

L'EXPLOITATION D'UNE ENQUÊTE PILOTE EN VUE DE L'AMÉLIORATION D'UNE ENQUÊTE INNOVANTE AUPRÈS DES MÉNAGES : LE CAS DE L'ENQUÊTE HISTOIRE DE VIE

Emmanuelle CRENNER^(), France GUERIN^(**), Frédérique HOUSEAUX^(*)*

^() Insee, Département de la Démographie
^(**) INED*

Introduction : mise en place et procédure d'un test d'exploitation

Procéder à un test d'exploitation n'est pas la règle à l'INSEE, et l'appellation n'a aucun caractère officiel. Le recours à cette opération, en lieu et place d'un dernier test classique avant la collecte, a été décidé par le comité de pilotage de l'enquête au moment du dernier test papier du questionnaire, en février 2001, soit deux ans avant la collecte. L'idée était de réaliser un véritable modèle réduit de l'enquête afin de pouvoir en tirer des enseignements plus riches que ceux d'un simple test. Il faut souligner tout d'abord que ce sont les spécificités de l'enquête en terme de questionnement et d'échantillon qui sont à l'origine de cette décision, et que notre expérience ne vise pas à remettre en cause la procédure classique et justifiée des tests CAPI à l'INSEE.

Quelles sont ces spécificités ?

Spécificités en termes de questionnement :

En premier lieu, il faut souligner la nouveauté du thème de l'enquête (cf. infra), abordé sous l'angle d'un questionnaire statistique. En effet, même si l'INSEE possède une expérience d'enquêtes sociologiques, aborder une question aussi complexe que le lien social dans un questionnaire aussi exhaustif que possible n'était pas sans susciter des interrogations sur la faisabilité, du côté de l'exploitabilité des questions notamment.

Même si l'enquête reprend un certain nombre de questions déjà utilisées dans des enquêtes précédentes, elle comprend aussi de nombreuses questions nouvelles. De plus, la plupart de ces questions nouvelles sont soit des questions subjectives, soit des questions ouvertes, donc des questions dont l'exploitabilité n'était pas garantie de prime abord.

De ce point de vue, un des buts du test d'exploitation était de recueillir (en plus des informations habituelles qu'apporte un test classique sur l'accueil et la compréhension des questions par les enquêtés) des informations sur les réelles potentialités d'exploitation de ces questions : seraient-elles capables d'apporter un éclairage statistique intéressant sur les phénomènes identitaires ? Serions-nous

capables d'exploiter les questions ouvertes ? Pourrions-nous accorder une réelle confiance aux questions subjectives, nombreuses dans le questionnaire ?

Ces interrogations, sur lesquelles la plupart des concepteurs peuvent se permettre de faire des paris peu risqués, se posaient à nous avec plus d'acuité étant donné le nombre de questions concernées et la place importante qu'elles seraient amenées à tenir dans les résultats de l'enquête.

L'opération pilote devait donc être réalisée sur un échantillon suffisamment large pour permettre quelques analyses statistiques suffisamment poussées pour pouvoir, le cas échéant, détecter les questions redondantes ou inexploitables.

Un intérêt non négligeable de cette opération était également la possibilité pour les chercheurs du groupe d'exploitation de l'enquête, de commencer à préparer quelques pistes d'exploitation, quelques programmes, et de gagner du temps au final une fois les fichiers définitifs mis à disposition.

Spécificités d'échantillonnage

Une des motivations premières de l'opération pilote a été le test de la méthode sondage. En effet, compte-tenu des sous-populations pour lesquelles nous souhaitions disposer de suffisamment de questionnaires pour réaliser des analyses approfondies, il était nécessaire de procéder à des surreprésentations. Or les variables dont nous avons besoin pour ces surreprésentations ne figurent pas, pour la plupart, dans les variables du recensement (pays de naissance des parents, pays de naissance de l'enquêté, existence d'une limitation dans les activités due à un problème de santé). Il nous était donc impossible de tirer notre échantillon dans les bases habituelles (échantillon maître). Cependant nous disposons de deux bases qui contiennent, pour l'une les pays de naissance de l'enquêté et de ses parents), et pour l'autre, l'existence d'une limitation dans les activités. Il s'agit des enquêtes Histoire Familiale (EHF) et Vie Quotidienne et Santé (VQS). L'échantillonnage s'annonçait donc relativement complexe, et l'UMS a souhaité que la méthode soit testée avant d'être appliquée à grande échelle.

Les caractéristiques de l'opération pilote

Une fois la décision prise par le comité de pilotage, l'opération pilote a donc été mise en place de la façon suivante : test du questionnaire sur un échantillon au 1/10^{ème} de l'enquête finale, soit 1300 fiches-adresses, qui ont donné au final 800 répondants environs, 5 DR concernées et 45 enquêteurs, pour un coût total d'environ 80 000 € soit environ 10 fois plus qu'un test habituel. Le souhait des concepteurs et du comité de pilotage était de s'approcher le plus possible des conditions de collecte. Cela comportait un certain nombre de modifications par rapport à la façon de gérer un test classique, qui n'ont pas toujours été clairement anticipées. Ainsi, par exemple, du tirage des zones de test : en règle générale, lors d'un test, les directions régionales de l'INSEE font appel à des enquêteurs particulièrement disponibles, ou réputés pour leur sérieux. Le choix de la zone d'enquête leur revient donc. Dans notre cas, il ne pouvait en être ainsi : tester la méthode d'échantillonnage supposait que nous imposions aux DR les zones à enquêter, ce qui n'a pas toujours été facile compte-tenu de la charge de travail des enquêteurs et de leur localisation. De même, le travail demandé aux enquêteurs l'a été sous la forme d'un nombre de fiches-adresses traitées, et non de questionnaires remplis ; leur rémunération quant à elle comportait la majoration habituelle des tests, mais également la rémunération des déchets, qui n'intervient normalement qu'en collecte définitive. Enfin il a été nécessaire également, afin de tester au plus juste la durée du questionnaire, de demander aux enquêteurs de ne pas faire plus de remarques en cours d'entretien qu'ils n'en auraient fait au cours d'une collecte normale, consigne qui allait à l'inverse de ce qui est en général demandé dans un test classique. Pour toutes ces raisons, il a été préféré l'appellation d'opération pilote plutôt que test d'exploitation qui laissait trop penser que les données devaient être traitées comme dans une procédure de test habituelle.

Une fois l'opération terminée, les données ont été apurées en DR et transmises au centre informatique de Lille qui a fourni aux concepteurs une table SAS exploitable. Nous avons ainsi pu tester l'efficacité

des surreprésentations prévues dans l'échantillon et préparer une exploitation similaire, en modèle réduit, à celle d'une enquête.

Après la description des objectifs de l'enquête, la suite de cette présentation expose les divers apports de cette opération pilote : sur le plan technique d'abord : amélioration de la qualité du questionnaire et de la méthode de sondage, et sur le plan de l'exploitation ensuite : préparation des programmes informatiques et défrichage du terrain des résultats envisageables.

1. L'enquête « Histoire de vie »

« Histoire de vie » est le nom officiel du projet couramment dénommé « Construction des Identités » (les connotations parfois policières du terme identité – contrôle d'identité, papiers d'identité – ont amené à son occultation pour le public). Ce n'est pas l'acception administrative du terme qui nous intéresse ici, mais sa dimension sociologique. L'identité désigne, dans l'analyse sociologique, la façon qu'a chaque individu de se construire une place dans la société qui permette à la fois son intégration et l'affirmation de son individualité propre. C'est une autre façon de dénommer ce qui constitue le lien social, car c'est bien de cela qu'il s'agit dans cette enquête : étudier le lien social tel qu'il est vécu par les individus.

Dans une analyse « traditionnelle » de l'intégration sociale, l'appartenance à une classe d'âge, à une catégorie sociale et, éventuellement, à un groupe religieux ou politique suffit à la fois à fournir à chaque individu les motivations de son attachement à la société et les paramètres de son individualisation relative : il trouve sa place, et il se différencie des autres, de ceux qui n'appartiennent pas aux mêmes catégories. Les mobilités géographiques (à l'intérieur ou hors des frontières des pays), sociales, politiques ou religieuses, les ruptures familiales, conduisent à une complexification de ces relations identitaires « simples ». L'enquête « Histoire de vie » se propose à la fois de mesurer comment se gère cette complexité et comment d'autres paramètres d'insertion dans la société permettent de constituer à la fois du lien et de la différenciation.

L'enquête vise donc à mesurer comment chaque individu construit sa relation aux autres et à la société en général en fonction de sa trajectoire personnelle (origines géographiques et sociales, mobilités, projets) et de l'appartenance à certains groupes sociaux (classe d'âge, situation professionnelle, situation familiale, état de santé, engagement éthique, pratiques culturelles...). Elle déterminera donc pour chaque individu l'existence d'un certain nombre de « vecteurs identitaires » auquel l'individu attache une importance variable, qui ne se limite pas à du temps passé ou à une rentabilité économique ou sociale, mais qui s'exprime aussi par un attachement, une affirmation.

L'enquête se devait donc de recueillir à la fois des données objectives (le temps passé à faire une activité sera recueilli), et d'autres subjectives (les attachements, affirmations, etc.). L'enquête favorise les premières, en les complétant par les secondes, qui sont essentielles dans ce domaine.

Les données objectives portent à la fois sur l'appartenance à un groupe (groupe des jeunes, des femmes, des chômeurs, des amateurs de musique, des militants politiques...) et sur les éléments descriptifs de « l'identité en actes » (porter des vêtements propres à son groupe d'âge, être inscrit dans une association de défense ou de représentation d'un groupe professionnel, porter un insigne distinctif...) qui permettent de s'assurer de la place qu'un élément d'identité tient dans l'image sociale que la personne se construit ou qui s'impose à elle.

Chaque individu est porteur de plusieurs identités. Toutes ne jouent pas le même rôle dans son intégration dans la société. L'objet de l'enquête est de mesurer à la fois les identités décrites positivement (identités revendiquées) et celles qui sont vécues négativement (assignations subies)¹.

¹ Etant entendu que tout ce qui est « subi » n'est pas forcément objectivement négatif (reconnaissance sociale) et qu'inversement ce qui est affirmé peut être objectivement négatif (auto exclusion).

1.1. Une enquête sociologique

L'enquête « Histoire de vie » s'inscrit dans la tradition des enquêtes sociologiques de l'INSEE, dans laquelle se sont inscrites ces dernières années, MGIS, l'enquête Education (1992), la série des enquêtes Emploi du temps, les enquêtes PCV « Bonheur au travail », « Contacts entre les personnes ». Comme les précédentes, mais aussi comme les enquêtes à fort contenu sociologique comme « Sortants du RMI » ou « Sans domicile », l'enquête « Histoire de vie », dont la formulation et la préparation doivent beaucoup aux préoccupations des chercheurs ayant travaillé sur ce domaine, entend participer à la réponse de l'Insee à la demande sociale.

1.2. Demandes explicites des partenaires de l'INSEE

1.2.1. Immigration

Le **Haut comité à l'intégration** souhaitait que l'INSEE réalise une nouvelle enquête du type MGIS. Plutôt que de renouveler l'opération à l'identique, on a préféré d'une part étudier la place de l'immigration comme un vecteur d'identité parmi d'autres et d'autre part se donner les moyens de mesurer comment les personnes immigrées associent cet élément de leur identité à d'autres qui sont communs à tous, immigrés ou non.

C'est également au titre de l'étude de l'intégration des populations immigrées que la **Direction de la population et des migrations** du **ministère de l'Emploi et de la solidarité** est associée à la préparation et au financement de l'enquête.

1.2.2. Discriminations

Le ministère l'Emploi et de la solidarité, et tout particulièrement en son sein le **Groupe d'étude et de lutte contre les discriminations (GELD)** et le **Service du droit des femmes et de l'égalité (SDFE)**, sont soucieux de recueillir des données susceptibles de justifier, définir et calibrer une politique de lutte contre les discriminations illégales (liées au sexe ou à l'origine géographique notamment). L'enquête « Histoire de vie » abordera ce thème –sans prétendre le traiter de façon exhaustive– à travers les « assignations identitaires subies » : on demandera à chaque personne si certains aspects de son identité lui valent des moqueries, injustices, évitements, et surtout si ces discriminations (on évitera d'employer le terme lui-même dans l'enquête) ont des conséquences sur son identité, sur ses choix (renoncer à certaines choses ou au contraire chercher à lutter contre ces « traitements inégaux »). On ne prétendra donc pas dans cette enquête mesurer objectivement l'intensité ou la nature de toutes les pratiques discriminatoires, mais analyser les situations où le vécu de la discrimination pèse sur les modalités d'intégration des personnes.

1.2.3. Identités professionnelles

La DARES, au **ministère de l'Emploi et de la solidarité** souhaitent pouvoir analyser l'évolution entre les générations de la part de l'identité professionnelle dans l'identité globale des individus : assiste-t-on, comme on le suppose, chez les jeunes générations, à un rééquilibrage entre les préoccupations liées au travail et les autres, en particulier chez les hommes ? Ce rééquilibrage conduit-il à une réduction de différences entre hommes et femmes ? L'enquête cherchera à répondre à ces questions, même si, comme toute étude transversale, elle ne permettra pas de distinguer les effets d'âge et de génération.

1.2.4. Pratiques culturelles et identité

Le **ministère de la Culture et de la communication** souhaite à travers le volet «activités hors travail» de l'enquête mesurer l'importance des activités de loisirs dans le processus de construction identitaire et l'auto-définition de soi. A ce titre l'enquête viendra enrichir la compréhension des pratiques culturelles des Français et des enjeux qui y sont attachés. Cette interrogation concerne prioritairement les activités à fort contenu identitaire, telles la tenue d'un journal intime, indicateur d'une construction identitaire plutôt introvertie, l'étude de la pratique de la généalogie, qui décrit plutôt une forte identité familiale, ou encore la pratique de chants/danses folkloriques ou la participation à des actions de sauvegarde du patrimoine.

1.2.5. Identités géographiques et langues régionales

On s'interroge souvent sur la «double culture» ou sur des conflits identitaires des populations immigrées. L'enquête se propose d'élargir ces questions aux personnes ayant connu des mobilités au sein du territoire : mobilités inter régionales, relations province / Paris, trajectoires rural / urbain pourront faire l'objet d'études, à partir de questions sur les lieux avec lesquels les personnes restent en contact.

Les « marqueurs de l'identité en acte » dans ce domaine seront les lieux choisis pour les vacances, les lieux envisagés pour la retraite ou une mobilité professionnelle future, ou encore les lieux où l'on envisage de se faire enterrer. Ils incluront la pratique des langues régionales ou étrangères.

Ces domaines sont des centres d'intérêt pour l'**INED** et de la **Délégation générale à la langue française**, dont une des principales missions est d'étudier et de valoriser les « langues de France ».

1.2.6. Santé et handicap

L'**INSERM** considère l'enquête « Histoire de vie » comme le moyen d'approfondir des dimensions peu étudiées à l'échelle des populations et pourtant d'un intérêt réel pour le développement de la Santé Publique. A ce titre, il participe activement à sa préparation.

L'analyse des liens entre des facteurs liés à la santé (type et origine de la maladie ou du handicap, conséquences en terme de perte d'autonomie), des facteurs socio-démographiques (âge, sexe, milieu social, ressources financières), des indicateurs de participation sociale (emploi, occupations hors travail, réseaux sociaux) et des facteurs plus subjectifs en termes d'appartenances revendiquées ou d'identités en actes (demande de reconnaissance sociale d'un handicap, participation associative) permettra d'identifier différents types de trajectoires de personnes malades et/ou handicapées et de mieux comprendre la place qu'elles occupent dans l'espace social. Cette approche innovante en terme d'identité, intégrant une dimension sur la discrimination perçue complète parfaitement les données issues de l'enquête HID en éclairant les modes d'intégration à la société française des personnes malades et handicapées.

1.3. Autres thèmes du débat public

1.3.1. Exclusion

L'une des manifestations de l'exclusion sociale est l'éloignement des groupes sociaux d'appartenance, et donc des vecteurs possibles d'identité : absence d'un groupe professionnel, d'une cellule familiale, d'une communauté amicale fournie et diversifiée. Les études sur l'environnement social des personnes en difficulté économique concluent plutôt à un « cumul de handicaps », les difficultés économiques ayant tendance à être doublées – plutôt que compensées - par un relatif isolement social. Dans ces

conditions, et si un tel lien existe malgré tout, comment ces personnes constituent-elles leur « lien social » ?

1.3.2. Crise des identités

Les thèmes de « crise des identités » nationales, professionnelles, politiques ou religieuses sont souvent évoqués pour décrire ou expliquer l'affaiblissement ou l'abandon de comportements synonymes d'intégration à certains groupes sociaux : baisse de la participation électorale ou attirance pour les votes extrêmes, affaiblissement de la pratique religieuse, du militantisme politique ou syndical. L'objet de l'enquête ne sera pas de décrire – encore moins en tendance – ces phénomènes de désaffection, mais bien de mesurer le plus objectivement possible ce que ces vecteurs « traditionnels » pèsent effectivement aujourd'hui dans la cohésion sociale. Des questions comme la part relative des identités publiques ou privées, de la taille et de la nature du groupe de référence (groupe social ou réseau familial ou électif, identités géographiques locale, départementale, régionale, nationale, continentale) seront abordées.

1.4. Questionnaire

Le questionnaire s'articule autour de 4 grandes parties.

1.4.1. Biographie

Sous la forme d'un tableau comportant des thèmes en colonne et des années en ligne, cette partie a comme objectif de faire décrire à la personne les principales étapes de sa vie en termes démographiques (naissance et décès d'enfants, décès des parents, formation et rupture de couple), résidentiels et professionnels. Des données démographiques et sociales sur les parents seront également collectées.

L'objectif consiste à récolter des informations sur la trajectoire de l'individu au sens large. On lui demandera aussi de déterminer les périodes de sa vie qu'il considère comme particulièrement bonnes ou particulièrement mauvaises, ce qui permettra d'évaluer le poids qu'il donne aux diverses dimensions de sa biographie et de lui faire décrire d'autres éléments centraux de son parcours qui n'auraient pas été cités.

1.4.2. Appartenances objectivées, identité(s) revendiquée(s)

Cette partie est déclinée par thèmes, dont chacun comporte des questions sur les pratiques « objectives » et d'autres sur les identités revendiquées.

Le premier type de questions comporte des indicateurs « classiques » d'appartenances à des groupes sociaux : familiaux, professionnels, associatifs, culturels. Des indicateurs « d'identité en acte » permettront d'approcher la place que ces appartenances jouent dans la vie de la personne.

Le second type de questions permet à la personne d'indiquer ce qui, selon elle, la caractérise, la définit, la différencie au sein de la société. Ces questions sont de nature subjectives et sont clairement présentée comme telles.

La question des engagements politiques, religieux ou éthiques sera abordée. Cependant, il ne s'agira pas de mesurer la répartition dans la population des opinions politiques ou religieuses, mais de mesurer la place que la politique, la religion ou les convictions de façon générale occupent dans la définition de la personne. Ce qui est pertinent pour l'étude, c'est par exemple de noter un fort degré de pratique religieuse ou d'engagement politique, sans forcément connaître la nature de la religion ou l'appartenance politique.

1.4.3. Assignations subies

Le thème de l'identité est vécu par certaines catégories de populations, notamment celles issues de l'immigration, mais aussi, selon les situations, par les femmes ou par les jeunes, sous le mode de l'assignation : « on me donne une identité de »...immigré, femme, jeune, et alors même que je ne revendique pas cette caractéristique, j'en subis les conséquences.

Même si ces « identités assignées » ne sont pas au cœur de nos préoccupations, elles sont abordées car elles ont des conséquences sur les modes d'intégration et sur la construction des identités « positives » (voir plus haut le paragraphe « discriminations »). Après avoir clairement indiqué à la personne interrogée que l'on s'intéresse désormais à la façon dont les autres la définissent (et non plus dont elle se définit), on lui demandera de décrire les situations au cours desquelles elle a eu le sentiment d'être « traitée différemment » pour des raisons liées à ses origines sa conditions sociales son sexe, son âge, *etc.* La description des circonstances des discriminations est abordée, mais on cherchera surtout à connaître l'influence de ce sentiment sur les comportements ultérieurs de la personne.

1.4.4. Conclusion : bilan

A la fin du questionnaire, on demandera à la personne de préciser et de hiérarchiser les « vecteurs de l'identité » qui ont été décrits dans les parties précédentes.

Cette conclusion est très importante, car un rapport à la société se construit par la volonté d'afficher certaines caractéristiques personnelles et de les revendiquer comme étant des façons de se définir, c'est-à-dire aussi de se différencier.

Dans l'exploitation, on pourra utiliser ces données « contrôlées » par celles des parties précédentes : ainsi l'appréciation d'une personne qui revendique un trait identitaire fort alors qu'aucune variable de comportement n'y correspond ne sera pas prise en compte de la même manière que si ce trait correspond à une appartenance « objective »².

1.5. Exploitation et résultats

Les premiers résultats de l'enquête sont prévus pour la fin de l'année 2003.

La première publication, réalisée par l'INSEE, portera sur la diversité et la répartition des modalités « dominantes » de l'identité : pour x% des individus, le principal lien social est la famille, pour x% le groupe professionnel, pour x% un engagement éthique, politique ou religieux, pour x% une activité culturelle ou de loisir. On analysera ces répartitions selon les descripteurs socio-démographiques « classiques », marqueurs identitaires repérés ou construits habituellement par la statistique publique (pays de naissance, nationalité, groupe socioprofessionnel, taille de commune, *etc.*), ainsi que pour d'autres marqueurs présents dans le débat public (militants ou membres d'associations, personnes handicapées, *etc.*).

Parmi les premiers résultats figureront aussi des analyses sur le poids des diverses dimensions dans des identités qui, de toutes façons, sont multiples : les résultats seraient alors plutôt du type : le lieu où l'on a passé son enfance / la région ou le pays d'origine des parents/ les engagements éthiques, politiques ou religieux / les engagements associatifs /... jouent un rôle significatif dans l'identité de x% des individus.

² Exemple cité par un chercheur : quand une personne déclare une identité arménienne alors que ni elle ni ses parents n'y sont nés, n'en ont la « nationalité », ne parlent arménien et n'ont aucune pratique religieuse du culte arménien, on est autorisé à considérer sa déclaration comme une identité purement « idéelle » dans l'analyse de son identité, et à ne pas la considérer pour certaines analyses.

Inversement, la mise en regard des appartenances objectives et revendiquées montrera dans quels cas les premières ne se retrouvent pas dans les secondes.

Des analyses de ce type pourront ensuite être systématisées, croisées entre elles et croisées avec les principaux descripteurs sociaux habituels dans des volumes de tableaux et dans des analyses des données du type ACP pour bâtir des typologies ou des représentations spatiales des constructions identitaires, à l'image des analyses faites, par exemple, à partir de l'enquête éducation.

Chacun des principaux partenaires de l'opération publiera, simultanément avec les premières publications de l'INSEE ou du moins dans des délais assez rapides, des premiers résultats sur les champs particuliers qui motivent leur participation à l'enquête. La cellule « statistiques et études sur l'immigration » de l'INSEE pourra également publier rapidement des résultats particulièrement centrés sur les populations immigrées.

Les exploitations « secondaires » seront ensuite réalisées à l'INSEE ou chez les partenaires de l'opération. Les différentes administrations parti-prenantes réaliseront les études nécessaires pour éclairer leurs activités. L'INED, le GELD et l'INSERM envisagent d'ores et déjà de mobiliser des chercheurs pour exploiter la source. Dans la mesure du possible (contraintes légales, politique de diffusion de l'INSEE, intérêt de la communauté des chercheurs), on facilitera l'exploitation de l'enquête par des chercheurs extérieurs à l'INSEE dans le cadre du groupe d'exploitation³.

2. Les apports en termes d'amélioration de la qualité de l'enquête finale

2.1 Apports sur la méthode d'échantillonnage

L'opération pilote nous a permis de tester la méthode de sondage de l'enquête.

La particularité de l'enquête « Construction des identités » en termes d'échantillonnage est l'utilisation d'une base de sondage individuelle issue d'échantillons d'enquêtes - les enquêtes Histoire Familiale (EHF) et Vie Quotidienne et Santé (VQS). L'échantillon n'est pas tiré dans l'échantillon maître (EM), comme c'est le cas habituellement pour les enquêtes auprès des ménages (hors Enquête Emploi en Continu).

Le recours à cette base de sondage était justifié par la nécessité de procéder à des sur représentations de certaines populations (les personnes nées à l'étranger, celles de parents nés à l'étranger et les personnes handicapées de moins de 60 ans). En effet ces catégories d'individus présentent un intérêt particulier en termes d'identité et doivent pouvoir faire l'objet d'études spécifiques dans le cadre de cette enquête. Or, ces groupes sont faiblement représentés dans la population générale. Il était donc nécessaire de les sur représenter afin d'améliorer la précision des analyses les concernant. Les informations nécessaires à cette sur représentation sont absentes du recensement de la population. Il fallait donc recourir à une autre base de sondage.

Les enquêtes Histoire Familiale (EHF) et Vie Quotidienne et Santé (VQS) ont porté sur environ un dixième de la population recensée et comprenaient des informations sur les lieux de naissance de l'individu et de ses parents (pour EHF) et le degré de handicap (pour VQS). Elles pouvaient donc constituer une base de sondage pour l'enquête Construction des identités.

Un des objectifs du test d'exploitation était de tester la base de sondage et l'échantillonnage, en essayant de mesurer l'efficacité des sur représentations de population attendues, de mettre en avant de possibles distorsions dues entre autres au mode de sélection de la personne interrogée et enfin de déceler d'éventuels problèmes de collecte liés à la base de sondage.

³ Les réunions du groupe d'exploitation sont ouvertes à tous. La participation aux travaux est soumise à l'approbation d'un projet de recherche. Pour tout renseignement, contacter Frédérique Houseaux 01.41.17.67.60 ou frederique.houseaux@insee.fr.

2.1.1 La méthode de sondage de l'opération pilote

Pour l'opération pilote d'avril-mai 2002, un échantillon d'un peu plus que 1300 fiches-adresses a été tiré uniquement parmi l'ensemble des individus répondants de l'enquête EHF et parmi les personnes limitées dans leurs activités pour des raisons de santé de moins de 60 ans ayant répondu à VQS⁴. Aucun logement n'a été tiré dans la BSLN, faute de base à disposition dans les zones hors EM.

On recherchait un individu dans chaque logement tiré. Cet individu était identifiable par son sexe son mois et son année de naissance. Si l'individu était retrouvé parmi les membres du ménage, il était interrogé. Sinon, on tirait un individu au hasard dans le ménage parmi les personnes de 18 ans et plus. On choisissait la personne dont la date d'anniversaire était la plus proche à partir de la date de l'enquête.

Les probabilités d'inclusion des personnes nées à l'étranger, de parent étranger ou jeunes handicapés ont été doublées.

2.1.2 Bilan de l'opération pilote

2.1.2.1 Efficacité de la sur représentation de catégories d'individus :

Un des objectifs de l'opération pilote était de vérifier que les sur représentations, effectuées lors du tirage de l'échantillon, se retrouvaient bien dans la population des répondants à l'enquête.

L'apport du test d'exploitation s'est avéré très important sur ce point. Il a tout d'abord permis de mettre en avant des problèmes d'ordre informatique qui ont annulé la sur représentation des « jeunes handicapés ». Sans la possibilité de procéder à des analyses statistiques de base sur ce point, nous aurions pu passer à côté de ces difficultés et renouveler l'erreur pour l'enquête, ce qui aurait été très dommageable. Le test d'exploitation a ensuite permis de confirmer la fiabilité de la méthode de sur représentation des immigrés et enfants d'immigrés, autre gros enjeux de la méthode de sondage de l'enquête.

Les « jeunes handicapés » :

Le nombre de personnes « jeunes handicapées » recherchées (19) est très inférieur à celui prévu au moment du tirage (80) (*tableau 1*). On peut donc presque considérer que la sur représentation de cette catégorie n'a pas été effectuée. Ceci n'est pas dû à la base de sondage, mais à une difficulté au moment du tirage des individus de l'échantillon qui a été identifiée a posteriori. On peut donc raisonnablement penser que la sur représentation sera effectué lors du tirage de l'enquête de 2003.

Cependant, malgré l'absence de sur représentation, le nombre et la proportion de « jeunes handicapés » parmi les répondants, au regard des variables de l'enquête, sont très supérieurs à ceux attendus. Une fois que l'on tient compte des probabilités d'inclusion, leur proportion, dans les 5 régions concernées, correspond au double de celle que l'on constate sur une population générale.

On peut donc se demander si la variable qui sert à repérer les « jeunes handicapés » dans l'enquête « construction des identités » ne serait pas comprise de manière différente par les enquêtés de celle de VQS qui sert à sélectionner les jeunes handicapés au moment de la constitution de la base de sondage. Pourtant la question de VQS n'est que légèrement différente mais la méthode d'enquête n'est pas comparable - il s'agissait alors d'une enquête auto-administrée.

On peut noter aussi que presque tous les handicapés de moins de 60 ans recherchés ont été retrouvés dans le ménage auquel ils appartenaient en 1999. Cette population semble donc peu mobile ce qui est plutôt rassurant quant aux chances de maintenir la sur représentations lors de l'enquête de 2003.

⁴ à l'exception des individus habitant dans des logements présents dans l'échantillon maître et dans les aires emploi et de ceux déjà interrogés dans l'enquête HID « ménages ».

Les immigrés en enfants d'immigrés :

Concernant les immigrés, sur les 215 recherchés, 70 ont été effectivement retrouvés dans le ménage auquel ils appartenaient en 1999. Cette catégorie semble assez mobile, mais la perte en proportion est finalement plutôt faible sur cette catégorie. En effet, ils étaient 19,8% dans l'échantillon et ils sont 16,6 % dans l'ensemble des répondants d'après les variables de l'enquête (*tableau Ibis*). De plus, si l'on tient compte des pondérations, on obtient 11% d'immigrés, contre 9% dans une population générale, ce qui est plutôt satisfaisant.

La perte des individus enfants d'immigrés entre le tirage de l'échantillon et la population des répondants est un peu plus importante que pour les immigrés, mais cette perte est compensée par les pondérations, puisqu'on obtient presque 14% d'enfants d'immigrés, contre 11% dans la population générale.

Ces résultats concernant les populations immigrées ou issues de l'immigration sont d'autant plus satisfaisants que le taux de réponse de ces populations est plus faible. En revanche, les motifs de non-réponse sont un peu différents. Pour les immigrés, il s'agit plus souvent d'enquêtes réalisées partiellement, de résidences occasionnelles ou secondaires ou de hors champs, alors que pour les enfants d'immigrés, il s'agit surtout de refus et d'impossibles à joindre. On peut penser qu'il ne s'agit pas du même type d'immigration.

Il semble que les DR choisies aient un effet sur ces résultats, car elles concentrent plus d'immigrés et d'enfants d'immigrés que les autres régions de France métropolitaine.

Enfin, le mode de sélection de l'individu interrogé influence assez peu les résultats sur les proportions d'individus immigrés, issus de l'immigration et « jeunes handicapés ». Aucun biais n'est particulièrement à constater.

Tableau 1 : Effectifs des différentes catégories selon la population

	Nés à l'étranger	De parent né à l'étranger	Handicapé	Autres	Total
	eff.	eff.	eff.	eff.	eff.
HORS LORRAINE ¹					
D'après l'échantillon tiré					
- Catégories d'individus recherchées parmi les répondants à l'enquête	113	115	10	381	619
D'après les variables de l'enquête « identités »					
- Ensemble des répondants	103	90	84	364	619
- Ménages identiques / RP99	93	74	72	305	525
dont :					
<i>individu retrouvé</i>	70	63	57	244	419
<i>individu non retrouvé</i>	23	11	15	61	106
- Ménages différent / RP99	10	16	12	59	94
- Individu kish (<i>ind. non retrouvé + ménage différent</i>)	33	27	27	120	200
Y COMPRIS LA LORRAINE					
D'après l'échantillon tiré, annoncé par l'UMS	273	285	80	805	1443
D'après les variables de l'enquête « identités »					
- Ensemble des répondants yc Lorraine	129	128	105	453	787
- Individu kish yc Lorraine	59	65	48	209	368

1 : Suite à un problème informatique, la variable qui sert à identifier la catégorie à laquelle appartient l'individu recherché est à blanc en Lorraine. Ces résultats sont présentés pour la majorité sur une population ne comprenant que les 4 autres DR ayant réalisé l'opération pilote. Les résultats comprenant la Lorraine ont surtout pour but de montrer la composition de l'échantillon d'enquête dont nous disposons effectivement.

Tableau Ibis : Proportions des différentes catégories selon la population :

	Nés à l'étranger	De parent né à l'étranger	Jeunes handicapés	Autres	Total
	%	%	%	%	Eff.
HORS LORRAINE ¹					
D'après l'échantillon tiré					
- Catégories d'individus recherchées parmi les répondants à l'enquête	18	19	2	62	619
D'après les variables de l'enquête « identités »					
- Ensemble des répondants	17	15	14	59	619
- Ménages identiques /RP99	18	14	14	58	525
dont :					
<i>individu retrouvé</i>	17	15	14	58	419
<i>individu non retrouvé</i>	22	10	14	58	106
- Ménages différent / RP99	11	17	13	63	94
- Individu kish (<i>ind. non retrouvé + ménage différent</i>)	17	14	14	60	200
- Ensemble des répondants (pondéré)	11	13	8	70	
Y COMPRIS LA LORRAINE					
D'après l'échantillon tiré, annoncé par l'UMS	19	20	6	56	1443
D'après les variables de l'enquête « identités »					
- Ensemble des répondants yc Lorraine	16	16	13	58	787
- Ensemble des répondants yc Lorraine (pondéré)	11	14	8	69	787
- Individu kish yc Lorraine	16	18	13	57	368
- Individu kish yc Lorraine (pondéré)	12	17	8	65	368
Population générale	9	11	4	76	

1 : Suite à un problème informatique, la variable qui sert à identifier la catégorie à laquelle appartient l'individu recherché est à blanc en Lorraine. Ces résultats sont présentés pour la majorité sur une population ne comprenant que les 4 autres DR ayant réalisé l'opération pilote. Les résultats comprenant la Lorraine ont surtout pour but de montrer la composition de l'échantillon d'enquête dont nous disposons effectivement.

2.1.2.2 Distorsions sur certaines catégories d'individus.

Le test d'exploitation a permis de mettre en avant des distorsions de l'échantillon principalement liées au mode de sélection de la personne interrogée dans le ménage, c'est-à-dire au fait que l'on cherche un individu précis dans un ménage.

Les marges de l'échantillon tiré montrent une sur représentation des personnes de plus de 60 ans et une sous représentation des moins de 30 ans. Ceci est lié à la composition de la base de sondage qui est légèrement plus âgée que la population générale. Cette distorsion est cependant encore plus importante parmi les répondants au test. (tableau 2)

Ceci est principalement dû au mode de sélection des personnes interrogées. En effet la population des personnes qui ont été retrouvées dans le ménage et interrogées (environ 53% de la population enquêtée) est beaucoup plus âgée que celle d'une population globale et encore plus que celle de l'échantillon de départ.⁵

⁵ Notons tout de même, qu'un problème informatique, a modifié les variables qui permettaient de retrouver les personnes tirées pour la DR de Lorraine. Ainsi, le pourcentage de personnes retrouvées devrait normalement être plus élevé.

Tableau 2 : Comparaison de la répartition par âge des répondants et de la population générale

	Répondants au test identités *			Estimation de pop. 1 ^{er} janv. 2002**	
	Retrouvés de EHF ou VQS	Tirage kish	ensemble	(18 ans et +)	(21 ans et +)
Age	%	%	%	%	%
moins 31 ans	4	18	11	21	17
31-40 ans	16	21	18	19	20
41 à 50 ans	21	20	20	18	19
51 à 60 ans	21	14	18	16	17
plus 60 ans	38	26	32	26	27

* pondération tenant compte uniquement des probabilités d'inclusion et du taux de réponse

** population des ménages ordinaires en France métropolitaine

On constate par ailleurs, qu'alors que 53% des personnes enquêtées ont été retrouvées dans le logement enquêté, seuls 8% des moins de 30 ans ont été retrouvés, contre plus de 80% des plus de 50 ans, avec une légère baisse autour de 60 ans.

On peut envisager plusieurs explications de ce biais :

La population des jeunes est plus mobile. On a donc plus de mal à la retrouver. Ils peuvent constituer un ménage qui a déménagé depuis 1999 ou avoir quitté le ménage dans lequel ils vivaient en 1999.

Il y a aussi plus de moins de 30 ans parmi les refus, les impossibles à joindre et les enquêtes réalisées partiellement que dans une population générale. N'ayant pas beaucoup évolué encore, ils ne voient peut-être pas l'intérêt de répondre à une telle enquête où on va leur demander de raconter leur vie. Ceci a pu augmenter les déchets, mais aussi les refus de répondre de la personne interrogée, même si quelqu'un d'autre du ménage avait accepté l'enquête. Des enquêteurs ont d'ailleurs signalé que dans les cas d'enfants vivant chez leurs parents, très souvent les parents acceptent de répondre mais pas l'enfant lorsqu'il est interrogé. Dans ces cas les parents ne comprenaient d'ailleurs pas pourquoi ils ne pouvaient pas répondre eux-même.

Enfin, la population retrouvée avait 18 ans et plus en 1999. Elle a donc vieilli de 3 ans depuis. La tranche des moins de 30 ans ne comprend donc plus d'individu de 18 à 20 ans et a été diminuée des individus de 27 à 29 ans. Il y a donc un effet mécanique qui diminue la part des plus jeunes et augmente celle des plus âgés. Seulement, lorsqu'on compare avec une population générale de 21 ans et plus, la sur représentation des moins de 30 ans diminue, mais reste très importante (voir tableau ci-dessus).

En revanche, la population des enquêtés choisis aléatoirement parmi les membres du ménage est assez proche d'une population générale. Heureusement elle représente presque la moitié de la population enquêtée, ce qui diminue l'effet de vieillissement induit par l'autre partie de la population. Mais si la DR de Lorraine avait eu son échantillon comme les autres DR, les conséquences du biais seraient encore plus importantes.

Ce biais de sélection a bien sûr aussi des conséquences sur les proportions de personnes en fonction de leur situation par rapport à l'emploi. Les retraités, retirés des affaires et les femmes au foyer sont plus nombreux que dans une population générale.

Tableau 3: Comparaison de la répartition par type d'occupation des répondants et de la population générale :

	Répondants au test identités *			EPCV oct 2001	
	Retrouvés de EHF ou VQS	Tirage kish	ensemble	(18 ans et +)	(21 ans et +)
Situation professionnelle	%	%	%	%	%
actifs occupés	44	51	48	53	54
chômeurs	5	4	5	7	6
étudiants	1	8	4	6	3
retraités	31	17	25	20	21
Retirés	7	4	6	4	4
femmes foyer	9	10	10	7	8
autres inactifs	2	5	3	4	4

*pondération tenant compte uniquement des probabilités d'inclusion et du taux de réponse

2.1.2.3 Difficultés de collecte :

Enfin, on cherchait aussi à étudier les effets de l'utilisation de cette base de sondage en termes de collecte.

Les principales difficultés rencontrées par les enquêteurs sont liées à la forte dispersion géographique de l'échantillon tiré.

Cette dispersion de l'échantillon est due à la composition de la base de sondage (enquêtes Famille et VQS) qui comprend elle-même des zones assez dispersées. Elle s'explique aussi par le fait que l'on a tiré l'échantillon en une seule phase, sans tenir compte de la possibilité qu'apporte l'enquête Famille de regroupements des individus au sein d'un nombre limité de zones de délégués du RP99.

Plusieurs dizaines de logements n'ont d'ailleurs pas été enquêtés faute d'enquêteur pour les réaliser. Il était difficilement envisageable de recruter un enquêteur pour 4 ou 5 FA.

Les longues distances entre les logements à enquêter ont compliqué le travail des enquêteurs. Les temps de déplacement étaient importants pour un nombre de FA parfois assez faible. Cela a donc souvent empêché les enquêteurs de réaliser plusieurs enquêtes le même jour.

Par ailleurs, pour la plupart des enquêteurs le fait de devoir se déplacer loin a impliqué plus de refus. En effet, les prises de contact en face à face étaient plus difficiles. Ils ont dû se replier sur le téléphone ou le courrier, situations dans lesquelles les ménages ont plus de facilité à refuser qu'en face d'une personne qui vient les voir. (voir note de compte rendu des réunions de bilan n°87/F170)

Lors des formations certains enquêteurs ont fortement réagi sur ce point car pour les déplacements, seuls les frais de kilométrage sont pris en compte et non le temps que cela leur prend. Cette question n'est pas propre à cette enquête, mais le problème est d'autant plus important que la dispersion géographique de l'échantillon est forte.

2.1.3 La méthode de sondage adoptée pour l'enquête

L'opération pilote a permis de relever certaines limites de la méthode de sondage adoptée. Suite à cette opération, une réflexion a été menée sur les modifications à introduire pour le tirage de l'échantillon de l'enquête de 2003.

Le plan de sondage de l'enquête vise à obtenir 8 000 individus répondants, ce qui conduit à un tirage de 13 500 individus environ, si on tient compte du taux de réponse obtenu lors de l'opération pilote.

Pour faire face à la dispersion géographique de l'échantillon, il a été décidé d'effectuer un tirage en deux phases. La première phase déterminera les zones de délégué au sein des strates des enquêtes EHF et VQS et la deuxième des individus au sein de cette zone. Cela permet d'avoir des enquêtes concentrées sur certaines zones de délégué pour limiter les déplacements des enquêteurs. De plus, pour les zones dont la superficie est supérieure à 40km², des communes géographiquement proches ont été sélectionnées avant le tirage des individus.

La base de sondage intermédiaire est constituée par tous les individus recensés dans les zones de délégués tirées pour les enquêtes VQS et EHF. Les individus des logements de l'échantillon maître et ceux des aires Emploi ont été enlevés au préalable de cette base. Ces zones constituent des unités primaires de 15 000 individus environ.

La base de sondage est une base d'individus, et les objectifs en terme de taille d'échantillon sont formulés par groupe d'individus.

Aussi, dans un premier temps, on a considéré l'ensemble des individus de chacune des zones de délégués sélectionnées précédemment et on a partitionné les zones en distinguant les 4 groupes suivants :

- Personnes handicapées de moins de 60 ans (tirées exactement et exclusivement parmi les individus de l'échantillon VQS)
- Immigrés (tirés exactement et exclusivement parmi les individus de l'échantillon EHF)
- Enfants d'immigrés (tirés exactement et exclusivement parmi les individus de l'échantillon EHF)
- Autres personnes

Les « autres personnes » constituent l'intégralité du complément dans la zone de délégué - mais pour des raisons liées au calcul des pondérations, on a éliminé au préalable de ce groupe tous les logements où se trouvent des individus ayant répondu à VQS et à l'EHF.

Le recours à l'ensemble de la zone de délégué évite un effet de grappe massif (le tirage des secteurs d'agents recenseurs est adapté pour de très grosses enquêtes telles EHF et VQS, mais beaucoup moins pour des enquêtes de l'ordre de 10 000 personnes).

Enfin on a tiré, dans chacun des 4 groupes, un nombre convenu d'individus par un tirage systématique par unité primaire. Cette phase a permis de contrôler les tailles d'échantillon obtenues dans chacun des différents groupes.

L'échantillonnage fait en sorte que les proportions d'individus par sexe, par tranche d'âge et par strate géographique soient voisines de celles qui prévalent dans la population.⁶

Par ailleurs, un échantillonnage aléatoire simple de logements neufs a été pratiqué dans les communes de la zone de délégué tirée.

2.2 Apports sur le contenu du questionnaire : l'exemple des questions sur les convictions.

2.2.1.1 Les objectifs de l'utilisation d'un test d'exploitation pour le contenu du questionnaire

Un des principaux objectifs de l'opération pilote était de tester de manière statistique le contenu du questionnaire de l'enquête « Construction des identités ».

⁶ Cette précaution était nécessaire pour corriger les particularités de l'échantillon EHF qui sur représente les femmes et certains départements où ont eu lieu des extensions sur les langues régionales.

Dans les tests plus « classiques » le questionnaire est le plus souvent testé en termes de passation et de qualité de l'entretien. La très petite taille des échantillons ne permet que rarement de le tester en termes de pertinence ou d'utilité statistique. On cherche plutôt à savoir le plus souvent si les questions sont claires pour les enquêtés, s'il n'y a pas de difficultés de compréhension, si les enquêteurs sont obligés de reformuler certaines questions ou modalités de réponse. On peut aussi tester l'enchaînement des thèmes et des questions, la pertinence des questions, aux yeux des enquêteurs et des enquêtés. Ces tests font surtout appel à l'avis et à l'expérience des enquêteurs. C'est principalement la confrontation avec « le terrain » que l'on cherche.

Même si cette forme de tests est incontournable et le plus souvent riche en enseignements, elle ne permet que rarement d'en tirer des conclusions d'ordre statistique.

Bien entendu, le test d'exploitation de l'enquête « construction des identités » a apporté des informations très utiles sur les difficultés de passation du questionnaire. Des améliorations ont pu être apportées à la seule vue des nombreuses remarques faites par les enquêteurs lors des réunions de bilan. Mais, ce test portant sur 1300 fiches-adresses et près de 800 répondants, les possibilités d'analyse étaient plus vastes. Des modifications du questionnaire ont pu être apportées suite à ces analyses statistiques, alors qu'elle n'auraient pas été envisagées sans cela.

Les principales questions auxquelles nous avons tenté de répondre dans le but d'améliorer le questionnaire étaient :

L'information statistique qu'apportent les questions posées est-elle utilisable compte tenu des objectifs de l'enquête ?

Certaines questions sont-elles redondantes avec d'autres informations déjà recueillies dans l'enquête ?

Pourrait-on fermer certaines questions ouvertes ?

Grâce aux études réalisées sur le fichier du test d'exploitation, des questions ont été modifiées, d'autres rajoutées et enfin certaines ont été supprimées. Une grande partie des questions ouvertes (questions totalement ouvertes et précision en clair de modalités « autre ») ont pu être transformées en questions fermées ou des modalités de réponse ont pu être affinées. La qualité du questionnaire a été améliorée et le temps de passage, jugé trop long lors des tests successifs, a pu être un peu réduit.

2.2.1.2 Les évolutions du questionnaire sur la partie « convictions »

Les études réalisées dans le but d'améliorer le questionnaire suite au test d'exploitation ont porté sur l'ensemble du questionnaire. Nous allons cependant présenter ici plus particulièrement les modifications apportées à la partie « convictions » de l'enquête. Compte tenu de la sensibilité de certaines de ces questions il nous a semblé nécessaire de leur apporter une attention particulière afin d'être certains de leur pertinence et de leur utilité. Mais surtout les apports du test d'exploitation sur ce thème sont assez représentatifs des possibilités que laissent un test d'exploitation sur le contenu du questionnaire.

Les analyses de cette partie du questionnaire ont permis de relever une redondance entre plusieurs questions, de confirmer l'utilité d'autres et de démontrer les possibilités d'exploitation d'une des questions ouvertes.

Les résultats présentés ici sont pondérés pour être représentatifs d'une population générale. Cette pondération tient compte des différentes probabilités d'inclusion des individus, du taux de réponses et les distorsions en termes d'âge et de sexe ont été aussi prises en compte.

2.2.2.1 Mariage religieux et sentiment religieux :

81% des personnes mariées se sont mariées religieusement, alors que seulement 74% expriment au moins un sentiment d'appartenance à une religion. (tableau 4)

L'intérêt de ces questions sur le mariage pourrait être de montrer qu'il n'y a pas de corrélation directe entre conviction religieuse et mariage religieux. Ce qui justifierait que le mariage religieux soit considéré comme un bon marqueur religieux. Or, on remarque très nettement que les réponses données aux questions sur les causes du mariage religieux et le sentiment religieux sont très liées.

Tableau 4 : Mariage religieux et sentiment religieux

% en ligne	Raison principale du mariage religieux					
	Conviction	Tradition	Cérémonie	Conjoint	Famille	Autre
Position par rapport à la religion						
Pratique régulière	79	19	3	0	0	0
Pratique occasionnelle	54	38	2	1	5	0
Sentiment d'appartenance	26	59	4	2	6	3
Ni pratique, ni sentiment d'appartenance	1	52	1	19	26	1
Rejet	2	66	0	0	32	0
NSP	40	0	0	0	60	0
Total	43	43	2	4	8	1

Champ : individus de 18 ans et plus.

Source : opération pilote de l'enquête « construction des identités »

Lorsqu'il n'y a pas de sentiment d'appartenance, voire même un rejet, le mariage religieux s'explique presque exclusivement par la tradition, l'envie de faire plaisir à la famille ou à son conjoint. Les questions sur le mariage religieux n'apportent donc pas d'information complémentaire sur le rapport à la religion de la personne.⁷

De plus, la question sur le sentiment religieux semble bien fonctionner. Les refus de répondre sont très faibles et la modalité «un rejet de la religion» semble apporter une amélioration nette de cette question. En effet, si on compare avec la même question à la seule différence de cette modalité, posée dans l'enquête PCV d'octobre 2001, on peut constater que les réponses pour les 3 premières modalités sont très proches, alors que la réponse « sans sentiment ni pratique » diminue au bénéfice des rejets et que les refus de répondre passent de 2% à 0,6%.

Compte tenu des améliorations apportées à cette question et de l'ensemble des informations qu'elle fournit, il a été décidé de supprimer les questions sur le mariage religieux.

⁷ Il faut noter cependant, que le terme « tradition » n'a semble-t-il pas toujours été compris de la même manière. En effet les personnes ayant une pratique religieuse, régulière ou occasionnelle ont pu l'entendre comme une tradition religieuse à respecter, ce qui relève aussi des convictions religieuses, alors que des personnes non pratiquantes ou sans aucune conviction religieuse ont pu l'interpréter comme une tradition sociale, sans lien avec la religion.

2.2.2.2 Proximité d'un parti politique, d'un mouvement ou d'une cause politique

Près de la moitié des individus interrogés déclarent se sentir proche d'un parti, d'un mouvement ou d'une cause politique. Une personne sur trois se déclare proche d'un parti politique et parmi les autres, une sur cinq déclare se sentir proche d'un mouvement ou d'une cause politique.

Les caractéristiques des personnes proches d'un parti politique ne sont pas tout à fait les mêmes que celles des individus proches d'un mouvement ou d'une cause, hors d'un parti. La proximité des partis concerne plus souvent les plus de 50 ans donc surtout les retraités et retirés des affaires. Elle concerne aussi un peu plus souvent les hommes. La proximité des autres mouvements et causes concentre plutôt des actifs, des étudiants (même si on y trouve également des retraités), des personnes entre 40 et 50 ans et autant de femmes que d'hommes.

Il n'y a qu'en terme de CS que les deux formes de proximité de la sphère politique se ressemblent. On rencontre plus souvent des cadres, des professions libérales et des professions intermédiaires dans les deux cas, et très peu d'ouvriers ou d'employés.

2.2.2.3 Transmission des convictions aux enfants :

Transmission des convictions religieuses :

Seuls 39% des individus ayant des enfants souhaitent leur transmettre leurs convictions religieuses.

Bien entendu, la volonté de transmission est assez liée à l'existence de convictions fortes. Plus la pratique religieuse est importante, plus la volonté de transmettre aux enfants est fréquente.

Tableau 5 : Transmission familiale des convictions religieuses et sentiment religieux

% en ligne	Volonté de transmission des convictions religieuses	
	Oui	Non
Position par rapport à la religion		
Pratique régulière	78	22
Pratique occasionnelle	57	43
Sentiment d'appartenance	28	72
Ni pratique, ni sentiment d'appartenance	2	98
Rejet	3	97
NSP	16	84
Total	39	61

Champ : individus de 18 ans et plus.

Source : opération pilote de l'enquête « construction des identités »

Seulement, on peut remarquer que, parmi les personnes ayant un sentiment d'appartenance à une religion mais aucune pratique, on rencontre plus souvent de personnes ne tenant pas particulièrement à transmettre leurs convictions à leurs enfants. Près de 30% de ces personnes le souhaiteraient tout de même.

Cette question pourrait donc être un bon indicateur d'identité religieuse pour cette catégorie qui représente tout de même plus du tiers des personnes de 18 ans et plus. Elle permet d'affiner l'information obtenue dans la question sur les convictions religieuses en distinguant le sentiment d'appartenance à une religion selon qu'il s'accompagne ou non d'une volonté de transmission aux enfants.

Transmission des convictions politiques :

Seuls 17% des individus ayant des enfants souhaitent leur transmettre leurs opinions politiques. (tableau 6)

Tableau 6 : Transmission familiale des convictions politiques et proximité d'un parti politique

% en ligne	Volonté de transmission des opinions politiques	
	Oui	Non
Proximité d'un parti politique		
Oui	33	67
Non	7	93
Total	17	83

Champ : individus de 18 ans et plus.

Source : opération pilote de l'enquête « construction des identités »

Tableau 7 : Transmission familiale des convictions politiques et proximité d'un mouvement hors parti

% en ligne	Volonté de transmission des opinions politiques	
	Oui	Non
Proximité d'une cause ou d'un mouvement politiques hors parti		
Oui	18	82
Non	4	96
NSP	13	87
Préfère ne pas répondre	0	100
Total	7	93

Champ : individus de 18 ans et plus.

Source : opération pilote de l'enquête « construction des identités »

Parmi les personnes attachant de l'importance au fait que leurs enfants partagent les mêmes opinions qu'eux on rencontre plus souvent de personnes proches d'un parti politique ou d'un mouvement ou une cause politique. Inversement, une très grande majorité des personnes ne se sentant pas proches d'un parti politique n'accordent pas d'importance au fait de transmettre leurs opinions à leurs enfants. Cependant, deux tiers des personnes proches d'un parti politique n'y accordent pas particulièrement d'importance non plus.

On pourrait donc considérer grâce à cette question deux niveaux d'attachement politique selon que la personne exprime une volonté de transmission familiale ou pas. Ces questions ont par conséquent été maintenues pour l'enquête.

Le même résultat apparaît concernant la proximité d'un mouvement ou d'une cause politique hors parti, mais le nombre de personnes concernées est très faible.

2.2.2.4 Appartenance à une classe sociale :

Il était demandé à chaque personne enquêtée si elle avait le sentiment d'appartenir à une classe sociale. Seules 8 personnes n'ont pas su répondre à cette question.

Une personne sur deux environ a déclaré avoir le sentiment d'appartenir à une classe sociale. (Tableau 8) Et parmi elles, seulement 9 personnes n'ont pas su définir de quelle classe sociale il s'agissait. Ce qui tend à penser que la question a un sens pour les enquêtés.

Tableau 8 : Sentiment d'appartenance a une classe sociale

	%	% cumulé
Oui	48	48
Non	51	99
NSP	1	100,00

Champ : individus de 18 ans et plus.

Source : opération pilote de l'enquête « construction des identités »

A âge, sexe, occupation professionnelle, catégorie socio professionnelle et revenus égaux, la seule chose qui caractérise vraiment les personnes ayant déclaré avoir le sentiment d'appartenir à une classe sociale est le fait de se sentir plus souvent proches d'un parti, d'un mouvement ou d'une cause politique. D'autre part, toutes choses égales par ailleurs, les retraités sont un peu moins nombreux que les autres à avoir ce sentiment d'appartenance à une classe sociale. Sinon, les répondants sont aussi nombreux à répondre oui ou non, quel que soit leur âge, leur sexe, leur catégorie socio professionnelle, leur occupation et le niveau de revenus de leur ménage.

Il était ensuite demandé aux personnes ayant le sentiment d'appartenir à une classe sociale, de quelle classe il s'agissait. Les réponses étaient données en clair.

A première vue, les réponses à la question sur le nom de cette classe sociale sont assez attendues et convenues. Elles expriment le plus souvent une catégorie socio professionnelle ou une profession ou un niveau économique. Pourtant elle apporte beaucoup d'information supplémentaire.

La première constatation que l'on peut faire en regardant la liste des réponses données est que la classe moyenne regroupe un nombre important d'individus : 151 personnes, soit près de 40% des personnes ayant répondu à cette question, soit encore une personne de 18 ans et plus sur cinq.

Même si cette réponse est assez attendue, elle ne correspond pas réellement à une catégorie socio professionnelle ni au niveau de revenu du ménage.

Même si on y rencontre principalement des personnes exerçant des professions intermédiaires et des employés, les professions intellectuelles, chefs d'entreprise, professions libérales et cadres supérieurs y sont également bien représentés. On retrouve aussi parmi ces personnes se déclarant de classe moyenne une majorité de personnes appartenant aux ménages au-dessus de la médiane des revenus. 20% d'entre eux font même partie du quart des ménages les plus riches. On ne peut donc pas non plus dire que la classe sociale déclarée est représentative du revenu du ménage et exprimerait un niveau économique.

De même, près de 30% des individus déclarant avoir le sentiment d'appartenir à une classe sociale, disent faire partie de la classe ouvrière. Or, seuls 44% de ces personnes sont effectivement ouvrières. Un quart semble s'identifier plus à la catégorie socio professionnelle de leur père qui était ouvrier alors qu'eux ne le sont pas et pour 5% d'entre eux, c'est la mère qui était ouvrière. Une personne sur dix se réfère à la catégorie socio professionnelle de son conjoint. Enfin, tout de même plus de 16% d'entre eux n'ont aucun lien avec la classe ouvrière dans leur entourage direct.

Tableau 9: Lien avec le milieu ouvrier parmi les personnes déclarant appartenir à la classe ouvrière

Personnes déclarant appartenir à la classe ouvrière	Effectif	%	Effectif cumulé	% cumulé
Etant elles-mêmes ouvrières	46	44	46	44
N'étant pas ouvrières, mais...				
...dont le père au moins était ouvrier	26	25	72	69
...dont seule ma mère était ouvrière	5	5	77	74
...dont le conjoint est ouvrier, mais pas les parents	11	10	88	84
...dont ni les parents ni le conjoint ne sont ouvriers	17	16	105	100.00

Champ : individus de 18 ans et plus.

Source : opération pilote de l'enquête « construction des identités »

Il semble aussi que ces différents types de références (ego, parents, conjoint, autre) varient avec l'âge et selon la catégorie socio professionnelle de la personne lorsqu'elle n'est pas ouvrière elle-même. Cependant, le faible nombre de personnes concernées ne permet pas de considérer ces effets comme significatifs. Il serait intéressant de vérifier s'ils se confirment sur une plus grande population.

D'autre part, en dehors des ouvriers, un très faible nombre d'individus s'identifient à une catégorie socio professionnelle. La deuxième référence serait plutôt le niveau de vie que la catégorie socio professionnelle. En effet, les autres individus se situent plutôt par rapport à ce qu'ils estiment être une moyenne. C'est ainsi que l'on peut distinguer un groupe de personnes se déclarant défavorisées ou pauvres et un autre se considérant plutôt aisées, privilégiées ou supérieures. Ces déclarations sont dans la majorité des cas assez proche du niveau de revenus du ménage et de sa place dans la distribution des revenus et concernent un nombre non négligeable de réponses (près de 6% des individus interrogés dans l'enquête et presque 12% des individus déclarant une classe sociale).

Même en se limitant à ces exemples, on constate que cette question nous apporte une information identitaire beaucoup plus fine que les indicateurs plus « classiques » comme la catégorie socio professionnelle et le revenu.

De plus, le codage des classes sociales est assez aisé. On arrive assez facilement à dégager une quinzaine de codes regroupant la majorité des intitulés (369 individus sur 396 réponses). Ce codage pourrait certainement être encore affiné avec une analyse textuelle sur une plus grande population. Le test d'exploitation a donc permis de montrer l'utilité de ces questions et d'affirmer les possibilités d'exploitations qu'elles apportaient. Ces questions ont par conséquent été maintenues.

3. Des pistes d'exploitation de l'enquête « Histoire de vie » à travers l'analyse des questions ouvertes

3.1. Les lieux de l'identité

Le parcours géographique des enquêtés constitue un facteur important dans la construction des identités individuelles. L'analyse de cette trajectoire en relation avec les espaces de vie, les lieux d'attachement ou encore les lieux de projet permet d'approcher la notion d'identité territoriale des populations. Une partie du questionnaire de l'enquête « Histoire de vie » est consacrée au recueil des lieux qui comptent pour chaque individu. En fin de cette partie, on pose la question ouverte suivante : « Si je vous demande d'où vous êtes, que répondez-vous ? ». On cherche à observer à travers cette question, d'une part, de quelle manière le sentiment d'appartenance est lié à l'histoire des individus, et d'autre part, l'échelle géographique à laquelle les personnes se définissent, du plus local au plus général.

Les résultats de l'enquête pilote indiquent que plus de la moitié (58%) des personnes interrogées (438) répondent à cette question en citant le nom d'une commune, 17% fournissent une région, administrative ou non (exemples : Provence, le Midi, l'Est, le Sud, la Normandie). L'échelle du département ou celle du pays (France ou pays étranger) sont cités à égalité par 10% des personnes interrogées dans le test. Quelques très rares réponses évoquent une échelle européenne ou même mondiale. Une part importante des personnes interrogées (12%) fournit un autre type de réponses qui ne se réfère pas à une localisation géographique précise mais à un type de lieux, d'espaces ou une caractéristique donnée. Parmi celles-ci, on relève : « banlieusard », « de la mer », « des montagnes », etc. ou une appartenance avec une référence à la nationalité : « citoyen français », « français d'Afrique du Nord », « français né en Espagne », « français d'Afrique du Nord », « d'origine algérienne », « pied noir », etc.

On observe aussi un ensemble de réponses (4%) qui sont plus imprécises et qui traduisent une difficulté à se définir : « d'en bas », « de nulle part », « d'ailleurs », « apatride », « de partout », « d'ici », « d'où je viens », « je ne sais plus », « moitié-moitié », « c'est compliqué ». Ce sont les personnes d'origine étrangère qui déclarent le plus souvent une appartenance double qui peut être déclinée à différentes échelles comme en témoignent les réponses suivantes « du Maroc ou de Mulhouse », « de Nice et du Maroc, moitié-moitié », mais aussi de deux régions « je suis moitié breton, moitié angevin ». Ces premiers résultats, obtenus à partir d'un échantillon restreint (800 enquêtés), démontrent déjà une grande diversité des sentiments d'appartenance géographique dont les proportions respectives dans l'ensemble de la population pourront être mesurées dans l'enquête. Ils suggèrent par ailleurs des pistes intéressantes d'exploitation de l'enquête finale à travers l'analyse de l'ensemble des réponses en relation avec les lieux cités dans la trajectoire, les lieux d'attachement et les lieux de projet pour tenter de comprendre comment se construit le sentiment d'appartenance géographique des populations. Ce type d'analyse pourrait aussi être enrichi par la prise en compte d'autres attributs identitaires que constitue par exemple la langue parlée.

3.2. L'analyse d'un matériau qualitatif : événements et « dessins » de la vie

L'intérêt de l'utilisation d'une grille biographique a été éprouvé dans plusieurs enquêtes (GRAB, 1999). Ce mode de recueil de l'information facilite le travail souvent fastidieux lié à la restitution des dates et des événements, par les enquêtés. Son utilisation dans le cadre d'une enquête CAPI constitue une première expérience que l'enquête pilote a permis d'éprouver avec succès. Le remplissage de la grille biographique débute par la mise en place de jalons constitués par les différents événements familiaux d'Ego : naissance, mise en couple éventuellement séparation ou divorce, naissance des enfants éventuellement leur décès, celui du conjoint et des parents. Les événements liés à la famille sont en effet les plus faciles à dater, et permettent de situer d'autres événements de la trajectoire qui leur sont liés, par exemple ceux de la trajectoire résidentielle, car ils constituent souvent des motifs de migration. Le lien avec le thème de l'enquête ne nécessitait pas un recueil exhaustif des logements occupés mais plus simplement celui des changements de communes de résidence. Par la trajectoire professionnelle on obtient la nature et les périodes d'activité mais aussi celles d'inactivité et de chômage. Une dimension plus subjective vient compléter l'ensemble de cette information. C'est le regard que porte Ego sur son histoire. Pour cela on lui demande d'effectuer un découpage de sa vie en différentes périodes relatives à l'aisance financière d'une part, et à une appréciation plus générale des années qui composent ces périodes, d'autre part.

La dernière colonne de la grille biographique a pour objectif de recueillir les événements « personnels ou historiques » les plus marquants de l'histoire de chacun. Le nombre d'événements cités est en moyenne de trois, avec une variation relativement importante entre les individus. Cette observation ne sera vraisemblablement pas modifiée par l'accroissement de la taille de l'échantillon. Le corpus constitué de l'ensemble des événements recueillis contient 2262 items dont les plus fréquents sont les événements démographiques classiques : « naissance » (541 occurrences), « décès » ou « mort » (492), « mariage » (167), plus loin derrière, sont cités les périodes de « guerre » (77) ou des événements tels que la « maladie » (57), un « cancer » (12), une « fausse couche » (12), un « divorce » (54), un

accident (44). Les événements « séparation » (25) ou « rupture » (11) opposés avec « rencontre » (25) ou « mise en couple » (12) apparaissent avec la même fréquence dans le corpus. De manière moins présente, on relève des événements historiques dont les plus fréquemment cités sont « mai 1968 » (16), et « l'élection de Mitterrand en 1981 » (15).

Le graphique suivant (**figure 1**) illustre l'analyse du vocabulaire associé aux événements, en relation avec les caractéristiques des répondants, par les méthodes de la statistique textuelle (Guérin-Pace, 1997). La démarche de traitement des données est la suivante : en premier lieu, on constitue un lexique de l'ensemble des mots présents dans le corpus « Événements », recueilli dans le tableau biographique. Ce lexique est ensuite croisé avec un ensemble de variables individuelles. Le tableau, soumis à une analyse factorielle des correspondances, (logiciel SPADT) est un tableau de contingence dans lequel figure en ligne chacun des mots du corpus et, en colonne, différentes variables caractéristiques des individus interrogés. Les variables retenues sont les suivantes : « âge » en quatre classes, le « sexe », et le « statut d'occupation » selon huit modalités. Le premier plan factoriel issu de cette ACM explique 60% de la variance du tableau. Sa lecture peut se faire par cadrans en suivant la courbe de l'âge et en interprétant les proximités entre mots et entre catégories. Ainsi, proches de la classe d'âge « 19 à 35 ans » et de la catégorie « étudiants », on relève les événements suivants : « bac », « rencontre », « rupture », « indépendance », « suicide », « coupe du monde » associés à des proches « ami », « parents », « grands-parents ». A proximité de la classe d'âge « 35-50 ans » on trouve des événements familiaux : « mariage », « naissance », « fausse couche », « divorce » mais aussi des événements liés à la politique ou l'actualité : « (élection de) Mitterrand », « chute du mur de Berlin », « attentats », « mai 1968 ». En progressant dans la variable âge, on recueille davantage d'événements liés à la santé : « maladie », « cancer », « handicap » puis proches de la catégorie « 65-90 ans » les termes « guerre », « petits-enfants » ou encore « opération ». Les mots qui participent le plus à la construction du deuxième axe forment une opposition entre des événements relatifs à Ego et des événements plus extérieurs. Le deuxième plan factoriel oppose d'une part les événements recueillis auprès des hommes, essentiellement politiques et professionnels, et ceux donnés par les femmes dans un registre plus familial avec la citation d'événements douloureux : « fausse couche », « maladie », « cancer », « séparation » et la présence de proches. Le quatrième axe est plus difficile à interpréter et la variable âge doit être utilisée avec prudence car elle concerne l'âge d'ego au moment de l'enquête et non pas l'âge au moment de l'événement. Cette dernière variable devra être introduite dans les analyses de l'enquête finale pour distinguer les effets liés à l'âge de ceux liés à la génération.

Tout au long de ce questionnaire nous avons tenté de recueillir quelques éléments plus subjectifs de la construction des identités individuelles. En particulier, en fin de questionnaire, nous avons cherché à établir une hiérarchisation des composantes de l'identité en demandant à chacun de préciser parmi les thèmes abordés dans l'enquête, ceux qui le caractérisaient le mieux. En conclusion du questionnaire, nous avons souhaité recueillir le regard porté par Ego sur sa trajectoire au moyen de différents dessins qui symbolisent des parcours de vie : une flèche dirigée vers le haut, une droite horizontale, un escalier, une flèche interrompue par une croix, idem mais suivie d'une descente, une alternance de flèches montantes et descendantes et enfin un dernier dessin qui représente une flèche dirigée vers le bas. Chacune des personnes enquêtées était invitée à choisir le dessin qui, pour elle, correspondait le mieux à son histoire de vie et à donner par la suite une courte explication du choix de ce dessin. Les résultats de l'enquête pilote montrent en premier lieu la très bonne acceptation de cette question par les répondants et démontre l'intérêt de varier les supports de recueil de l'information dans les enquêtes. On a ensuite observé les différentes interprétations qui étaient données aux dessins en fonction des motifs de choix qui leur étaient associés.

L'objectif de l'analyse textuelle étant de mettre à jour les différenciations du corpus, indépendamment des effectifs, son application sur les résultats du test est adaptée et permet d'entrevoir les explorations futures qui pourront être envisagées. Cette première lecture des résultats démontre une grande richesse des interprétations et du contenu associé aux dessins proposés. Si le premier dessin (flèche dirigée vers le haut) évoque pour presque tous une « ascension » ou une « évolution », une « progression » mais aussi une « amélioration », une « construction », ou encore une « vie réussie », le contenu donné à ce choix diffère d'un individu à l'autre. Pour certains, il reflète une satisfaction dans le domaine familial,

pour d'autres, c'est une réussite professionnelle qui est évoquée ou une situation financière qui s'améliore au cours de la vie.

La **figure 2** illustre la richesse du vocabulaire des motifs en relation avec le choix des dessins. Selon le même principe, on a constitué un lexique de l'ensemble des mots présents dans le corpus constitué des motifs et construit un tableau de contingence dans lequel figure en ligne chacun des mots du lexique et, en colonne, les modalités de la variable qui correspondent au choix du dessin. Le premier axe de ce graphique oppose essentiellement le dessin 6 (alternance de flèches montantes et descendantes) aux autres dessins. Le vocabulaire qui caractérise ce dessin est très homogène et s'exprime le plus souvent par « des hauts et des bas » ou en « dents de scie », sans autre explication sur ce qui constitue les hauts et les bas.

A l'opposé sur l'axe, le dessin 2 (droite horizontale) se détache, entouré d'adjectifs tels que « droit », « régulier », « stable », « calme » et souvent interprétés en termes de « vie normale », « sans gros problème », parfois qualifiée de « banale », « toujours pareil », « un long fleuve tranquille ». L'axe 2 oppose cette vision linéaire de la vie avec celle plus chaotique, constituée de « paliers », d'« échelons », d'« escalier » et qui se rapporte le plus souvent à la trajectoire professionnelle. On relève à proximité des dessins 4 et 5 qui contiennent tous les deux une rupture dans la trajectoire, les événements « décès », « santé », « maladie », « cassure », « divorce ». Cette analyse pourrait être approfondie en introduisant, en plus des caractéristiques des répondants, des éléments significatifs de leur trajectoire. Par ailleurs, l'analyse des réponses « autres » en réponse au choix du dessin a justifié l'introduction d'un dessin supplémentaire et confirme à nouveau l'utilité de l'exploitation d'une enquête pilote par des méthodes adaptées à la particularité de l'échantillon.

Bibliographie

[1] Groupe de Réflexion sur l'Approche Biographique : *Biographie d'enquêtes : bilan de 14 collectes biographiques*, INED-IRD, 340 p. 1999.

[2] Guérin-Pace France : « La statistique textuelle : un outil exploratoire en sciences sociales », *Population*, 4, pp. 865-888.

FIGURE 2



