 50 ans ne sont pas prises en compte, ce qui rend impossible toute description de l'évolution sur le long terme.

Une deuxième enquête de même nature a été conduite en 1994, dans le cadre du cycle européen des *Fertility and Family Surveys* (enquêtes sur la fécondité et la famille). Touchant 23 pays, cette opération a donné lieu à plus de soixante-dix projets de travaux de comparaison internationale, en cours de réalisation, à partir de fichiers de données standardisés. Des informations sur ces travaux sont disponibles, à l'adresse <http://www.unece.org/ead/pau/ffs>

avec le niveau de diplôme, et décroissant pour les hommes. Les valeurs des indicateurs de fécondité peuvent diverger légèrement en moyenne entre hommes et femmes (en raison de la mortalité différentielle et des migrations) mais, surtout, les comportements des hommes sont davantage diversifiés que ceux des femmes : on compte davantage d'hommes sans enfant ou pères de

## Encadré 2

### Les enquêtes auprès des personnes vivant en communauté

Les personnes vivant en cité universitaire ou en foyer de travailleurs ont été incluses dans des échantillons particuliers ; comme les personnes vivant en domicile ordinaire, elles ont rempli un questionnaire complémentaire dans le cadre du recensement de la population. Ont ainsi été collectés 1 900 bulletins.

L'extension de l'enquête au cas des personnes vivant en maison de retraite ou détenues en prison s'est en revanche traduite par la mise en place de deux opérations spécifiques.

*Les maisons de retraite : un bulletin simplifié, rempli à l'occasion d'une autre enquête*

Les 6 000 hospices et maisons de retraite de France hébergent quelque 400 000 personnes, dont la moitié ont plus de 85 ans, et les trois quarts sont des femmes. Les tests réalisés dans les maisons de retraite ont montré qu'il était nécessaire d'améliorer la collecte sur deux plans : la simplification du questionnaire et l'assistance aux personnes interrogées par l'agent recenseur. Cela risquait de compliquer la collecte pour les agents recenseurs. On a donc décidé de réaliser la partie « maisons de retraite » de l'enquête non pas dans le cadre du recensement, mais à l'issue des entretiens de l'enquête « Handicaps-incapacités-dépendance » (HID)<sup>(1)</sup>, réalisés à la fin 1998, notamment auprès de 5 000 personnes hébergées en maison de retraite. Si la personne en était d'accord, et si elle était apte à répondre elle-même, l'enquêteur poursuivait l'entretien avec certaines questions de l'Étude de l'histoire familiale dans sa version simplifiée (celles qui ne figuraient pas déjà dans le questionnaire HID). Plus de 3 000 bulletins ont ainsi été remplis.

*L'enquête auprès des détenus : une opération entièrement nouvelle*

Le recensement incluant les personnes incarcérées, on a décidé de les inclure également dans l'enquête. On connaît fort mal, en effet, la démographie de la population carcérale (50 000 personnes, dont 4 % de femmes). On a des raisons de penser qu'elle est proche de celle des groupes les plus démunis, mais aucune étude d'ensemble ne permet de s'en assurer pour l'instant. L'administration pénitentiaire a soutenu sans réserve la réalisation d'un test, qui s'est avéré largement positif, puis celle de l'enquête elle-même. Après le test, l'enquête a été limitée aux hommes, et les bulletins ont été légèrement modifiés pour que l'entretien soit adapté à la situation des hommes incarcérés, mais sans insérer de questions spécifiques au monde carcéral. Les bulletins ont été remplis par entretiens, en toute confidentialité, et ont été confiés à l'Insee sans être vus par le personnel des établissements. Au total, l'enquête a eu lieu dans 28 établissements et 1 700 bulletins ont été remplis<sup>(2)</sup>.

<sup>(1)</sup> Cf. l'article de Pierre Mormiche : « L'enquête HID de l'Insee », *Courrier des statistiques* n° 87-88, Insee, décembre 1998.

<sup>(2)</sup> De nombreuses publications ont exposé les résultats de ce volet de l'enquête (voir bibliographie de l'enquête), en particulier Francine Cassan, (sous la coordination de), *L'histoire familiale des hommes détenus*, Synthèses, n° 59, Insee, janvier 2002.

nombreux enfants, en raison de la plus grande fréquence pour les hommes du célibat, d'une part, des remariages, d'autre part. Enfin, les enfants vivant le plus souvent avec leur mère après une éventuelle rupture du couple parental, ils vivent plus fréquemment avec un beau-père qu'avec une belle-mère. Du point de vue des adultes, la résidence avec des enfants du conjoint se développe pour les hommes, mais reste rare pour les femmes.

Des tests réalisés à partir de 1993 ont montré que le taux de réponse masculin était égal au taux féminin, et qu'il n'existait pas de différence notable dans la qualité des réponses. La seule difficulté qui est apparue concerne les jeunes hommes célibataires, qui ne voient pas toujours l'intérêt de répondre à un questionnaire sur la famille. C'est pour contrer ce phénomène que l'on a rebaptisé le questionnaire « Étude de l'histoire familiale », au lieu de « Étude sur les familles » : même si l'on n'a pas fondé soi-même de famille, chacun a une histoire familiale.

Aux accords de genre près, le questionnaire est identique pour les hommes et les femmes. Or, les deux membres d'un couple ont le plus souvent des histoires familiales très proches. Aussi, pour éviter de demander deux fois la même information, les échantillons d'hommes et de femmes ont-ils été disjointes : dans un même logement, ont donc été enquêtés soit des hommes soit des femmes, en pratique tous les hommes ou toutes les femmes du ménage âgés de 18 ans ou plus.

### **3) *Plus de limite d'âge***

L'extension aux personnes de plus de 64 ans se justifie par le caractère rétrospectif de l'enquête Famille. On peut ainsi récapituler les mouvements démographiques de la France en remontant aussi loin que possible, au-delà du *baby boom*. Dans le volet linguistique, il est même possible de remonter plus haut, puisque les personnes les plus âgées indiqueront en quelle langue leur parlaient leurs parents quand elles avaient elles-mêmes cinq ans, c'est-à-dire à la veille de la guerre de 1914-1918.

La suppression de la limite d'âge supérieure n'est cependant pas sans risques : biais de sélection dus à la mortalité, problèmes de mémoire, incapacité physique ou mentale à remplir soi-même un questionnaire, qui viennent compliquer la collecte.

### **4) *Enfants, beaux-enfants, petits-enfants, arrière-petits-enfants***

Comme dans les versions précédentes, le questionnaire commence par des questions sur les enfants de la personne enquêtée, y compris les enfants adoptés. Pour chaque enfant, on demande le sexe, la date et le lieu de naissance ainsi que, pour les enfants adoptés, la date de l'accueil de l'enfant. Le cas échéant, les âges au départ du domicile parental et éventuellement au décès (avec une précision sur les enfants mort-nés, que les personnes ont à cœur de mentionner) permettent de retracer les périodes pendant lesquelles la

personne a élevé ses enfants. Les questions sur les enfants adoptés permettent de savoir – c'est une première – qui sont les adoptants, d'un point de vue social et démographique, et comment se sont diffusées les diverses pratiques de l'adoption depuis la Grande Guerre (le lieu de naissance de l'enfant permettant de séparer adoption nationale et adoption internationale)<sup>(6)</sup>.

Viennent ensuite des questions sur les beaux-enfants, entendus ici comme les enfants d'un conjoint (ou d'un ex-conjoint) que la personne a élevés. Pour les beaux-enfants, les questions sont identiques à celles des enfants, permettant ainsi, pour certaines études, d'assimiler enfants et beaux-enfants élevés.

Puis trois questions viennent compléter l'histoire parentale : l'accueil d'autres enfants, la présence de petits-enfants et arrière-petits-enfants avec leur nombre et l'âge de l'aîné d'entre eux, ce qui permettra enfin de savoir à quel âge en moyenne l'on devient grand-parent ou arrière-grand-parent, une donnée qui n'est accessible pour l'instant qu'au moyen de simulations statistiques.

### ***5) Périodes de vie en couple, avec ou sans mariage***

Les questions sur l'histoire conjugale se concentrent sur les unions de fait, avec ou sans mariage. Après une question sur la situation actuelle, on demande les dates de début et de fin des première et dernière périodes de vie en couple, ainsi que les dates éventuelles de mariage, de rupture, de divorce et de décès du conjoint. Sur la recommandation de la Commission nationale de l'informatique et des libertés, on a limité l'interrogation à deux unions au maximum, la première et la dernière, la notion d'union étant restreinte pour le passé aux unions ayant donné lieu à une corésidence d'au moins six mois.

Des questions sur les caractéristiques du dernier conjoint (date et lieu de naissance, mariage éventuel avant l'union, profession) permettent de situer socialement les couples rompus. D'un point de vue d'étude démographique, il aurait été préférable de poser des questions sur le premier conjoint, pour les personnes ayant vécu plusieurs unions, mais les tests ont montré la réticence des personnes à évoquer trop précisément un conjoint qui a été « remplacé » par un autre dans l'histoire conjugale.

Les questions sur les périodes de vie en couple et sur les enfants et beaux-enfants livrent une description précise des diverses situations familiales (familles « monoparentales » et « recomposées »<sup>(7)</sup>). Aussi mettent-elles sur le même plan les cohabitations sans mariage et les couples mariés, en incluant éventuellement les enfants n'habitant pas avec le parent ou le beau-parent. Leur dimension rétrospective permet, en outre, de retracer la succession des types de famille dans lesquels les personnes ont pu vivre.

---

<sup>(6)</sup> Voir l'article de Juliette Halifax, « Les familles adoptives en France », dans cet ouvrage.

<sup>(7)</sup> Voir par exemple sur ces sujets les articles, dans cet ouvrage, d'Élisabeth Algava sur « Les familles monoparentales, des caractéristiques liées à leur histoire matrimoniale », et celui de Corinne Barre « 1,6 million d'enfants vivent dans une famille recomposée ».

### **6) Origines et parcours d'insertion sociale**

Aux questions classiques sur le nombre d'enfants qu'a eus la mère de la personne enquêtée, le rang de naissance de celle-ci et la profession de ses parents, sont venues s'ajouter de nouvelles, portant sur le lieu de naissance des parents et leur décès éventuel. Le bulletin comporte également des questions sur les principales étapes de la vie professionnelle et sociale : âge à la fin des études, au départ du domicile parental et au premier emploi, interruptions d'activité professionnelle ou périodes de chômage.

### **7) Le volet linguistique**

Des questions sur les langues parlées en famille (et au-delà avec des proches) visent à décrire la diversité linguistique du pays et la dynamique de l'intégration nationale tout au long du XX<sup>e</sup> siècle. Il s'agit d'une grande première, car la statistique publique n'avait jamais abordé ce thème à l'échelle nationale. Grâce aux données recueillies, très attendues des historiens et linguistes, l'Insee peut enfin assurer avec les pays étrangers un minimum de comparaison dans le domaine des langues parlées et de leur transmission familiale (encadré 3)<sup>(8)</sup>.

Le questionnaire met délibérément sur le même plan les langues régionales et les langues nationales, car ces dernières sont souvent des dialectes, tandis que les premières peuvent être des prolongements frontaliers de langues nationales.

S'agissant des immigrés, une information explicative essentielle, non directement recueillie par l'enquête mais qui est récupérée dans les bulletins individuels du recensement, est la date d'arrivée en France. En l'absence de cette donnée, on risquerait d'imputer à un effet culturel un comportement qui dépend surtout de l'ancienneté du séjour en France.

Une autre information déterminante, qui vaut pour tout le monde, est le désir éventuel, si l'on n'y vit plus, de retourner dans sa région ou son pays d'origine. Elle fait l'objet de la dernière question de l'enquête.

### **8) Enquête Famille et recensement**

Le couplage au recensement permet d'interroger un échantillon de grande taille pour un coût marginal très faible<sup>(9)</sup>. Comme pour les éditions précédentes, et pour des raisons de simplicité de collecte, on a choisi de retenir des secteurs entiers d'agent recenseur<sup>(10)</sup>, dont certains seraient entièrement dédiés à l'enquête auprès des hommes et les autres à l'enquête auprès des

---

<sup>(8)</sup> Outre la partie VIII consacrée aux résultats du volet linguistique, un ouvrage en collaboration entre l'Ined et la DGLFLF (Délégation générale à la langue française et aux langues de France) est à paraître sur ces sujets : *Le siècle des langues. Matériaux pour une histoire de la transmission des langues dans la France du XX<sup>e</sup> siècle*.

<sup>(9)</sup> L'enquête Famille porte uniquement sur la métropole mais elle est comparable à l'enquête Famille réalisée dans un des Dom, l'île de la Réunion (encadré 4).

### Encadré 3

#### Langues et parlers : une longue attente

La question des langues parlées en famille a été abordée au XIX<sup>e</sup> siècle, lorsque Victor Duruy, ministre de l'Instruction publique de Napoléon III, avait demandé aux instituteurs et secrétaires de mairie de dresser le bilan des enfants francophones et non francophones dans leur commune (1863)<sup>(1)</sup>. Signalées par Jacques Revel et Michel de Certeau (*Une politique de la langue : la Révolution française et les patois*, Paris, 1975), les données n'ont été exploitées qu'en 1977 par l'historien américain Eugen Weber. Les résultats sont très agrégés (on distingue les communes selon que le français est parlé par tous les enfants, une partie ou aucun).

Pour les périodes plus récentes, on dispose de résultats régionaux ou locaux, issus d'enquêtes de l'Insee ou de sondages commandités par les régions, les départements ou des associations locales. Il n'est toutefois pas possible d'en tirer des estimations à l'échelle nationale, les méthodologies mises en œuvre n'étant pas comparables<sup>(2)</sup>.

Le volet linguistique de l'enquête Famille 1999 permet donc de produire une information attendue de très longue date. Il trouve ses fondements dans l'enquête Efforts d'éducation des familles, conduite par l'Insee en 1992 sur un échantillon limité (5 300 parents d'élèves). Incluant des questions sur les langues, cette enquête a en effet montré qu'il était possible de questionner la population française sur le sujet et de calculer des taux de passage au français en une génération pour l'ensemble des langues nationales et régionales. L'enquête MGIS (Mobilité géographique et insertion sociale des immigrés) a confirmé la même année la pertinence de ces questions pour mesurer le rythme d'intégration culturelle des familles d'origine étrangère.

(1) Quelques années plus tard, le ministère de la Guerre, à la demande de Duruy, élargissait l'enquête aux conscrits.

(2) À titre d'exemple, la question n'a été traitée qu'une seule fois dans le cadre du recensement : c'était par la direction régionale de l'Insee-Alsace, en 1962.

femmes. Comme pour les éditions précédentes, il a été décidé de réaliser l'enquête dans un logement sur cinquante.

Les agents recenseurs concernés ont ainsi distribué à chaque ménage retenu dans l'échantillon, en même temps que les documents du recensement, autant de bulletins de l'enquête Famille que de personnes majeures du sexe correspondant, ainsi qu'une notice de présentation (voir annexe 1) exposant les objectifs de l'enquête, des résultats des éditions précédentes, la justification des questions sur le passé et quelques instructions pour remplir les bulletins.

(10) Pour le recensement, l'unité de collecte est le district (330 000 au total). Chacun des 115 000 agents recenseurs recrutés par les mairies pour l'exécution du RP 99 avait en charge un ou plusieurs districts. En moyenne, un secteur d'agent recenseur correspondait à une population d'environ 225 ménages et 500 personnes.



#### Encadré 4

##### **L'enquête Famille Dom : un autre protocole, mais qui préserve la comparaison avec l'enquête de métropole**

L'enquête Famille associée au recensement porte uniquement sur la métropole, mais est comparable à l'enquête Famille, plus complète, introduite dans un des Dom, celui de l'île de La Réunion.

À la demande des directions régionales des Antilles et de la Réunion, une enquête Famille spécifique a été conçue pour les Dom. Celle-ci s'est déroulée selon la formule classique d'une enquête ménages déconnectée du recensement, dès novembre-décembre 1997 pour la Réunion. Mais elle n'a pas été réalisée aux Antilles.

Administré par un réseau d'enquêteurs, le questionnaire Dom est sensiblement plus détaillé que le questionnaire métropole (une trentaine de pages). Il répond aux standards internationaux fixés par le programme mondial des enquêtes DHS (*Demographic and Health Surveys*), tout en étant cohérent avec les enquêtes Ined-Insee sur les situations familiales de 1986 et 1994.

Le détail de l'opération n'entre pas dans le champ de cette présentation, mais on soulignera un point important : le questionnaire métropole est intégralement inclus dans le questionnaire Dom. Ce dernier a été rédigé de manière à préserver la possibilité de publier des résultats France entière en combinant les données des Dom à celles de métropole.

Plusieurs articles ont été écrits à partir de cette enquête (voir bibliographie de l'enquête). Dans cet ouvrage, l'article de Didier Breton, « Fécondité et naissances précoces dans un département d'outre-mer : comparaisons entre l'Île de la Réunion et la métropole » utilise également ces données.

L'enquête Famille est certes indirectement nominative puisque les identifiants utilisés pour le recensement (numéro de district, rang de l'immeuble dans le district, rang du logement dans l'immeuble) sont reportés sur les bulletins collectés. Ceux-ci ne contiennent toutefois ni nom ni adresse. En outre et pour mieux assurer la confidentialité des réponses, ils ont été directement envoyés à l'Insee par les délégués du recensement, sans transiter par les mairies.

#### **9) L'utilisation des informations issues du recensement**

Le fichier de l'enquête Famille ne se limite donc pas aux informations contenues dans les bulletins de l'enquête. Il comprend également l'essentiel des informations provenant du recensement. L'appariement a été effectué à partir des identifiants utilisés pour le recensement, de la date de naissance et du sexe. Le taux d'appariement final atteint les 97 %, c'est-à-dire que sur les 380 000 questionnaires de l'enquête Famille recueillis, 368 000 ont pu être appariés avec les bulletins de recensement correspondant à la même personne et à son ménage.

Ces informations en amont sont bien sûr très utiles. Elles précisent la composition du ménage, ce qui permet notamment de mettre en parallèle les enfants qu'a eus la personne interrogée et ceux qui restent à sa charge, et de connaître l'ensemble des personnes qui résident avec elle. Surtout, elles fournissent les principales variables explicatives de l'exploitation : état matrimonial légal, lieu de résidence, niveau d'études et diplôme, catégorie socioprofessionnelle à deux chiffres, indicateur de nationalité, date d'arrivée en France pour les immigrés – et cela aussi bien pour la personne interrogée que pour l'ensemble des autres personnes du ménage. De fait, si l'on excepte la raison sociale de l'établissement employeur et certains éléments descriptifs du logement, c'est presque la totalité des informations du recensement qui intéressent l'enquête Famille. C'est pourquoi les ménages inclus dans le champ de cette enquête ont été systématiquement retenus dans le champ de l'exploitation « complémentaire » du recensement de la population<sup>(11)</sup>.

En outre, les données issues du recensement permettent de recueillir des informations sur les personnes qui ont refusé de remplir le bulletin de l'enquête (point développé ci-dessous)<sup>(12)</sup>.

### 10) L'échantillon

Avec un logement sur cinquante et compte tenu du taux de non-réponse tel qu'il avait été prévu (10 %), la taille de l'échantillon *a priori* pouvait être de 470 000 femmes ou 430 000 hommes majeurs. Il fallait éviter que l'élargissement de l'enquête aux hommes ne conduise à diviser par deux la partie féminine de l'échantillon, qui reste, en tout état de cause, au centre des analyses. D'un autre côté, l'échantillon masculin devait être suffisamment important pour faire l'objet d'analyses précises, que ce soit dans le volet démographique ou le volet linguistique de l'enquête. On a donc adopté une solution intermédiaire : le partage hommes/femmes ne se faisant pas à l'équilibre mais dans un rapport de 3/5.

Cet échantillon – il s'agissait là encore d'une première – devait en outre être conçu de sorte à être représentatif au niveau régional. À cet effet, on s'est basé sur des taux de sondage moyens de 1/170 pour les hommes et 1/110 pour les femmes, avec forcément un minimum de 5 000 hommes et 8 500 femmes s'agissant des régions pour lesquelles ces taux de sondage conduisaient à des effectifs inférieurs (sauf pour la Corse et le Limousin). Dans certaines zones, il a été jugé nécessaire, en vue de l'exploitation du volet linguistique de l'enquête, d'augmenter la taille des échantillons obtenus

---

<sup>(11)</sup> Cf. l'article de Mylène Chaleix et Olivier Haag : « Le plan d'exploitation du RP 99 », *Courrier des statistiques* n° 89, mars 1999 (dans cet article, il est question d'exploitation légère et d'exploitation lourde ; la terminologie a depuis lors évolué, et l'on ne dit plus « légère » et « lourde », mais « principale » et « complémentaire »).

<sup>(12)</sup> L'Insee n'a pas souhaité qu'elle soit obligatoire, cela en raison du caractère privé et rétrospectif des informations demandées. L'enquête a reçu le label d'intérêt général du Cnis (Conseil national de l'information statistique) et a fait l'objet d'un avis favorable de la Cnil (Commission nationale de l'informatique et des libertés).

par application de ces règles, surtout pour les hommes. Il s'agit du Nord, de l'Alsace, de la Corse, du Pays catalan, du Pays basque et de la Bretagne.

Dans la pratique, a d'abord été sélectionné dans chaque région ou partie de région (en cas d'extension « langues ») un échantillon représentatif de districts, par tirage systématique après classement en fonction des caractéristiques sociodémographiques des quartiers selon une méthode mise au point par Nicole Tabard. Les zones de délégué contenant ces districts ont été incluses dans l'échantillon de l'enquête. À l'intérieur de chacune, les secteurs d'agent recenseur retenus, soit pour l'enquête auprès des hommes, soit pour l'enquête auprès des femmes, ont été sélectionnés avec des probabilités d'inclusion spécifiques à chaque région ou partie de région.

Le taux de participation ayant été plus faible que prévu, on a recueilli *in fine* 380 000 bulletins, auprès de 145 000 hommes et 235 000 femmes.

### 11) Les non-réponses

Le taux de non-réponse à l'enquête, ou plus précisément le taux de personnes éligibles à l'enquête ayant rendu leur bulletin de recensement mais pour lesquelles le questionnaire de l'enquête EHF est manquant, s'élève finalement à 20,6 %. Ce taux est comparable au taux d'échec d'autres enquêtes non-obligatoires, mais très supérieur à celui de l'enquête Famille de 1990 (3 %), qui était obligatoire. Les 96 000 personnes qui ont rempli leur bulletin de recensement mais pour lesquelles on ne dispose pas de l'enquête EHF ne se répartissent pas « au hasard » dans la population. Ainsi les hommes ont un peu moins souvent répondu que les femmes (taux de participation de 78,8 % pour les hommes contre 79,8 % pour les femmes). Comme dans toutes les enquêtes, les personnes de moins de 25 ans ou de plus de 85 ans, les immigrés arrivés en France après 1975, les personnes vivant seules, les célibataires, ont moins participé que les autres. Les personnes n'ayant pas complètement rempli leur bulletin de recensement (pas d'activité professionnelle ni de diplôme déclaré) ont également plus souvent refusé l'enquête<sup>(13)</sup>.

Une modélisation de la probabilité de non-réponse à l'enquête selon différents critères confirme la plus faible participation des personnes âgées, des célibataires ne vivant pas en couple, des personnes nées à l'étranger, des personnes n'ayant pas déclaré de diplôme. Par contre, elle montre que les jeunes ont autant participé que les autres « toutes choses égales par ailleurs ». Elle confirme aussi des disparités importantes d'une région à l'autre, le taux de non réponse variant de 16 à 36 % suivant les régions. Ces probabilités différentes de participer à l'enquête ont été prises en compte dans les pondérations finales, grâce aux méthodes de post-stratification.

---

(13) Pour en savoir plus, voir la documentation des fichiers de l'enquête, en particulier la note méthodologique sur l'ampleur de la non-réponse, de Nicolas Razafindratsima.

## ***12) Le calendrier d'exploitation***

Avant de pouvoir produire des résultats fiables, il a fallu bien sûr tenir compte, en plus des pondérations décrivant le plan de sondage, des corrections pour non-réponses partielles et totales à partir des informations issues du recensement, et du diagnostic sur le degré de validité des résultats. Plusieurs étapes ont ainsi été nécessaires jusqu'à la diffusion du fichier définitif.

En juin 2000, on a disposé d'un premier fichier apuré et pondéré (la première pondération tenant compte uniquement du plan de sondage), dont le contenu était limité aux informations recueillies au travers du questionnaire de l'enquête EHF uniquement, sans appariement avec le recensement. Ce fichier a été utilisé pour produire des résultats provisoires.

En août 2001 a été rendu disponible un premier fichier d'étude, intégrant les variables prises en compte dans le cadre de l'exploitation « principale » du recensement (soit presque toutes, à l'exception de la PCS et des variables « ménages-familles »), des variables d'étude sur les individus et les enfants et les informations recueillies au travers du volet linguistique de l'enquête. Ce fichier a été pondéré en tenant compte des non-réponses par sexe, âge et taille du ménage (pondération provisoire).

Enfin, à l'été 2002, le groupe d'exploitation de l'enquête a pu travailler sur le fichier définitif, incluant les variables de l'exploitation complémentaire du recensement, les variables d'étude à l'échelle du ménage et les pondérations définitives (cf. ci-après). Il a été mis à la disposition du public à l'été 2003.

## ***13) Pondération et post-stratification***

Dans le fichier définitif, une palette de pondérations est proposée, utilisables suivant les objectifs de recherche. Une première pondération correspond donc simplement à l'inverse du taux de sondage. La seconde pondération proposée est le résultat d'une post-stratification par sexe et âge. Une pondération plus complexe a ensuite été élaborée sur la base d'une post-stratification selon les 7 variables suivantes : sexe et groupes d'âges, situation de couple, diplôme, pays de naissance, date d'arrivée en France, région de résidence, tranche de l'unité urbaine. Ces pondérations peuvent aussi être ajustées pour chaque région, afin de permettre des études régionales à partir de l'enquête<sup>(14)</sup>.

## ***14) De nombreuses possibilités de vérification...***

Quel crédit accorder aux déclarations recueillies dans un questionnaire auto-administré<sup>(15)</sup> ? Toute enquête de la statistique publique repose sur une relation de confiance. On suppose *a priori* que les déclarations des intéressés

---

<sup>(14)</sup> Pour en savoir plus, voir la documentation des fichiers de l'enquête, en particulier la note méthodologique sur le redressement, l'échantillonnage et les pondérations de Laurent Toulemon.

reflètent correctement leur situation. Mais cela ne suffit pas, et les questionnaires doivent receler des possibilités de vérification interne, permettant de recouper les données collectées et de les valider. Ainsi, l'enquête Famille aborde trois calendriers, celui des naissances, celui des unions et celui de l'insertion sociale. Si le premier est complet, le second est simplifié par rapport à la chronologie détaillée qu'on peut développer dans les enquêtes démographiques par enquêteurs, et le troisième reste sommaire. Mais l'essentiel est dans la mise en rapport de ces trois calendriers : il faut surtout pouvoir situer le contexte matrimonial et social de l'arrivée des enfants.

De nombreuses informations peuvent faire l'objet d'une validation interne par comparaison des réponses à l'échelle globale. Par exemple, le nombre de couples peut être mesuré à partir des réponses des hommes comme à partir des réponses des femmes. De même, le nombre d'arrière-grands-parents peut être déduit des réponses directes à la question « avez-vous des arrière-petits-enfants », ou bien estimé à partir du nombre de personnes qui déclarent à la fois avoir des petits-enfants et des parents vivants, en tenant compte du nombre de frères et sœurs de ces personnes.

En sus de ces vérifications internes, on a pu confronter les résultats issus de l'enquête à des informations connues par ailleurs : le nombre annuel de mariages et de naissances de l'état civil par exemple<sup>(16)</sup>. Bien que compliquées par les migrations, les décès et le fait que certains hommes peuvent ne pas avoir reconnu leurs enfants, ces vérifications ont permis cependant de contrôler assez précisément la qualité des résultats obtenus.

### **15) ... y compris pour le volet linguistique**

Il en va de même pour l'évolution de la pratique des langues. En effet, on interroge toutes les générations sur les langues reçues de la génération précédente et sur celles transmises à la génération suivante. Ainsi, le comportement d'une même génération fait l'objet d'une triple description, par elle-même et par les deux générations qui l'encadrent. On peut contrôler de cette façon un risque souligné avec force par les linguistes et les dialectologues consultés, à savoir que les pratiques linguistiques sont souvent sous-déclarées ou surdéclarées. Si le taux de perte d'une génération à l'autre peut dessiner une pente douteuse, trop forte ou trop faible selon le cas, la « pente des pentes » réunissant les générations sera fiable<sup>(17)</sup>. Elle permettra même d'estimer les biais de déclaration.

---

(15) Il s'agit d'une des questions traitées par l'article d'Éva Lelièvre et de Magali Mazuy, dans cet ouvrage, qui compare les réponses des mêmes personnes au questionnaire auto-administré de l'enquête EHF et à celui répondu en face à face avec un enquêteur à l'enquête Biographies et Entourage.

(16) Un tel travail de validation interne et externe est mené dans le document de travail de Magali Mazuy et de Laurent Toulemon, *Études de l'histoire familiale. Premiers résultats de l'enquête en ménages*, Ined (coll. Dossiers et recherches), n° 93, février 2001.

(17) Voir à ce sujet François Héran, « Les langues et la statistique publique », *Ville-École-Intégration Enjeux* n° 130, septembre 2002.

### ***16) Un groupe d'exploitation largement ouvert***

Les données recueillies ont permis et permettent de nombreuses études, tant sur chacun des différents événements démographiques décrits dans les bulletins que sur les histoires familiales dans toute leur complexité, à partir de la mise en regard des différents calendriers qui y sont abordés. Compte tenu de la taille de l'échantillon, ces études peuvent porter sur la situation actuelle et l'évolution des comportements démographiques dans différents groupes sociaux ou dans différentes régions. Peuvent également être réalisées des analyses de plus long terme, en tirant parti des enquêtes Famille précédentes. Sans oublier les comparaisons internationales : le domaine démographique se prête particulièrement bien à ce genre d'exercice, et de nombreux pays disposeront, dans le cadre des recensements à venir, de données plus ou moins similaires.

D'une façon générale, la préparation de l'enquête a donné lieu à une large consultation des chercheurs de l'Ined. Pour la partie sur les langues, exploitée en liaison avec la Délégation à la langue française et aux langues de France, avaient également été mis à contribution de nombreux organismes : le Haut Conseil de la francophonie et le Fonds d'action sociale pour les langues nationales, les délégations aux affaires culturelles ou linguistiques (il existe une telle délégation dans une quinzaine de nos départements) pour les langues régionales.

Les partenaires sont donc nombreux et divers. Aussi, dès 2000, il a été décidé de constituer, à l'Ined, un groupe d'exploitation de l'enquête, qui s'est chargé de la mise à disposition des données et de l'organisation des études à l'extérieur de l'Insee. Ce groupe a d'emblée été largement ouvert. Depuis sa première réunion en septembre 2000, il s'est réuni 8 fois. Ces réunions furent l'occasion de présenter à l'ensemble des personnes travaillant sur l'enquête les recherches en cours et d'en discuter. Les chercheurs de l'Ined en composent la plus grande partie, et l'Insee y participe également de manière importante (division « Enquêtes et Études Démographiques », « Redistribution et Politiques Sociales » et plusieurs Directions régionales). Des chercheurs de l'EHESS, de Paris I, de Paris II, du Cepremap, de Lyon II, de Strasbourg II, de Versailles Saint-Quentin-en-Yvelines, de la Cnav, et de la Drees, y participent aussi. Des lettres d'information ont été diffusées au groupe à raison de deux ou trois par an, et l'ensemble des informations relatives au groupe d'exploitation de l'enquête est disponible sur le site [www-ehf.ined.fr](http://www-ehf.ined.fr). Un des objectifs de ce groupe d'exploitation fut la réalisation du présent ouvrage collectif.