



Enquêtes en ligne : peut-on  
extrapoler les comportements et  
les opinions des internautes à la  
population générale ?

**CAHIER DE RECHERCHE N° 273**

**DÉCEMBRE 2010**

**Régis Bigot  
Patricia Croutte  
Fanette Recours**

**Département « Conditions de vie et Aspirations des Français »**

**Dirigé par Régis Bigot**

*Cette recherche a bénéficié d'un financement au titre de la subvention recherche attribuée au CRÉDOC.*

## SOMMAIRE

---

<b>Synthèse des principaux résultats .....</b>	<b>5</b>
<b>Introduction .....</b>	<b>15</b>
<b>PARTIE 1 : ETAT DES LIEUX ET PERSPECTIVES.....</b>	<b>18</b>
<b>I. La diffusion des enquêtes en ligne : limites des modes traditionnels et avantages de l'outil Internet .....</b>	<b>19</b>
1. Les modes traditionnels : un taux de refus de plus en plus élevé.....	19
2. Un mode économique, rapide et sans limite géographique.....	20
3. Avantage de l'auto-administré .....	21
4. Des nouvelles possibilités d'enquête .....	22
<b>II. Processus d'échantillonnage et collectes des données pour les enquêtes en ligne .....</b>	<b>24</b>
1. Deux types d'échantillonnage .....	24
2. Les échantillons probabilistes.....	25
3. Les échantillons empiriques – ou non probabilistes .....	27
<b>III. Arbitrer entre les différentes erreurs .....</b>	<b>30</b>
1. L'erreur d'échantillonnage .....	30
2. L'erreur de couverture .....	30
3. L'erreur de non-réponse.....	34
4. L'erreur de mesure, d'observation .....	37
<b>IV. Un ajustement des erreurs est-il possible ? .....</b>	<b>44</b>
1. La pondération traditionnelle .....	44
2. Scores de propension ou probabilité d'inclusion.....	46
3. Pondération fondée sur une enquête de référence .....	48
4. Les enquêtes mixtes.....	49
<b>V. Les pistes d'améliorations des études en ligne.....</b>	<b>51</b>
1. L'aspect visuel .....	51
2. La transparence .....	52
3. La qualité des données .....	53
4. Les incitations financières.....	54

<b>PARTIE 2 : Expérimentation Peut-on extrapoler les résultats des internautes à la population française ? .....</b>	<b>55</b>
<b>Principe de l'expérimentation .....</b>	<b>56</b>
<b>I. Les internautes en France .....</b>	<b>57</b>
1. Le fossé numérique et son évolution .....	57
2. Le profil sociodémographique des internautes en 2009 .....	59
3. Les facteurs déterminants pour expliquer, en 2009, l'accès à Internet .....	66
<b>II. La spécificité des modes de vie et des opinions des internautes.....</b>	<b>71</b>
<b>III. Le redressement de l'échantillon d'internautes.....</b>	<b>74</b>
1. Le principe du redressement.....	74
2. Les effets du redressement .....	76
<b>IV. L'échantillon d'internautes redressé : des Français comme les autres ? .....</b>	<b>78</b>
1. TIC, équipement mais aussi moral économique ou mœurs : les variables où le redressement est insuffisant .....	81
2. Les variables où le redressement a été nécessaire et où il s'avère suffisant .....	85
3. Les variables pour lesquelles le redressement n'est pas nécessaire .....	93
<b>Conclusion .....</b>	<b>95</b>
<b>Bibliographie .....</b>	<b>97</b>
<b>Annexes.....</b>	<b>102</b>

## REMERCIEMENTS

Nous remercions chaleureusement les participants au groupe de travail sur les « Enquêtes en ligne et en mode mixte auprès des ménages » — créé en 2010 à l’initiative de Marie-Hélène Amiel (INSEE) et co-animé par le CRÉDOC — pour leurs réflexions, le partage de leurs connaissances, leurs retours d’expérience sur les méthodologies d’enquêtes et les perspectives qu’ils ouvrent à chaque réunion : Séverine Arnaud (INSEE), François Beck (INPES), Jean Chiche (CEVIPOF), Thomas Denoyelle (INSEE), Arnaud Gautier (INPES), Lucie Gonzalez (DREES), Dominique Guédès (INSEE), Romain Guignard (INPES), Stéphane Legleye (INED), Antoine Moreau (IPSOS), Jean-Baptiste Richard (INPES), Laurent Toulemon (INED), Daniel Verger (INSEE) et Josiane Warszawski (INSERM).

Nous tenons également à remercier Benoît Riandey (INED), président du groupe « Enquête, Modèles et applications » de la Société Française de Statistique, ainsi que Anne-Marie Dussaix (ESSEC, SFdS) et Anne Gayet (AID, SFdS) pour leurs nombreuses initiatives visant à mobiliser et réunir les chercheurs, statisticiens, instituts d’études publics et privés, ainsi que tous les acteurs cherchant à faire progresser la connaissance autour des enquêtes en ligne.

Tous nous ont aidé à mieux comprendre les enjeux liés aux enquêtes par Internet. Ils retrouveront dans cette recherche beaucoup d’idées qu’ils nous ont données. Nous restons, bien entendu, seuls responsables des erreurs et imperfections que pourrait contenir le rapport.

## **Online-Surveys: Is it possible to extrapolate behaviors of Internet-users to the general population**

### **Summary:**

Every year the increasing number of surveys makes them a key-issue of public debates. Simultaneously Internet has become a major source of data collecting because it provides cheap and fast access to a relatively large group of respondents. However, this new mode of data collection arouses a number of questions. In the first place one may ask if Internet is an appropriate tool in order to determine reliably behaviors and opinions of the population as a whole (coverage errors).

The following report is divided in two parts:

In a first part we will give a short overview on the state of art of methodological debates on online surveys.

We finally attempt to discuss the importance of coverage errors for the results of online surveys. More specifically we will outline the risk of questioning only people having access to Internet ignoring all those who don't (one in three persons has no domestic Internet-access in France). Furthermore we will test the efficiency of statistical methods which allow to adjust and to correct coverage errors. Our experimental material is extracted from the CREDOC survey on live conditions and aspirations of the French population. The paper concludes that the under-coverage cannot be adjusted statistically in any case and for a certain number of tested items (especially regarding cultural practices, leisure time activities, standard of living, indicator of well-being of the household, patrimony).

## Synthèse des principaux résultats

---

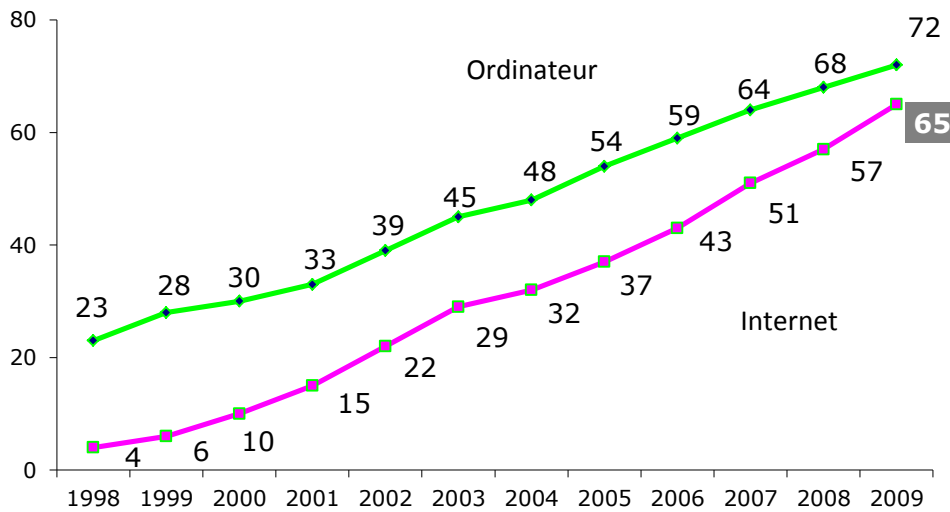
Depuis quelques années, les études en ligne sur Internet connaissent un véritable engouement. A priori plus simples à mettre en œuvre, moins coûteuses, elles séduisent de plus en plus de chercheurs, d'instituts de sondage et de commanditaires d'études. Il est vrai que ce mode d'enquête représente plusieurs avantages. Un coût financier qui peut être moins élevé, son mode auto-administré susceptible de diminuer les effets de désirabilité sociale (se montrer sous son meilleur jour), et la simplicité d'utilisation facilitant la passation des questionnaires (gestion des filtres, rotation aléatoire des items, utilisation du son, des images, de la vidéo...). Mais ces enquêtes sont aussi confrontées à un inconvénient de taille lorsqu'on est intéressé par l'inférence en population générale : leur représentativité.

### **L'erreur de couverture, principal problème évoqué à propos des études en ligne**

En France, 72% de la population disposent d'un ordinateur et 65% ont accès à Internet à leur domicile. Ces chiffres sont élevés et en croissance, mais il reste encore un nombre important de personnes qui ne peuvent pas être interrogées par Internet : il y a donc un risque d'erreur de couverture. Un risque d'autant plus gênant qu'on observe des différences sociales très marquées dans la probabilité d'être connecté. Les internautes français sont plus jeunes, plus diplômés, plus aisés et sont davantage représentés chez les cadres et professions intellectuelles supérieures. En fixant les effets d'âge, de diplôme, de taille du foyer, de profession et catégorie socioprofessionnelle, de revenus et de taille d'agglomération à l'aide d'une régression logistique, on calcule qu'une personne âgée de 70 ans ou plus a presque cinq fois moins de chance de disposer chez elle d'un accès à Internet qu'une personne de 40-59 ans ; une personne qui vit seule a deux fois moins de chance d'être équipée qu'une personne qui vit dans un foyer de deux personnes ; une personne qui vit dans un foyer de trois personnes a quatre fois plus de chance d'être équipée ; entre ceux qui vivent seuls et ceux qui vivent dans des foyers de trois personnes ou plus, la probabilité d'être connecté à Internet varie presque de 1 à 8 ; la probabilité pour une personne issue de l'enseignement supérieur d'être connectée à domicile est trois fois plus élevée qu'un titulaire du Bepc ; les catégories aisées ont trois fois plus de chance d'avoir Internet que les classes moyennes inférieures et enfin, une personne résidant dans une agglomération de moins de 2000 habitants a quasiment deux fois moins de chance qu'un habitant d'une grande agglomération d'être équipé.



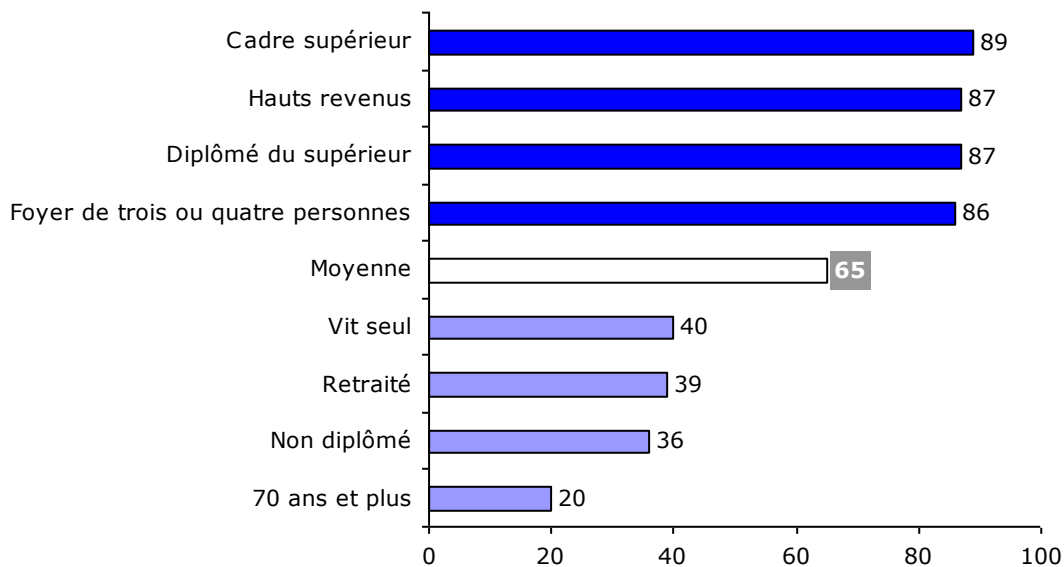
**Graphique 1 - Taux de pénétration de l'ordinateur et d'Internet**



Source : CRÉDOC, enquêtes « Conditions de vie et Aspirations des Français », vague de janvier de chaque année.

**Graphique 2 - Le « fossé numérique » en 2009**

(Proportion d'individus disposant d'une connexion Internet à domicile, en %)



Source : CRÉDOC, enquête « Conditions de vie et Aspirations des Français », début 2009.

### Les internautes ont des opinions et des modes de vie spécifiques

En plus des spécificités sociodémographiques, les attitudes des internautes divergent de celles des non-internautes sur un grand nombre de sujets. Sur 191 variables de l'enquête « Conditions de vie et aspirations des Français » du CRÉDOC, le fait de disposer d'une



connexion à Internet ou pas se traduit — toutes choses égales par ailleurs — par des différences d'opinion ou de comportement dans 37% des cas. Les régressions logistiques testées montrent en effet que la variable « Internaute/non-internaute » explique de nombreuses différences d'attitude à âge, niveau de diplôme, niveau de revenu, profession, taille du foyer et lieu de résidence comparables. L'influence est plus ou moins marquée selon les sujets d'étude :

- ✓ Les attitudes vis-à-vis des nouvelles technologies, des innovations techniques et du progrès scientifique sont bien évidemment complètement liées au fait d'avoir accès à Internet à domicile.
- ✓ Moins intuitif : un très grand nombre de taux d'équipements des ménages sont différents selon que l'on interroge un internaute ou un non-internaute (possession d'une automobile, d'un lave vaisselle, d'un congélateur...). La spécificité demeure alors même que la taille du foyer a été neutralisée : ce n'est donc pas un effet d'économie d'échelle qui joue ici indirectement.
- ✓ La plupart des variables s'intéressant au niveau de vie et au patrimoine révèlent une forte corrélation avec Internet (alors même que les effets du revenu, de la profession, de l'âge et du niveau de diplôme ont été fixés).
- ✓ Les internautes ont des opinions sur les mœurs assez marquées : ils sont davantage favorables à l'union homosexuelle, l'homoparentalité et le travail des femmes ;
- ✓ Une fois sur deux, les indicateurs ayant trait aux caractéristiques du logement sont, également, liés à l'accès à Internet. C'est le cas par exemple du type de logement, de la taille de celui-ci ou du mode de chauffage principal.

Pour les autres thèmes, la corrélation avec Internet semble moindre. Elle semble même quasi nulle pour les questions relatives à la santé ou à l'environnement.

**Tableau 1 : Corrélation entre le fait d'être équipé d'Internet ou pas et différents comportements ou opinions**

	Nombre de variables	Effet 'Internet' intrinsèque	
		P<0.05	%
Attitudes vis-à-vis des TIC	3	3	100%
Equipped du ménage	15	14	93%
Niveau de vie	8	7	88%
Situation personnelle et familiale	16	12	75%
Opinions en matière de mœurs	5	3	60%
Logement	13	7	54%
Loisirs, pratiques culturelles	21	7	33%
Emploi, chômage, situation professionnelle	10	3	30%
Moral économique	21	6	29%
Opinions sur les politiques familiales	20	5	25%
Jugements sur les politiques sociales	13	2	15%
Environnement	7	1	14%
Inquiétudes et préoccupations	19	1	5%
Santé	10	0	0%
Opinions sur la société	10	0	0%
	191	71	37%

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

### **L'erreur totale dépend également des erreurs d'échantillonnage, de non-réponse et de mesure**

Si le risque d'erreur de couverture est très souvent évoqué au sujet des enquêtes en ligne, d'autres erreurs peuvent entrer en ligne de compte — comme dans tout type de sondage d'ailleurs. Elles sont tout aussi importantes les unes que les autres. L'erreur d'échantillonnage correspond à la différence entre l'estimation calculée à partir d'une enquête par échantillon et la « vraie » valeur qui aurait été obtenue si un recensement exhaustif auprès de la population avait été effectué dans les mêmes conditions. On parle d'erreur de non-réponse complète (si le répondant sélectionné ne répond pas) ou d'erreur de non-réponse partielle (s'il ne répond pas à toutes les questions). Si les caractéristiques des répondants diffèrent de celles des non-répondants, l'erreur de non-réponse risque d'être importante et d'entacher sérieusement les résultats. Enfin, l'erreur de mesure peut être liée à plusieurs facteurs : la bonne ou mauvaise foi des répondants, l'influence de l'enquêteur, les erreurs de saisie, de codification, de programme informatique, une mauvaise compréhension, la formulation de la question, l'ordre des questions...

Ces erreurs ne sont pas propres aux enquêtes en ligne. Elles existent dans les modes d'enquêtes plus classiques. Il s'agit d'évaluer si elles sont, avec l'outil Internet, plus ou moins grandes.

L'erreur d'échantillonnage est inhérente à tout sondage, quel que soit le mode d'enquête (téléphone, face-à-face, auto-administré...).

Les erreurs de couverture et de non-réponses vont dépendre du type d'enquête en ligne mis en place, selon qu'elles sont menées selon le paradigme probabiliste (effectuées à partir d'un tirage aléatoire d'un échantillon d'individus qui correspondent à l'ensemble de la population cible) ou pas. Aujourd'hui, la plupart des *access panels* proposés par les instituts de sondage ne sont pas constitués de manière parfaitement aléatoire : tout l'enjeu consiste alors à comprendre comment les panélistes ont été recrutés.

Les erreurs de couverture sont quasi inexistantes dans les enquêtes ciblées probabilistes, c'est-à-dire lorsque l'on ne cherche pas à être représentatif de la population en entier mais d'un sous-groupe seulement, entièrement joignable par Internet. Le risque se pose surtout pour les enquêtes ciblées qui se prétendent généralistes.

Enfin, l'erreur de mesure fluctue également selon les modes : la présence d'un enquêteur peut entraîner des effets de « désirabilité sociale » (l'enquêté donne une réponse qu'il pense être celle attendue par l'enquêteur), le téléphone peut souffrir des effets de mémoire (les dernières modalités proposées seraient davantage choisies), Internet soulève le risque du « *satisficing* » (à cause d'une faible implication des répondants, les premières réponses pourraient cette fois-ci être davantage sélectionnées afin de finir de remplir plus vite le questionnaire).

Ces erreurs sont inhérentes à la théorie des sondages, ce qui n'empêche pas celle-ci de s'appuyer sur des bases théoriques solides. L'important est de pouvoir mesurer et contrôler ces erreurs.

### **Les ajustements possibles pour limiter l'erreur**

Il est possible d'essayer de se rapprocher de la réalité par des traitements statistiques visant à corriger les erreurs ou du moins pour tenter d'en donner une appréciation. Concernant les enquêtes par Internet, l'idée consiste à « redresser » l'échantillon des internautes pour le rendre comparable à la population générale. Pour cela, plusieurs méthodes existent : la pondération traditionnelle à l'aide de données auxiliaires (ensemble de variables mesurées dans l'enquête dont on connaît également la répartition dans la population : par exemple, on s'assure que l'échantillon comporte autant de ruraux, d'urbains, de jeunes, de seniors, de riches, de pauvres, d'ouvriers, de cadres, de diplômés, de non-diplômés que dans l'ensemble de la population) ; la pondération à

partir de scores de propension, qui consiste à modéliser une variable indiquant si une personne participe ou non à l'enquête (modèle de régression logistique où la variable indicatrice est la variable dépendante et les variables attitudinales sont les variables explicatives) ; la pondération fondée sur une enquête de référence, c'est-à-dire à partir d'un échantillon probabiliste mené en parallèle qui permettrait d'ajuster les valeurs.

Aujourd'hui, les chercheurs sont assez partagés — et plutôt dubitatifs — sur l'efficacité de ces traitements statistiques concernant les enquêtes en ligne réalisées à partir d'accès *panels* non-probabilistes : l'erreur de couverture, de non réponse, le risque lié à l'autosélection (lorsque l'enquêté décide lui-même d'intégrer un panel ou de participer à telle ou telle enquête spécifique, sans que l'on connaisse ses motivations à le faire) rendent très incertaines les chances de réussite de ces redressements, puisqu'on ne contrôle pas tous les paramètres.

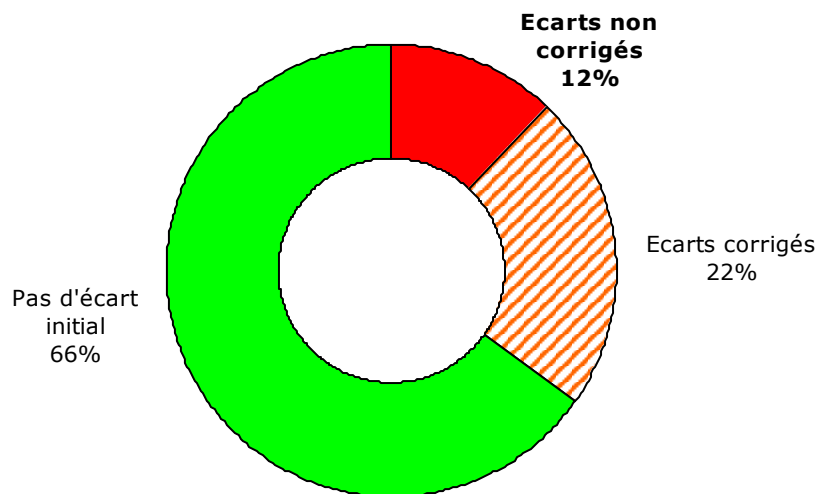
A noter que, dans certains cas, les enquêtes en mode mixte (associant une enquête par Internet, par téléphone ou en face-à-face) peuvent être utilisées pour compenser, en quelque sorte, les faiblesses du mode Internet. Elles peuvent permettre aussi tout simplement d'augmenter le taux de réponse.

### **Résultats d'une expérience où l'on cherche à corriger l'erreur de couverture**

À partir de l'enquête « Conditions de vie et Aspirations des Français », un redressement de la population des internautes a été proposé afin que sa structure socio-démographique soit le plus proche possible de celle de la population française (il s'agit donc d'un calcul de pondération a posteriori classique). Bien que l'enquête n'ait pas été menée en ligne (elle se déroule en face-à-face), nous avons souhaité voir si les réponses des internautes, après redressement, différaient des réponses de la population générale. Après avoir déterminé les critères les plus déterminants de la probabilité d'être connecté à Internet (à l'aide d'une régression logistique), une pondération par calage sur marge a été mise en place sur ces critères.

Notons tout d'abord que, sur les 196 variables analysées, dans environ un tiers des cas (34%), les réponses de l'échantillon des internautes non redressé diffèrent de celles apportées par l'échantillon global, représentatif de l'ensemble de la population. Après redressement, encore 12% des variables montrent des spécificités entre l'échantillon redressé des internautes et celui de la population totale.

**Graphique 3 : les écarts d'attitudes entre les internautes et la population générale avant et après redressement**

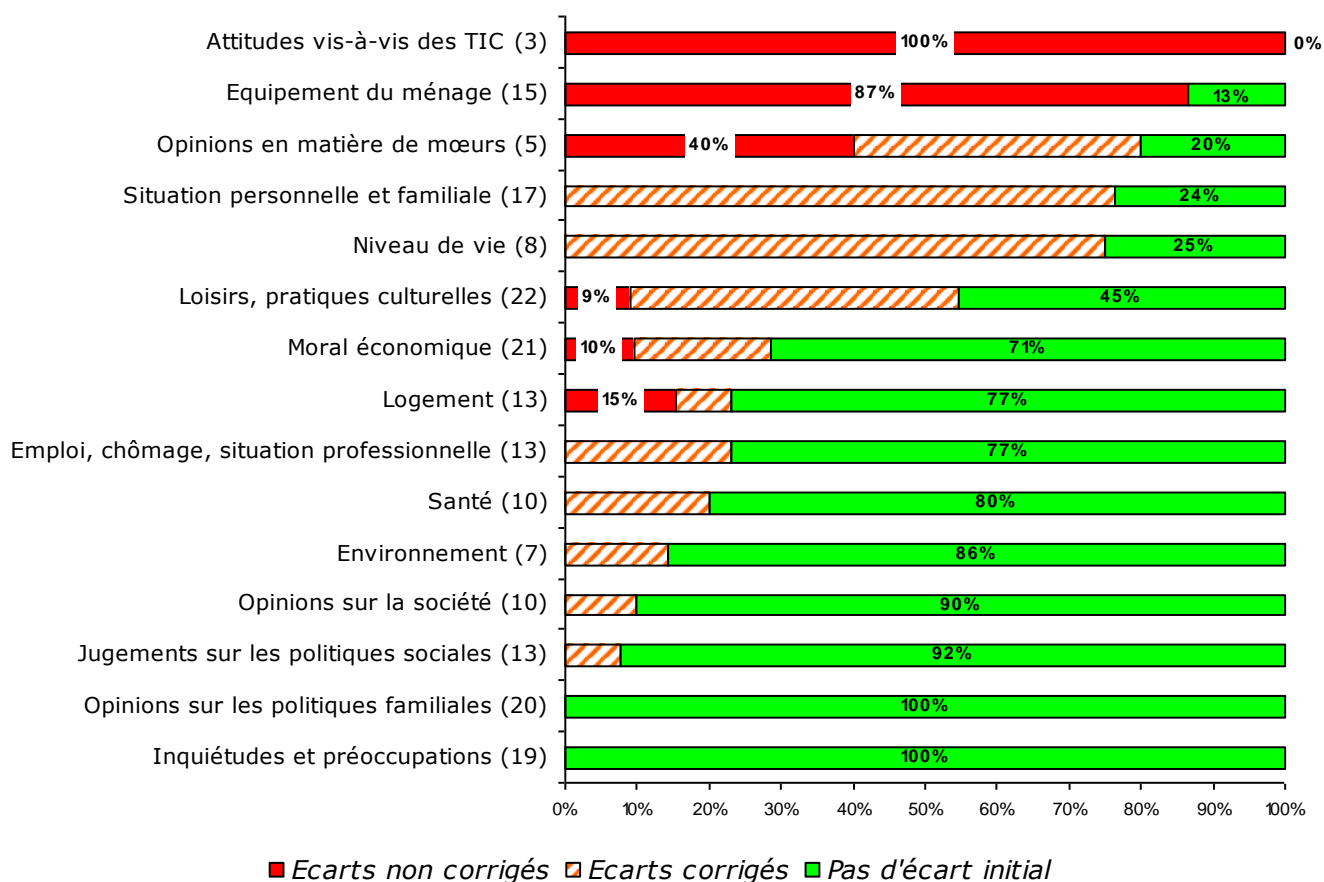


Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

Dans le détail :

- ✓ 12% des questions laissent apparaître des spécificités des internautes, spécificités qui ne peuvent pas être corrigées par redressement. Cela concerne les attitudes sur les technologies de l'information et, plus généralement, l'équipement du ménage, puis certaines opinions en matière de mœurs, des caractéristiques du logement, des indicateurs liés au moral économique ou aux loisirs et pratiques culturelles ;
- ✓ 22% des résultats ont pu être corrigés : les internautes se distinguaient initialement, mais le redressement permet de retrouver des réponses identiques à celles observées en population générale. Il s'agit des jugements sur les politiques sociales, des opinions sur la société, des attitudes par rapport à l'environnement, des perceptions de son état de santé et des opinions sur le système de soins, du niveau de vie et du patrimoine ;
- ✓ 66% des questions étaient déjà similaires entre les internautes et l'ensemble de la population, sans qu'il soit nécessaire de procéder à un redressement. Il s'agit surtout des sujets d'inquiétudes ou de préoccupations, d'opinions sur les politiques familiales ou les politiques sociales, et de jugements portés sur le fonctionnement de la société.

**Graphique 4 - Les écarts d'attitudes entre les internautes et la population générale, avant et après redressement**



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

La repondération en post-stratification se révèle donc insuffisante dans un certain nombre de cas. C'est aussi ce que relève la littérature. Peut-être des repondérations se basant sur des scores de propension ou à partir d'une enquête de référence – que nous n'avons pas testées dans ce cahier de recherche – pourraient mieux corriger certains effets. Cet exercice pourrait être expérimenté dans un prochain travail de recherche, en s'appuyant à nouveau sur la richesse de l'enquête « Conditions de vie et aspirations » ou en mobilisant d'autres grandes enquêtes sociétales analysant des thèmes variés.

## Introduction

---

D'année en année, les sondages se multiplient, occupant une place grandissante dans les débats publics. Au point que la commission des lois du Sénat a décidé de créer en son sein, le 14 octobre 2009, une mission d'information sur les sondages en matière électorale. Le recours à Internet pour recueillir des informations se généralise : la rapidité de la collecte des données et la faiblesse des coûts de terrain militent pour la mise en œuvre de telles enquêtes. De même, dans la mesure où les taux de participation aux enquêtes sont en nette diminution, les études en ligne peuvent venir en complément des modes traditionnels — face-à-face, téléphone, auto-administré papier — qui suscitent de la lassitude de la part des ménages de plus en plus sollicités. D'après les chiffres du Syntec<sup>1</sup>, la moitié des instituts interrogés déclarent réaliser des études quantitatives sur Internet en 2009. En moyenne, elles représentent 10% de leur chiffre d'affaires. Elles contribuent pour 31% du total des études quantitatives tandis que les études téléphoniques et les tests en salle perdent du terrain depuis 2008. Ce chiffre est peut-être même sous-estimé car toutes les sociétés d'études (surtout celles qui réalisent beaucoup d'enquêtes en ligne) ne sont pas membres du Syntec et parce que pour certains types d'études (tests de concepts, tests de publicité...), le recueil par Internet a pratiquement supplanté tous les autres modes de recueil<sup>2</sup>.

Face à cet essor des enquêtes en ligne, les publications concernant leur méthodologie n'ont cessé de se multiplier. Mick P. Couper et Peter V. Miller (2008)<sup>3</sup> notent que les premiers articles publiés sur les enquêtes en ligne sont apparus en 1996 et que, depuis lors, les investigations ont proliféré. Le site WebSM.org, dédié à la méthodologie des enquêtes en ligne, comporte une bibliographie sur le sujet systématiquement mise à jour auprès de cent sources différentes avec à peu près 80 unités ajoutées chaque mois.

Comme beaucoup d'instituts de sondage, d'organismes de recherche et producteurs de statistiques officielles, le CRÉDOC s'interroge sur le bien-fondé des enquêtes en ligne. Des doutes subsistent : en particulier, quelle validité peut-on accorder à ces enquêtes dont une partie de la population, celle qui n'est pas équipée d'un accès à Internet, est *de facto* exclue ? En juin 2010, on dénombre encore 29% de la population qui ne disposent

---

<sup>1</sup> Syntec représente près de 1250 groupes et sociétés françaises spécialisées dans les professions de l'Ingénierie, des Services Informatiques, des Études et du Conseil, de la Formation Professionnelle.

<sup>2</sup> « L'Essentiel sur les Etudes Marketing », *Marketing études, Le guide des études marketing, média et opinion*, Tarsus, Semo, 16<sup>ème</sup> édition, 2010

<sup>3</sup> Mick P. Couper, Peter V. Miller, « Web survey methods », Introduction, *Public Opinion Quarterly*, Volume 72, Issue 5, 2008, p. 831-835, <http://poq.oxfordjournals.org/content/72/5.toc>

d'aucune connexion à Internet<sup>4</sup>. Le doute s'accroît à la lecture de certains résultats d'enquêtes en ligne qui montrent des divergences étonnantes. Citons-en deux :

- Exemple 1 : enquête par téléphone selon la méthode des quotas, réalisée les 24 et 25 août 2010. A la question « Vous savez que le gouvernement procède actuellement à des reconduites en Roumanie de Roms dont les camps ont été démantelés cet été. Êtes-vous favorable ou opposé à ces reconduites ? », 48% des personnes interrogées répondent qu'elles sont favorables. Exactement à la même période, les 24 et 25 août 2010, une autre enquête réalisée cette fois Internet pose une question très similaire, mais conduit à des résultats très différents. La question « Êtes-vous favorable ou opposé aux actions suivantes mises en œuvre ces dernières semaines ? L'expulsion de Roms sans papier vers la Roumanie » recueille 65% de réponses positives.
- Exemple 2 : d'après l'INSEE (source « Enquête permanente sur les conditions de vie », volet Participation culturelle et sportive auprès de 8000 ménages interrogés en face-à-face selon un échantillonnage probabiliste), 26% des personnes de 15 ans et plus sorties du système éducatif en France métropolitaine ont lu une bande dessinée au cours des douze mois précédant l'enquête, réalisée en 2003. En 2005, une enquête réalisée auprès d'un échantillon de 1 027 internautes « représentatifs de la population » indique que 67% des 15-64 ans ont lu une bande-dessinée dans l'année. En 2008, l'enquête sur les « Pratiques culturelles des Français » du Ministère de la Culture et de la communication, réalisée en face-à-face selon la méthode des quotas auprès de 5004 individus, conclut à 29% de lecteurs de BD parmi les 15 ans et plus.

Dans ces deux exemples, les résultats obtenus par Internet semblent très décalés avec ceux obtenus par d'autres modes d'interrogation. Certes, les décalages ne s'expliquent pas uniquement par le mode de collecte : le plan l'échantillonnage, la période d'enquête et la formulation des questions entrent aussi en ligne de compte. Mais il y a tout de même de quoi être troublé.

Le développement des enquêtes par courrier et téléphone avait lui-même donné lieu à de nombreuses études pour comparer leurs performances à celles du face-à-face. Les problématiques d'échantillonnage, de représentativité, de comparabilité entre les modes

---

<sup>4</sup> Régis Bigot et Patricia Croutte, *La diffusion des technologies de l'information et de la communication dans la société française*, 2010, CRÉDOC, 2010, [http://www.cqiet.org/documents/2010\\_12\\_10\\_Rapport\\_CRÉDOC\\_2010.pdf](http://www.cqiet.org/documents/2010_12_10_Rapport_CRÉDOC_2010.pdf)



n'ont cessé d'alimenter les débats. Aujourd'hui, les études en ligne viennent remettre en question les modes devenus depuis traditionnels

Parmi les inconvénients des études en ligne, l'erreur de couverture est peut-être celle qui est la plus souvent citée. Même si depuis une douzaine d'années la progression de l'équipement Internet est quasiment linéaire, cela n'empêche pas un certain nombre d'inégalités de perdurer. En France, l'âge, le niveau de diplôme, les revenus ou la taille du foyer sont par exemple des facteurs qui jouent encore fortement sur l'acquisition d'Internet.

Après une première partie qui tentera de dresser un rapide état des lieux de la recherche sur les études en ligne, nous tenterons d'évaluer, dans une deuxième partie, l'ampleur de l'erreur de couverture tout en testant l'efficacité d'une procédure statistique de correction de cette erreur (la repondération par post-stratification telle qu'elle est le plus souvent pratiquée dans les *access panels*). Nous utiliserons comme matériau d'expérimentation l'enquête du CRÉDOC sur les « Conditions de vie et les aspirations des Français » réalisée au début 2009 auprès d'un échantillon représentatif de la population française âgée de plus de 18 ans, interrogée selon la méthode des quotas.

PARTIE 1 : ETAT DES LIEUX  
ET PERSPECTIVES

# I. La diffusion des enquêtes en ligne : limites des modes traditionnels et avantages de l'outil Internet

---

D'après les chiffres du Syntec Etudes Marketing & Opinion, pour la première fois sur le marché français, la part des études quantitatives réalisées sur Internet a dépassé la part des études téléphoniques, et elle se rapproche de plus en plus de la part du face à face. Ainsi, les études en ligne ont contribué pour 31% du total des études quantitatives. Il faut dire que ce mode d'enquête offre des avantages non négligeables que nous allons détailler dans un premier temps.

## 1. Les modes traditionnels : un taux de refus de plus en plus élevé

Pour Don A. Dillman et al. (2009)<sup>5</sup>, la méthodologie d'enquête est entrée dans une période « turbulente ». A partir de l'historique des enquêtes menées aux États-Unis, ils montrent que dans les années 60, le face à face dominait, avec des enquêteurs qui persuadaient les individus de répondre, tandis que le courrier et le téléphone étaient très mal perçus depuis un sondage qui avait donné Alf Landon gagnant aux élections contre Roosevelt en 1936. De plus, les connections téléphonique étaient de mauvaise qualité et coûteuses. Quant au courrier, il n'existait alors pas de liste exhaustive des foyers. La décennie suivante a vu néanmoins ces deux modes se développer, compte tenu de la modernisation du téléphone, des photocopies, etc. Le téléphone a nécessité quelques changements : les individus interrogés doivent décider vite si elles souhaitent répondre, on évite les questions ouvertes... Globalement, le téléphone est utilisé lorsqu'un échantillon national est souhaité, le courrier si une liste adéquate existe et le face à face quand l'omission est intolérable.

Depuis les années 90 cependant, les enquêtes par téléphone deviennent plus difficiles à mener. Elles sont considérées de plus en plus biaisées parce que le nombre de personnes qui refusent de répondre augmente<sup>6</sup>. Elles deviennent de moins en moins représentatives

---

<sup>5</sup> Don A. Dillman, Jolene D. Smyth, Leah Melani Christian, *Internet, mail, and mixed-mode surveys, The tailored Design Method*, Third Edition, 2009, 499 p.

<sup>6</sup> Romuald Le Lan (2009) montre qu'entre 1992 et 2006, « les taux de réponse aux enquêtes de l'INSEE auprès des ménages ont baissé, principalement parce que les ménages ont été de plus en plus difficiles à joindre ». Par exemple, le taux de réponse de l'enquête Loyers et Charges a baissé de dix points entre 1998 et 2003. Il s'est stabilisé autour de 70% depuis 2004, mais certainement grâce à des modifications dans la collecte (informatisation, partie en face à face...). De même, entre 1992 et 2006, les parts des répondants aux enquêtes Logement, Patrimoine, Budgets de famille et Statistiques sur les Ressources et les Conditions de Vie ont diminué (baisse entre cinq et dix points ; pour l'enquête Logement par exemple, baisse du taux de réponse d'environ 87% en 1992 à 78% en 2006).

Romuald Le Lan, « Enquêtes ménages : vers la fin de la baisse des taux de réponse ? », Courrier des Statistiques n°128, INSEE, septembre-décembre 2009.

de la population générale. Parallèlement à cet effet de lassitude de réponse aux enquêtes, les annuaires se révèlent moins fiables qu'avant. L'apparition du téléphone mobile, ses effets de substitution avec le téléphone fixe, l'apparition d'Internet et de la téléphonie par ADSL, la multiplication des opérateurs télécom, les listes rouge et orange, obligent les instituts à remettre en question cette technique. Dans le même temps, les enquêtes en ligne apparaissent, permettant de grandes économies, à la fois en termes de coûts et en termes de délais.

## **2. Un mode économique, rapide et sans limite géographique**

Les trois avantages les plus souvent mis en évidence lorsque l'on parle des études en ligne sont les suivants : les coûts de recueil sont déconnectés de la taille d'échantillon, les délais sont plus rapides, les limites géographiques sont abolies.

### ***Des coûts moindres***

A priori, certains coûts diminuent nettement avec les enquêtes en ligne : comparés aux enquêtes par courrier, les coûts de papier, d'impression, d'enveloppes, de timbres... sont supprimés. Comparés aux enquêtes qui ne sont pas en mode auto-administré, ce sont les coûts des enquêteurs qui disparaissent, ainsi que les coûts des communications pour les enquêtes par téléphone. Certes, de nouveaux coûts interviennent pour l'équipement, le *web-design*, la technologie (programmation, serveurs, maintenance...). De même, les coûts vont dépendre de l'échantillonnage choisi : ils peuvent être plus élevés pour les enquêtes basées sur des listes où les participants reçoivent une invitation à répondre par mail, pour les enquêtes en autosélection non restreinte avec des bannières sur des sites très fréquentés, pour les échantillons pré-recrutés ou lorsque des incitations financières sont proposées. Une chose est sûre : le coût moyen par interview diminue lorsque le nombre de personnes enquêtées augmente. Plus l'échantillon est grand, plus l'avantage comparatif des enquêtes en ligne augmente : on économise des ressources budgétaires, on augmente la précision des estimations, et on peut cibler des sous-groupes de plus en plus étroits.

### ***Des délais plus rapides***

Le deuxième avantage souvent évoqué à propos des études par Internet est qu'elles seraient les plus rapides à mener. Le face à face et le téléphone demandent une plus grande organisation : les enquêteurs doivent être sélectionnés et formés, un superviseur est nécessaire... De plus, ces modes nécessitent des interventions dans des moments restreints, sur des tranches horaires bien définies, ce qui limite le nombre de répondants par journée. Or ces limites dans le temps n'existent plus avec l'enquête en ligne. L'enquêté répond à n'importe quel moment de la journée... ou de la nuit. Arnaud Gautier

*et al.* (2010)<sup>7</sup> indiquent notamment que, lors du baromètre Santé conduit auprès des médecins généralistes par l'INPES en 2009, 53% des médecins ayant accepté de répondre sur Internet après un premier refus téléphonique ont rempli le questionnaire entre minuit et 8h du matin.

Les délais de saisie sont également supprimés. Les résultats « bruts » sont même quasiment disponibles en temps réel. La diffusion de l'enquête, la collecte des données et le traitement de l'information sont donc plus rapides que dans les modes traditionnels (Caroline Bayart, 2009<sup>8</sup>).

### ***Pas de limite géographique***

Les enquêtes peuvent se dérouler facilement dans plusieurs pays ; elles peuvent être plus rapidement proposées en plusieurs langues. Par rapport aux enquêtes en face à face, cet avantage est incontestable. Elles permettent également de dépasser la barrière de plus en plus fréquente des digicodes à l'entrée des immeubles.

### **3. Avantage de l'auto-administré**

Comme le rappelle Edith D. de Leeuw (2008)<sup>9</sup>, il existe surtout deux formes dans la collecte des données : celles avec un enquêteur et celles auto-administrées. Le face à face et le téléphone relève de la première, le courrier et Internet de la seconde. Lorsque l'enquêteur est absent, il ne peut pas jouer un rôle dans le processus d'enquête, que celui-ci soit positif ou négatif. Il n'est pas là pour motiver les personnes interrogées, pour éventuellement clarifier certaines questions mais il ne risque pas non plus d'avoir une influence quelconque sur les réponses. Le mode auto-administré est plus anonyme et réduit la tendance que peuvent avoir les individus à vouloir se montrer sous leur meilleur jour face à un interlocuteur. De plus, ils peuvent choisir leur propre rythme de réponse.

Lorsque les questions sont posées oralement, cela demande une certaine capacité de mémoire, ce qui n'est pas le cas lorsque celles-ci sont écrites. Par téléphone, les répondants ont donc tendance à davantage choisir les dernières réponses plutôt que les premières dans les listes de modalités. Néanmoins, l'auteure signale que les utilisateurs

---

<sup>7</sup> Arnaud Gautier, François Beck, Valérie Blinneau, Isabelle Evrard et Julien Mousques, « Recours à une enquête web pour augmenter la participation à une enquête téléphonique : exemple du Baromètre santé médecins généralistes 2009 », Communication au sixième colloque francophone sur les sondages, organisé par la Société Française de Statistique à Tanger, les 23-25 mars 2010.

<sup>8</sup> Caroline Bayart, *Le potentiel du web pour les enquêtes de mobilité urbaine*, thèse Université Lumière Lyon II, décembre 2009

<sup>9</sup> Edith D. de Leeuw, « Choosing the Method of Data Collection », p. 113-135, in Edith D. de Leeuw, Joop J. Hox, Don A. Dillman, *International Handbook of Survey Methodology*, Taylor & Francis Group, 2008, 549 p.

d'Internet peuvent être plus impatientes et ne pas lire la liste entière de modalités proposées.

D'un autre côté, la familiarité avec le mode d'enquête joue aussi un rôle important. En général, les individus sont habitués aux interactions en face à face, tandis que dans l'auto-administré, certains peuvent n'être pas très à l'aise avec l'écriture par exemple. Edith D. de Leeuw note cependant que les nouvelles générations sont à l'aise avec Internet et que l'écriture « SMS » ou « chat » enlève les contraintes grammaticales au bénéfice d'une plus grande liberté d'expression, puisqu'il s'agit d'écrire comme l'on parle.

#### **4. Des nouvelles possibilités d'enquête**

Par rapport au courrier, qui est le mode qui lui ressemble le plus, Internet permet de nouvelles possibilités. Par exemple, les filtres sont automatiquement gérés, ce qui évite des confusions. De même, les items peuvent être proposés en rotation aléatoire, ce qui peut limiter certains biais. Anne-Marie Côté et Danielle Laroche (2009)<sup>10</sup> ont constaté, lors du recensement canadien de 2006, que les questionnaires remplis en ligne étaient plus complets que les questionnaires papier. Ces différences étaient, selon elles, en partie attribuables au profil des répondants (niveau d'éducation élevé) et aux particularités de l'application Internet (messages de validation, rubrique d'aide...).

Internet permet de mettre en place des procédures automatiques de validation, de rappels, d'alertes, de relances, ce qui peut être particulièrement utile pour les enquêtes reposant sur le recueil par un carnet alimentaire sur une semaine ou pour les enquêtes sur l'emploi du temps. De plus, c'est un mode qui permet d'intégrer facilement, rapidement et à moindre coût des images, des sons, des vidéos. Des aides et des instructions en ligne le rendent également plus complet qu'un simple questionnaire papier.

Plusieurs chercheurs affirment que les études en ligne sont plus efficaces pour traiter les sujets sensibles (avantage de l'auto-administré) mais les avis ne vont pas tous dans le même sens. De même, certains disent que les réponses ouvertes sont plus riches sur Internet (David R. Schaefer et Don A. Dillmann, 1998<sup>11</sup>, cités par Katja Lozar Manfreda et Vasja Vehovar, 2008<sup>12</sup>), d'autres assurent que la richesse des réponses dépend des

---

<sup>10</sup> Anne-Marie Côté, Danielle Laroche, « L'Internet : un nouveau mode de collecte au Recensement », Symposium 2008, Collecte des données: défis, réalisations et nouvelles orientations, Statistique Canada, 2009

<sup>11</sup> David R. Schaefer, Don A. Dillmann, « Development of a standard e-mail methodology : Results of an experiment », Public Opinion Quarterly, n°62, 1998, p.378-397

<sup>12</sup> Katja Lozar Manfreda, Vasja Vehovar, "Internet Surveys", p. 264-284, in Edith D. de Leeuw, Joop J. Hox, Don A. Dillman, International Handbook of Survey Methodology, Taylor & Francis Group, 2008, 549 p.

interviewés et sont peut-être plus abondantes lorsqu'il s'agit de personnes dont les attitudes sont plus marquées (Michael Bosnjak, 2001, cité par Katja Lozar Manfreda et Vasja Vehovar, 2008).

Un autre avantage consiste dans la possibilité de contrôler le processus, de mesurer la rapidité de la réponse. Il est désormais possible de savoir sur quelles questions les individus mettent le plus de temps à répondre, à partir de quel endroit se sont arrêtées les personnes qui ont abandonné le questionnaire en cours de route, ce qui était impossible à savoir dans les enquêtes courrier. Neil Malhotra (2008)<sup>13</sup> montre qu'il est important de comprendre la relation entre le temps de réponse et la qualité des données et d'inclure ce temps comme variable de contrôle dans les modèles statistiques.

Les qualités de réponse peuvent également être observées en combinant les études en ligne à des technologies un peu plus poussées. Mirta Galesic et al. (2008)<sup>14</sup>, par exemple, évoquent la technologie pour suivre le mouvement des yeux (Eye-tracking), devenue de plus en plus facile à utiliser ces dernières années. Ils ont enregistré le mouvement des yeux des répondants à un sondage en ligne. Cela permet de voir, entre autres, si les répondants choisissent plus souvent les premières réponses parce qu'ils les regardent plus longtemps que les dernières. Ils ont aussi regardé combien de temps les enquêtés passaient à rechercher la définition de certains mots, selon l'effort nécessaire pour les consulter, ainsi que d'autres comportements<sup>15</sup>...

---

<sup>13</sup> Neil Malhotra, 2008, Completion time and response order effects in web surveys, *Public Opinion Quarterly*, Volume 72, Issue 5, p. 914-934, <http://poq.oxfordjournals.org/content/72/5.toc>

<sup>14</sup> Mirta Galesic, Roger Tourangeau, Mick P. Couper, Frederick G. Conrad, 2008, Eye-tracking data, *New Insights on Response Order Effects and Other Cognitive Shortcuts in Survey Responding*, *Public Opinion Quarterly*, Volume 72, Issue 5, p. 914-934, <http://poq.oxfordjournals.org/content/72/5.toc>

<sup>15</sup> Nous reviendrons sur quelques-uns de leurs résultats dans la partie V de ce rapport.

## II. Processus d'échantillonnage et collectes des données pour les enquêtes en ligne

---

Comme le notent Mick P. Couper et Peter V. Miller (2008), la particularité des enquêtes en ligne, c'est leur diversité. Tandis que l'on peut aisément connaître les modalités de recrutement par téléphone (annuaire ou tirage aléatoire de numéros), courrier ou face à face, il est plus difficile de savoir comment l'internaute a été sélectionné (liens laissés sur certains sites, mails, *access panels*, pré-recrutement par courrier ou téléphone...). Cette seconde partie tente de faire le point sur les diverses méthodes mises en place.

### 1. Deux types d'échantillonnage

Mick P. Couper (2000)<sup>16</sup> distingue les enquêtes probabilistes des autres.

**Tableau 5 – Les différents types d'enquêtes en ligne**

Méthodes non-probabilistes	Méthodes probabilistes
Les sondages comme divertissement	Les enquêtes type « sortie des urnes » (tous les visiteurs d'un site Internet)
Les enquêtes en autosélection non restreintes	Les enquêtes basées sur des listes (mails d'étudiants d'une université...)
Les panels de volontaires	Les enquêtes en mode mixte (complétées par du face à face, du téléphone, du courrier...)
	Les panels d'internautes pré-recrutés
	Les panels avec recrutement probabiliste dans l'ensemble de la population (équipement fourni)

Source : d'après Mick P. Couper (2000)

Les enquêtes en ligne probabilistes sont effectuées sur un échantillon aléatoire d'individus qui correspondent à l'ensemble de la population cible. Chaque individu doit avoir exactement la même chance que les autres de participer à l'enquête. La représentativité n'est pas garantie puisque la non-réponse est possible (on ne peut pas forcer les individus à répondre) mais comme le processus de recrutement est maîtrisé<sup>17</sup>, la non-réponse peut être contrôlée, estimée et corrigée.

---

<sup>16</sup> Mick P. Couper, « Web Survey : A Review of Issues and Approaches », *The Public Opinion Quarterly*, Volume 64, N°4, décembre 2000, p. 464-494

<sup>17</sup> En effet, la population cible est ici connue. Il est donc possible de savoir qui répond, contrairement aux cas d'*access panels* de volontaires.



Les sondages empiriques s'opposent aux sondages probabilistes en ce sens où il n'est plus possible de déterminer a priori quelle est la probabilité qu'a chaque individu de la population d'appartenir à l'échantillon (Pascal Ardilly, 2006<sup>18</sup>). Le but est de se rapprocher au maximum d'un tirage aléatoire, en essayant de respecter la répartition constatée dans la population mère.

## **2. Les échantillons probabilistes**

Dans la mesure où tout le monde n'a pas accès à Internet, il existe deux approches pour constituer un échantillonnage probabiliste. La première revient à limiter l'échantillon aux internautes, ce qui réduit la population d'intérêt. Le résultat n'est alors absolument pas généralisable à l'ensemble de la population (c'est l'exemple de la bande dessinée citée en introduction de ce rapport). La seconde approche consiste en l'utilisation de méthodes alternatives (par exemple les enquêtes en mode mixte) pour atteindre un plus large échantillon de la population.

### ***Les enquêtes ciblées : « sortie des urnes », « listes » et « panels d'internautes pré-recrutés »***

Nous appelons « Enquêtes ciblées » les enquêtes qui ne sont pas censées être représentatives de la population générale française, mais d'une population réduite. On cherche à étudier certaines populations spécifiques. C'est le cas par exemple des visiteurs d'un site Web, de l'ensemble des étudiants dans un établissement ayant tous une adresse email, de l'ensemble des salariés d'une entreprise, etc. Il s'agit des deux méthodes mises en évidence par Mick P. Couper (2000) : les enquêtes type « sorties des urnes » et les enquêtes basées sur des listes. Ces méthodes sont souvent utilisées pour des enquêtes de satisfaction, d'évaluation d'un site... Elles ne doivent pas être comparées à des résultats en population générale (elles le sont d'ailleurs rarement).

Certaines enquêtes se font également à partir de panels d'internautes pré-recrutés, c'est-à-dire que des individus sont tirés aléatoirement dans la population générale (par exemple, avec une génération aléatoire de numéros de téléphone), les personnes qui sont équipées d'Internet et qui acceptent de répondre font alors partie du panel. Mick P. Couper (2000) les classe dans la catégorie des enquêtes probabilistes à cause du mode de recrutement qui les différencie, selon lui, des *access panels*. Des contrôles sont effectués pour s'assurer qu'aucune personne qui n'aurait pas été sélectionnée par tirage aléatoire ne puisse intégrer le panel. Il faut cependant avoir en tête qu'il s'agit alors d'un

---

<sup>18</sup> Pascal Ardilly, *Les techniques de sondage*, Paris, Editions Technip, 2006, 675 pages

panel d'internautes (et non d'un échantillon représentatif de la population générale) et mentionner clairement cet aspect méthodologique pour mettre en garde les lecteurs non-initiés.

### ***Les enquêtes en population générale : « panels probabilistes » et « modes mixtes »***

Certains instituts tentent de construire des panels en population générale recrutés sur un mode probabiliste. Mick P. Couper et Peter V. Miller (2008) évoquent le Knowledge Networks aux États-Unis et le CentERdata's MESS panel aux Pays-Bas.

Le Knowledge Networks<sup>19</sup> recrute les ménages en sélectionnant de façon aléatoire des adresses résidentielles à partir d'un fichier de l'ensemble des adresses, ce qui leur permet de toucher les populations difficilement accessibles par d'autres canaux (les personnes ne possédant qu'un téléphone mobile, les Afro-américains<sup>20</sup>, les Latinos-Américains et les jeunes adultes). Comme près de trois américains sur dix n'ont pas accès à Internet à domicile, on leur fournit un ordinateur netbook avec un accès Internet. Les ménages ne peuvent pas intégrer l'*access panel* de leur propre initiative et ils ne peuvent pas participer à plusieurs enquêtes par semaine.

Le CentERdata est un institut de recherche spécialisé dans l'analyse et la collecte de données, notamment à travers les études en ligne. En 2006, il a reçu un financement pour le projet MESS (Measurement and Experimentation in the Social Sciences), afin d'établir un nouveau panel en ligne : le LISS panel, basé sur un échantillon probabiliste<sup>21</sup>. Il est constitué de 5000 ménages, comprenant 8000 personnes. L'échantillon est réalisé à partir de la base d'adresses de l'Institut National De Statistiques et inclut les ménages n'ayant pas d'accès Internet (on leur fournit là aussi un ordinateur et un accès Internet). Toutes les personnes de l'échantillon sont contactées par des moyens traditionnels (courrier, téléphone ou face à face). Celles qui n'ont pas été incluses dans l'échantillon initial ne sont pas autorisées à participer, afin d'éviter les phénomènes d'autosélection (nous reviendrons sur ce point ultérieurement).

D'autres panels recrutés selon une approche probabiliste sont en train d'être développés. Selon Mick P. Couper (2000), ces panels sont la seule méthode qui a le potentiel pour obtenir un échantillon aléatoire de la population entière et pas seulement de ceux qui ont

---

<sup>19</sup> Pour plus de précisions, consulter leur site : <http://www.knowledgenetworks.com/knpanel/index.html>

<sup>20</sup> Selon la terminologie ayant cours outre-Atlantique.

<sup>21</sup> Pour plus de précisions, consulter leur site : <http://www.centerdata.nl/en/MESS>

un accès Internet. Cependant, les coûts sont assez élevés (c'est d'ailleurs la raison pour laquelle on en fait des panels).

Enfin, Mick P. Couper (2000) évoque la possibilité des enquêtes en mode-mixte, c'est-à-dire dans lesquelles les répondants ont le choix du mode de réponse (les personnes n'ayant pas de connexion Internet, par exemple, peuvent choisir le téléphone ou le courrier, etc.). Ces enquêtes ont fait et font encore l'objet de nombreuses publications : Don A. Dillman et al. (2009), Colm O'Muircheartaigh (2009)<sup>22</sup>, Reinder Banning and Barry Schouten (2009)<sup>23</sup>, Mark Pierzchala (2008)<sup>24</sup>, Edith D. de Leeuw (2005)<sup>25</sup>, etc. Le mélange des modes offre la possibilité de compenser les inconvénients d'un mode avec les avantages offerts par un autre, par exemple en combinant les modes moins chers avec des modes qui conduisent à des erreurs de sondage plus petits. Elles sont de plus en plus utilisées pour réduire la non-réponse, mais posent d'importants problèmes de comparabilité, puisque, comme le rappellent Annette Jäckle, Caroline Roberts et Peter Lynn (2008)<sup>26</sup>, les différents modes 1) fournissent un accès à différents types de personnes, 2) attirent différents types de répondants, 3) obtiennent des réponses différentes.

### **3. Les échantillons empiriques – ou non probabilistes**

#### ***Les sondages de « divertissement »***

Certains sondages sont déposés sur des sites Internet. Ceux-ci n'ont pas une prétention de représentativité. Ces sondages ne peuvent être en aucun cas généralisés et leurs résultats ne doivent pas être publiés en tant que résultats officiels. Ce ne sont que des « divertissements » représentatifs d'une partie de la population fréquentant un site Internet en particulier. Certains journaux utilisent ce type de sondages. Par exemple, le 13 décembre 2010, les personnes se rendant sur le site du *Figaro* pouvaient répondre par

---

<sup>22</sup> Colm O'Muircheartaigh, « Adapting to the Challenges of Household Survey Data Collection in a Multimode Environment », Modernisation of Statistics Production (MSP2009), 2-4 November 2009

<sup>23</sup> Reinder Banning and Barry Schouten, « A mixed-mode follow-up of panel refusers in the Dutch LFS », Discussion paper (09035), Statistics Netherlands, 2009

<sup>24</sup> Mark Pierzchala, « Expériences en matière d'enquêtes multimodales, Symposium 2008, Collecte des données : défis, réalisations et nouvelles orientations », La série des symposiums internationaux de Statistique Canada : recueil, 2008

<sup>25</sup> Edith D. de Leeuw, « To Mix or Not to Mix Data Collection Modes in Surveys », Journal of Official Statistics, Vol. 21, No. 2, 2005, pp. 233-255

<sup>26</sup> Annette Jäckle, Caroline Roberts, Peter Lynn, « Assessing the Effect of Data Collection Mode on Measurement », ISER Working Paper Series, No. 2008-08, February 2008

« oui » ou « non » à la question suivante : « Faut-il abandonner les 35h comme durée légale du temps de travail hebdomadaire ? ». A 17h27, 76,11% des votants avaient répondu « oui ». On imagine que les réponses auraient pu être différentes si la question avait été posée sur le site de *Libération*. Autre exemple : en juillet 2010, la presse reprenait les résultats d'une enquête selon laquelle en France, 74% des salariés se disaient prêts à travailler pendant leurs vacances. Or ces résultats étaient issus de votes enregistrés sur un site Internet fréquenté par un public au profil très spécifique (des cadres d'entreprises privées). Dans le premier cas comme dans le deuxième, les populations susceptibles de répondre sont donc très spécifiques et aucunement représentatives de la population générale.

### ***Les sondages en autosélection non restreints***

Cette approche utilise des invitations ouvertes sur des portails, des sites web fréquemment visités ou des sites dédiés aux enquêtes. Souvent, ces enquêtes n'ont pas un accès restreint, c'est-à-dire que tout le monde peut y répondre et ce, autant de fois qu'il le souhaite. C'est, en quelque sorte, un élargissement des sondages de divertissement. Elles leur ressemblent, à ceci près qu'elles se veulent généralistes, ce que ne font pas les premières. Mick P. Couper (2000) affirme que c'est sans doute là la forme la plus répandue d'enquêtes en ligne et selon lui, c'est également la forme la plus menaçante pour les instituts d'études parce qu'elles revendiquent parfois une légitimité scientifique. Les biais de sélection sont très forts. Par exemple, Jelke Bethlehem (2009<sup>27</sup>, 2010<sup>28</sup>) décrit un sondage réalisé pour élire le livre de l'année 2005 aux Pays-Bas. Une pré-sélection de livres était soumise au vote, mais les internautes pouvaient ajouter un livre de leur choix ne figurant pas dans la liste initiale. Plus de 90 000 personnes ont participé à l'enquête et le gagnant a été une traduction de la Bible (qui ne figurait pas dans la liste initiale), grâce à une campagne lancée par des sociétés bibliques, un radiodiffuseur chrétien, un journal chrétien, etc.

### ***Les Accés panels***

Un *accès panel* est une base de données de répondants volontaires qui sont sollicités pour des études en ligne ou pour des études par téléphone (Edith D. de Leeuw, 2008). Les avantages et inconvénients des panels ont été souvent décrits (par exemple, Daniel

---

<sup>27</sup> Jelke Bethlehem, « Peut-on établir des statistiques officielles à partir d'enquêtes en ligne reposant sur le principe de l'autosélection ? », Symposium 2008, Collecte des données : défis, réalisations et nouvelles orientations, La série des symposiums internationaux de Statistique Canada : recueils, 2009

<sup>28</sup> Jelke Bethlehem, « Selection Bias in Web Surveys », International Statistical Review, 78, 2, 2010, p. 161-188

Kasprzyk, Greg Duncan, Graham Kalton, M.P. Singh, 1989<sup>29</sup>). Ce qui est nouveau, c'est qu'Internet permet de les sélectionner à moindre coût. La façon dont les *access panels* sont constitués est un critère de qualité essentiel, selon s'ils reposent sur un échantillonnage aléatoire (par exemple, recrutement sur téléphone par tirage aléatoire du numéro) ou s'ils reposent sur l'autosélection (bannières ou invitations sur des sites). Les analyses statistiques n'ont de sens que sur les premiers. Dans les panels non probabilistes, l'échantillon reste de convenance, quel que soit le nombre de répondants.

Et pourtant, les recrutements en autosélection sont nombreux. Certains recrutent sur plusieurs centaines de sites par pays et considèrent donc avoir un processus de sélection très rigoureux mais il n'est pas possible de savoir si les personnes qui acceptent de répondre sur ces sites ont bien des comportements similaires aux personnes qui n'y répondent pas. L'argument avancé par les gérants d'*access panel* est que, comme toutes les catégories sociales sont présentes sur Internet, il suffit de bien constituer l'échantillon à partir de critères bien choisis. Mais l'échantillonnage ne correspond pas réellement à un échantillonnage par quotas où l'enquêteur tenterait tant bien que mal de suivre les consignes. Il s'agit d'un échantillonnage de volontaires, méthode « statistiquement » très difficile à défendre (Pascal Ardilly, 2006) : « *La seule exploitation raisonnable d'une enquête auprès de volontaires (...) consiste à utiliser des pondérations égales, au moins par sous-populations, c'est-à-dire à admettre une homogénéité de comportement entre volontaires et « non-volontaires » qui manifestement n'a aucune réalité* ».

Certes, l'accès aux enquêtes est très contrôlé (les adresses électroniques erronées et les doublons sont vérifiés, les panélistes fournissant des réponses contradictoires sont écartés, les critères comportementaux d'échantillonnage sont parfois très nombreux, etc.). Il reste que, la plupart du temps, l'entrée dans le panel est à l'initiative de l'enquêté : le risque de biais lié à l'autosélection est très élevé.

---

<sup>29</sup> Daniel Kasprzyk, Greg Duncan, Graham Kalton, M.P. Singh, « Panel surveys », John Wiley and Sons, New York, 1989

### III. Arbitrer entre les différentes erreurs

---

Rappelons que la théorie des sondages est un ensemble d'outils statistiques permettant l'étude d'une population au moyen de l'examen d'une partie de celle-ci. Pour cela, on réalise un plan de sondage qui est une méthode de sélection des individus au sein d'une population. On a vu que la méthode pouvait être probabiliste (les individus sont sélectionnés selon une méthode aléatoire) ou empirique (le choix est raisonné, afin d'avoir certaines catégories en proportions identiques dans l'échantillon et dans la population). On distingue quatre sources essentielles d'erreur dans les enquêtes par sondage, quel que soit le mode utilisé : l'erreur d'échantillonnage, l'erreur de couverture, l'erreur causée par la non-réponse et l'erreur de mesure. L'erreur totale étant la somme de ces quatre erreurs, il s'agit, dans cette partie, de mesurer l'importance de chacune de ces erreurs dans les enquêtes en ligne.

#### 1. L'erreur d'échantillonnage

L'erreur d'échantillonnage existe dans toutes les enquêtes par sondage et est inhérente au fait que l'on n'interroge qu'un échantillon et non pas toute la population. Elle est mesurable par le biais, la variance (ou l'écart-type) et/ou l'erreur quadratique moyenne. Pascal Ardilly (2006) montre que l'on peut y assimiler certaines erreurs résultant « des imperfections de la base de sondage » comme les doubles comptes et le vieillissement de l'information auxiliaire.

#### 2. L'erreur de couverture

L'erreur de couverture est celle qui est la plus souvent mentionnée lorsque l'on évoque les limites des enquêtes en ligne (Katja Lozar Manfreda et Vasja Vehovar, 2008). En effet, toute étude en ligne qui vise à représenter la population générale connaît un problème de couverture (sauf si l'on équipe les individus non-équipés) puisque la population n'est pas entièrement équipée. En France, près des deux tiers des personnes interrogées en 2009 pour l'enquête « Conditions de vie et Aspirations des Français » disent disposer, à leur domicile, d'une connexion à Internet (65% exactement), avec des inégalités d'équipements et de pratiques, mais nous reviendrons sur le cas français en deuxième partie de ce cahier de recherche. C'était le même problème au début des enquêtes par téléphone (et c'est un peu le même aujourd'hui pour les enquêtes par téléphone à partir d'annuaires, compte tenu de l'importance croissante des listes orange

et des listes rouges, notamment sur les mobiles). Sur Internet, les adresses sont multiples et non standardisées : il est donc extrêmement difficile aujourd'hui de générer des adresses au hasard, et le risque est grand, lorsqu'on se livre à ce type de pratique, d'être référencé en tant qu'émetteur de *spam* (publicité non sollicitée).

S'il existait une liste complète de tous les membres de la population cible et que cette liste comprenne les informations pour les contacter, il n'y aurait pas de différences entre les modes d'enquêtes (Edith D. de Leeuw, 2008). À l'inverse, lorsqu'il s'agit d'étudier la population générale et qu'il n'existe pas de registre de la population, le face à face a le plus grand potentiel puisqu'il suffit d'aller taper aux portes. Mais il est très coûteux. Le mode par téléphone est bien si le taux de couverture en téléphone est élevé. Pour être sûr d'inclure tout le monde, il est possible de faire un tirage aléatoire des numéros. Le courrier requiert un échantillon explicite avec noms et adresses. Souvent, on utilise les annuaires téléphoniques pour récupérer les adresses, mais on ne peut alors pas atteindre les personnes qui n'ont pas le téléphone ou qui figurent sur listes rouges ou orange.

Pour les études en ligne, nous pouvons séparer les cas selon les différents types vus dans la partie précédente. Dans le cas des sondages à partir de listes (établissements, étudiants), soit le problème du taux de couverture est inexistant (par définition), soit ce taux est très élevé. Dans le cas des panels d'internautes pré-récrutés, le risque de non couverture est également évité si l'on considère bien qu'il ne s'agit pas d'un panel représentatif de la population générale mais d'un panel d'internautes. En revanche, si les résultats sont extrapolés à la population générale, l'erreur de couverture est élevée. D'après Jelke Bethlehem (2009, 2010), son importance est déterminée par deux facteurs : la taille relative de la population sans Internet (l'erreur diminue donc à mesure que la couverture Internet augmente) et le contraste entre les moyennes de la population Internet et de la population sans Internet (plus la moyenne diffère, plus l'erreur est importante).

Dans le cas des panels Internet recrutés aléatoirement dans toute la population, l'avantage est que l'information sur les non-répondants peut être obtenue à chaque étape, permettant un examen détaillé des erreurs de non réponses et des biais d'attrition (Mick P. Couper, 2000). Les données peuvent alors être pondérées pour compenser l'erreur de couverture, de non-réponse et d'échantillonnage. Les probabilités de sélection sont connues à chaque étape.

Quand il n'existe pas de procédure propre de recrutement, on ne peut pas corriger les biais (Dirk Sikkel, 2008<sup>30</sup>). Par exemple, avec des participants volontaires recrutés sur un

---

<sup>30</sup> Dirk Sikkel, « Panel Surveys », p. 479-499, in Edith D. de Leeuw, Joop J. Hox, Don A. Dillman, International Handbook of Survey Methodology, Taylor & Francis Group, 2008, 549 p.

site Internet, il n'est pas possible de connaître les probabilités d'inclusion pour chacun des participants. On est alors en présence d'autosélection. Jelke Bethlehem (2009, 2010) propose une formule pour calculer l'erreur d'autosélection, ce qui permet de comparer les erreurs éventuelles de divers plans d'enquête. Il compare deux enquêtes, l'une issue d'un échantillonnage probabiliste et l'autre à partir de volontaires sur Internet. Il conclut que l'erreur d'autosélection de l'enquête en ligne (malgré les 170 000 répondants) peut être 13 fois supérieure à celle de l'enquête probabiliste de moindre envergure. Néanmoins, Jelke Bethlehem fait de fortes hypothèses dans le calcul de son erreur (sa formule dépend du coefficient de corrélation, c'est-à-dire du taux de corrélation entre la participation à l'enquête et le comportement des individus, mais ne connaissant pas ce dernier, il lui alloue la plus forte valeur possible). Dans la mesure où ce coefficient de corrélation est majoré, le résultat est évidemment très défavorable pour l'enquête en ligne.

Des études menées aux États-Unis et au Canada ont montré que les populations ayant accès à Internet diffèrent des populations sans accès à Internet (Gera E. Nagelhout et al., 2010)<sup>31</sup>, les premières étant plus jeunes, plus souvent employées, faisant moins souvent partie d'un groupe minoritaire, ayant plus d'éducation, et ayant un revenu plus élevé [Edith D. de Leeuw (2005), Scott Fricker et al. (2005)<sup>32</sup>, Cunningham et al.(2006)<sup>33</sup>, Jacqueline Stoddard et al. (2006)<sup>34</sup>]. Gera E. Nagelhout et al. (2010) ont comparé une étude en ligne (l'International Tobacco Control - ITC) à une étude téléphonique, sur le thème de la consommation de tabac aux Pays-Bas<sup>35</sup>. Les Pays-Bas

---

<sup>31</sup> Gera E Nagelhout, Marc C Willemsen, Mary E Thompson, Geoffrey T Fong, Bas van den Putte, Hein de Vries, « Is web interviewing a good alternative to telephone interviewing? Findings from the International Tobacco Control (ITC) Netherlands Survey », BMC Public Health, 2010

<sup>32</sup> Scott Fricker, Mirta Galesic, Roger Tourangeau, Ting Yan, « An experimental comparison of web and telephone surveys », Public Opinion Quarterly, 69, 2005, p.370-392

<sup>33</sup> John Cunningham, Peter Selby, Kypros Kypri, Keith Humphreys, « Access to the Internet among drinkers, smokers and illicit drug users: Is it a barrier to the provision of interventions on the World Wide Web? » Informatics for Health and Social Care, 31, 2006, p. 53-58

<sup>34</sup> Jacqueline Stoddard, Erik Augustson, « Smokers Who Use Internet and Smokers Who Don't, Data from the Health Information and National Trends Survey (HINTS) », Nicotine & Tobacco Research, 8(Suppl 1), 2006, p. 77-85.

<sup>35</sup> Les répondants web ont été tirés de TNS NIPObase, une grande base de données en tirage probabiliste avec plus de 140 000 répondants néerlandais potentiels sur le web qui ont indiqué leur volonté de participer à une recherche sur une base régulière. On ne peut pas demander à participer, ce qui entraîne un faible nombre de répondants professionnels et inattentifs. Les membres du panel Web sont recrutés par téléphone ou par courrier, mais pas par Internet. Une procédure de sélection choisit les répondants qui satisfont aux critères d'inclusion pour l'enquête ITC Pays-Bas : être un fumeur mensuel de cigarettes et avoir fumé au moins 100 cigarettes dans sa vie. Les quotas sur le sexe, la région géographique, la taille du ménage et l'éducation ont été déterminés à partir d'une enquête néerlandaise menée en 2007 sur les habitudes de fumer (DCSSH) avec un échantillon représentatif des fumeurs néerlandais âgés de 15 ans et plus. Une annonce e-mail au sujet de l'enquête en ligne a été envoyée à 2 331 répondants fumeurs de TNS NIPObase (cette base contient plus de 80 000 répondants). Un échantillon a été tiré au hasard à partir des numéros de téléphone de la base de données. Les deux derniers chiffres de chaque numéro tiré au hasard ont été automatiquement remplacés par deux chiffres aléatoires. Au total, 28 563 numéros ont été appelés par les enquêteurs. Lorsque l'intervieweur joint une personne au téléphone, il demande combien de personnes dans le ménage répondent aux critères d'inclusion (personnes ayant fumé au moins une fois par mois et qui ont fumé au moins 100 cigarettes dans leur vie). Dans les foyers où plus d'une personne répondent aux critères, l'intervieweur sélectionne une personne selon la méthode de la date d'anniversaire.



sont un pays où une très grande partie de la population a accès à Internet (85%). Néanmoins, il est prouvé que les fumeurs sont moins susceptibles d'avoir accès à Internet et que ceux qui y ont accès l'utilisent moins souvent que les non-fumeurs. Il a également été constaté que les fumeurs sont plus susceptibles de préférer les enquêtes en ligne. Pour évaluer la couverture de l'échantillon en ligne, les personnes interrogées par téléphone qui avaient accès à Internet ont été comparées aux répondants au sondage téléphonique sans accès à Internet. Pour évaluer la couverture de l'échantillon téléphonique, les répondants sur Internet qui ont une ligne de téléphone fixe ont été comparés aux répondants sur Internet sans ligne fixe. Les auteurs concluent qu'il existe des différences de qualité des données, mais elles sont petites. Comme l'accès à Internet ne cesse d'augmenter et que les téléphones fixes ne cessent de diminuer, le degré d'erreur de couverture dans les enquêtes en ligne devrait diminuer tandis que le degré d'erreur de couverture dans les enquêtes téléphoniques devrait augmenter. Néanmoins, les auteurs notent que leur étude peut avoir surestimé le problème de l'erreur de couverture puisqu'ils ne retenaient que l'accès Internet à la maison comme indicateur de la possibilité de participer à des sondages en ligne.

Nombreuses sont les expériences de ce type. Thorsten Faas (2004)<sup>36</sup>, par exemple, compare les résultats de trois enquêtes réalisées avec des méthodologies très différentes lors de la préparation des dernières élections fédérales allemandes. La première utilise un échantillon représentatif de la population allemande, la deuxième est une enquête représentative des utilisateurs d'Internet, et la troisième s'appuie sur les répondants d'une enquête en ligne en libre accès. La comparaison des trois révèle des différences considérables sur la structure des populations échantillonnées selon l'âge, le sexe et l'éducation, mais aussi pour les variables clés (intention de vote, intérêt pour la politique). De même, Bryan D. Rookey et al. (2008)<sup>37</sup> comparent aux États-Unis une enquête en ligne à une enquête par courrier. Ils montrent que les panélistes Internet sont plus susceptibles d'être mariés, travaillant à temps plein et plus jeunes que les panélistes courrier. Au demeurant, les panélistes assignés aux deux modes sont de toute façon différents de la population générale, avec des revenus plus élevés, des plus hauts diplômes, étant plus susceptibles d'être mariés et de sexe masculin.

---

<sup>36</sup> Thorsten Faas, « Online or Not Online?, A Comparison of Offline and Online Surveys Conducted in the Context of the 2002 German Federal Election », *Bulletin de méthodologie sociologique* n°82, 2004, p. 42-57, <http://bms.revues.org/index1086.html#ftn2>

<sup>37</sup> Bryan D. Rookey, Steve Hanway, Don A. Dillman, « Does a probability-based household panel benefit from assignment to postal response as an alternative to Internet-only? », *Public Opinion Quarterly*, Vol. 72, No. 5, 2008, p. 962-984

### 3. L'erreur de non-réponse

On distingue l'erreur de non-réponse de l'erreur de couverture par le fait que les non-répondants sont sélectionnés dans l'échantillon mais non mesurés, tandis que les non-couverts n'ont aucune chance d'être dans l'échantillon (on ne connaît pas leur adresse ou leur numéro de téléphone) et donc ne peuvent pas être mesurés (Edith D. de Leeuw, 2008). Il existe trois sources de non-réponses : le contact ne peut pas être fait, le refus est explicite, la personne est en incapacité de coopérer. Edith D. de Leeuw rappelle que la réponse à une enquête peut être influencée par plusieurs facteurs : le sujet du questionnaire, sa longueur, l'organisme qui réalise ou commande l'enquête, le nombre de rappels ou d'autres caractéristiques de conception. En général, la non-réponse augmente dans le temps. Edith D. de Leeuw et Wim De Heer (2002)<sup>38</sup> ont montré que le taux de réponse a décliné internationalement. Ils ont analysé les données de 16 pays sur la période 1980-1998 et ont mis en évidence l'augmentation du non contact et de la non-réponse. Les taux de réponse sont rarement étudiés dans les enquêtes sur Internet. Les premiers résultats de telles probabilités basées sur des enquêtes en ligne sont prometteurs (Vasja Vehovar, Zenel Batagelj, Katja Lozar Manfreda, Metka Zatelel, 2002<sup>39</sup>) bien que des études montrent que les taux sont généralement plus faibles qu'avec les autres modes d'enquête (Hisako Matsuo et al. 2004<sup>40</sup>, Mick P. Couper 2000, Michael D. Kaplowitz et al., 2004<sup>41</sup>). En général, les *access panels* ont des meilleurs taux de réponse, mais il s'agit de volontaires. Pour calculer le taux de réponse pour les *access panels*, il faudrait donc aussi considérer la non-réponse initiale pour composer le panel. Pour les enquêtes non scientifiques sur Internet où l'on reçoit une invitation à remplir le questionnaire sur un portail, les taux de réponse ne peuvent pas être calculés.

Les méthodes ne diffèrent pas seulement dans les non-réponses mais aussi dans l'opportunité d'atteindre les personnes qui ne sont pas à domicile et les personnes réticentes. Selon le mode utilisé, on peut avoir plus ou moins d'information sur les non-répondants et sur le pourquoi et le comment de la non-réponse. Le face à face permet davantage de convaincre ou d'obtenir des informations sur les non-répondants. Dans le téléphone, il est plus difficile de convaincre ; par contre, il est plus facile de tenter à

---

<sup>38</sup> Edith D. de Leeuw et Wim De Heer, Trends in household survey nonresponse : A longitudinal and international comparison. In Robert M. Groves, Don A. Dillman, John L. Eltinghe, Roderick .J.A. Little (Eds.), Survey nonresponse, New York: Wiley, 2002.

<sup>39</sup> Vasja Vehovar, Zenel Batagelj, Katja Lozar Manfreda, Metka Zatelel, Nonresponse in web surveys, in Robert M. Groves, Don A. Dillman, John L. Eltinghe, Roderick .J.A. Little (Eds), Survey Nonresponse, New York: Wiley, 2002.

<sup>40</sup> Hisako Matsuo, Kevin P. McIntyre, Terry Tomazic, Barry Katz, The online survey : Its contributions and potential problems, ASA Proceedings of the Section on Survey Research Methods, 27, 2004, p. 3998-4000

<sup>41</sup> Michael D. Kaplowitz, Timothy D. Hadlock, and Ralph Levine, "A Comparison of Web and Mail Survey Response Rates", Public Opinion Quarterly, 68, 2004, p.98-101

nouveau de joindre ceux qui n'étaient pas à domicile au moment de l'appel (et moins coûteux). Puisqu'il n'y a pas d'enquêteur par mail et courrier, les possibilités de persuasion sont faibles. Des recherches ont montré que la personnalisation des courriers a des effets positifs sur le taux de réponse. De plus, ceux-ci ont moins de problèmes quant à l'accès des personnes hors domicile. Mais il n'est pas possible de détecter des raisons de non-réponses, à l'exception des *access panels* qui connaissent les caractéristiques des non-répondants. Cela permet de connaître la non-réponse selon les sujets. Par contre, l'information reste à néant sur la non-réponse pré-panel.

Dans les sondages du type « sortie des urnes » ou « liste », si les répondants sont censés être représentatifs des visiteurs du site (ou de l'établissement...), l'erreur de non-réponse peut être importante puisque ceux qui acceptent de répondre sont peut-être différents des autres. La couverture peut donc être correcte mais il s'agit d'être très prudent quant au taux de réponse. Dans le cas des études sur les visiteurs d'un site Internet, les taux de réponse peuvent être très faibles et rien ne garantit que les comportements des répondants et des non-répondants soient similaires. Dans le cas des panels pré-recrutés d'utilisateurs Internet, si le taux de couverture n'est pas une préoccupation majeure, la non-réponse peut arriver à plusieurs étapes : lors du recrutement (pas de renseignement sur les personnes qui refusent de participer : à ce stade cependant, même effet de non-réponse entre ce mode et le mode téléphonique), mais elle peut être augmentée par la suite si par exemple l'interviewé déclare ne pas avoir de connexion Internet ou d'adresse mail valide. Enfin, parmi ceux qui acceptent dans un premier temps, il n'est pas sûr qu'ils répondent lors de l'envoi du questionnaire. Toutefois, dans ce dernier cas, on peut estimer ce biais de non-réponse et éventuellement tenter de le corriger. Le taux de non-réponse est donc mesurable ici, pas comme dans le cas non probabiliste de bénévoles.

Pour étudier l'erreur de non-réponse, Gera E. Nagelhout et al. (2010)<sup>42</sup> ont comparé les répondants de deux échantillons (enquête en ligne et enquête téléphonique) aux caractéristiques démographiques des répondants fumeurs de Statistics Netherlands (CBS) de 2006 à 2007. Les statistiques du CBS sont les statistiques officielles nationales des Pays-Bas. Le CBS mène des entretiens individuels sur un échantillon de 10 000 personnes. La non-réponse associée aux différentes erreurs de couverture a été mise en évidence dans les deux modes : en ligne et en téléphone, des différences s'observent à l'égard de l'état matrimonial et du niveau d'éducation, mais pas en ce qui concerne le sexe. C'est dans l'enquête téléphonique qu'a été mise en évidence la plus grande

---

<sup>42</sup> Gera E Nagelhout, Marc C Willemsen, Mary E Thompson, Geoffrey T Fong, Bas van den Putte, Hein de Vries, « Is web interviewing a good alternative to telephone interviewing? Findings from the International Tobacco Control (ITC) Netherlands Survey », BMC Public Health, 2010

différence avec la population néerlandaise des fumeurs à l'égard du niveau d'éducation, avec 12% de répondants instruits en moins que dans la population néerlandaise de fumeurs.

Selon Robert M. Groves (2006), il existe cinq situations possibles de non réponse : 1) la propension à répondre est distincte et non corrélée avec le sujet de l'enquête ; 2) la raison est commune au sujet de l'enquête et à la non-réponse, par exemple si le sujet est « intéressé » (enquête auprès des utilisateurs d'un lac de loisirs avec des taux de réponse plus élevés pour les utilisateurs dont les maisons sont sur le lac). Cependant, on peut éliminer l'effet avec une pondération ; 3) il existe une raison de causalité directe (dans les enquêtes Emploi du temps, concernant le temps passé à la maison, ceux qui y sont peu ont moins de chance d'être interrogés, ce qui conduit à une surestimation du temps passé à la maison) ; 4) il existe un lien entre la non-réponse et l'erreur de mesure. Le niveau de propension à répondre détermine l'ampleur d'une erreur de mesure ; 5) situation sans erreur de non-réponse, malgré des liens évidents avec la variable d'enquête. La prescription la plus courante pour les chercheurs étant de réduire les taux de non-réponse, il n'est pas rare de repondérer, d'utiliser des scores de propension, des méthodes de calage pour réduire les erreurs de non-réponse corrélées avec les variables d'enquête et les erreurs de couverture. Ces techniques d'ajustement nécessitent des hypothèses sur les réponses des répondants et des non-répondants (ce qui est invérifiable). Pour les deux premiers modèles, les difficultés sont moindres. Pour les trois derniers, cela est plus délicat. Nous y reviendrons dans la partie suivante.

Pour évaluer les erreurs de non-réponse, plusieurs méthodes sont possibles (Robert M. Groves, 2006<sup>43</sup>). Tout d'abord, il est possible de comparer les taux de réponse entre les sous-groupes (si les taux de réponse sont similaires dans les sous-groupes, les chercheurs disent qu'il n'existe aucune preuve d'erreur de non-réponse, ce qui est une hypothèse extrêmement forte, voire intenable selon l'auteur). Ensuite, il est possible de comparer avec des estimations semblables provenant d'autres sources (point fort : l'estimation est indépendante de l'enquête ; point faible : les variables clefs du sondage n'existent souvent pas dans l'enquête externe). On peut également étudier la variation dans l'enquête actuelle (non réponse des études de suivi), mais l'information ne sera pas directe sur les non-répondants à l'enquête. Enfin, des pondérations peuvent mesurer la quantité de l'erreur qui pourrait être éliminée par un ajustement post-enquête. Chacune de ces techniques est intéressante mais certaines peuvent augmenter l'erreur de couverture, et certaines sont limitées sur l'information de l'erreur. Comme les taux de

---

<sup>43</sup> Robert M. Groves, Nonresponse rates and nonresponse bias in household surveys, *Public Opinion Quarterly*, Vol. 70, No. 5, Special Issue, 2006, p. 646-675

réponse diminuent, les chercheurs sont confrontés de plus en plus à mesurer les erreurs de non-réponse afin d'informer sur l'évaluation des estimations de l'enquête. En raison de la diversité des propriétés des différentes techniques, Robert M. Groves (2006) explique qu'il est sage d'étudier ces erreurs en utilisant plusieurs méthodes simultanément. L'auteur résume les résultats d'un grand nombre d'études qui présentent des estimations de l'erreur de non-réponse. Dans les 30 articles qu'il étudie, le taux moyen de non-réponse est de 35%. Le pourcentage moyen absolu de l'erreur relative est de 8,7%. La corrélation entre le taux de non-réponse et l'erreur de non-réponse est de 0,33. L'auteur en conclut que le taux de non-réponse n'est pas un très bon indicateur de l'ampleur de l'erreur de non-réponse. Les deux dernières décennies ont clairement montré que les chercheurs ont le pouvoir de modifier la propension à répondre des personnes qu'ils échantillonnent. Dans ce cas, il faut savoir si une augmentation des taux de réponse est bonne ou mauvaise.

#### **4. L'erreur de mesure, d'observation**

Cette erreur tient au fait que « la valeur que l'on recueille lors de l'enquête peut être une valeur  $Y_i^*$  différente, voire très différente, de la vraie valeur  $Y_i$  qui est attachée à l'individu  $i$  » (Pascal Ardilly, 2006). Elle peut être due à la bonne ou mauvaise foi des répondants, à l'influence de l'enquêteur, à des erreurs de saisie, de codification, de programme informatique, de mauvaise compréhension, de formulation ou d'ordre des questions...

##### ***L'outil informatique***

Avec les sondages en ligne, des erreurs propres à l'outil informatique peuvent apparaître : selon la définition de l'écran de chacun, tout le monde n'a peut-être pas la même aisance pour dérouler le questionnaire ; il en va de même selon le débit de la connexion, selon la dextérité informatique de la personne interrogée.

##### ***La longueur du questionnaire***

Les durées des enquêtes n'ont pas le même effet selon le mode utilisé, ce qui peut également entraîner des erreurs de mesure. Selon Edith D. de Leeuw (2008), le face-à-face est le mode qui a le plus de potentiel : il peut durer plus longtemps que les autres

modes. D'après Ronald Czaja et Johnny Blair (2005)<sup>44</sup>, les enquêtes par Internet doivent être relativement courtes : 10-15 minutes. Mais des enquêtes plus longues sur Internet peuvent être mises en place avec succès pour des cibles spécifiques, des membres de panels.

Internet permet de savoir combien de temps les interviewés passent sur chaque réponse. C'est ainsi que l'on peut par exemple mettre en évidence des effets de fatigue : on peut parfois observer une variance des réponses plus faible en fin de questionnaire.

### ***La durée de l'enquête dans sa globalité***

Pour les enquêtes non probabilistes (visiteurs d'un site), l'un des problèmes clé (Mick P. Couper, 2000) est d'identifier le meilleur moment pour inviter le visiteur à remplir le questionnaire. Poser la question dès le début peut permettre de toucher à la fois ceux qui restent longtemps sur le site et les autres. De même, il faut compter des délais d'enquête assez longs pour laisser le temps aux internautes non réguliers de répondre. Certes, Internet permet de réduire les délais de recueil des informations, mais il peut être intéressant de recueillir l'avis des répondants « tardifs », qui peuvent présenter des profils spécifiques.

### ***Absence/présence de l'enquêteur***

En comparant le courrier d'un côté et le téléphone et face à face de l'autre, Edith D. de Leeuw avait déjà montré en 1992<sup>45</sup> qu'il est plus difficile d'obtenir les réponses avec le mode courrier. En revanche, lorsque les individus répondent, les résultats sont de meilleure qualité, surtout s'il s'agit de questions sur des sujets sensibles. D'après l'auteur, les questionnaires auto-administrés, lorsque les personnes interrogées ont le contrôle, apportent des réponses plus consistantes. Cela dépend néanmoins du type de questions : pour les petites questions ouvertes, les différences sont faibles. Pour les questions plus complexes, l'assistance est nécessaire pour aller plus loin dans les détails. L'absence d'enquêteur rend les enquêtes en ligne beaucoup moins flexibles quand les questionnaires sont complexes (Edith D. de Leeuw, 2008).

---

<sup>44</sup> Ronald Czaja, Johnny Blair, *Designing surveys, A guide to decisions and procedures*, 2005, Thousand Oaks : Sage,  
[http://books.google.fr/books?id=GCWZVZXEf2gC&printsec=frontcover&dq=Ronald+Czaja+et+Johnny+Blair&source=bl&ots=dFqLgMSvIa&sig=e2y3NdDBmQwaNepWuB-47Q9s3B8&hl=fr&ei=7noLTc\\_5JcicOrWziIUJ&sa=X&oi=book\\_result&ct=result&resnum=1&sqi=2&ved=0CBgQ6AEwAA#v=onepage&q&f=false](http://books.google.fr/books?id=GCWZVZXEf2gC&printsec=frontcover&dq=Ronald+Czaja+et+Johnny+Blair&source=bl&ots=dFqLgMSvIa&sig=e2y3NdDBmQwaNepWuB-47Q9s3B8&hl=fr&ei=7noLTc_5JcicOrWziIUJ&sa=X&oi=book_result&ct=result&resnum=1&sqi=2&ved=0CBgQ6AEwAA#v=onepage&q&f=false)

<sup>45</sup> Edith D. de Leeuw, « Data quality in mail, telephone and face-to-face surveys », Amsterdam TT-Publikaties, 1992, Retrieved from <http://www.xs4all.nl/~edithl>

## **La désirabilité sociale**

La désirabilité sociale reflète la tendance des personnes interviewées à vouloir donner une bonne image d'eux-mêmes ou à répondre de la façon la plus proche possible de ce qu'ils considèrent être la norme sociale. En général, on considère que la désirabilité sociale est plus faible en mode auto-administré, notamment si le sujet d'enquête est sensible. Selon Edith D. de Leeuw (2008), les modes avec enquêteurs produisent plus de désirabilité sociale et moins de réponses consistantes mais plus de détail aux questions ouvertes. Les différences entre le face à face et le téléphone sont faibles, le face à face étant légèrement meilleur que le téléphone. Frauke Kreuter et al. (2008)<sup>46</sup>, en voulant savoir si les différents modes d'auto-administration varient dans leurs effets sur ce qui est socialement souhaitable de répondre, concluent que l'enquête en ligne a augmenté la communication d'informations sensibles par rapport aux autres modes classiques<sup>47</sup>.

Cependant, Herschel Knapp et Stuart A. Kirk, 2003<sup>48</sup>, eux, n'avaient trouvé aucune différence entre les modes. De même, Didier Fripiat et Nicolas Marquis (2010)<sup>49</sup>, dans leur revue de littérature, montrent que d'autres auteurs tempèrent cette analyse car la confiance serait moindre dans l'outil Internet, avec un sentiment de moindre maîtrise des données personnelles qui y sont renseignées.

## **La conception du questionnaire**

De toute façon, comme le note Stéphane Ganassali et Jean Moscarola (2002)<sup>50</sup>, le mode n'est pas le seul facteur à prendre en compte, la façon dont sont présentées les questions est encore plus déterminante. Dans un travail de recherche du CRÉDOC, Claire Piau (2004)<sup>51</sup> a montré par exemple que le passage de questions ouvertes à des questions fermées conduit à des variations de résultats parfois très fortes et que les variations de réponses causées par les changements de formulation sont encore plus

---

<sup>46</sup> Frauke Kreuter, Stanley Presser, Roger Tourangeau, « Social desirability bias in CATI, IVR, and Web surveys, The effects of mode and question sensitivity », *Public Opinion Quarterly*, Vol. 72, No. 5, 2008, p. 847-865

<sup>47</sup> Frauke Kreuter et al. examinent les effets de mode sur la déclaration des informations potentiellement sensibles par un échantillon de diplômés universitaires. Les répondants ont été assignés au hasard à l'un des trois modes de collectes (données téléphoniques assistées par ordinateur (CATI), reconnaissance vocale interactive (RVI), et Web). Ils ont ensuite été interrogés sur les deux attributs désirables et indésirables de leur expérience scolaire. Les dossiers de l'Université ont été utilisés pour évaluer l'exactitude des réponses et examiner les différences dans les erreurs de non réponse selon le mode.

<sup>48</sup> Herschel Knapp, Stuart A. Kirk, « Using Pencil and Paper, Internet and Touch-Tone Phones for Self-Administered Surveys: Does Methodology Matter? », *Computers in Human Behavior* 19, 2003, p. 117-34

<sup>49</sup> Didier Fripiat, Nicolas Marquis, « Les enquêtes par Internet en sciences sociales : un état des lieux », *Population*, Volume 65, n°2, 2010, p. 309-338

<sup>50</sup> Stéphane Ganassali, Jean Moscarola, « Protocoles d'enquête et efficacité des sondages par Internet », *Journées E-Marketing AFM / AIM - Nantes*, septembre 2002

<sup>51</sup> Claire Piau, *Quelques expériences sur la formulation des questions d'enquête à partir du matériau « Aspirations et Conditions de vie des Français »*, Cahier de recherche n°206, CRÉDOC, 2004, 65 p.

marquées chez les répondants les moins instruits, les moins favorisés. Ces personnes perçoivent encore plus la situation d'enquête comme un examen et s'efforcent souvent de trouver la « bonne » réponse. Les exemples sont nombreux à ce sujet. Sur Internet, l'apport visuel est un avantage de taille. Mais il peut aussi conduire à de lourdes erreurs : Kenyon, Couper et Tourangeau (2001)<sup>52</sup> montrent que la réponse à la question « Combien de fois avez-vous pris un repas hors de votre domicile depuis le premier mars dernier » varie significativement selon que le visuel l'accompagnant était un *fastfood* ou un restaurant gastronomique.

La présence ou non du « Ne sait pas » est également un choix important à faire. En face à face ou au téléphone, l'enquêteur, souvent, ne propose pas l'option, mais peut la cocher si l'interviewé ne parvient pas à choisir. La faire figurer sur Internet peut exagérer les réponses évasives, mais cacher l'option peut aussi avoir des effets pervers. On retrouve ici le reproche classique fait par Pierre Bourdieu aux sondages lorsqu'il écrit que « *l'opinion publique n'existe pas* ».

Vera Toepel et al. (2008)<sup>53</sup> ont souhaité tester les différences de formats, sur les panels Center Data et LISS. Par exemple, si plusieurs éléments sont réunis sur un seul écran, cela nécessite moins d'actions physiques (frappes au clavier ou clics de souris). Ceci suggère que plus les informations sont nombreuses sur un écran unique, plus le temps nécessaire pour remplir le questionnaire pourra être réduit. Cependant, aucune différence significative dans la durée moyenne entre formats (1, 5, et 10 éléments par écran) n'a été mise en évidence.

Pour appréhender l'erreur de mesure, Gera E. Nagelhout et al. (2010) ont comparé les échantillons web et téléphoniques sur le tabagisme, à l'aide des tests du Chi<sup>2</sup> et de Mann-Whitney (attitude envers les fumeurs, intention d'arrêter de fumer, auto-efficacité pour cesser de fumer), ainsi que des t-tests et des Levene F-tests (nombre de cigarettes fumées par jour, nombre de tentatives précédentes de cesser de fumer, et durée avant la première cigarette après le réveil). Ces analyses ont été menées en excluant les répondants qui n'avaient pas de téléphone fixe ou qui n'avaient pas accès à Internet, afin de contrôler les différences d'erreur de couverture. En outre, des analyses de régression ont été mises en place pour tester les effets de modes sur les réponses et contrôler le sexe, l'état matrimonial, le niveau d'éducation et l'âge. Les interactions entre le mode et le niveau d'instruction ont été testées. Les auteurs ont mis en évidence des différences

---

<sup>52</sup> Kristin Kenyon, Mick P. Couper, Roger Tourangeau, « Experiments on Visual Effects in Web Surveys », AAPOR Annual Conference, Montreal, Canada, 2001.

<sup>53</sup> Vera Toepel, Marcel Das, Arthur Van Soest, « Effects of design in web surveys comparing trained and fresh respondents », *Public Opinion Quarterly*, Vol. 72, No. 5 2008, p. 985-1007



sur trois des six variables liées au tabagisme. Bien que ces différences soient significatives statistiquement, elles restent mineures selon les auteurs : l'écart moyen de l'utilisation des catégories de réponses entre le web et l'enquête téléphonique a été de 4%. Les répondants sur le Web ont utilisé l'expression « Ne sait pas » plus souvent que les répondants par téléphone, certainement parce que les enquêteurs téléphoniques ne proposent pas cette option à voix haute. Les auteurs émettent aussi une autre hypothèse : il n'est pas impossible que les répondants sentent moins de pression sur Internet pour donner une réponse quand ils n'en ont aucune idée. En outre, les répondants téléphoniques ont été plus négatifs sur le tabagisme mais peut-être à cause d'une désirabilité sociale plus forte.

### ***Le satisficing***

D'après Katja Lozar Manfreda et Vasja Vehovar (2008), l'emploi de mots spécifiques n'introduit pas plus de biais qu'avec les autres modes. Par contre, certains ont montré que les questions sur Internet doivent être plus concises et bien présentées parce que les internautes lisent rarement le texte attentivement mais plutôt en le survolant (Lorenz Gräf, 2002<sup>54</sup>). Le risque est alors de faire face au « satisficing », ce que Didier Frippiat et Nicolas Marquis (2010)<sup>55</sup> traduisent par « faible implication du répondant ». Ce phénomène qualifie le fait que les répondants optent pour la première des réponses qui leur paraît acceptable en ne cherchant pas plus loin.

L'hypothèse est que répondre à une enquête demande un grand travail cognitif : il faut comprendre et interpréter le sens des questions, faire appel à sa mémoire, émettre un jugement... Bien que de nombreux répondants effectuent ce travail sans difficulté, d'autres répondants tendent à prendre des raccourcis cognitifs pour réduire l'effort nécessaire. Cette implication du répondant peut être influencée par trois facteurs : la motivation, la capacité et la difficulté de la tâche (Dirk Heerwegh, Geert Loosveldt, 2008<sup>56</sup>).

La crainte avec un sondage par Internet, c'est que l'effet de « *satisficing* » conduise plus souvent les répondants à choisir la réponse « Ne sait pas » et qu'ils se différencient moins sur les échelles d'évaluation. Dirk Heerwegh et Geert Loosveldt (2008) ont mis en place une expérience sur un échantillon d'étudiants de première année de la Katholieke

---

<sup>54</sup> Lorenz Gräf, Assessing Internet questionnaires: The online pretest lab. In: Bernad Batinic, Ulf Dietrich Reips, Michael Bosnjak, Andreas Werner, eds, Online social sciences, Seattle, WA: Hogrefe & Huber, 2002

<sup>55</sup> Didier Frippiat, Nicolas Marquis, « Les enquêtes par Internet en sciences sociales : un état des lieux », Population, Volume 65, n°2, 2010, p. 309-338

<sup>56</sup> Dirk Heerwegh, Geert Loosveldt, « Face-to-face versus web surveying in a high-internet-coverage population, Differences in response quality », Public Opinion Quarterly, Vol. 72, No. 5 2008, p. 836-846

Universiteit Leuven. Les données suggèrent que les répondants au sondage web ont fait moins attention aux questions, car le sondage en ligne a été terminé plus rapidement que le sondage en face à face, ce qui impliquerait un moins bon niveau de qualité des données. Le taux moyen de « Ne sait pas » est effectivement plus élevé. Bien que la différence soit faible, le sondage en ligne a un taux de différenciation (nombre de points de différentes échelles d'évaluation utilisées) un peu plus faible que le face à face, principalement dû à une sélection accrue des réponses du milieu. De même, le taux de non-réponses est plus élevé dans le sondage en ligne. Cela suggère que les répondants au sondage en ligne sont plus enclins au « satisficing » que les répondants au sondage face à face.

### ***L'expérience des répondants***

Avec la prolifération des *access panels* dans les enquêtes en ligne, une question importante est de savoir si les individus qui ont pris l'habitude de répondre à des enquêtes ont des comportements différents des autres. C'est même là, selon Alexandre Sagakian (2010)<sup>57</sup> le point de départ de la remise en cause des *access panels* : en 2006, un institut lance la même étude administrée à deux échantillons identiques d'un même panel et les résultats sont différents pour le lancement d'un nouveau produit. Les plus expérimentés étaient plus négatifs sur le nouveau concept. L'auteur décrit alors trois types de « mauvais panélistes » : les fraudeurs (3 à 5%), les professionnels (5 à 20%) et les inattentifs (10 à 15%). Les premiers peuvent être détectés mais les seconds et troisième peuvent poser de sérieux problèmes dans l'interprétation des résultats.

Vera Toepel et al. (2008)<sup>58</sup> montrent également que les répondants formés semblent être plus porté au « satisficing »<sup>59</sup>. Cependant, les résultats divergent dans la littérature. Certains voient un biais, d'autres non. Le sens varie aussi : Terrence Coen et al. (2005)<sup>60</sup> comparent les répondants fréquents et peu fréquents et constatent que les réponses des intervenants fréquents sont plus en phase avec le comportement réel des

---

<sup>57</sup> Alexandre Sagakian, « Vers la fin des access panels online ? », Séminaire SFdS – Enquêtes en ligne, Jeudi 3 décembre 2009, Institut Henri Poincaré, Paris, <http://www.sfds.asso.fr/ressource.php?fct=ddoc&i=527&PHPSSESSIONID=baee6c2b98018cde75f74cab538c289c>

<sup>58</sup> Vera Toepel, Marcel Das, Arthur Van Soest, « Effects of design in web surveys comparing trained and fresh respondents », *Public Opinion Quarterly*, Vol. 72, No. 5 2008, p. 985–1007

<sup>59</sup> Pour étudier les effets du plan sur les répondants formés et « frais », l'auteur a utilisé deux panels de ménages en ligne administrés par CentERdata. Les deux panels sont représentatifs de la population des 16 ans et plus aux Pays-Bas. Ainsi, le CentERpanel était composé de répondants formés (avec une durée moyenne de 6 ans et 8 mois, écart-type de 4 ans), tandis que le panel LISS qui a démarré en 2007 était composé de répondants entièrement nouveaux.

<sup>60</sup> Terrence Coen, Jacqueline Lorch, Linda Piekarski, « The Effects of Survey Frequency on Panelists' Responses », ESOMAR retrieved July 27, 2005, 2007 on [www.websm.org](http://www.websm.org)

consommateurs que ceux des intervenants moins fréquents. Ce résultat est en contraste avec la vision traditionnelle selon laquelle l'expérience n'est pas souhaitable en ce qui concerne les erreurs de mesure (les auteurs citent William H. Williams 1970<sup>61</sup>; William H. Williams et Colin L. Mallows 1970<sup>62</sup>; Henk Meurs, Leo vanWissen et Jacqueline Visser, 1989<sup>63</sup>; Thomas F. Golob 1990<sup>64</sup>; Julia Brannen 1993<sup>65</sup>; Nancy A. Mathiowetz, Tamra J. Lair, 1994<sup>66</sup>; J. Patrick Sharpe, David G. Gilbert, 1998<sup>67</sup>; Larry M. Bartels 1999<sup>68</sup>; Patrick Sturgis, Nick Allum et Ian Brunton-Smith 2007<sup>69</sup>).

Par ailleurs, dans l'expérience de Vera Toepel et al. (2008), les panélistes formés semblaient utiliser la proximité des éléments comme une indication de leur sens, peut-être au détriment de la lecture attentive de chaque élément. Les panélistes « frais », quant à eux, étaient peut-être motivés par la nouvelle expérience de participer à une enquête et donc étaient davantage amenés à lire chaque article de plus près. L'auteur note également une différence dans la durée moyenne de l'enquête (composée de trois essais) entre les panels : 436 secondes pour le panel entraîné et 576 secondes pour le panel d'interviewés « frais ». En revanche, aucune différence significative n'a pu être mise en évidence dans la non-réponse entre les répondants formés et « frais », ni sur l'effet des catégories de réponse. Contrairement à Keith A. Choquette, Michie N. Hesselbrock (1987)<sup>70</sup>, les auteurs n'ont trouvé aucune preuve de différences dans les biais de désirabilité sociale entre les deux types de répondants.

---

<sup>61</sup> William H. Williams, « The Systematic Bias Effects of Incomplete Responses in Rotation Samples », *Public Opinion Quarterly* 33, 1970, p. 593-602

<sup>62</sup> William H. Williams, Colin L. Mallows, « Systematic Biases in Panel Surveys Due to Differential Nonresponse », *Journal of the American Statistical Association* 65, 1970, p. 1338-49

<sup>63</sup> Henk Meurs, Leo vanWissen, Jacqueline Visser, « Measurement Biases in Panel Data », *Transportation* 16, 1989, p. 175-94

<sup>64</sup> Thomas F. Golob, « The Dynamics of Household Travel Time Expenditures and Car Ownership Decisions », 1990, *Transportation Research* 24A, p. 443-63

<sup>65</sup> Julia Brannen, « The Effects of Research on Participants: Findings from a Study of Mothers and Employment », *Sociological Review*, 41, 1993, p.328-46

<sup>66</sup> Nancy A. Mathiowetz, Tamra J. Lair, « Getting Better? Changes or Errors in the Measurement of Functional Limitations », *Journal of Economic & Social Measurement*, 20, 1994, p. 237-62

<sup>67</sup> J. Patrick Sharpe, David G. Gilbert, « Effects of Repeated Administration of the Beck Depression Inventory and Other Measures of Negative Mood States », *Personal Individual Differences*, 24, 1998, p. 457-63

<sup>68</sup> Larry M. Bartels, « Panel Effects in the American National Election Studies », *Political Analysis*, 8, 1999, p. 1-20

<sup>69</sup> Patrick Sturgis, Nick Allum, Ian Brunton-Smith, « Attitudes over Time: The Psychology of Panel Conditioning », In *Methodology in Longitudinal Surveys*, ed. P. Lynn, 2007, p. 1-13, Chichester: Wiley

<sup>70</sup> Keith A. Choquette, Michie N. Hesselbrock, « Effects of Retesting with the Beck and Zung Depression Scales in Alcoholics », *Alcohol and Alcoholism*, 22, 1987, 277-83

## IV. Un ajustement des erreurs est-il possible ?

---

Une fois ces limites posées, il est néanmoins possible d'essayer de se rapprocher de la réalité, de chercher des pistes pour corriger ces erreurs ou du moins pour tenter d'en donner une appréciation.

Il a été longuement débattu dans la littérature de savoir si oui ou non il est possible de procéder à des ajustements sur des échantillons d'internautes afin de représenter à la fois les internautes et les non-utilisateurs (Bryan D. Rookey et al., 2008<sup>71</sup>). Ces procédures supposent, par exemple, que quelqu'un qui est plus âgé et a accès à Internet, a des opinions similaires à quelqu'un dans la même catégorie d'âge qui n'a pas accès à Internet. Bien qu'il existe des avantages à la post-stratification (Paul P. Biemer, Sharon L. Christ, 2008<sup>72</sup>), des recherches récentes suggèrent que l'ajustement des résultats obtenus à partir d'un seul segment de la population (par exemple, les utilisateurs d'Internet qui se portent volontaires pour participer à des panels) sont insuffisantes pour la plupart des recherches qui visent à décrire les opinions et les comportements du public (Vasja Vehovar et Katja Lozar Manfreda 1999<sup>73</sup>; Sunghee Lee 2006a, 2006b<sup>74</sup>; Mick P. Couper et al. 2007<sup>75</sup>). L'expérience que nous présenteront dans la deuxième partie de cette recherche va dans le même sens.

### 1. La pondération traditionnelle

Les redressements, ou calculs de repondérations, essaient d'améliorer la qualité des estimations d'enquête à l'aide de données auxiliaires. Ces données auxiliaires sont un ensemble de variables mesurées dans l'enquête dont on connaît également la répartition dans la population. On peut alors juger de la représentativité de l'enquête sur ces critères et éventuellement calculer des coefficients de pondération qui permettent de caler les résultats sur la réalité. L'hypothèse faite est que par effet d'entraînement, si

---

<sup>71</sup> Bryan D. Rookey, Steve Hanway, Don A. Dillman, « Does a probability-based household panel benefit from assignment to postal response as an alternative to Internet-only? », *Public Opinion Quarterly*, Vol. 72, No. 5, 2008, p. 962-984

<sup>72</sup> Paul P. Biemer, Sharon L. Christ, « Weighting Survey Data », In *International Handbook of Survey Methodology*, eds. Edith, de Leeuw, Joop, J. Hox, and Don, A. Dillman. New York, 2008, Taylor & Francis Group

<sup>73</sup> Vasja Vehovar, Katja Lozar Manfreda, « Web Surveys: Can Weighting Solve the Problem? », *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 1999, p. 962-67

<sup>74</sup> Sunghee Lee, « Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys », *Journal of Official Statistics* 22, 2006a, p. 329-49 ; Sunghee Lee, « An Evaluation of Nonresponse and Coverage Errors in a Prerecruited Probability Web Panel Survey », *Social Science Computer Review*, 24, 2006b, p. 460-75

<sup>75</sup> Mick P. Couper, Arie Kapteyn, Matthias Schonlau, Joachim Winter, « Noncoverage and Nonresponse in an Internet Survey », *Social Science Research*, 36, 2007, p. 131-48

l'enquête est représentative sur les variables auxiliaires, elle le sera aussi sur l'ensemble. C'est également le principe des enquêtes par quotas.

Didier Fripiat et Nicolas Marquis (2010)<sup>76</sup> rappellent que cette procédure, utilisée fréquemment, se révèle parfois insuffisante compte tenu des erreurs de couverture et d'autosélection (Jelke Bethlehem, Ineke Stoop, 2007<sup>77</sup>). Les différences mises en évidence par Thorsten Faas (2004)<sup>78</sup> qui compare les résultats de trois enquêtes conduites avec des méthodologies très différentes lors de la préparation des dernières élections fédérales allemandes (voir précédemment), persistent même avec des ajustements par rapport au sexe et à l'âge. L'autosélection contient donc un grand potentiel de partialité. Le problème se pose particulièrement lorsque les variables à partir desquelles on effectue la pondération ne sont pas suffisamment corrélées avec les variables sur lesquelles portent la recherche (Thorsten Faas et Harald Schoen, 2006<sup>79</sup>). Ainsi, Geert Loosveldt et Nathalie Sonck (2008)<sup>80</sup> notent qu'une pondération par l'âge, le genre et le niveau d'instruction n'a eu qu'un impact mineur sur les résultats de l'enquête et n'a pas rendu l'échantillon tout-venant comparable à un échantillon constitué de manière aléatoire parce que les associations entre les variables (par exemple entre le niveau d'instruction et l'intérêt pour la politique) continuaient de différer dans les deux échantillons.

Bryan D. Rookey et al. (2008), également, montrent que les différences persistent après les traditionnels ajustements de pondération. Pour arriver à cette conclusion, les auteurs comparent les réponses combinées d'une double interrogation courrier et Internet, et les réponses Internet pondérées uniquement. Avant la pondération, le courrier et le web différaient par cinq points ou plus de pourcentage entre les cellules dans 30 des 92 éléments de l'enquête. Après pondération des réponses, le courrier et le web diffèrent dans le même ordre de grandeur (soit au moins cinq points de pourcentage) sur 34 des 92 mêmes articles. Les réponses les plus différentes (au moins cinq points de pourcentage) semblent être distribuées dans tout le questionnaire et non pas concentrées dans une région particulière ou limitée à un type particulier de la question.

---

<sup>76</sup> Didier Fripiat, Nicolas Marquis, « Les enquêtes par Internet en sciences sociales : un état des lieux », *Population*, Volume 65, n°2, 2010, p. 309-338

<sup>77</sup> Jelke Bethlehem, Ineke Stoop, « Online panels. A paradigm theft? », in Trotman et al. (eds), *The Challenges of a Changing World*, Southampton, Association for Survey Computing, 2007, p. 113-131

<sup>78</sup> Thorsten Faas, « Online or Not Online?, A Comparison of Offline and Online Surveys Conducted in the Context of the 2002 German Federal Election », *Bulletin de méthodologie sociologique* n°82, 2004, p. 42-57, <http://bms.revues.org/index1086.html#ftn2>

<sup>79</sup> Thorsten Faas et Harald Schoen, « Putting a questionnaire on the web is not enough : A comparison of online and offline surveys conducted in the context of the German federal election 2002 », *Journal of Official Statistics*, 22(2), 2006, p. 177-190

<sup>80</sup> Geert Loosveldt, Nathalie Sonck, « An evaluation of the weighting procedures for an online access panel survey », *Survey Research Methods*, 2(2), 2008, p. 93-105

## 2. Scores de propension ou probabilité d'inclusion

Récemment, des scores de propension ont été suggérés pour réduire les erreurs de couverture, d'autosélection et de non-réponse (Sunghee Lee, 2006<sup>81</sup>). On obtient des scores de propension en modélisant une variable qui indique si une personne participe ou non à l'enquête [Jelke Bethlehem (2009<sup>82</sup>, 2010<sup>83</sup>)]. La méthode, généralement attribuée à Paul R. Rosenbaum et Donald B. Rubin (1983)<sup>84</sup>, a été développée pour réduire les biais de sélection dans les statistiques d'analyses obtenues à partir d'études d'observation (par exemple, la relation entre le tabagisme et le cancer du poumon) qui ne permettaient pas la généralisation à toute la population (Mick P. Couper, 2000<sup>85</sup>). On utilise habituellement un modèle de régression logistique où la variable indicatrice est la variable dépendante et les variables attitudinales sont les variables explicatives (Jelke Bethlehem, 2009, 2010). On suppose que ces variables attitudinales expliquent pourquoi une personne participe ou non. L'ajustement du modèle de régression logistique se résume à estimer la probabilité de participer (score de propension), compte tenu des valeurs des variables explicatives.

### Les scores de propension (Jelke Bethlehem, 2009)

Chaque personne  $k$  dans la population est supposée avoir une certaine probabilité inconnue  $K$  de participer à l'enquête, pour  $k = 1, 2, \dots, N$ . Soit  $r_1, r_2, \dots, r_N$ , des variables indicatrices, où  $r_k = 1$  si la personne  $k$  participe à l'enquête, et  $r_k = 0$  sinon. Par conséquent,

Le score de propension ( $X$ ) est la probabilité conditionnelle qu'une personne ayant les caractéristiques  $X$  participe à l'enquête, à savoir

On suppose que dans les strates définies par les valeurs des caractéristiques observées  $X$ , toutes les personnes ont la même propension à répondre. Le score de propension est exprimé avec un modèle logistique :

<sup>81</sup> Sunghee Lee, « Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys », *Journal of Official Statistics* 22, 2006a, p. 329–49

<sup>82</sup> Jelke Bethlehem, « Peut-on établir des statistiques officielles à partir d'enquêtes en ligne reposant sur le principe de l'autosélection ? », *Symposium 2008, Collecte des données : défis, réalisations et nouvelles orientations, La série des symposiums internationaux de Statistique Canada : recueil*, 2009

<sup>83</sup> Jelke Bethlehem, « Selection Bias in Web Surveys », *International Statistical Review*, 78, 2, 2010, p. 161–188

<sup>84</sup> Paul R. Rosenbaum, Donald B. Rubin, « The central role of the propensity score in observational studies for causal effects », *Oxford Journals, Mathematics & Physical Sciences, Biometrika*, Vol 70, Issue 1, 1983, p. 41-55

<sup>85</sup> Mick P. Couper, « Web Survey : A Review of Issues and Approaches », *The Public Opinion Quarterly*, Volume 64, N°4, décembre 2000, p. 464-494

Jelke Bethlehem (2009, 2010) indique que deux conditions sont nécessaires à l'application de la pondération fondée sur les scores de propension. Premièrement, on doit disposer de variables auxiliaires pertinentes, c'est-à-dire pouvant expliquer si une personne est disposée ou non à participer à l'enquête en ligne. Deuxièmement, on doit disposer des valeurs de ces variables pour les participants et les non-participants. On peut alors calculer une probabilité d'inclusion pour chaque individu. La pondération se fera en fonction de l'inverse de la probabilité d'inclusion, de telle manière qu'un individu ayant moins de chance de posséder une connexion à Internet et/ou de s'inclure dans un panel se voit octroyer un poids plus important (Adriaan Hoogendoorn et Jacco Daalmans, 2009<sup>86</sup>).

Ainsi, à partir d'une référence à un échantillon de qualité, des poids sont calculés. Comme dans toute pondération, il faut que les variables auxiliaires soient bien choisies et disponibles. Mais il faut alors être très transparent sur les procédures de pondération et le pouvoir prédictif du modèle (Edith D. de Leeuw<sup>87</sup>, 2008). Les instituts de sondage utilisent largement ces méthodes. C'est une caractéristique clef de l'approche de Harris Interactive. Utilisant des enquêtes téléphoniques en parallèle, Harris Interactive estime la propension à faire partie de l'échantillon en ligne en se basant sur un vecteur de covariables mesurées dans les deux modes (Mick P. Couper, 2000<sup>88</sup>).

Sunghee Lee (2006)<sup>89</sup>, Sunghee Lee et Vaillant (2009)<sup>90</sup> ainsi que Matthias Schonlau et al. (2009)<sup>91</sup> offrent une bonne introduction à la technique de l'ajustement par la probabilité d'inclusion (Didier Frippliat et Nicolas Marquis, 2010<sup>92</sup>). La question est donc de savoir si une série de manipulations sur l'échantillon de convenance permet de lui attribuer les mêmes qualités qu'un échantillon tiré à partir d'une population connue et identifiée. Bien que novatrice par rapport à la post-stratification classique, cette méthode n'en reste pas moins problématique à plusieurs égards (Didier Frippliat et Nicolas Marquis, 2010) : « *avant tout, elle suppose que le chercheur est en mesure de constituer*

---

<sup>86</sup> Adriaan Hoogendoorn, Jacco Daalmans, « Nonresponse in the Recruitment of an Internet Panel Based on Probability Sampling », *Survey Research Methods*, Vol 3, No 2, 2009

<sup>87</sup> Edith D. de Leeuw, « Choosing the Method of Data Collection », p. 113-135, in Edith D. de Leeuw, Joop J. Hox, Don A. Dillman, *International Handbook of Survey Methodology*, Taylor & Francis Group, 2008, 549 p.

<sup>88</sup> Mick P. Couper, « Web Survey : A Review of Issues and Approaches », *The Public Opinion Quarterly*, Volume 64, N°4, décembre 2000, p. 464-494

<sup>89</sup> Sunghee Lee, « Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys », *Journal of Official Statistics* 22, 2006a, p. 329-49

<sup>90</sup> Sunghee Lee, Richard Vaillant, « Estimation for volunteer panel web surveys using propensity score adjustment and calibration adjustment », *Sociological Methods Research*, 37(3), 2009, p. 319-343

<sup>91</sup> Matthias Schonlau, Arthur van Soest, Arie Kapteyn, and Mick Couper, « Selection Bias in Web Surveys and the Use of Propensity Scores », *Sociological Methods & Research*, 2009, 37, p. 291-318

<sup>92</sup> Didier Frippliat, Nicolas Marquis, « Les enquêtes par Internet en sciences sociales : un état des lieux », *Population*, Volume 65, n°2, 2010, p. 309-338

*un échantillon représentatif à partir de la population de référence qu'il aura identifiée. Cet échantillon est par ailleurs considéré comme exempt de tout biais, alors qu'il est généralement d'une taille bien moindre que l'échantillon constitué par Internet* », avec des résultats certainement moins précis. De plus, seulement certaines caractéristiques sont prises en compte pour mettre en place ces procédures de pondérations mais ce ne sont peut-être pas les seules pour lesquelles les deux échantillons (internauts et non internauts) diffèrent (Katja Lozar Manfreda et Vasja Vehovar, 2008<sup>93</sup>). Cela correspond également aux conclusions de Gerty J. L. M. Lensvelt-Mulders et al. (2009)<sup>94</sup> : pour savoir si les membres d'un panel diffèrent ou non de répondants qui seraient sélectionnés à partir d'un échantillon aléatoire, les auteurs ont comparé les résultats d'un panel de volontaires Internet à un sondage en ligne basé sur un échantillon aléatoire sur la même population. L'échantillon aléatoire a été mené auprès d'une base de données nationale hollandaise sur 2500 répondants qui ont reçu une lettre d'invitation pour se connecter sur Internet. Les non-répondants sont relancés par téléphone. 1347 personnes ont répondu. L'échantillon de volontaires sur Internet était mené auprès d'un large panel en utilisant la méthode des quotas pour être représentatif sur le sexe, la profession et l'âge (496 répondants). Les auteurs concluent que les différences de composition des deux populations sur le sexe, la profession et l'âge sont faibles, compte tenu de l'utilisation de la méthode des quotas. En revanche, d'autres différences apparaissent, petites mais significatives. Les scores de propension mis en place pour corriger l'erreur de couverture n'empêchent pas certaines différences de demeurer entre les deux échantillons. Il reste donc des erreurs d'autosélection.

### **3. Pondération fondée sur une enquête de référence**

En l'absence de variables auxiliaires pertinentes, on peut envisager de mener une enquête de référence reposant sur un petit échantillon probabiliste (Jelke Bethlehem 2009<sup>95</sup>, 2010<sup>96</sup>) : « *On recueille les données selon un mode différent de l'enquête en ligne. Dans l'hypothèse d'une non-réponse nulle ou ignorable, cette enquête de référence produit des estimations sans biais de la répartition des variables auxiliaires au sein de la population* ». Matthias Schonlau et al. (2003) soulignent que le succès de cette méthode

---

<sup>93</sup> Katja Lozar Manfreda, Vasja Vehovar, « Internet Surveys », p. 264-284, in Edith D. de Leeuw, Joop J. Hox, Don A. Dillman, *International Handbook of Survey Methodology*, Taylor & Francis Group, 2008, 549 p.

<sup>94</sup> Gerty J. L. M. Lensveldt-Mulders, Peter J. Lugtig, Marianne Hubregtse, « Separating Selection Bias and Non-coverage in Internet Panels using Propensity Matching », *Survey practice*, 2009.

<sup>95</sup> Jelke Bethlehem, « Peut-on établir des statistiques officielles à partir d'enquêtes en ligne reposant sur le principe de l'autosélection ? », *Symposium 2008, Collecte des données : défis, réalisations et nouvelles orientations*, La série des symposiums internationaux de Statistique Canada : recueil, 2009

<sup>96</sup> Jelke Bethlehem, « Selection Bias in Web Surveys », *International Statistical Review*, 78, 2, 2010, p. 161-188



repose sur deux hypothèses : les variables webographiques peuvent expliquer l'écart entre les répondants à l'enquête en ligne et les autres membres de la population cible, et l'enquête de référence n'est pas entachée d'une non-réponse non ignorable. En pratique, il n'est pas facile de satisfaire à ces conditions (Jelke Bethlehem 2009, 2010).

La méthode de l'enquête de référence présente aussi un inconvénient. Jelke Bethlehem (2007)<sup>97</sup> montre qu'elle, « *réduit l'erreur des estimations au prix d'une variance plus élevée* » car si la taille de l'enquête de référence est faible, la variance de l'estimateur sera moins précise. Marc N. Elliott et Amelia Haviland (2007)<sup>98</sup> décrivent une méthodologie utilisée pour combiner un échantillon de convenance avec un échantillon probabiliste afin de produire un estimateur ayant une erreur quadratique moyenne plus faible que les estimateurs fondés uniquement sur un échantillon probabiliste. L'idée est que, compte tenu du faible coût pour disposer de grands échantillons de convenance (en autosélection sur des sites ou en panels pré-recrutés), le couplage avec un échantillonnage probabiliste permettrait de diminuer les erreurs. Il s'agirait de réserver une petite partie du budget de l'échantillon probabiliste pour créer un gros échantillon de convenance supplémentaire. Or, même en calant sur marges l'échantillon de convenance pour réduire les erreurs, les auteurs montrent que les résultats sont décevants et que, quelle que soit la taille de l'échantillon de convenance, il ne peut pas beaucoup améliorer l'erreur. Par ailleurs, l'erreur des échantillons de convenance n'étant pas connue (par définition), il faut l'estimer avec l'échantillon probabiliste et celui-ci doit être vraiment conséquent pour bien estimer l'erreur.

#### **4. Les enquêtes mixtes**

Nous pouvons considérer les enquêtes en mode mixte comme une piste, également, pour corriger les erreurs de couverture, de non-réponse et de mesure. Une approche en mode mixte peut réduire l'erreur de couverture, lorsque les gens qui n'ont pas accès à un mode ont accès à un autre mode (Gera E. Nagelhout et al., 2010). Par exemple, l'utilisation d'une approche en mode mixte dans les études multi-pays, avec un mode au sein de chaque pays, peut réduire l'erreur de couverture quand un pays a davantage accès à Internet ou un autre est plus équipé du téléphone. Ou bien, l'erreur de non-réponse peut être réduite en utilisant une approche en mode mixte dans laquelle les répondants peuvent choisir leur mode d'entrevue. Le problème de la peur des *spams* par exemple

---

<sup>97</sup> Jelke Bethlehem, « Reducing the bias of web survey based estimates », Document de travail 07001, Voorburg, Pays-Bas: Statistics Netherlands, 2007

<sup>98</sup> Marc N. Elliott, Amelia Haviland, « Utilisation d'un échantillon de convenance électronique comme complément à un échantillon probabiliste », Techniques d'enquête, Volume 33, N°2, 2007, p. 233-238, Statistique Canada, n°12-001-X au catalogue

peut être évité en mode mixte avec des lettres papier en avance ou des appels téléphoniques pour présenter l'étude.

Néanmoins, compte tenu de tout ce que nous avons vu jusqu'à présent, il convient de rester vigilant à la comparabilité des résultats. D'après Mick P. Couper et Peter V. Miller (2008), si l'objectif des enquêtes en mode mixte impliquant Internet est d'augmenter le taux de réponse global, les données récentes ne sont pas très encourageantes. En revanche, si l'objectif est une réduction des coûts grâce à l'ajout de questionnaires en ligne ou une couverture plus large, elles sont prometteuses (D. Rookey et al., 2008). Mais elles ne semblent pas être une solution au problème de non-réponse.

## V. Les pistes d'améliorations des études en ligne

---

### 1. L'aspect visuel

D'après Katja Lozar Manfreda et Vasja Vehovar (2008), le succès des invitations générales dépend de leur lieu (sur quels sites, groupes de discussion etc.), de leur visibilité, du moment auquel elles sont postées. Il peut s'agir de fenêtres pop-up, de bannières ou d'un simple lien hypertexte. Les premières sont plus visibles mais peuvent irriter. Envoyer des invitations individualisées aux participants est plus efficace que des invitations générales. Il est recommandé d'utiliser de multiples contacts et de les planifier (préavis, invitation principale, relances...). Pour les préavis d'enquête, il a été montré que des courriers papiers engendraient plus de retours que des courriers Internet. Katja Lozar Manfreda et Vasja Vehovar (2008) expliquent que la page d'introduction est importante pour motiver à remplir entièrement l'enquête ; il faut qu'une communication soit possible avec l'institut d'étude ou de recherche, que l'on puisse laisser des commentaires, échanger des idées, ajouter des questions ouvertes pour que les répondants puissent parler de leurs propres expériences, partager les résultats avec les répondants (ce qui motive pour la prochaine enquête), assurer de la confidentialité de l'enquête. L'aspect visuel est très important, surtout en l'absence d'enquêteur pour motiver les répondants.

Suite à son expérience menée avec la technologie de l'Eye-tracking, Mirta Galesic et al. (2008) proposent quelques recommandations visuelles. Les répondants passent souvent plus de temps à lire la partie supérieure des listes plutôt que la partie inférieure. Par ailleurs, les options qui sont d'abord cachées (menus déroulants...) sont moins regardées que les options visibles dès le début. Enfin, pour la définition des concepts, même s'ils encombrant l'écran, ils gagneraient à être clairement affichés si l'on souhaite qu'ils soient lus. En effet, les auteurs expliquent qu'augmenter l'effort de l'enquêté diminue les chances qu'il aille lire l'information. Les chercheurs doivent donc tout faire pour motiver les répondants à être attentifs à tous les aspects des questions, par exemple en rendant le questionnaire court et intéressant. En outre, comme certains répondants vont probablement prendre des raccourcis cognitifs, il est sage de compenser ces effets de stratégie avec des rotations aléatoires, par exemple, ou en rendant les options immédiatement visibles.

De même Vera Toepel et al. (2008)<sup>99</sup> rappellent que plusieurs travaux donnent à penser que la conception d'un sondage en ligne a un impact très significatif sur l'erreur de mesure (voir, par exemple, Mick P. Couper, Michael W. Traugott et Mark J. Lamias 2001<sup>100</sup>; Dillman et Christian 2002; Leah M. Christian et Don A. Dillman 2004<sup>101</sup>; Roger Tourangeau, Mick P. Couper et Frederick Conrad 2004, 2007<sup>102</sup>; Don A. Dillman 2007<sup>103</sup>). La conception du questionnaire en ligne est peut-être plus importante que dans les autres modes car il existe de nombreux outils disponibles et de nombreuses variations possibles dans la façon dont l'enquête apparaît à l'écran.

## 2. La transparence

En tout cas, la transparence est nécessaire pour faciliter l'évaluation scientifique de la méthode (Mick P. Couper et Peter V. Miller, 2008). Pour cela, Mario Callegaro et Charles Disogra (2008)<sup>104</sup> proposent un ensemble de normes de divulgation pour les taux de réponse dans les panels probabilistes, ainsi que des mesures possibles pour pouvoir également évaluer la performance des recrutements non probabilistes. Ils rappellent que plusieurs associations ont publié des lignes directrices pour les paramètres de calcul des enquêtes en ligne : l'ADM (Arbeitskreis Deutscher Markt) a publié les « normes d'assurance qualité pour les enquêtes en ligne » en 2001<sup>105</sup>, de même qu'EFAMRO (European Federation of Associations of Market Research Organisations) en 2004, basé sur l'ESOMAR (European Society for Opinion and Marketing Research), organisation mondiale qui œuvre pour une meilleure recherche dans les marchés, les consommateurs et les sociétés et qui a publié plusieurs documents sur les études conduites en utilisant Internet. Des méthodes pour calculer les taux de réponse y sont proposées, ainsi que des

---

<sup>99</sup> Vera Toepel, Marcel Das, Arthur Van Soest, « Effects of design in web surveys comparing trained and fresh respondents », *Public Opinion Quarterly*, Vol. 72, No. 5 2008, p. 985-1007

<sup>100</sup> Mick P. Couper, Michael W. Traugott, Mark J. Lamias, « Web Survey Design and Administration », *Public Opinion Quarterly*, 65, 2001, p. 230-53

<sup>101</sup> Leah M. Christian, Don A. Dillman, « The Influence of Graphical and Symbolic Language Manipulations to Self-Administered Questions », *Public Opinion Quarterly*, 68, 2004, p. 57-80

<sup>102</sup> Roger Tourangeau, Mick P. Couper, Frederick Conrad, « Spacing, Position, and Order. Interpretive Heuristics for Visual Features of Survey Questions », *Public Opinion Quarterly*, 68, 2004, p. 368-93 ; Roger Tourangeau, Mick P. Couper, Frederick Conrad, « Color, Labels, and Interpretive Heuristics for Response Scales », *Public Opinion Quarterly*, 71, 2007, p. 91-112

<sup>103</sup> Don. A. Dillman, *Mail and Internet Surveys. The Tailored Design Method*, 2007, Hoboken, NJ : Wiley

<sup>104</sup> Mario Callegaro, Charles Disogra, « Computing response metrics for online panels », *Public Opinion Quarterly*, Volume 72, Issue 5, 2008, p. 1008-1032, <http://poq.oxfordjournals.org/content/72/5.toc>

<sup>105</sup> Sur le site de l'ADM (<http://www.adm-ev.de/index.php?id=2&L=1>), Richtlinie für Online-Befragungen, [http://rat-marktforschung.de/fileadmin/user\\_upload/pdf/R09\\_RDMS.pdf](http://rat-marktforschung.de/fileadmin/user_upload/pdf/R09_RDMS.pdf)

définitions précises : le taux de réponse cumulée, le taux de vue, le taux de participation, le taux d'achèvement... (Gunther Eysenbach, 2004<sup>106</sup>).

Ainsi, Mario Callegaro et Charles Disogra (2008) proposent des formules pour calculer le taux de recrutement (dans le cas des panels avec recrutement probabiliste), le taux de profil lors de la connexion (pour les panels de volontaires et les panels probabilistes, il s'agit de remplir son profil sociodémographique), le taux d'absorption (qualité des mails –mauvaise adresse, boîte aux lettres pleine, erreur de réseau, filtre anti-spam...-), le taux d'achèvement (proportion des personnes ayant répondu à l'enquête en ligne parmi les personnes ayant été invitées à participer), le taux d'abandon, le taux d'achèvement du dépistage (pour trouver une population spécifique ou rare), le taux d'admissibilité, le taux de réponse cumulée, le taux d'attrition (définition différente de celle de panels en longitudinal pour lesquels il s'agit des membres qui ne peuvent pas être interrogés d'une vague à l'autre, tandis que dans les *access panels*, l'attrition est définie comme le pourcentage des membres qui quittent le groupe dans un laps de temps défini). Tous ces taux peuvent aider à juger de la qualité d'une étude. Les auteurs concluent qu'il est plus important d'étudier et de comprendre les erreurs de non-réponse que de poursuivre exclusivement des taux de réponse élevés. Comme cela a été discuté par Robert M. Groves (2006)<sup>107</sup> et Groves et Peytcheva (2008)<sup>108</sup> les taux de réponse seuls ne sont pas un très bon indicateur de l'ampleur de l'erreur de non-réponse. Selon les auteurs encore, les forts coûts pour augmenter le taux de réponse et la tendance à la baisse de ces coûts de réponse n'est pas une raison suffisante pour ne mettre en place plus que des méthodes non probabilistes. En outre, les données auxiliaires de la base de sondage peuvent être utilisées pour améliorer les estimations d'enquête et pour des ajustements de pondération. Il faut donc recueillir autant de données auxiliaires que possible afin d'améliorer les estimations et de compenser la non-réponse. Le suivi de l'attrition est également important.

### **3. La qualité des données**

Pour améliorer la qualité des données, diverses méthodes peuvent être mises en place. Anne-Marie Côté et Danielle Laroche (2009) par exemple, expliquent que dans le cadre de l'institut Statistique Canada, des études qualitatives en salle et chez les répondants ont été mises en place pour mieux comprendre les résultats obtenus par Internet. Cela a

---

<sup>106</sup> Gunther Eysenbach, « Improving the Quality of Web Surveys: The Checklist for Reporting Results from Internet E-Surveys (Cherries) », *Journal of Medical Internet Research* 6(3) : e34, 2004

<sup>107</sup> Robert M. Groves, « Nonresponse Rates and Nonresponse Bias in Household Surveys », *Public Opinion Quarterly*, 70(5), 2006, p. 646–75

<sup>108</sup> Robert M. Groves, Emilia Peytcheva, « The Impact of Nonresponse Rates on Nonresponse Bias. A Meta-Analysis », *Public Opinion Quarterly*, 72(2), 2008, p. 167–189

notamment permis de se rendre compte des problèmes avec les menus déroulants, de la non visibilité de toutes les cases à cocher, etc.

Neil Malhotra (2008)<sup>109</sup>, quant à lui, conseille de regarder les temps de réponse pour filtrer les participants ou leurs données. En effet, l'auteur a étudié un échantillon aléatoire de la population adulte des États-Unis sur le *Web*. Comprendre la relation entre le temps de réponse et la qualité des données est important, car les incitations financières peuvent avoir des effets pervers sur les données. Bien que les entreprises de collecte de données sur Internet font de leur mieux pour éliminer les panélistes qui ont une trop faible implication, l'auteur note qu'il en demeure dans l'échantillon. Ces résultats ont des implications importantes pour les concepteurs, les administrateurs et les utilisateurs des sondages sur le Web. Même si le délai de traitement, en soi, ne peut pas être un critère de filtrage optimal (car il est seulement un indicateur de « satisficing », sélectionnant le sous-groupe de répondants qui ont les plus faibles compétences cognitives), les chercheurs devraient inclure ce temps comme variable de contrôle dans les modèles statistiques.

#### **4. Les incitations financières**

Selon Katja Lozar Manfreda et Vasja Vehovar (2008)<sup>110</sup>, dans les enquêtes en ligne, comme pour les autres modes d'enquête, les incitations financières augmentent les taux de réponse et motivent à répondre au questionnaire en entier. Il est nécessaire de faire encore des recherches sur le type d'incitation utilisée et son effet sur les erreurs de sondage et les coûts d'enquête pour les enquêtes en ligne. Bien sûr, l'impact dépend du type d'incitations utilisées (offertes à tous les répondants, en tirage au sort...). Les incitations peuvent poser problème sur Internet si les participants essaient de répondre plusieurs fois, avec des faux noms et des fausses réponses. De plus, cela peut accroître les erreurs d'autosélection : l'échantillon final peut être biaisé avec une sur-représentation de certaines personnes (les plus bas revenus, les plus jeunes, ceux qui ont davantage de temps...). Dans certains pays, les incitations financières pour les enquêtes ne sont pas autorisées.

---

<sup>109</sup> Neil Malhotra, 2008, Completion time and response order effects in web surveys, *Public Opinion Quarterly*, Volume 72, Issue 5, p. 914-934, <http://poq.oxfordjournals.org/content/72/5.toc>

<sup>110</sup> Katja Lozar Manfreda, Vasja Vehovar, "Internet Surveys", p. 264-284, in Edith D. de Leeuw, Joop J. Hox, Don A. Dillman, *International Handbook of Survey Methodology*, Taylor & Francis Group, 2008, 549 p.

PARTIE 2 : Expérimentation  
Peut-on extrapoler les  
résultats des internautes à la  
population française ?

## Principe de l'expérimentation

---

La première partie de cette recherche nous a rappelé que les questions soulevées par les enquêtes en ligne sont extrêmement nombreuses et qu'on ne peut pas répondre en une phrase à la question de savoir si les enquêtes en ligne sont fiables ou pas. Tout dépend du sujet d'étude, du souhait que l'on a de pouvoir extrapoler les résultats à la population générale ou pas, du protocole de recrutement des internautes interrogés, des précautions mise en œuvre pour limiter les erreurs de couverture, de non-réponse, de mesure ou d'échantillonnage, des traitements statistiques envisagés pour corriger ces erreurs, etc.

Nous chercherons, dans cette deuxième partie, à approfondir la question de l'erreur de couverture et de son éventuelle correction. Peut-on, en interrogeant uniquement des internautes, en tirer des conclusions que l'on peut extrapoler à l'ensemble de la population ? Les internautes sont-ils finalement si différents de l'ensemble de la population ? Il y a 10 ans, lorsque les internautes étaient très peu nombreux, la réponse à cette question aurait certainement été positive. Mais aujourd'hui, alors que ce sont les non-internautes qui sont devenus minoritaires dans le corps social, le problème se pose-t-il encore ? Et si les internautes présentent encore des spécificités, est-il possible de « corriger statistiquement » les réponses recueillies auprès des internautes, afin qu'elles ressemblent le plus possible à celles qui seraient enregistrées si l'on interrogeait un échantillon constitué à la fois d'internautes et de non-internautes ? Y a-t-il des sujets sur lesquels on peut plus facilement se « risquer » à réaliser des enquêtes par Internet sans craindre d'obtenir des résultats biaisés ? Quels sont les thèmes sur lesquels, au contraire, il convient d'être très prudent ?

C'est à ces questions que nous allons tenter d'apporter quelques éléments de réponse.

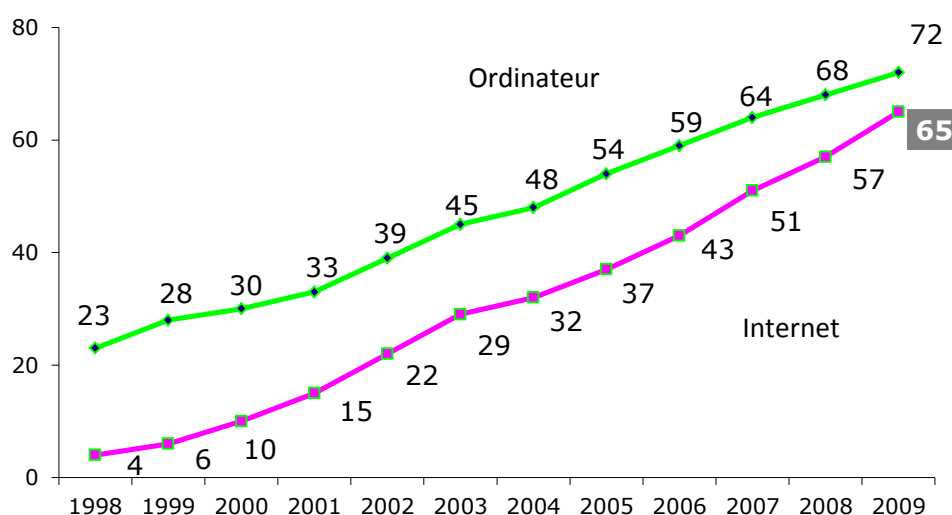


## I. Les internautes en France

---

Internet s'est installé dans une majorité de foyers : près des deux tiers des personnes interrogées en 2009 pour l'enquête « Conditions de vie et Aspirations des Français » disent disposer, à leur domicile, d'une connexion à Internet (65% exactement, voir Graphique 6). La progression de cet équipement, corrélée à celle de l'ordinateur, est spectaculaire : en 2000, seule une personne sur dix bénéficiait d'une connexion à Internet à son domicile.

**Graphique 6 - Taux de pénétration de l'ordinateur et d'Internet**



Source : CRÉDOC, enquêtes « Conditions de vie et Aspirations des Français », vague de janvier de chaque année.

### 1. Le fossé numérique et son évolution

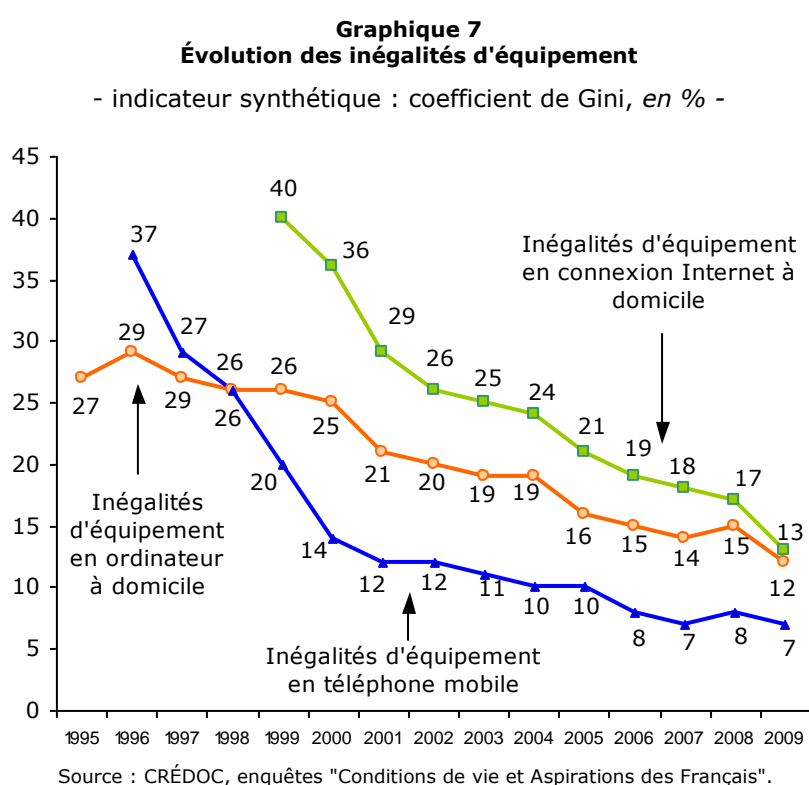
Si les taux d'équipement progressent, qu'en est-il des inégalités d'équipement ? Depuis une quinzaine d'années, le CRÉDOC a mis en place un indicateur des inégalités d'accès aux nouvelles technologies<sup>111</sup>, reposant sur le calcul de coefficients de Gini. L'analyse suivante<sup>112</sup> propose de quantifier les inégalités existantes, de les hiérarchiser et, aussi, de les mettre en perspective.

<sup>111</sup> David ALIBERT, Régis BIGOT et David FOUCAUD, *La dynamique des inégalités en matière de nouvelles technologies (Méthodes d'approche - Analyse évolutive)*, Cahier de recherche du CRÉDOC n°217, novembre 2005, <http://www.CRÉDOC.fr/pdf/Rech/C217.pdf>

<sup>112</sup> Cette analyse du fossé numérique est régulièrement actualisée dans le cadre du baromètre de diffusion des technologies de l'information et de la communication en France, réalisé par le CRÉDOC pour le compte de l'ARCEP (Autorité de régulation des communications électroniques et des postes) et du CGIET (Conseil général de l'industrie, de l'énergie et des technologies).

La méthode de calcul du coefficient de Gini ne sera pas détaillée ici : le lecteur pourra se reporter au rapport du CRÉDOC sur le fossé numérique pour davantage d'informations<sup>113</sup>. Nous rappellerons simplement que plus le coefficient de Gini est élevé et plus fortes sont les inégalités : un coefficient de Gini égal à zéro signifie qu'il n'y a pas d'inégalités, que le bien est également réparti dans la population, tandis qu'un coefficient de Gini égal à 100% signifie qu'un seul groupe détient tout l'équipement, les autres groupes en étant totalement privés.

Voici l'allure qu'ont pris les inégalités d'équipement en ordinateur, en connexion à Internet et en téléphone mobile depuis 1995 (Graphique 7).



Trois conclusions principales se dégagent :

- ✓ Après une pause en 2008, **la tendance à la baisse des inégalités reprend**. Cette baisse est particulièrement marquée pour l'ordinateur (- 3 points) et l'accès à Internet à domicile (- 4 points).

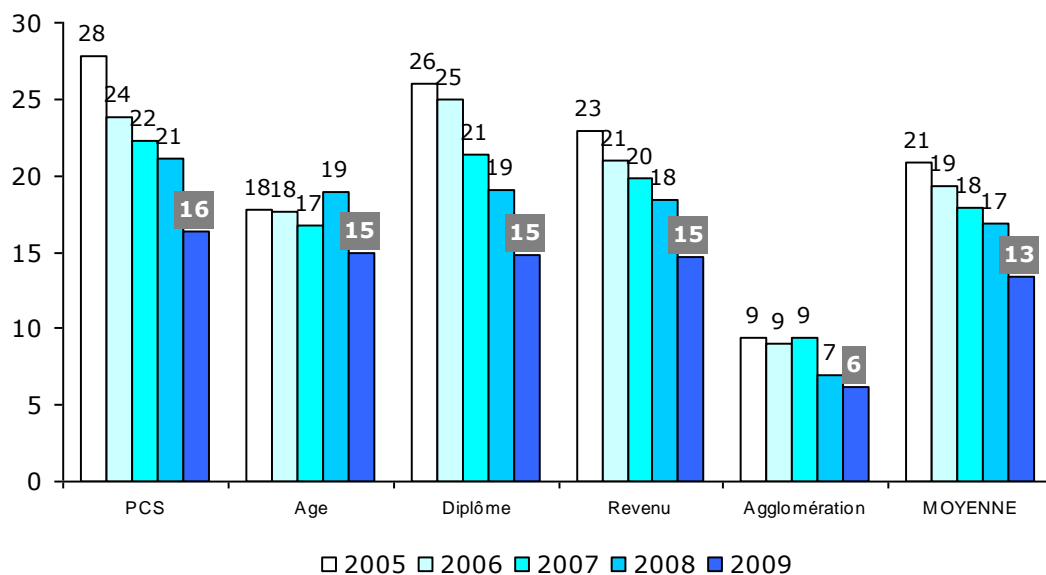
<sup>113</sup> op. cit.

- ✓ Les inégalités sont **plus élevées** pour l'équipement **en Internet et en ordinateur** à domicile (les coefficients de Gini en 2009 sont respectivement de 13% et 12%) que pour le **téléphone mobile** (coefficient égal à 7%).
- ✓ La diffusion des nouvelles technologies est **mieux répartie** aujourd'hui qu'elle ne l'était il y a encore quelques années. Pour Internet, les inégalités sont deux fois moins flagrantes aujourd'hui qu'en 2002 (13% contre 26%) et trois fois moins qu'il y a juste 10 ans (40% en 1999).

Dans le détail, s'agissant de **l'accès à Internet à domicile**, on note très peu d'inégalités selon le lieu de résidence (6%). Les écarts les plus importants se mesurent selon la PCS (16%), l'âge (15%), le diplôme (15%) et les revenus (15%). Tous ces indicateurs sont en baisse et l'indicateur synthétique recule, au final, de 4 points entre 2008 et 2009 (passant de 17 à 13%).

**Graphique 8 - Inégalités d'accès à Internet à domicile, selon plusieurs critères sociodémographiques**

- indicateur : coefficient de Gini, en % -



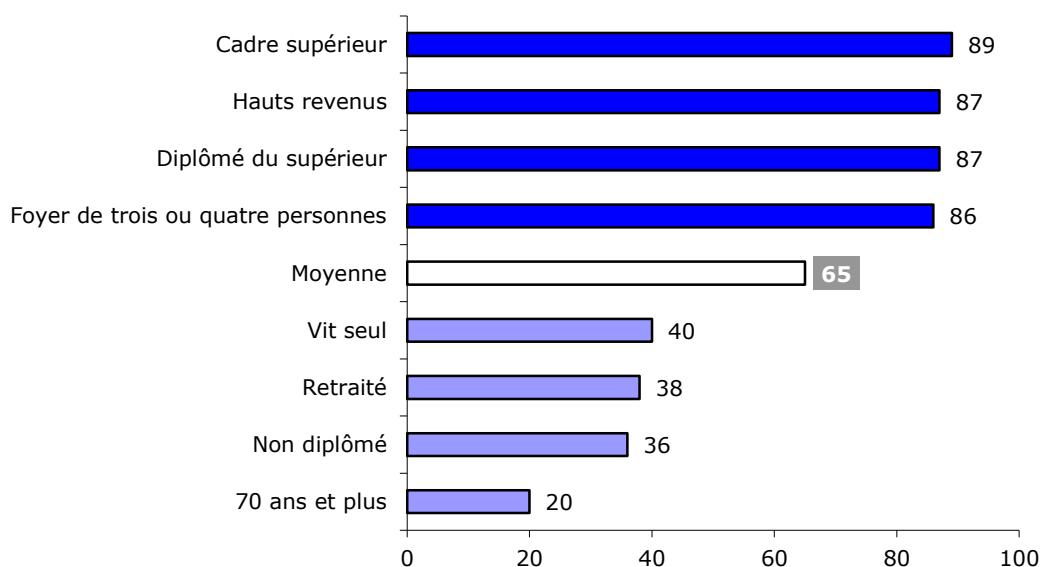
Source : CRÉDOC, enquêtes "Conditions de vie et Aspirations des Français"

## 2. Le profil sociodémographique des internautes en 2009

La progression spectaculaire de l'équipement en Internet à domicile s'accompagne donc **d'une réduction du « fossé numérique »**. Pour autant, l'accès à Internet ne se fait toujours pas de façon homogène dans tous les groupes de population ; certains restent en marge de la révolution numérique. À l'exception du genre, **tous les autres descripteurs** socio-économiques témoignent d'écarts significatifs.

### Graphique 9 – Le fossé numérique en 2009

(en % d'individus disposant d'une connexion Internet à domicile)

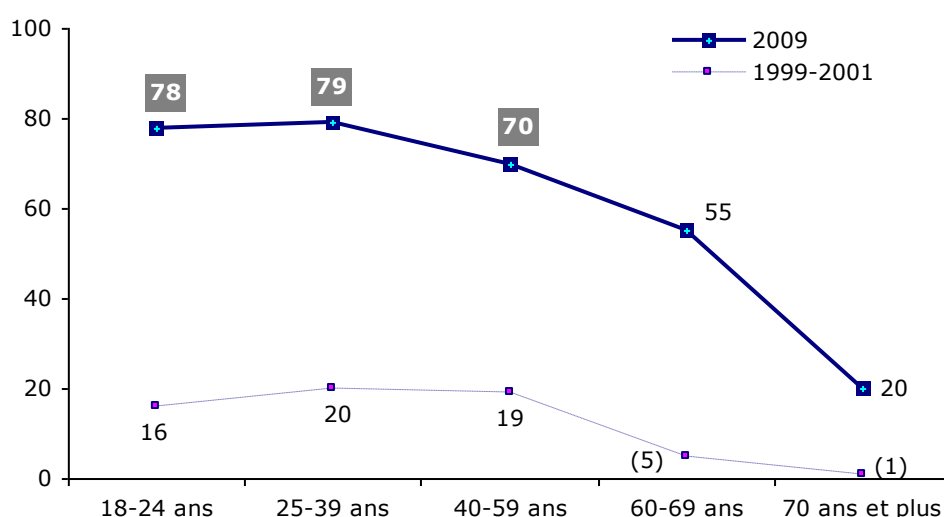


Source : CRÉDOC, enquête « Conditions de vie et Aspirations des Français », début 2009.

- **L'âge** est certainement la variable la plus corrélée à l'équipement en Internet : entre les plus jeunes qui, quatre fois sur cinq disposent d'une connexion à domicile, et les plus âgés qui ne sont que 20% à pouvoir se connecter chez eux, l'écart est considérable.

### Graphique 10 – Le taux d'accès à Internet à domicile quatre fois plus élevé avant 40 ans qu'après 70 ans

(en % d'individus disposant d'une connexion Internet à domicile)

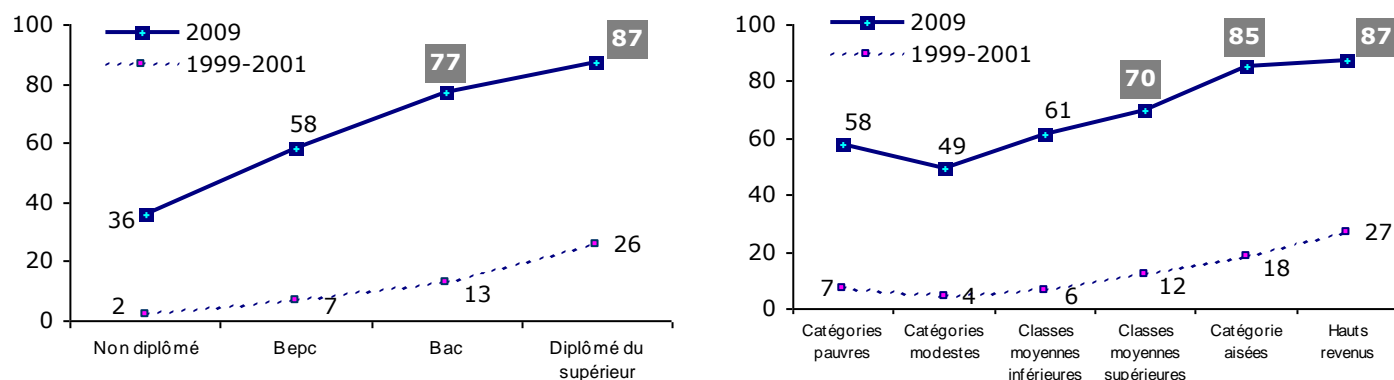


Source : CRÉDOC, enquêtes « Conditions de vie et Aspirations des Français ».

- **Le niveau de diplôme** : 87% des diplômés du supérieur, contre seulement un tiers des non-diplômés, sont équipés (Graphique 11).
- Le **niveau de revenus** est également déterminant : 87% des hauts revenus sont équipés, contre une personne sur deux parmi les catégories modestes (Graphique 11).

**Graphique 11 – Le taux d'accès à Internet à domicile croît régulièrement avec le niveau de diplôme et le niveau de revenus**

(en % d'individus disposant d'une connexion Internet à domicile)



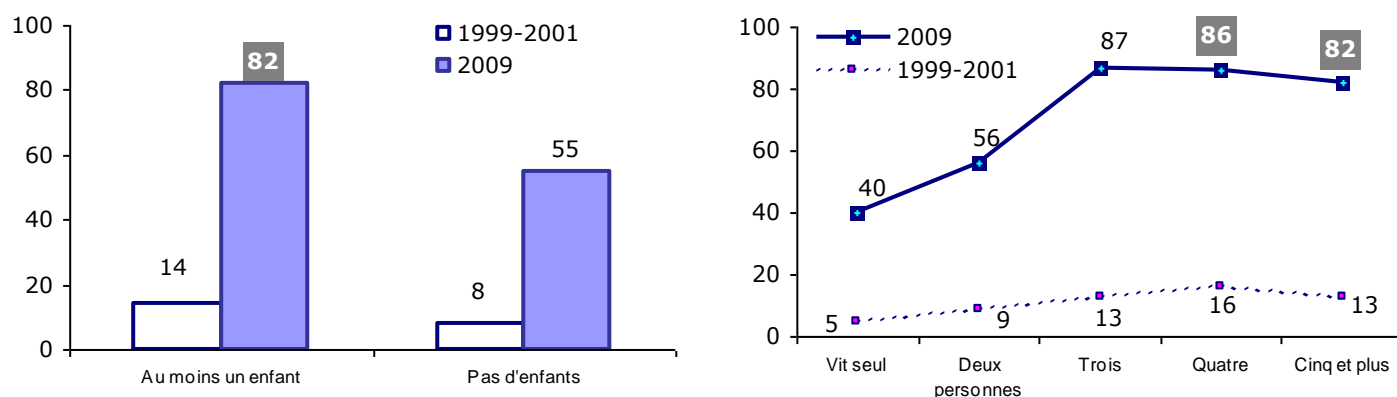
Source : CRÉDOC, enquêtes « Conditions de vie et Aspirations des Français ».

**Note** : la typologie de catégories de revenus repose sur la répartition de la population en déciles de niveau de vie (catégories pauvres : 1<sup>er</sup> décile ; catégories modestes : 2<sup>ème</sup> et 3<sup>ème</sup> déciles ; classes moyennes inférieures : 4<sup>ème</sup> à 6<sup>ème</sup> déciles ; classes moyennes supérieures : 7<sup>ème</sup> et 8<sup>ème</sup> déciles ; catégories aisées : 9<sup>ème</sup> décile ; hauts revenus : 10<sup>ème</sup> décile).

- **La présence d'un enfant dans le foyer** augmente la probabilité d'être connecté (82% dès qu'un enfant est présent, contre 55% pour les foyers sans enfants, voir Graphique 12). D'ailleurs, en fonction du **nombre de personnes dans le foyer**, qui traduit le plus souvent la présence d'enfants, le taux de connexion varie entre 40% (personnes seules) et 86% (foyers de quatre personnes, Graphique 12).

### Graphique 12 – Le taux d'accès à Internet à domicile est lié à la présence d'enfants dans le foyer

(en % d'individus disposant d'une connexion Internet à domicile)

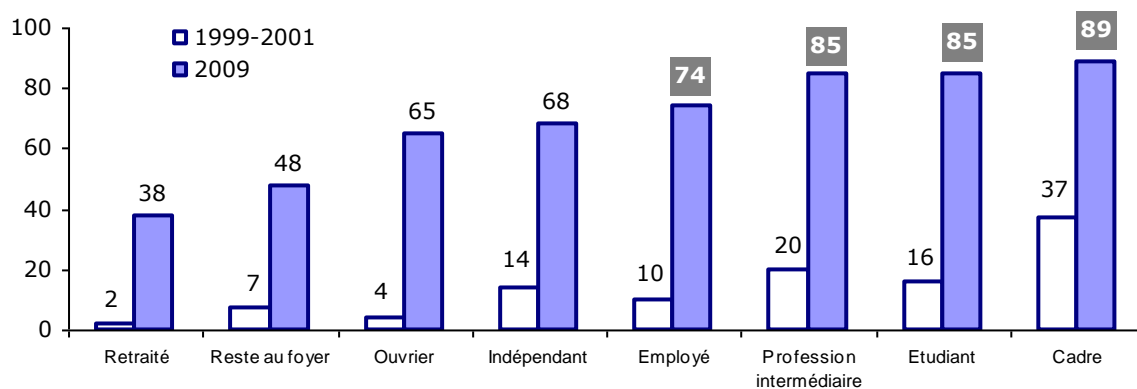


Source : CRÉDOC, enquêtes « Conditions de vie et Aspirations des Français ».

- o Enfin, la **profession catégorie sociale** met également en lumière des écarts conséquents : les cadres et les professions intellectuelles supérieures sont, de loin, les mieux équipés (89%) ; les retraités ferment le ban (38%, Graphique 13).

### Graphique 13 – Le taux d'accès à Internet à domicile selon la PCS de la personne interrogée

(en % d'individus disposant d'une connexion Internet à domicile)



Source : CRÉDOC, enquêtes « Conditions de vie et Aspirations des Français ».

**Tableau 2**  
**Taux d'accès à Internet à domicile en 2009**

*(en %)*

		Est équipé d'Internet	N'est pas équipé d'Internet	Total
<b>Sexe</b>	Homme .....	67	33	100
	Femme .....	62	38	100
<b>Age</b>	18 - 24 ans .....	78	22	100
	25 - 39 ans .....	79	21	100
	40 - 59 ans .....	70	30	100
	60 - 69 ans .....	55	45	100
	70 ans et plus .....	20	80	100
<b>Enfant dans le foyer</b>	Oui .....	82	18	100
	Non .....	55	45	100
<b>Nombre de personnes dans le logement</b>	Un .....	40	60	100
	Deux .....	56	44	100
	Trois .....	87	13	100
	Quatre .....	86	14	100
	Cinq et plus .....	82	18	100
<b>Diplôme</b>	Aucun, Cep .....	36	64	100
	Bepc .....	58	42	100
	Bac .....	77	23	100
	Diplôme du supérieur .....	87	13	100
<b>Profession</b>	Indépendant .....	68	32	100
	Cadre .....	89	11	100
	Profession intermédiaire .....	85	15	100
	Employé .....	74	26	100
	Ouvrier .....	65	35	100
	Reste au foyer .....	48	52	100
	Retraité .....	38	62	100
Etudiant .....	85	15	100	
<b>Revenus par unité de consommation</b>	Catégories pauvres .....	58	42	100
	Catégories modestes .....	49	51	100
	Classes moyennes inférieures ...	61	39	100
	Classes moyennes supérieures ..	70	30	100
	Catégories aisées .....	85	15	100
	Hauts revenus .....	87	13	100
<b>Lieu de résidence</b>	Moins de 2 000 habitants .....	58	42	100
	De 2 000 à 20 000 habitants ....	57	43	100
	De 20 000 à 100 000 habitants .	61	39	100
	Plus de 100 000 habitants .....	69	31	100
	Paris et agglo. parisienne .....	80	20	100
<b>ZEAT</b>	Region parisienne .....	80	20	100
	Nord .....	64	36	100
	Est .....	69	31	100
	Bassin Parisien .....	53	47	100
	Ouest .....	63	37	100
	Sud Ouest .....	62	38	100
	Centre Est .....	62	38	100
	Mediterranee .....	64	36	100
<b>Total .....</b>	<b>65</b>	<b>35</b>	<b>100</b>	

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : début 2009, 87% des diplômés du supérieur disposent, à domicile, d'un accès à Internet. C'est le cas de 36% des non-diplômés.

**Note** : la typologie de catégories de revenus repose sur la répartition de la population en déciles de niveau de vie (catégories pauvres : 1<sup>er</sup> décile ; catégories modestes : 2<sup>ème</sup> et 3<sup>ème</sup> déciles ; classes moyennes inférieures : 4<sup>ème</sup> à 6<sup>ème</sup> déciles ; classes moyennes supérieures : 7<sup>ème</sup> et 8<sup>ème</sup> déciles ; catégories aisées : 9<sup>ème</sup> décile ; hauts revenus : 10<sup>ème</sup> décile).

Une autre façon d'analyser la situation consiste à étudier qui sont, en 2009, les internautes et qui sont les non internautes, en comparant leurs profils respectifs avec l'ensemble de la population (Tableau 3). Les informations similaires relatives à la période 1999-2001 se trouvent au Tableau A 2, p. 104.

Chez les personnes qui disposent, chez elles, d'une connexion à Internet, on relève une sur-représentation des 25-39 ans (34%, + 6 points par rapport à l'ensemble de la population) et des familles avec enfants (44%, + 9 points). Les diplômés du supérieur sont plus nombreux qu'en moyenne (33% contre 24% dans l'ensemble de la population), de même que les cadres, professions intermédiaires et employés (48%, + 10 points par rapport à l'ensemble de la population).

Les personnes non équipées sont plus âgées qu'en moyenne (33% ont 70 ans ou plus), 43% vivent seules (25% en moyenne). Les retraités (41%, + 18 points par rapport à l'ensemble de la population) et les personnes au foyer (16%, + 5 points) sont sur-représentés. Les personnes peu ou pas diplômées sont largement majoritaires (79%, contre 56% dans la population totale).



**Tableau 3**  
**Structure de la population en 2009, selon qu'elle dispose ou pas d'Internet à domicile**

*(en %)*

		Est équipé d'Internet	N'est pas équipé d'Internet	Ensemble de la population
<b>Sexe</b>	. Homme .....	49	44	48
	. Femme .....	51	56	52
<b>Age</b>	. 18 - 24 ans .....	15	8	13
	. 25 - 39 ans .....	34	16	28
	. 40 - 59 ans .....	36	28	33
	. 60 - 69 ans .....	10	15	12
	. 70 ans et plus .....	5	33	15
<b>Enfant dans le foyer</b>	. Oui .....	44	17	35
	. Non .....	56	83	65
<b>Nombre de personnes dans le logement</b>	. Un .....	16	43	25
	. Deux .....	27	40	32
	. Trois .....	25	7	19
	. Quatre .....	20	6	15
	. Cinq et plus .....	11	5	9
<b>Diplôme</b>	. Aucun, Cep .....	11	36	20
	. Bepc .....	33	43	36
	. Bac .....	23	12	19
	. Diplôme du supérieur .....	33	9	24
<b>Profession</b>	. Indépendant .....	5	5	5
	. Cadre .....	10	(2)	7
	. Profession intermédiaire .....	18	6	13
	. Employé .....	20	13	18
	. Ouvrier .....	15	15	15
	. Reste au foyer .....	8	16	11
	. Retraité .....	14	41	23
. Etudiant .....	10	3	8	
<b>Revenus par unité de consommation</b>	. Catégories pauvres .....	8	10	9
	. Catégories modestes .....	14	26	18
	. Classes moyennes inférieures ....	23	28	25
	. Classes moyennes supérieures ..	19	14	17
	. Catégories aisées .....	12	4	9
	. Hauts revenus .....	12	3	9
<b>Lieu de résidence</b>	. Moins de 2 000 habitants .....	23	30	26
	. De 2 000 à 20 000 habitants ....	15	21	17
	. De 20 000 à 100 000 habitants ..	13	15	13
	. Plus de 100 000 habitants .....	31	26	29
	. Paris et aggl. parisienne .....	19	9	15
<b>ZEAT</b>	. Region parisienne .....	21	10	17
	. Nord .....	7	7	7
	. Est .....	9	8	9
	. Bassin Parisien .....	15	24	18
	. Ouest .....	14	15	14
	. Sud Ouest .....	11	12	11
	. Centre Est .....	11	12	12
	. Mediterranee .....	12	12	12
<b>Total .....</b>	<b>100</b>	<b>100</b>	<b>100</b>	

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

Les chiffres entre parenthèses reposent sur des effectifs faibles, ils sont à considérer avec précaution.

**Lecture** : début 2009, 43% de ceux qui n'ont pas, à domicile, un accès à Internet vivent seuls, contre 25% de l'ensemble de la population en moyenne et 16% des personnes équipées.

**Note** : la typologie de catégories de revenus repose sur la répartition de la population en déciles de niveau de vie (catégories pauvres : 1<sup>er</sup> décile ; catégories modestes : 2<sup>ème</sup> et 3<sup>ème</sup> déciles ; classes moyennes inférieures : 4<sup>ème</sup> à 6<sup>ème</sup> déciles ; classes moyennes supérieures : 7<sup>ème</sup> et 8<sup>ème</sup> déciles ; catégories aisées : 9<sup>ème</sup> décile ; hauts revenus : 10<sup>ème</sup> décile).

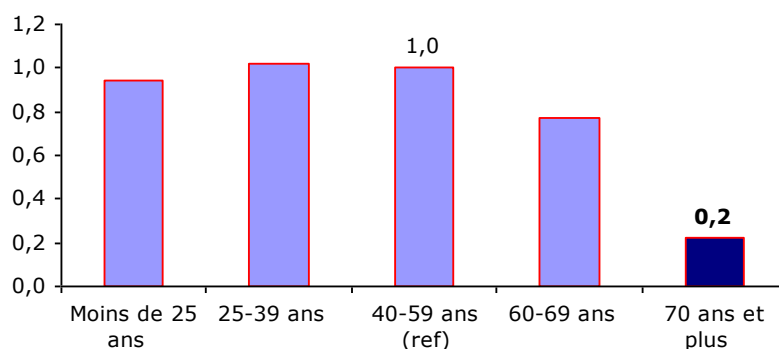
### 3. Les facteurs déterminants pour expliquer, en 2009, l'accès à Internet

Afin de déterminer quels sont **les facteurs les plus déterminants** qui expliquent l'accès à Internet à domicile, nous avons utilisé des régressions logistiques qui permettent d'isoler les effets intrinsèques des différents paramètres.

Au final, six variables se révèlent avoir un effet propre sur le fait d'avoir, ou pas, accès à Internet à domicile. Il s'agit de l'âge, du diplôme, de la taille du foyer, de la profession et catégorie socioprofessionnelle, des revenus et de la taille d'agglomération (Tableau 4). On retiendra que :

- Toutes choses égales par ailleurs, une personne âgée de 70 ans ou plus a presque cinq fois moins de chance de disposer chez elle d'un accès à Internet qu'une personne de 40-59 ans, prise comme modalité de référence (Graphique 14).

**Graphique 14 – Effet intrinsèque de l'âge sur la probabilité d'avoir accès à Internet à domicile**  
- odds ratios issus de régressions logistiques -

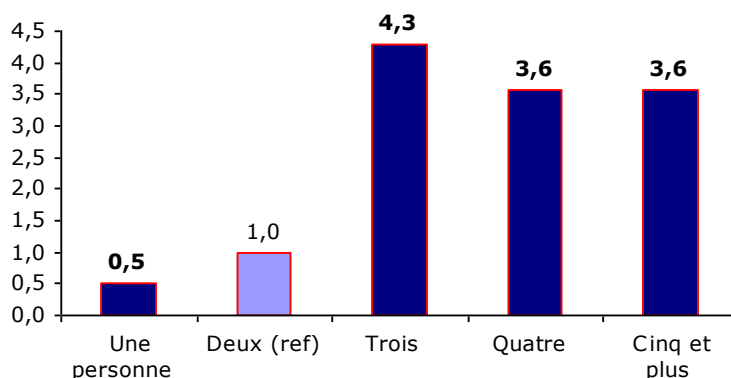


Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

Note : les histogrammes en noir indiquent les modalités significativement différentes de la modalité de référence.

- Une personne qui vit seule a deux fois moins de chance d'être équipée qu'une personne qui vit dans un foyer de deux personnes ; une personne qui vit dans un foyer de trois personnes a quatre fois plus de chance d'être équipée (Graphique 15). Entre ceux qui vivent seuls et ceux qui vivent dans des foyers de trois personnes ou plus, la probabilité d'être connecté à Internet, toutes choses égales par ailleurs, varie presque de 1 à 8.

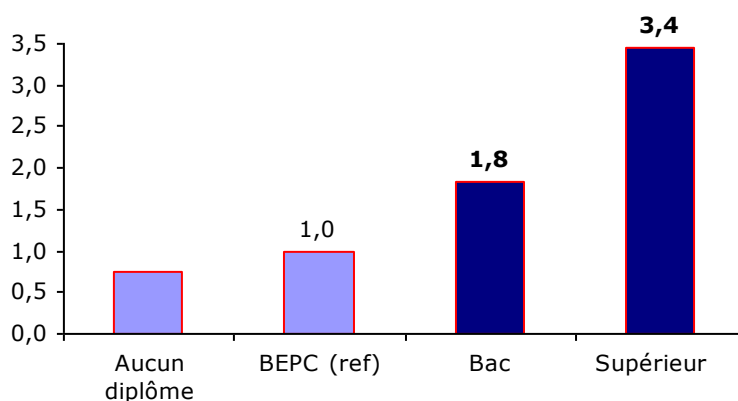
**Graphique 15 – Effet intrinsèque de la taille du foyer sur la probabilité d’avoir accès à Internet à domicile**  
- odds ratios issus de régressions logistiques -



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.  
Note : les histogrammes en noir indiquent les modalités significativement différentes de la modalité de référence.

- La probabilité pour une personne issue de l’enseignement supérieur d’être connectée à domicile est trois fois plus élevée qu’un titulaire du Bepc, à âge, revenus ou structure familiale comparables (Graphique 16).

**Graphique 16 – Effet intrinsèque du diplôme sur la probabilité d’avoir accès à Internet à domicile**  
- odds ratios issus de régressions logistiques -



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.  
Note : les histogrammes en noir indiquent les modalités significativement différentes de la modalité de référence.

- Un étudiant a deux fois plus de chance d’avoir Internet à domicile qu’un employé, pris comme modalité de référence (Graphique 17).

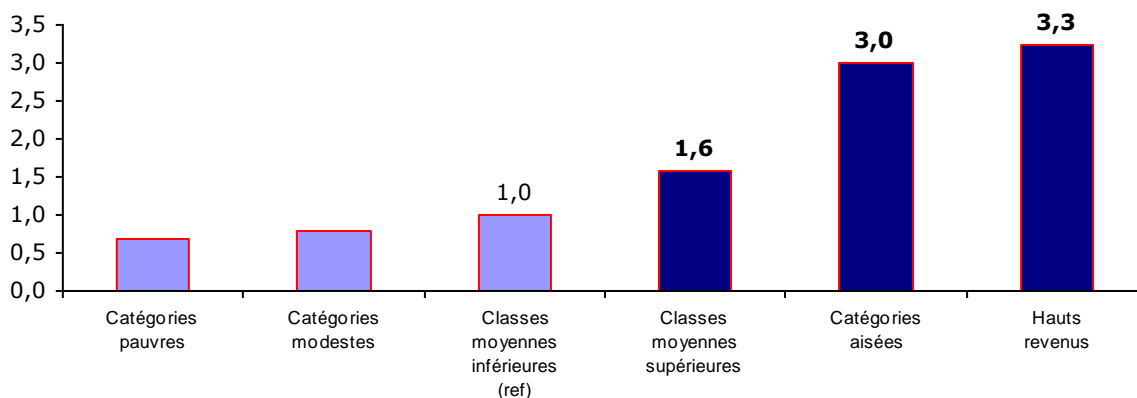
**Graphique 17 – Effet intrinsèque de la PCS sur la probabilité d’avoir accès à Internet à domicile**  
 - odds ratios issus de régressions logistiques -



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.  
 Note : les histogrammes en noir indiquent les modalités significativement différentes de la modalité de référence.

- Toutes choses égales par ailleurs, les catégories aisées ont trois fois plus de chance d’avoir Internet que les classes moyennes inférieures, choisies ici comme modalité de référence (Graphique 18).

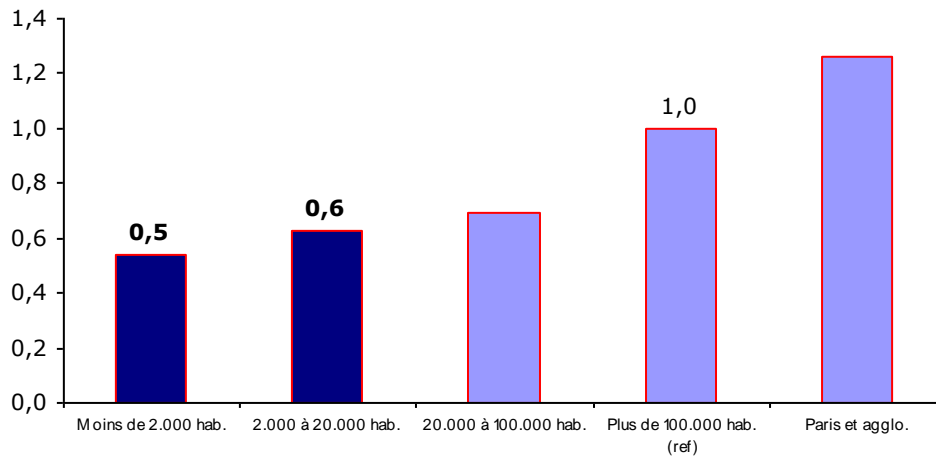
**Graphique 18 - Effet intrinsèque des revenus sur la probabilité d’avoir accès à Internet à domicile**  
 - odds ratios issus de régressions logistiques -



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.  
 Note : les histogrammes en noir indiquent les modalités significativement différentes de la modalité de référence.  
**Note** : la typologie de catégories de revenus repose sur la répartition de la population en déciles de niveau de vie (catégories pauvres : 1<sup>er</sup> décile ; catégories modestes : 2<sup>ème</sup> et 3<sup>ème</sup> déciles ; classes moyennes inférieures : 4<sup>ème</sup> à 6<sup>ème</sup> déciles ; classes moyennes supérieures : 7<sup>ème</sup> et 8<sup>ème</sup> déciles ; catégories aisées : 9<sup>ème</sup> décile ; hauts revenus : 10<sup>ème</sup> décile).

- Une personne résidant dans une agglomération de moins de 2 000 habitants a quasiment deux fois moins de chance, toutes choses égales par ailleurs, qu'un habitant d'une grande agglomération de province d'être équipé en Internet.

**Graphique 19 – Effet intrinsèque du lieu de résidence sur la probabilité d'avoir accès à Internet à domicile**  
- odds ratios issus de régressions logistiques -



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.  
Note : les histogrammes en noir indiquent les modalités significativement différentes de la modalité de référence.

**Tableau 4**  
**Effets des différentes variables sur l'accès à Internet à domicile**

Résultats de régressions logistiques		Odds ratio	Signif (en %)
<b>Age</b>	Moins de 25 ans .....	0,94	80,2
	25-39 ans .....	1,02	91,9
	40-59 ans .....	1,0	Ref.
	60-69 ans .....	0,77	33,2
	<b>70 ans et plus</b> .....	<b>0,22</b>	<b>0,0</b>
<b>Nombre de personnes dans le logement</b>	<b>Une personne</b> .....	<b>0,51</b>	<b>0,0</b>
	Deux .....	1,0	Ref.
	<b>Trois</b> .....	<b>4,29</b>	<b>0,0</b>
	<b>Quatre</b> .....	<b>3,58</b>	<b>0,0</b>
	<b>Cinq et plus</b> .....	<b>3,57</b>	<b>0,0</b>
<b>Diplôme</b>	Aucun diplôme (ou Cep) .....	0,75	6,7
	BEPC, technique (inf. au bac) .....	1,0	Ref.
	<b>Bac, niveau bac</b> .....	<b>1,84</b>	<b>0,0</b>
	<b>Diplôme du supérieur</b> .....	<b>3,44</b>	<b>0,0</b>
<b>Profession et catégorie sociale</b>	Indépendant .....	0,95	85,8
	Cadre .....	0,96	89,4
	Profession intermédiaire .....	1,18	48,7
	Employé .....	1,0	Ref.
	Ouvrier .....	0,91	63,9
	Reste au foyer .....	0,68	8,6
	Retraité .....	0,90	70,3
	<b>Etudiant</b> .....	<b>2,09</b>	<b>3,2</b>
<b>Revenus par unité de consommation</b>	Catégories pauvres .....	0,68	10,1
	Catégories modestes .....	0,78	16,2
	Classes moyennes inférieures .....	1,0	Ref.
	<b>Classes moyennes supérieures</b> .....	<b>1,59</b>	<b>1,0</b>
	<b>Catégories aisées</b> .....	<b>3,00</b>	<b>0,0</b>
	<b>Hauts revenus</b> .....	<b>3,25</b>	<b>0,0</b>
<b>Taille d'agglomération</b>	<b>Moins de 2.000 habitants</b> .....	<b>0,54</b>	<b>0,0</b>
	<b>2.000 à 20.000 habitants</b> .....	<b>0,63</b>	<b>1,0</b>
	20.000 à 100.000 habitants .....	0,69	5,2
	Plus de 100.000 habitants .....	1,0	Ref.
	Paris et agglomération parisienne .....	1,26	25,0

Source : CRÉDOC, Enquête « Conditions de Vie et Aspirations des Français », début 2009.

**Note** : la typologie de catégories de revenus repose sur la répartition de la population en déciles de niveau de vie (catégories pauvres : 1<sup>er</sup> décile ; catégories modestes : 2<sup>ème</sup> et 3<sup>ème</sup> déciles ; classes moyennes inférieures : 4<sup>ème</sup> à 6<sup>ème</sup> déciles ; classes moyennes supérieures : 7<sup>ème</sup> et 8<sup>ème</sup> déciles ; catégories aisées : 9<sup>ème</sup> décile ; hauts revenus : 10<sup>ème</sup> décile).

**Lecture** : toutes choses égales par ailleurs, **les membres des catégories aisées ont 3 fois plus de chance de disposer d'Internet à leur domicile que les membres des classes moyennes inférieures**, pris ici comme modalité de référence.

Modalité de référence : Ref.

## II. La spécificité des modes de vie et des opinions des internautes

Dans cette section, nous avons cherché à savoir si les internautes ont des modes de vie, des attitudes et des opinions spécifiques. Si la réponse est positive, on a déjà une idée du risque que l'on prend en excluant d'une enquête les non-internautes.

L'expérience consiste à savoir si un certain nombre d'attitudes, de comportements et d'opinions sont liées au fait d'être internaute ou pas, et cela « toutes choses égales par ailleurs ». Nous avons donc testé plusieurs modèles de régressions logistiques dans lesquels le fait d'accéder ou pas à Internet chez soi est une variable explicative des comportements et des opinions, au même titre que le sexe, l'âge, le niveau de diplôme, la PCS, le niveau de vie et le lieu de résidence.

Sur 191 variables, on relève **un effet propre à Internet dans 37% des cas**. Toutes choses égales par ailleurs — c'est-à-dire en contrôlant le sexe, l'âge, le niveau de diplôme, le niveau de revenu, la profession, la taille du foyer et le lieu de résidence —, le fait de disposer d'un accès à Internet est, plus d'une fois sur trois, lié à la variable étudiée (71 fois sur 191, Tableau 5).

**Tableau 5 : influence intrinsèque d'Internet sur différents thèmes de l'enquête**

	Nombre de variables	Effet 'Internet' intrinsèque	
		P<0.05	%
Attitudes vis-à-vis des TIC	3	3	100%
Équipement du ménage	15	14	93%
Niveau de vie	8	7	88%
Situation personnelle et familiale	16	12	75%
Opinions en matière de mœurs	5	3	60%
Logement	13	7	54%
Loisirs, pratiques culturelles	21	7	33%
Emploi, chômage, situation professionnelle	10	3	30%
Moral économique	21	6	29%
Opinions sur les politiques familiales	20	5	25%
Jugements sur les politiques sociales	13	2	15%
Environnement	7	1	14%
Inquiétudes et préoccupations	19	1	5%
Santé	10	0	0%
Opinions sur la société	10	0	0%
	191	71	37%

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

Selon les thèmes abordés, la prégnance de l'effet Internet est plus ou moins nette (Graphique 20).

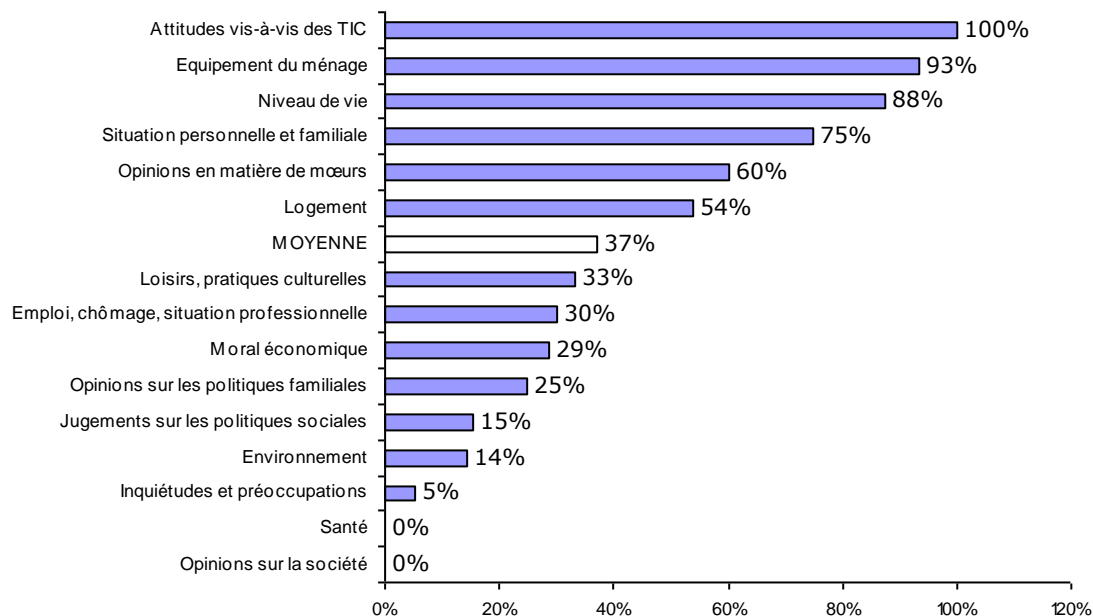
- ✓ Les trois questions relatives aux **attitudes vis-à-vis des nouvelles technologies** et des innovations techniques sont liées à l'accès à Internet à domicile, tout comme 14 des 15 questions relatives à **l'équipement** du ménage. Bien entendu, disposer d'Internet chez soi a un effet intrinsèque sur la possession ou l'usage d'un micro ordinateur, d'un téléphone fixe ou mobile. Mais on trouve aussi des liens avec l'équipement en appareil photo numérique, en lave vaisselle ou en congélateur ! Et, à âge, lieu de résidence et revenus comparables, disposer d'un accès à Internet est également lié à la possession d'une voiture, au fait d'avoir conduit au cours des douze derniers mois et même au kilométrage effectué au cours de l'année écoulée.
- ✓ La plupart des variables s'intéressant au **niveau de vie et au patrimoine** laissent apparaître un fort pouvoir explicatif d'Internet (7 fois sur 8) : le lien est très net avec la possession d'un produit d'assurance vie ou de valeurs mobilières ; il est non négligeable avec la probabilité de percevoir une prestation par la caisse d'allocations familiales, la possession de biens immobiliers ou d'un produit d'épargne liquide.
- ✓ Trois fois sur quatre (12 questions sur 16), la **situation personnelle** est liée, toutes choses égales par ailleurs, à l'accès à Internet. C'est le cas par exemple pour le statut matrimonial, la présence et l'âge des enfants dans le foyer ou encore la situation professionnelle du conjoint.
- ✓ Certaines opinions sur les **mœurs** ne sont pas sans lien avec l'accès à Internet : les prises de position sur l'union homosexuelle, l'adoption par les couples de même sexe et le travail des femmes.
- ✓ Une fois sur deux, les indicateurs ayant trait au **logement** sont, également, liés à l'accès à Internet. C'est le cas par exemple du type de logement, de la taille de celui-ci ou du mode de chauffage principal.

Pour les autres thèmes, la corrélation avec Internet semble moindre. Elle est, en tout cas, moins forte qu'en moyenne (Graphique 20). Enfin, pour les questions relatives à la santé ou à l'environnement, Internet semble n'avoir aucun lien direct ou indirect, tel qu'on peut le mesurer statistiquement au travers des régressions logistiques.



Le détail des résultats peut se lire en annexe, à partir du Tableau A 3, pages 108 et suivantes.

**Graphique 20 : influence intrinsèque d'Internet sur différents thèmes de l'enquête**



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

En résumé, ces premiers résultats indiquent que les internautes présentent de nombreuses particularités qui leur sont propres. Interroger uniquement des internautes comporte un risque de sous-représentation ou de sur-représentation de certaines attitudes, comportements et opinions. Dans de nombreux domaines d'analyse des modes de vie et des opinions, exclure les non-internautes de l'échantillonnage risque d'introduire des erreurs de mesure. Si certains thèmes semblent, *a priori*, moins problématiques que d'autres, les études et les recherches portant sur les nouvelles technologies, sur l'équipement des ménages, sur le niveau de vie, sur la situation patrimoniale, sur la situation familiale, sur les opinions en matière de mœurs et sur le logement présentent des risques.

Nous allons maintenant expérimenter une technique de correction de ces décalages, fondée sur le redressement d'une population d'internautes afin qu'ils ressemblent le plus possible, dans leur structure sociodémographique, à l'ensemble de la population (pondération par post-stratification).

### III. Le redressement de l'échantillon d'internautes

---

#### 1. Le principe du redressement

Nous souhaitons nous placer **dans la situation hypothétique où nous n'aurions interrogé**, dans l'enquête de début 2009 sur les « Conditions de vie et les aspirations des Français », **que les internautes** (65% de l'échantillon interrogé en face-à-face, soit 1286 personnes sur 2006).

La population des internautes, bien que de plus en plus nombreuse, reste particulière. Pour neutraliser ces particularités, nous avons **redressé** ce sous-échantillon d'internautes afin qu'il soit le plus proche possible de la population française. Par population française, nous entendons ici l'ensemble de la population interrogée dans l'enquête (les 2006 individus âgés de 18 ans et plus).

Pour cela, nous utilisons comme critères de redressement les six variables repérées au chapitre précédent, grâce aux régressions logistiques, comme déterminantes de l'accès à Internet : l'âge, la taille du foyer, le diplôme, la PCS, le niveau des revenus et la taille d'agglomération.

Le redressement, qui a pour objectif de rendre comparable – sur les six critères retenus – la sous-population des internautes à celle de la population totale a été effectué grâce au logiciel Spad (Tableau 6). Plusieurs essais ont été menés et, comme souvent, un arbitrage a dû être opéré entre la qualité du redressement (autrement dit l'écart subsistant entre les pourcentages demandés et les pourcentages obtenus) et la dispersion des poids (l'écart entre le poids le plus faible et le plus élevé).

On notera que le profil du sous-échantillon d'internautes dans le Tableau 6 diffère très peu de celui commenté dans le Tableau 3 (profil non redressé *versus* profil redressé).

**Tableau 6**  
**Objectifs du redressement de l'échantillon d'internautes**

*(en %)*

		% mesurés sur le sous- échantillon des internautes	% attendus	% obtenus
<b>Age</b>	. 18 - 24 ans .....	15	13	13
	. 25 - 39 ans .....	34	28	28
	. 40 - 59 ans .....	36	33	33
	. 60 - 69 ans .....	11	12	12
	. 70 ans et plus .....	5	15	13
<b>Nombre de personnes dans le logement</b>	. Un .....	17	25	24
	. Deux .....	28	32	32
	. Trois .....	25	19	20
	. Quatre .....	19	15	15
	. Cinq et plus .....	11	9	9
<b>Diplôme</b>	. Aucun, Cep .....	12	20	19
	. Bepc .....	33	36	36
	. Bac .....	23	19	19
	. Diplôme du supérieur .....	33	24	25
<b>Profession</b>	. Indépendant .....	4	5	5
	. Cadre .....	9	7	7
	. Profession intermédiaire .....	18	13	13
	. Employé .....	20	18	17
	. Ouvrier .....	15	15	15
	. Reste au foyer .....	8	11	10
	. Retraité .....	15	23	25
. Etudiant .....	10	8	8	
<b>Revenus par unité de consommation</b>	. Catégories pauvres .....	8	9	8
	. Catégories modestes .....	15	18	18
	. Classes moyennes inférieures ...	24	25	24
	. Classes moyennes supérieures ..	18	17	18
	. Catégories aisées .....	12	9	9
	. Hauts revenus .....	11	9	9
<b>Lieu de résidence</b>	. Moins de 2 000 habitants .....	23	26	25
	. De 2 000 à 20 000 habitants ....	15	17	17
	. De 20 000 à 100 000 habitants .	13	13	13
	. Plus de 100 000 habitants .....	30	29	30
	. Paris et aggl. parisienne .....	19	15	16
<b>Total .....</b>	<b>100</b>	<b>100</b>	<b>100</b>	

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : début 2009, 33% des internautes sont diplômés de l'enseignement supérieur, contre 24% seulement de l'ensemble de la population.

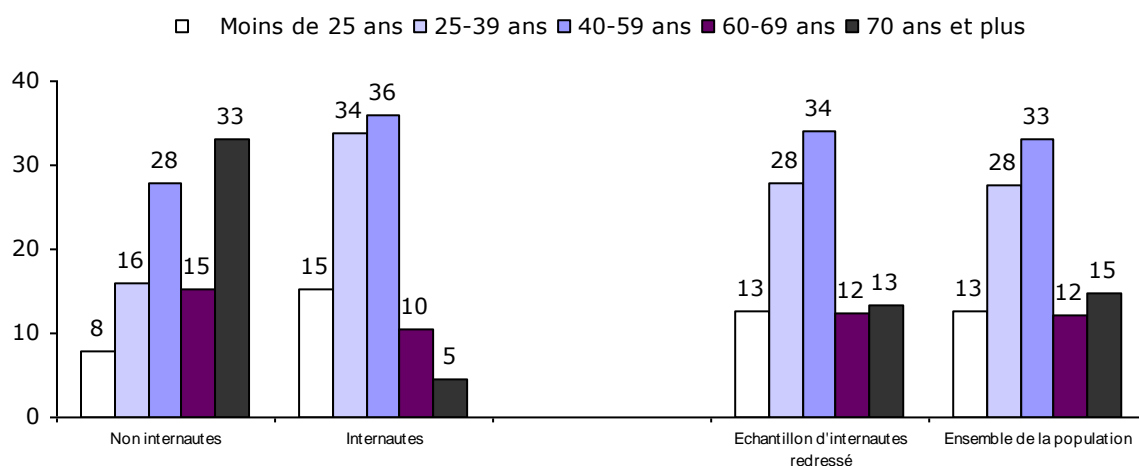
**Note** : la typologie de catégories de revenus repose sur la répartition de la population en déciles de niveau de vie (catégories pauvres : 1<sup>er</sup> décile ; catégories modestes : 2<sup>ème</sup> et 3<sup>ème</sup> déciles ; classes moyennes inférieures : 4<sup>ème</sup> à 6<sup>ème</sup> déciles ; classes moyennes supérieures : 7<sup>ème</sup> et 8<sup>ème</sup> déciles ; catégories aisées : 9<sup>ème</sup> décile ; hauts revenus : 10<sup>ème</sup> décile).

## 2. Les effets du redressement

Les effets du redressement sont très nets, sans pour autant parvenir à gommer tous les écarts (Tableau 6). Pour trois des critères, il reste, après redressement, un écart d'environ 2 points entre les pourcentages demandés (ceux de la population totale) et les pourcentages obtenus.

- Les internautes sont beaucoup plus jeunes que la population totale. Après redressement, la part des 70 ans et plus reste légèrement inférieure à ce qu'elle devrait être (13% contre 15%, Graphique 21).

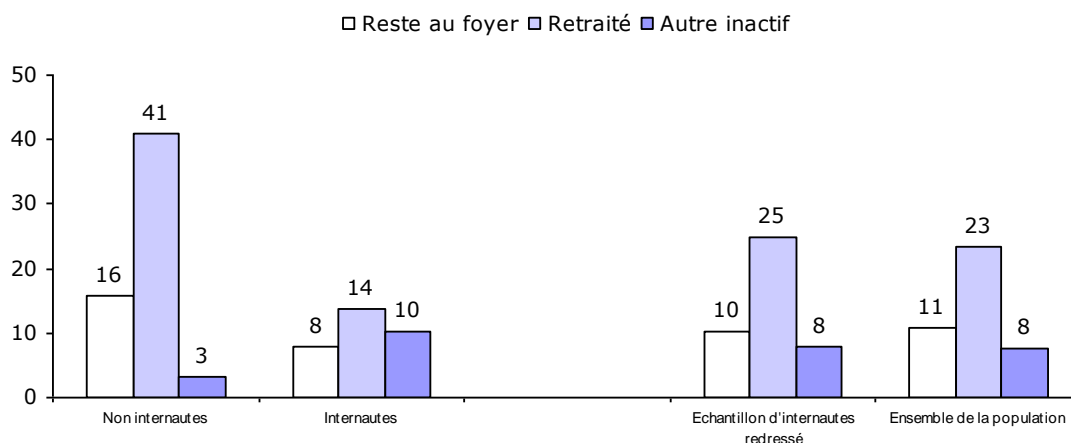
**Graphique 21 : effet du redressement sur l'âge**



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

- S'agissant de la PCS, c'est surtout le poids des inactifs qui pose problème. Le redressement corrige la situation de façon trop prononcée : les retraités passent, en effet, de 14% à 25%, alors qu'ils représentent 23% de la population totale (Graphique 22).

**Graphique 22 : effet du redressement sur la PCS, le cas des inactifs**

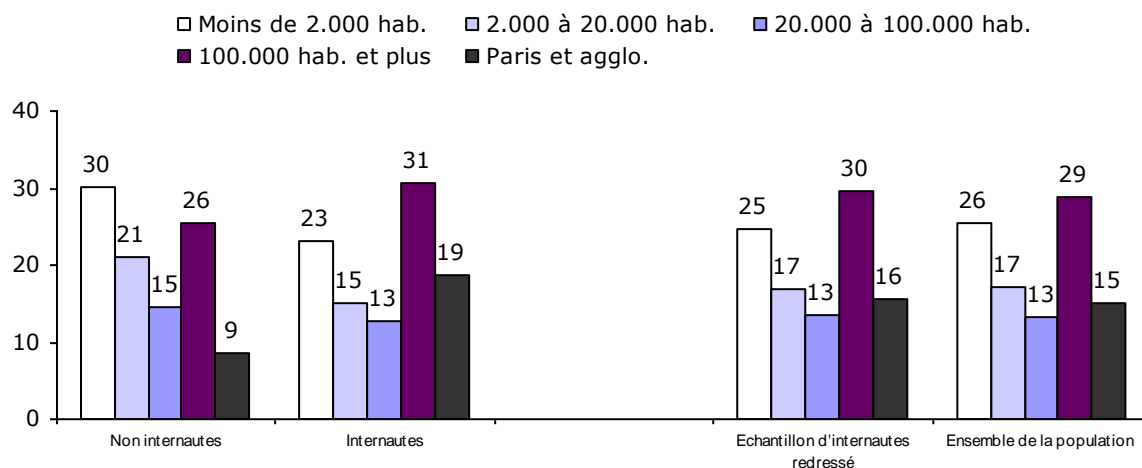


Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

On atteint là les limites de l'exercice de redressement : on ne pourrait pas, en effet, obtenir davantage de personnes âgées de 70 ans et plus, ce qui est souhaitable, sans augmenter la part des retraités, déjà trop importante.

- o Enfin, les internautes se recrutent davantage dans les zones urbaines. Le redressement corrige partiellement la donne : il subsiste une légère sur-représentation des personnes résidant à Paris ou dans les grandes agglomérations de province (46% contre 44%, Graphique 23).

**Graphique 23 : effet du redressement sur la taille d'agglomération**



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

Pour les autres critères, le redressement remplit parfaitement sa fonction. Le détail des différents profils se trouve en annexe (Graphique A 1 pour la situation des actifs et Graphique A 2 pour la taille du foyer p. 105 ; Graphique A 3 pour le niveau de diplôme et Graphique A 4 pour les revenus p.106).

## IV. L'échantillon d'internautes redressé : des Français comme les autres ?

Pour appréhender à la fois l'ampleur et la nature des risques d'erreurs liés à l'interrogation des seuls internautes, nous avons comparé, à partir de 196 variables, les réponses de quatre groupes distincts :

- les internautes
- les non internautes
- les internautes « redressés »
- l'ensemble de la population (internautes + non-internautes).

Nous avons ensuite calculé les écarts entre les pourcentages de réponses apportées à 196 questions. Pour déterminer si les écarts entre les échantillons sont significatifs ou pas, nous avons appliqué les modes de calculs utilisés dans les modèles probabilistes<sup>114</sup>, en retenant le seuil de significativité de 5%.

Les résultats sont synthétisés dans le Tableau 7. On constate que les différences entre internautes et non-internautes sont extrêmement fréquentes (dans 67% des cas). Dans environ un tiers des cas (34%), les réponses de l'échantillon des internautes non redressé diffèrent de celles apportées par l'ensemble de la population. Enfin, et surtout, **dans 12% des cas, l'échantillon des internautes redressé apporte des réponses qui diffèrent de la population totale.**

**Tableau 7 : les écarts d'attitudes entre les différents échantillons (non-internautes, internautes, internautes « redressés », population totale)**

	Nombre d'indicateurs	Ecart d'attitudes entre les internautes et les non-internautes		Ecart d'attitudes entre les internautes et la population générale		Ecart d'attitudes entre l'échantillon des internautes "redressés" et la population générale	
Attitudes vis-à-vis des TIC	3	3	100%	3	100%	3	100%
Équipement du ménage	15	14	93%	13	87%	13	87%
Opinions en matière de mœurs	5	5	100%	4	80%	2	40%
Logement	13	7	54%	2	15%	2	15%
Moral économique	21	12	57%	5	24%	2	10%
Loisirs, pratiques culturelles	22	17	77%	12	55%	2	9%
Situation personnelle et familiale	17	15	88%	13	76%	0	0%
Niveau de vie	8	7	88%	6	75%	0	0%
Emploi, chômage, situation professionnelle	13	8	62%	3	23%	0	0%
Santé	10	7	70%	2	20%	0	0%
Environnement	7	4	57%	1	14%	0	0%
Opinions sur la société	10	3	30%	1	10%	0	0%
Jugements sur les politiques sociales	13	9	69%	1	8%	0	0%
Opinions sur les politiques familiales	20	9	45%	0	0%	0	0%
Inquiétudes et préoccupations	19	11	58%	0	0%	0	0%
	196	131	67%	66	34%	24	12%

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

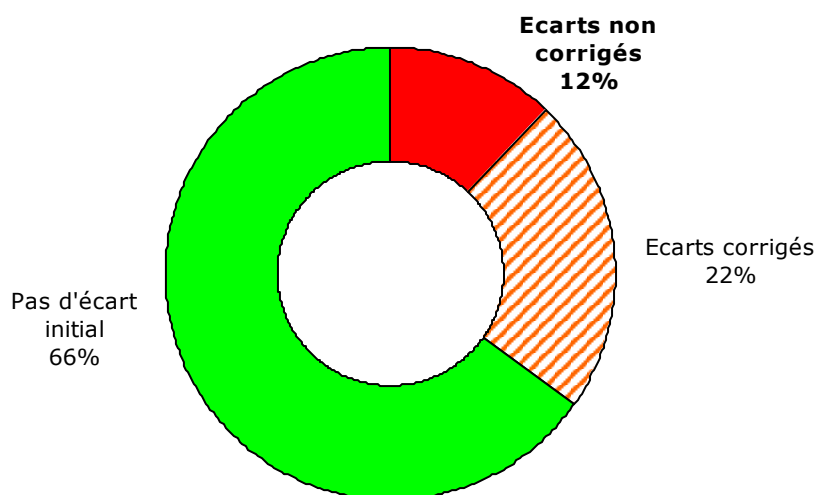
<sup>114</sup> Cela revient à faire l'hypothèse que l'échantillonnage de l'enquête « Conditions de vie et aspirations » remplit les conditions d'un sondage probabiliste. Or, il convient de rappeler que cette enquête est réalisée selon la méthode des quotas, avec une stratification en deux temps : découpage de l'échantillon global en 9 grandes régions (ZEAT), puis, au sein de ces ZEAT, quotas de sexe, d'âge (6 critères), de PCS (12 critères) et de taille d'agglomération (9 critères).

Au final, les 196 variables peuvent être réparties en trois catégories :

- ✓ 24 questions laissent apparaître des spécificités des internautes, spécificités qui ne peuvent pas être corrigées par redressement.
- ✓ 43 questions pour lesquelles le redressement a été nécessaire et suffisant : les internautes se distinguaient, mais, après redressement, l'échantillon redressé apporte des réponses identiques à celles de la population totale.
- ✓ 129 variables où les réponses des internautes sont comparables à celles de l'ensemble de la population, sans qu'il soit nécessaire de procéder à un redressement.

Le détail des résultats est consigné en annexe, à partir du Tableau A 3, pages 107 et suivantes.

**Graphique 24 : les écarts d'attitudes entre les internautes et la population générale avant et après redressement**



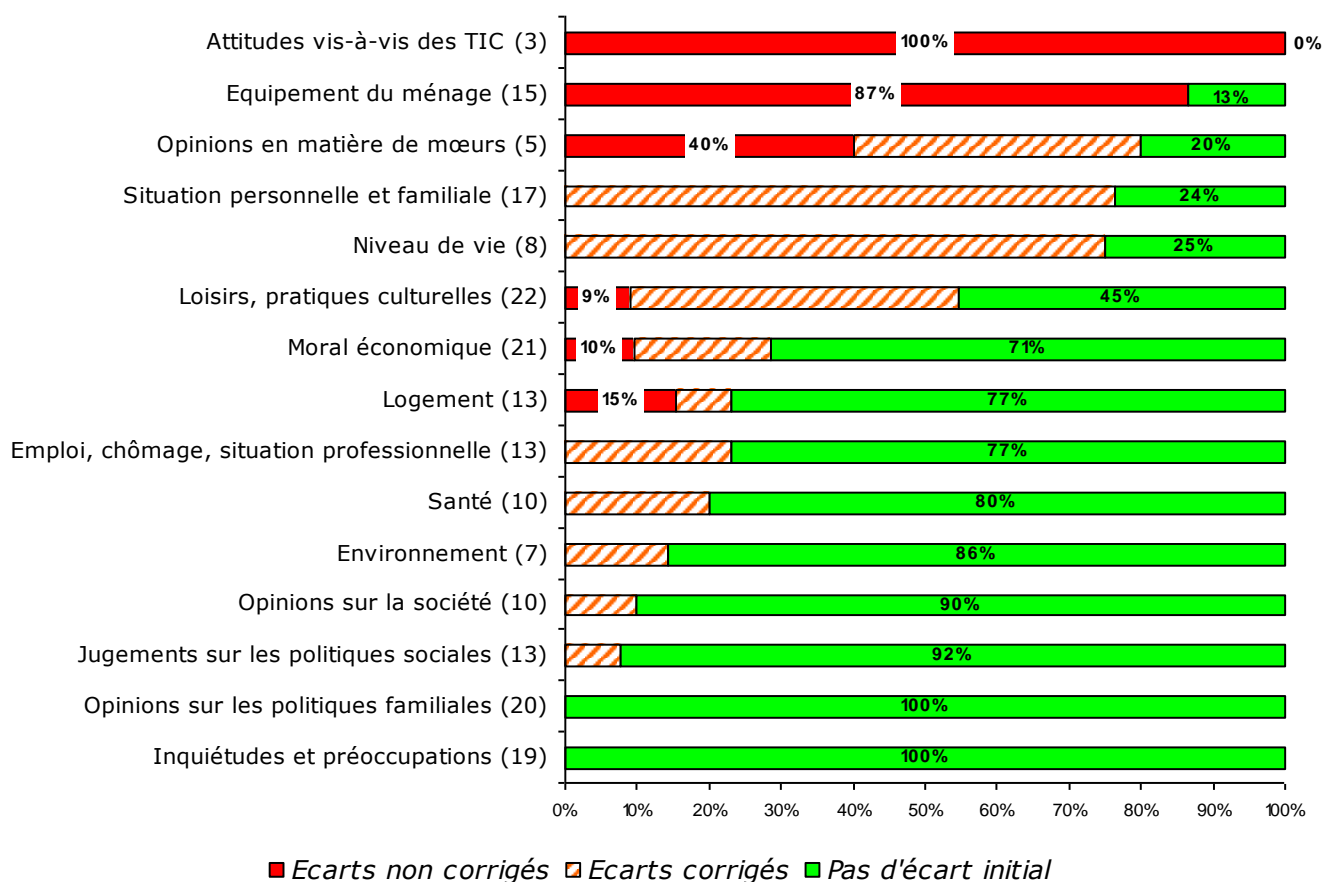
Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

Nous allons maintenant étudier plus en détail quels domaines sont concernés par les risques d'erreur de mesure (Graphique 25). Sur la quinzaine de champs d'investigation analysés :

- ✓ Deux thèmes semblent à l'abri des erreurs : il s'agit des inquiétudes et préoccupations (19 questions), ainsi que des opinions sur les politiques familiales (20 questions). Même sans redressement, les réponses des internautes sont très proches de celles de l'ensemble de la population.

- ✓ Pour sept autres thèmes, certains écarts existent mais ils sont corrigés après redressement. Il s'agit des jugements sur les politiques sociales (13 questions), des opinions sur la société (10 questions), des attitudes par rapport à l'environnement (7 questions), des perceptions de son état de santé et des opinions sur le système de soins (10 questions), de la situation professionnelle (13 questions), du niveau de vie et du patrimoine (8 questions) et de la situation personnelle et familiale (17 questions).
- ✓ Six thèmes comportent davantage de risques car, malgré le redressement, des écarts subsistent entre les réponses des internautes et celles de l'ensemble de la population. Le plus problématique concerne les attitudes sur les technologies de l'information (3 questions) et, plus généralement, l'équipement du ménage (13 des 15 questions). Certaines opinions en matière de mœurs (2 des 5 questions), des caractéristiques du logement (2 des 13 questions), des indicateurs liés au moral économique (21 questions) ou aux loisirs et pratiques culturelles (2 des 22 questions) sont spécifiques chez les internautes et le redressement ne permet pas de corriger les écarts.

**Graphique 25 - Les écarts d'attitudes entre les internautes et la population générale, avant et après redressement**



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.



## 1. TIC, équipement mais aussi moral économique ou mœurs : les variables où le redressement est insuffisant

On s'intéresse ici à la vingtaine de questions pour lesquelles le redressement n'aura pas été suffisant. La plupart ont trait à l'équipement et aux TIC, mais on constate également que l'interrogation des internautes sur le moral économique, les pratiques de loisirs ou les opinions en matière de mœurs peut se révéler risquée : sur ces thèmes également, les internautes ont des prises de positions spécifiques qui ne sont pas corrigées par redressement.

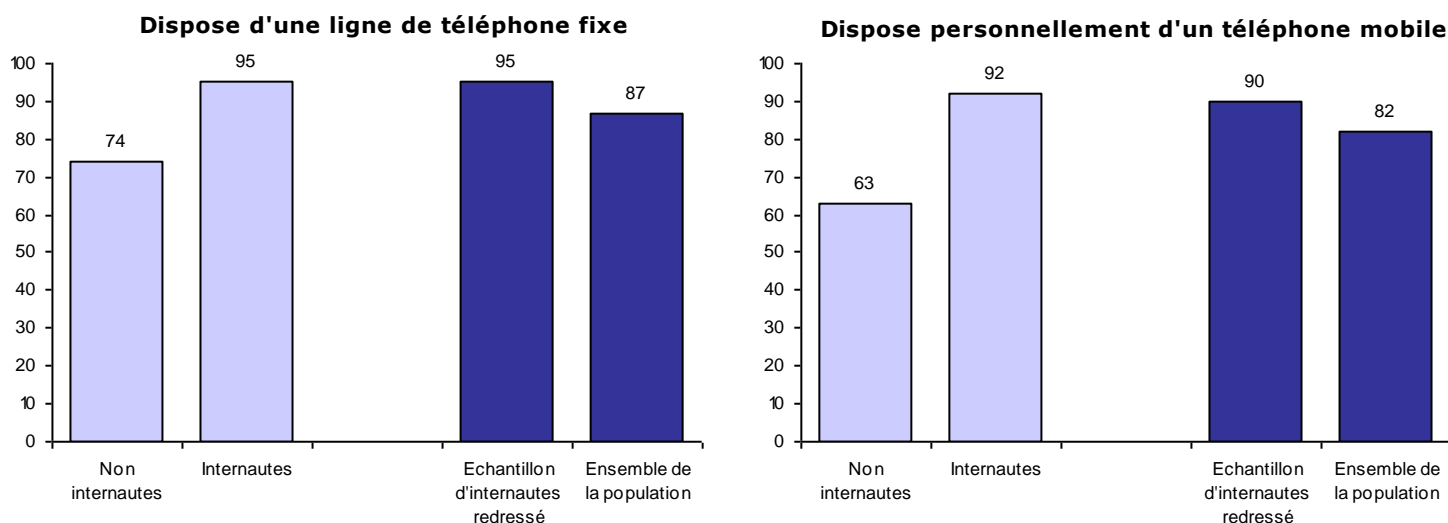
On notera que toutes ces variables avaient été repérées au chapitre précédent : toutes choses égales par ailleurs, l'accès à Internet avait, sur ces questions, un rôle explicatif propre.

### 13 QUESTIONS PORTENT SUR L'ÉQUIPEMENT DU MENAGE

C'est une évidence : même après redressement, le taux d'équipement chez les internautes reste significativement supérieur pour ce qui concerne le micro-ordinateur. Le redressement parvient à peine à diminuer légèrement le taux de multi-équipement.

Pour le téléphone, qu'il soit fixe ou mobile, les internautes affichent, même après redressement, des taux d'équipement supérieurs de 8 points à la moyenne nationale (Graphique 26).

**Graphique 26 : le redressement ne parvient pas à corriger les écarts d'équipement en téléphonie**



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

On constate que les niveaux d'équipement sont bien plus élevés pour toute une série de biens qui ne s'apparentent pas forcément à des produits à fort contenu technologique (Tableau 8) : le sur-équipement s'étend aussi bien au lave vaisselle (+ 7 points), qu'au magnétoscope (+ 5 points), à l'appareil photo numérique (+ 7 points) ou au lecteur de DVD (+ 6 points).

Tous les indicateurs ayant trait à l'automobile font état du même avantage qui persiste en faveur des internautes malgré le redressement : ils ont plus souvent l'usage d'une voiture (+ 4 points), ils ont plus souvent conduit au cours de l'année écoulée (+ 4 points), le kilométrage affiché est plus élevé ...

Le redressement corrige une partie du différentiel entre les internautes et la population totale, mais il ne réussit pas à gommer tous les écarts (voir aussi en annexe Tableau A 3, p. 107).

**Tableau 8 : les 13 variables d'équipement du ménage non suffisamment corrigées par redressement**

		Non internautes	Internautes	Echantillon d'internautes redressé	Ensemble de la population
<b>Lave-vaisselle ?</b>	oui	34	61	58	51
	non	66	39	42	49
<b>Magnétoscope ?</b>	oui	53	64	65	60
	non	47	36	35	40
<b>Lecteur DVD ?</b>	oui	68	92	90	84
	non	32	8	10	16
<b>Appareil photo numérique ?</b>	oui	30	82	76	63
	non	70	18	24	37
<b>Four micro-ondes ?</b>	oui	82	94	93	89
	non	18	6	7	11
<b>Micro-ordinateur à domicile</b>	oui, un seul	20	70	76	53
	oui, plusieurs	1	29	24	19
	non, aucun	79	0	1	28
<b>Téléphone fixe</b>	oui	74	95	95	87
	non	26	5	5	13
<b>Nb de postes téléphoniques</b>	aucun	26	5	5	13
	un seul	55	57	57	57
	plusieurs	19	37	38	31
<b>Téléphone mobile</b>	oui	63	92	90	82
	non	37	8	10	18
<b>Voiture</b>	oui	67	84	82	78
	non	33	16	18	22
<b>A conduit ?</b>	oui	60	82	78	74
	non, a permis	16	4	6	8
	n a pas le permis	24	14	15	17
<b>Nb voitures</b>	une	54	44	48	47
	plusieurs	13	40	34	30
	aucune	33	16	18	22
<b>Nb kilomètres parcourus</b>	moins de 3000 km	11	10	10	10
	3000 ... 9000 km	18	16	16	17
	9000 ... 19000 km	20	28	28	25
	19000 km et plus	10	24	22	19
	n.s.p., nr	2	3	3	3
	non concerne	40	18	22	26

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

### 3 QUESTIONS DECRIVENT LES ATTITUDES VIS-A-VIS DES NOUVELLES TECHNOLOGIES

Même après redressement (Tableau 9 et Tableau A 4 p. 103), les internautes restent plus enthousiastes vis-à-vis des découvertes scientifiques (+ 5 points), ils estiment plus souvent que la diffusion de l'informatique est souhaitable (+ 4 points) et ils sont davantage attirés par des produits comportant des innovations technologiques (+ 5 points).

**Tableau 9 : les 3 variables d'opinions sur les TIC non suffisamment corrigées par redressement**

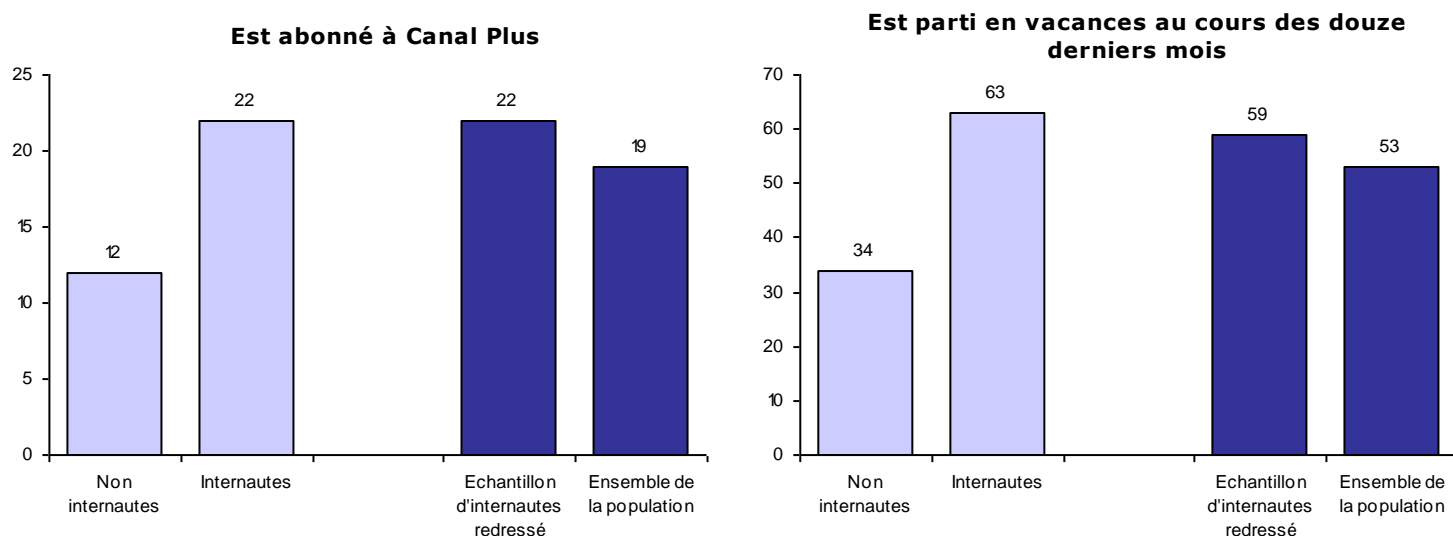
		Non internautes	Internautes	Echantillon d'internautes redressé	Ensemble de la population
<b>Les découvertes scientifiques améliorent-elles votre vie quotidienne ?</b>	un peu	61	54	54	56
	beaucoup	21	36	34	31
	pas du tout	17	9	10	12
<b>La diffusion de l'informatique ...</b>	souhaitable	24	44	41	37
	inévitable	57	48	50	51
	regrettable	18	8	9	11
<b>Est attiré par innovations technologiques</b>	très attiré	5	15	14	12
	assez attiré	18	35	32	29
	peu attiré	34	34	35	34
	pas du tout attiré	43	16	19	26

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

### 2 QUESTIONS ONT TRAIT AUX PRATIQUES CULTURELLES, AUX LOISIRS

Les internautes sont plus souvent qu'en moyenne abonnés à Canal +, et le redressement n'y change rien (Graphique 27 et Tableau A 5 p. 110). S'agissant des départs en vacances, la correction est insuffisante (+10 points pour les internautes avant redressement, + 5 points après).

**Graphique 27 : le redressement ne parvient pas à corriger certains écarts de pratiques culturelles**



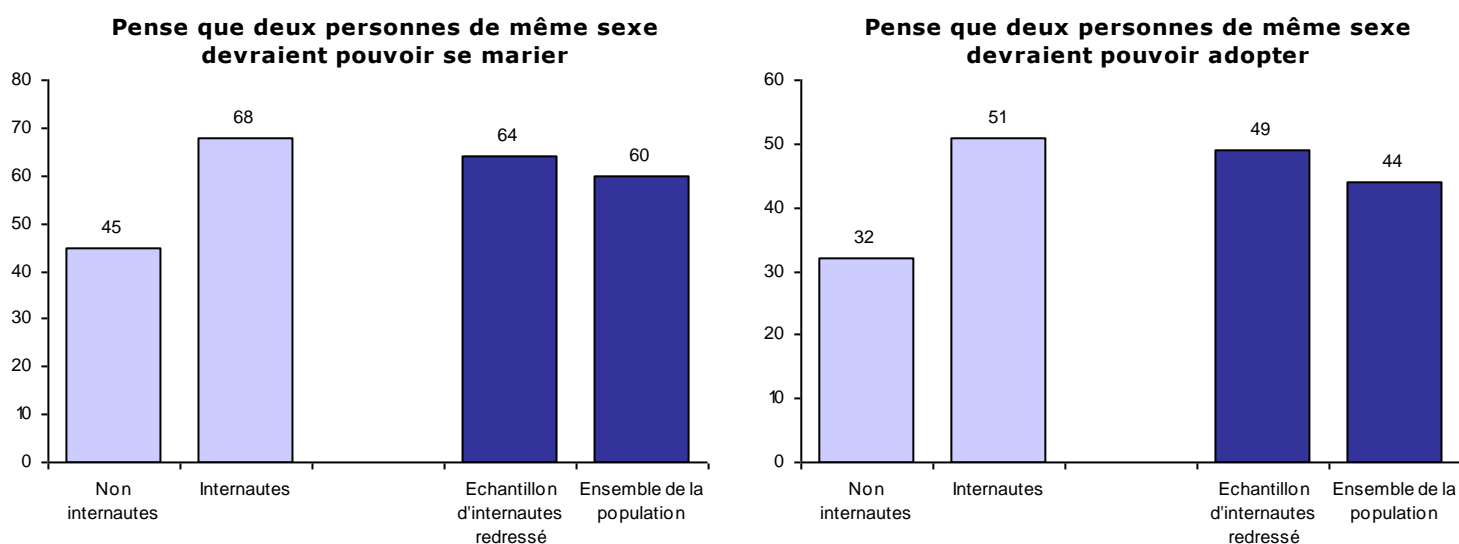
Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

## 2 QUESTIONS MESURENT L'ACCEPTATION DU MARIAGE HOMOSEXUEL ET DE L'HOMOPARENTALITE

Les internautes font montre, s'agissant des couples homosexuels, d'une plus grande ouverture d'esprit. Cette particularité n'est que partiellement corrigée par le redressement (Graphique 28 et Tableau A 6 p.103).

Ils sont, en effet, davantage partisans du mariage homosexuel (+ 8 points avant redressement, + 4 points après) et plus souvent favorables à l'idée que deux personnes de même sexe devraient adopter un enfant (+ 7 points avant redressement, + 5 points après).

**Graphique 28 : le redressement ne parvient pas à corriger les opinions sur les couples homosexuels**



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

## 1 QUESTION S'INTERESSE AU LOGEMENT

Les internautes habitent moins fréquemment dans des logements anciens, et le redressement ne corrige pas cet écart (- 5 points pour des logements achevés avant 1962, après redressement, voir Tableau A 7 p. 103).

## 1 QUESTIONS CONCERNE LE MORAL ECONOMIQUE

Alors que les internautes bénéficient d'un équipement automobile bien supérieur à la moyenne, ils déclarent, dans le même temps, davantage de restrictions à ce sujet : 45%

se plaignent (+ 10 points par rapport aux non internautes). Le redressement accentue même légèrement la tendance (46% des internautes de l'échantillon redressé disent se restreindre, + 5 points par rapport à la moyenne de la population, voir Tableau A 8 p. 103).

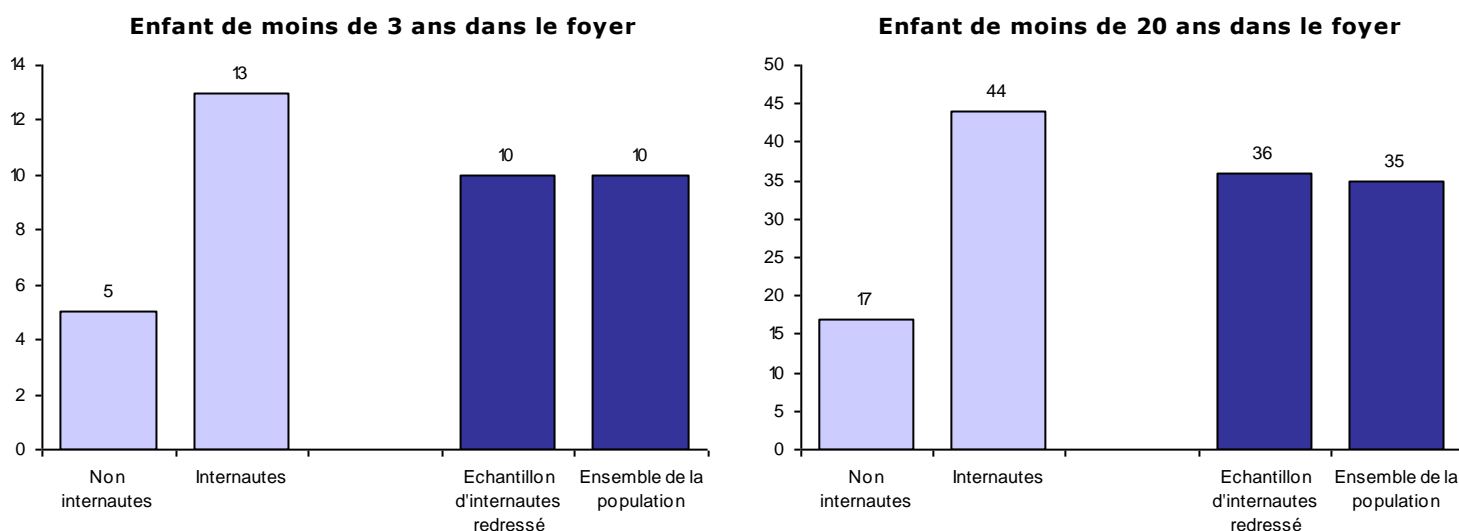
## 2. Les variables où le redressement a été nécessaire et où il s'avère suffisant

### 13 QUESTIONS CONCERNANT LA SITUATION PERSONNELLE ET FAMILIALES

Beaucoup de variables ayant trait à la situation personnelle et familiale sont corrigées par le redressement parce qu'elles sont plus ou moins liées aux variables directement utilisées pour le redressement (comme l'âge, la taille du foyer ou la PCS).

Alors que les internautes vivent moins souvent seuls et ont plus souvent la charge d'enfants, le redressement permet de corriger parfaitement tous les paramètres qui en découlent (type de ménage, nombre d'enfants que l'enquêté a eu ...), y compris l'âge des enfants présents dans le ménage (Graphique 29).

**Graphique 29 : le redressement corrige la présence et l'âge des enfants dans le foyer**

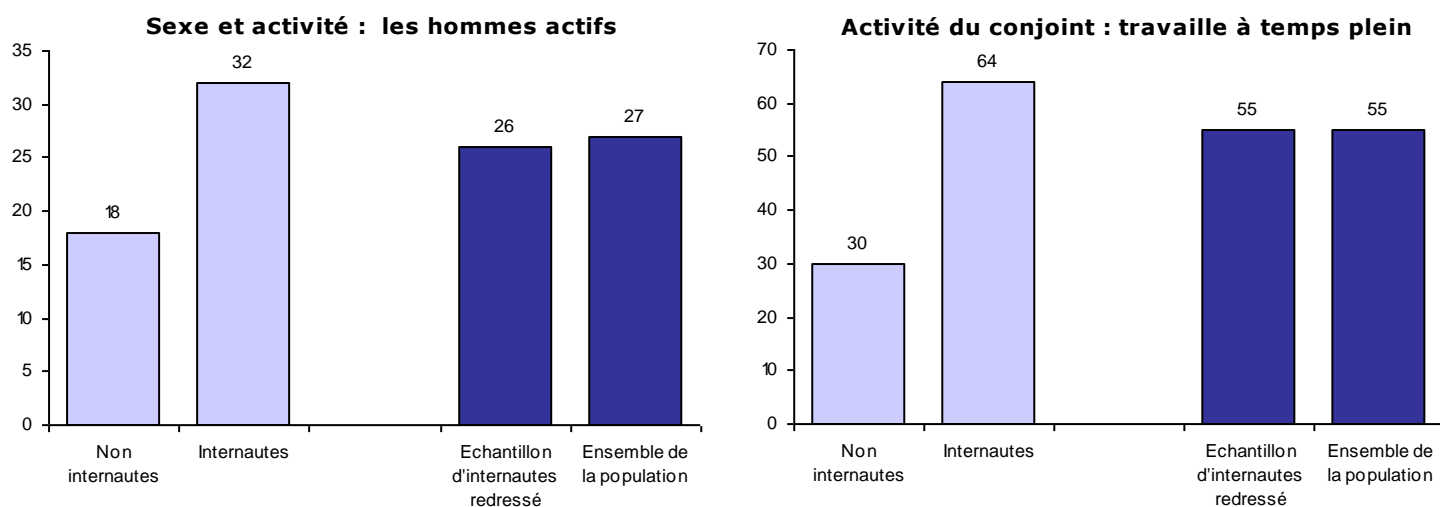


Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

Pareillement, les variables croisant l'âge et le sexe ou l'âge et la région de résidence sont corrigées grâce au redressement.

Enfin, d'autres indicateurs sur l'activité de la personne interrogée (sexe x activité) ou de son conjoint ne présentent plus d'écarts avec la population totale après redressement (Graphique 30). Toutes les variables concernées sont décrites en annexe (Tableau A 9, p.114).

**Graphique 30 : le redressement corrige la situation de la personne interrogée et de son conjoint**

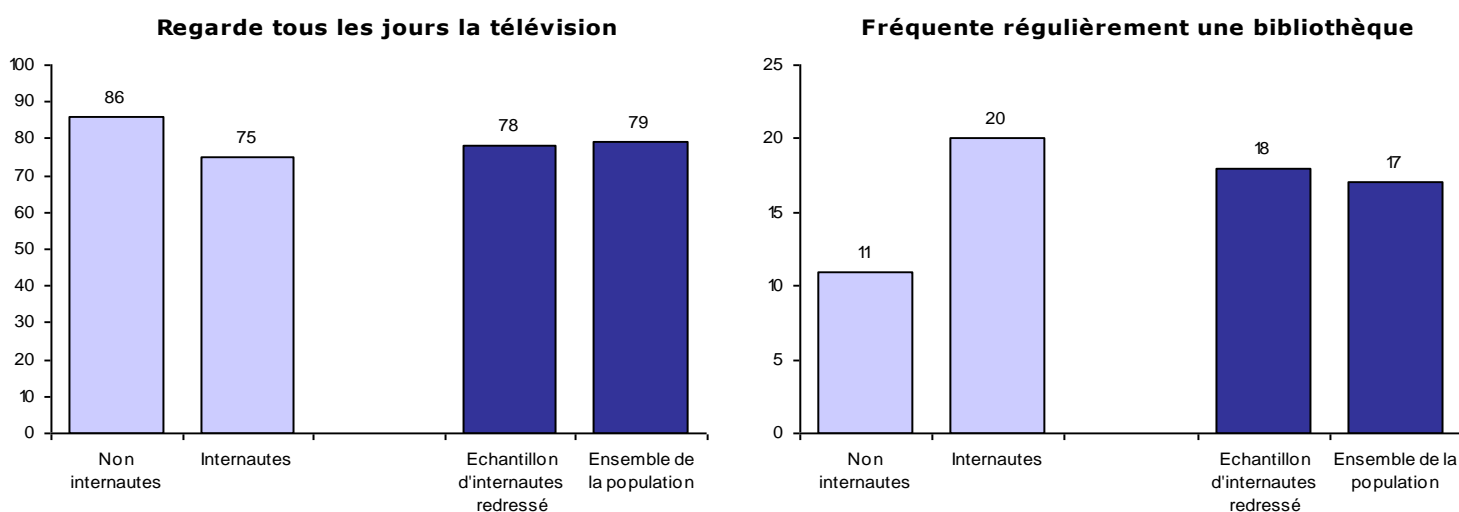


Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

### 10 QUESTIONS CONCERNENT LES PRATIQUES CULTURELLES, LES LOISIRS

Les internautes ont une sociabilité plus étendue, plus diversifiée que l'ensemble de la population : ils reçoivent plus régulièrement des invités chez eux, ils adhèrent plus fréquemment à toutes sortes d'associations et fréquentent, de façon plus assidue, cinéma, bibliothèque ou équipement sportif (Tableau A 5, p.110). En revanche, ils passent moins de temps devant le petit écran. Toutes ces spécificités sont corrigées par le redressement. Le Graphique 31 illustre les corrections opérées sur la pratique télévisuelle et la fréquentation d'une bibliothèque.

**Graphique 31 : le redressement corrige nombre de pratiques culturelles et de loisirs**



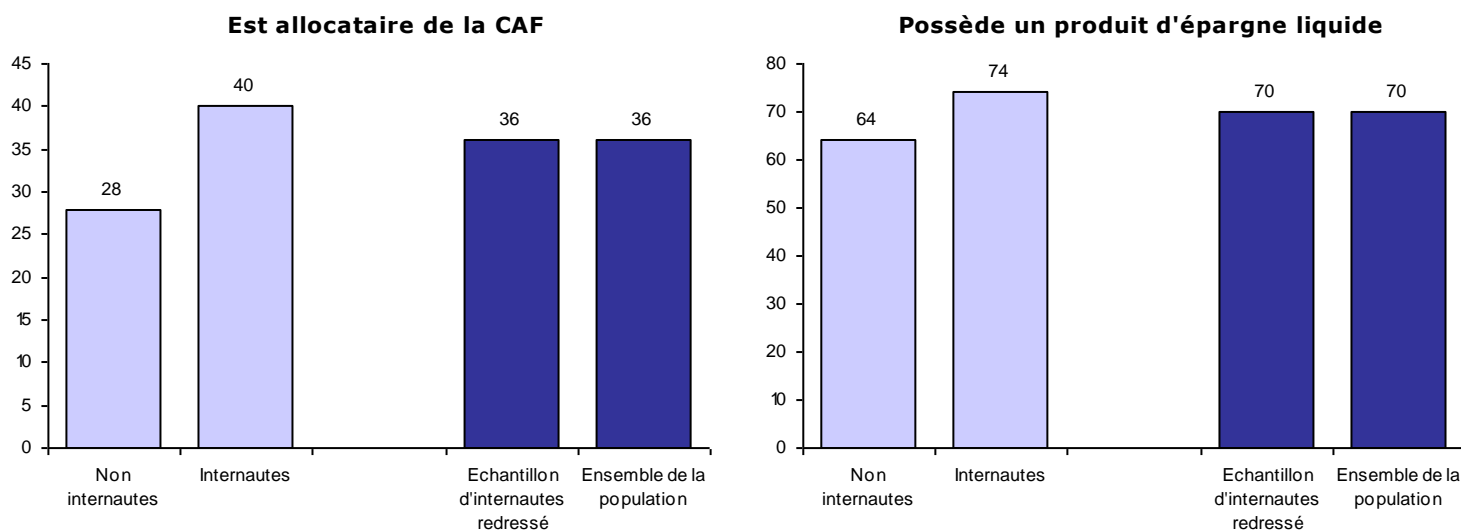
Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

## 6 QUESTIONS CONCERNENT LE NIVEAU DE VIE ET LE PATRIMOINE

S'agissant du niveau de vie, nous savons que les internautes bénéficient de revenus supérieurs à la moyenne. Ils sont également, plus souvent qu'en moyenne, détenteurs de valeurs mobilières, de produits d'épargne liquide ou d'assurance vie (Tableau A 10, p.115). Le fait qu'ils aient plus souvent des enfants à charge les rend, dans le même temps, plus susceptibles de bénéficier d'une prestation versée par la CAF, mais rarement pour un minimum social ou une prestation logement.

De même que le redressement corrige la distribution des revenus du foyer, il corrige également tous ces écarts (Graphique 32).

**Graphique 32 : le redressement corrige tous les biais liés au niveau de vie**



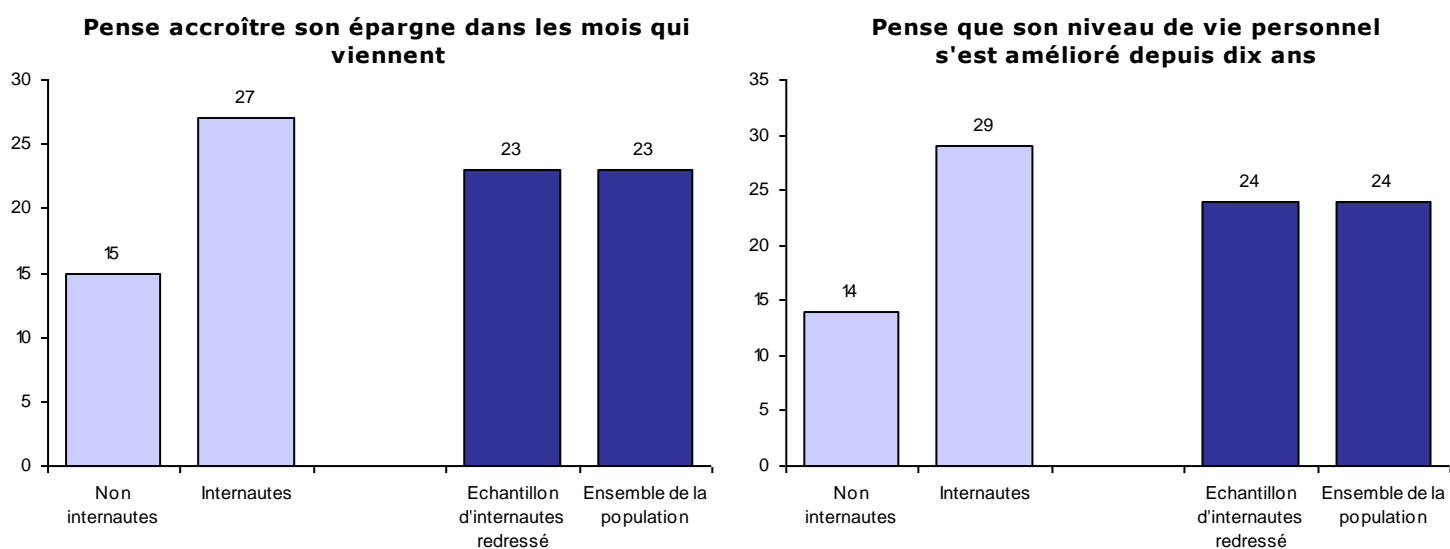
Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

## 4 QUESTIONS CONCERNENT LE MORAL ECONOMIQUE

A la fois parce qu'ils sont plus jeunes et que leur situation financière est bonne, les internautes se révèlent plus optimistes quant à la perception de l'évolution de leur niveau de vie. Ils sont également davantage convaincus que leurs conditions de vie vont prochainement s'améliorer ; ils pensent pouvoir épargner plus souvent qu'en moyenne et portent un regard moins critique sur le passage à l'Euro.

Toutes ses spécificités disparaissent après redressement (Tableau A 8, p. 113 et Graphique 33). Rappelons qu'il subsiste néanmoins, dans ce thème, deux questions pour lesquelles le redressement est inopérant : les restrictions déclarées à propos de l'automobile et des produits de beauté.

**Graphique 33 : le redressement corrige quasiment tous les indicateurs liés au moral économique**

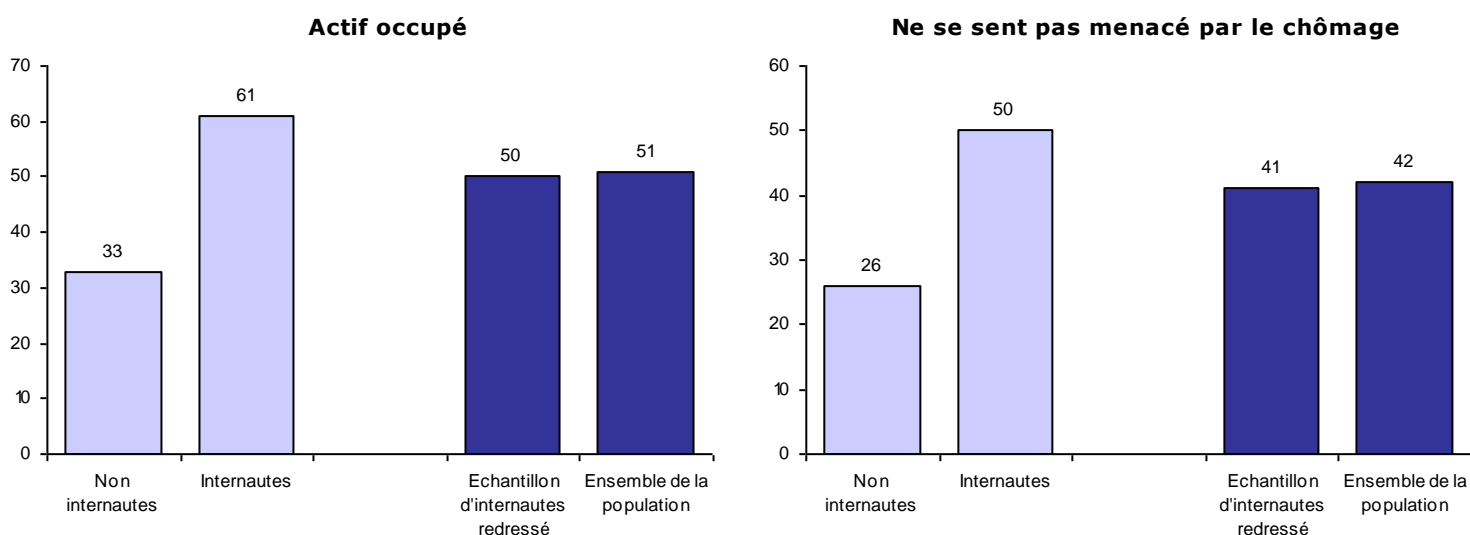


Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

### 3 QUESTIONS CONCERNENT LA SITUATION PROFESSIONNELLE

Très liées à la profession catégorie sociale, les deux variables plus ou moins détaillées décrivant la situation d'emploi sont corrigées par le redressement, de même que le sentiment, pour la personne interrogée, d'être menacé par le chômage (Graphique 34 et Tableau A 11 p. 116).

**Graphique 34 : après redressement, plus aucun écart ne subsiste s'agissant de l'emploi, du chômage et de la situation professionnelle**



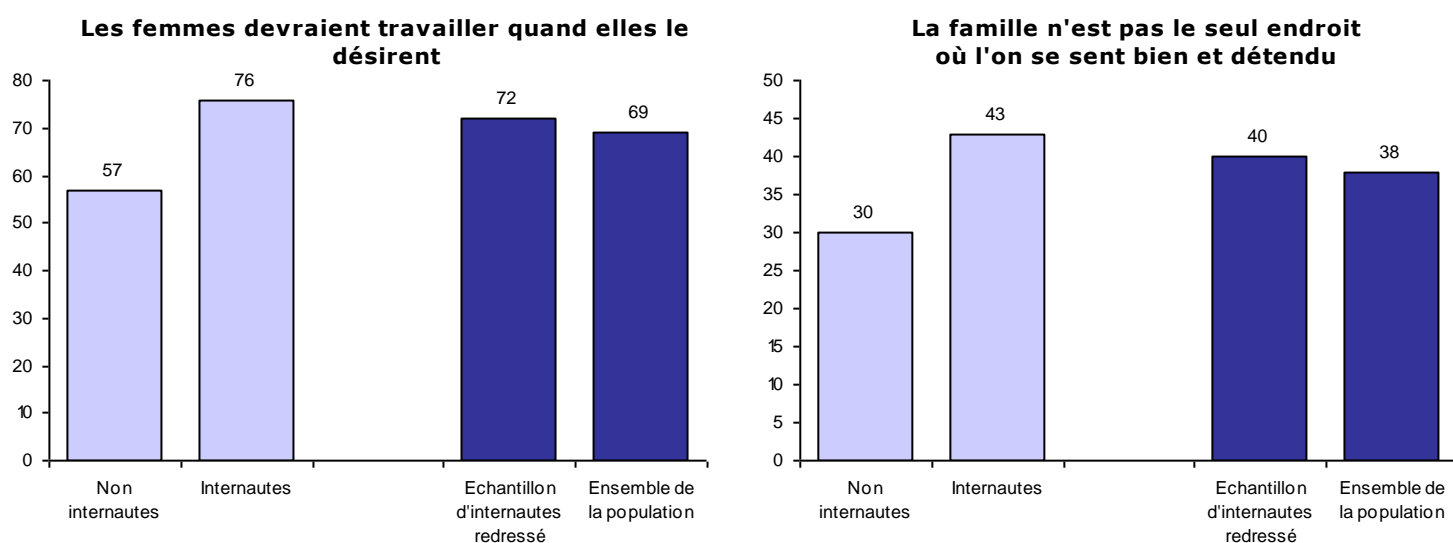
Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.



## 2 QUESTIONS CONCERNENT LES OPINIONS SUR LA FAMILLE ET LES MOEURS

Plus jeunes, les internautes sont plus favorables au travail des femmes et moins centrés sur la famille : ils acceptent plus volontiers l'idée qu'on puisse se sentir bien et détendu ailleurs que dans la cellule familiale (Tableau A 6, p.111). Le redressement supprime ces écarts (Graphique 35). Rappelons que les prises de positions libérales des internautes sur le mariage homosexuel ou l'adoption par des personnes de même sexe ne sont, elles, pas totalement corrigées par le redressement.

**Graphique 35 : le redressement tempère, sur deux indicateurs, le trop grand modernisme des internautes en matière d'opinions sur la famille et les mœurs**

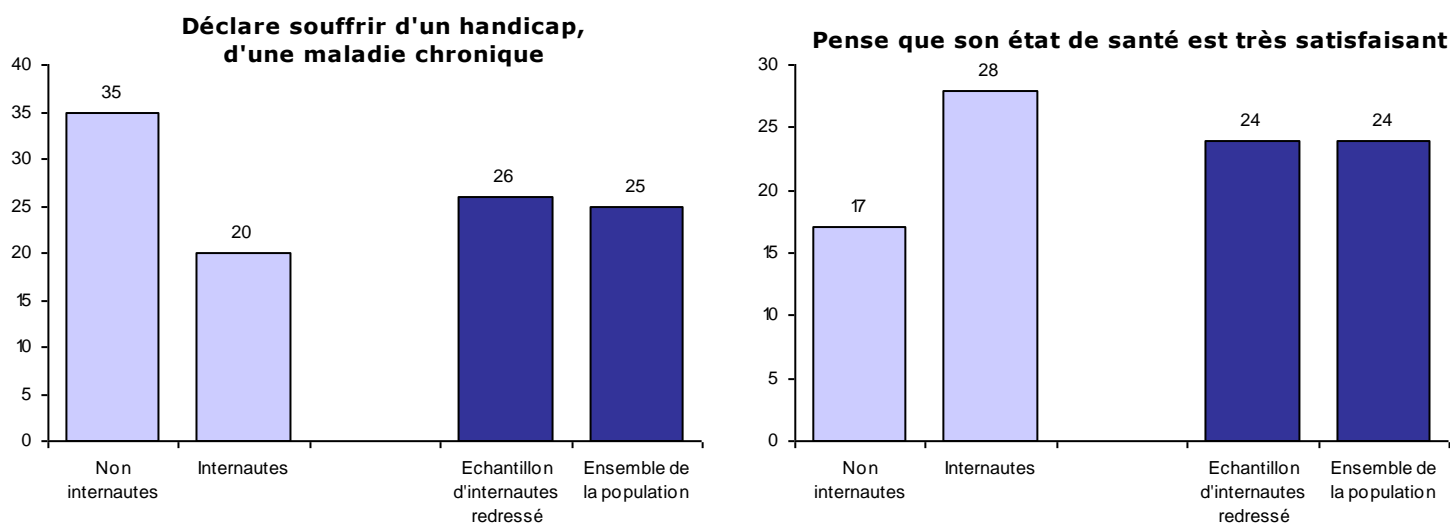


Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

## 2 QUESTIONS CONCERNENT LA SANTE

Les internautes se révèlent beaucoup plus satisfaits de leur état de santé que les non internautes et que la population générale. Ils souffrent moins souvent d'une maladie chronique ou d'un handicap. Le redressement corrige parfaitement ces écarts (Graphique 36 et Tableau A 12, p. 117).

**Graphique 36 : le redressement corrige la déclaration d'un handicap et l'état de santé ressenti**

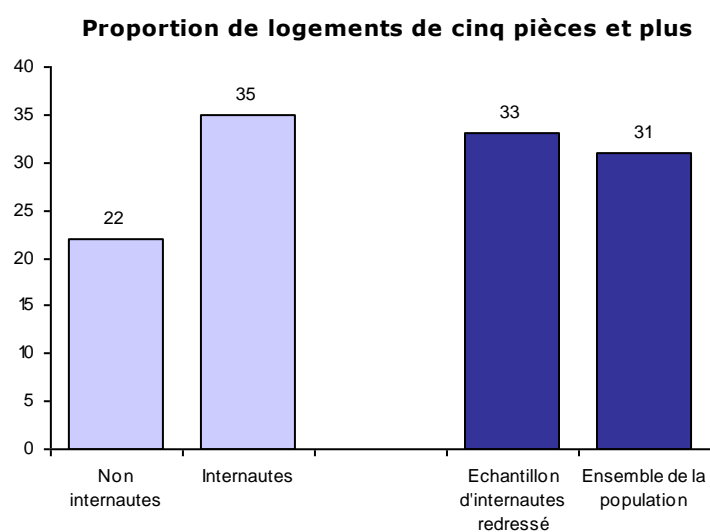


Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

### 1 QUESTION CONCERNE LE LOGEMENT

Les internautes résident dans des logements de plus grande taille : 35% bénéficient de logements de cinq pièces et plus (contre 22% des non internautes). Après redressement, la distribution du nombre de pièces du logement occupé coïncide avec celle mesurée dans l'ensemble de la population (Graphique 37 et Tableau A 7 p.103).

**Graphique 37 : le redressement corrige le nombre de pièces du logement**

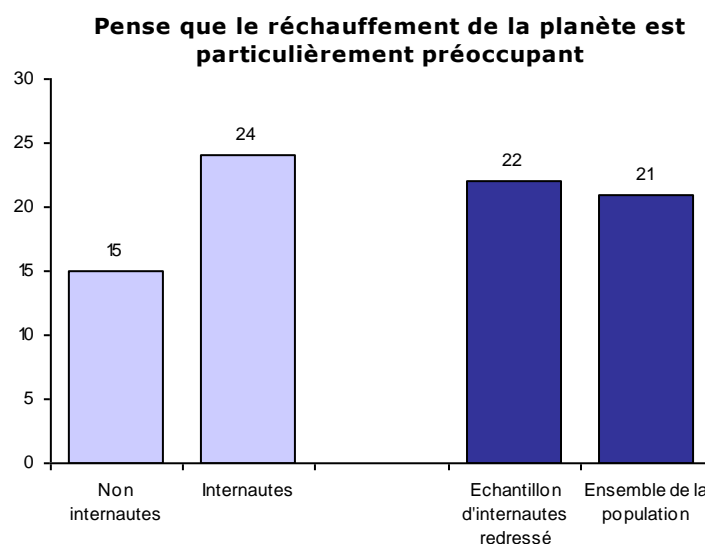


Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

## 1 QUESTION CONCERNE L'ENVIRONNEMENT

Le réchauffement de la planète est une problématique qui inquiète particulièrement les internautes. Après redressement, cette spécificité disparaît (Graphique 38 et, en annexe, Tableau A 13 p. 103).

**Graphique 38 : le redressement corrige la sensibilité à l'égard du réchauffement de la planète**

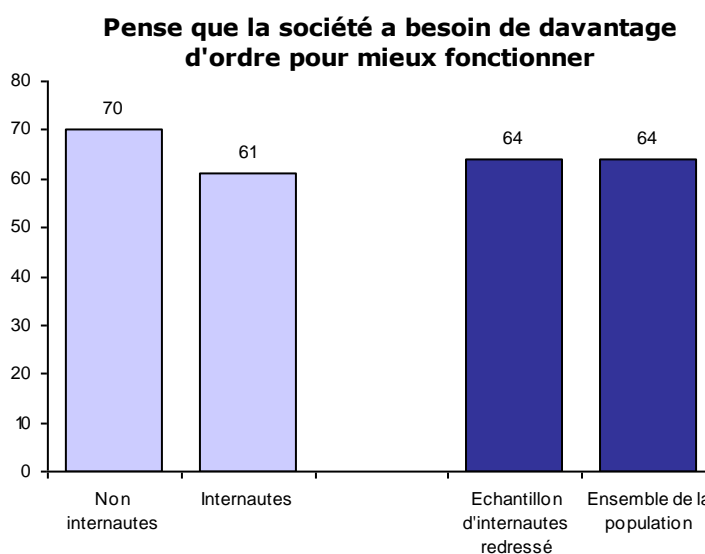


Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

## 1 QUESTION CONCERNE LES OPINIONS SUR LA SOCIÉTÉ

Les internautes pensent moins souvent qu'en moyenne que la société a besoin de davantage d'ordre pour mieux fonctionner ; cette spécificité est totalement corrigée par le redressement (Graphique 39 et Tableau A 14 p.119).

**Graphique 39 : le redressement corrige le besoin ressenti de davantage d'ordre dans la société**

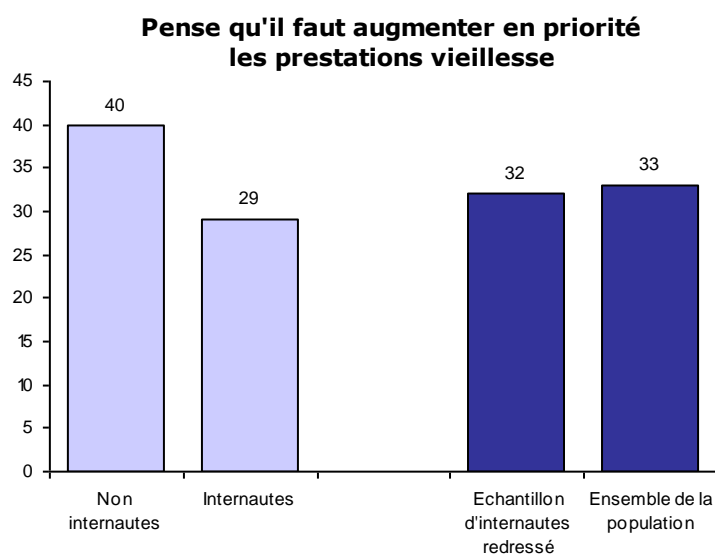


Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

## 1 QUESTION CONCERNE LES JUGEMENTS SUR LES POLITIQUES SOCIALES

Lorsqu'il s'agit d'élire, parmi différentes prestations sociales, celle qu'il convient d'augmenter en priorité, les internautes ont moins d'appétence que la moyenne pour les prestations vieillesse. Après redressement, le taux de citations des prestations vieillesse est corrigé (Graphique 40 et, en annexe, Tableau A 15 p. 120).

**Graphique 40 : le redressement corrige la citation des prestations vieillesse comme aide à augmenter à priorité**



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

### 3. Les variables pour lesquelles le redressement n'est pas nécessaire

Il faut aussi se souvenir que, pour deux variables sur trois (129 sur 196), les réponses des internautes avant redressement ne diffèrent pas sensiblement de celles de l'ensemble de la population. Pour certains thèmes, l'interrogation d'internautes pourrait être envisagée sans trop de craintes car elle ne semble pas, *a priori*, générer trop d'écarts dans les attitudes, les opinions ou les comportements (Graphique 25, p. 80).

S'agissant par exemple des opinions sur **les politiques familiales**, les internautes ne se distinguent jamais de l'ensemble de la population (Tableau A 17 p. 122). Ils ont les mêmes préférences que l'ensemble de la population pour toutes les questions abordant l'accueil de la petite enfance (lorsqu'il s'agit d'élire le mode de garde préférable pour un enfant de moins d'un an ou pour un enfant un peu plus grand, quand ils évoquent la durée la plus adéquate pour le congé parental ou l'aide à apporter aux parents d'un petit enfant). Pour tout ce qui concerne les enfants plus grands (aide pour les familles ayant un adolescent à charge, montants et mode de distribution de l'Allocation de rentrée scolaire) et les jeunes adultes (aides à favoriser ou à réduire pour les 18 ans et plus), les internautes affichent exactement les mêmes positions que la population générale. Le constat est identique pour les objectifs à assigner aux politiques familiales ou les types d'aides à favoriser.

Pour les questions relatives aux **politiques sociales** et au traitement de la pauvreté (Tableau A 15, p. 120), les internautes se conforment également à l'ensemble de l'opinion. En effet, les internautes ont, le plus souvent, le même avis que la population générale sur les causes de la pauvreté, les efforts que doivent faire les pouvoirs publics en la matière, sur les mesures à mettre en œuvre pour la combattre. Ils assignent au RSA les mêmes objectifs que l'ensemble des Français et partagent la même perception, plutôt positive, de ce nouveau dispositif.

D'ailleurs, les opinions portées sur **le fonctionnement de la société** (que ce soit les institutions, le système éducatif, les services publics ou le système d'aide et de protection sociale) ne sont pas influencées par le fait d'avoir accès chez soi à une connexion à Internet. Et le regard porté sur le fonctionnement de la justice en France ou les éventuels besoins de transformation de la société française est le même pour les internautes et pour l'ensemble de la population (Tableau A 14, p. 103).

Il ne semble pas non plus y avoir de spécificité propre aux internautes sur toutes les questions ayant trait aux **inquiétudes ou aux préoccupations** (Tableau A 16, p. 121).

Ils se disent aussi sensibles qu'en moyenne aux risques du chômage, inquiets des risques d'une guerre ou d'un accident de centrale nucléaire, préoccupés par les maladies graves ou la pauvreté en France ou dans le monde.

Le thème de **l'environnement** (Tableau A 13, p.103) est, le plus souvent (à 86%), épargné par des écarts résultant du fait de posséder Internet ou pas. Les internautes ne sont ni plus ni moins sensibles à l'environnement que l'ensemble de la population, ils souhaitent dans les mêmes proportions payer des taxes dédiées à l'environnement, adhèrent aussi souvent que les autres à des associations oeuvrant pour la défense de l'environnement...

Les questions relatives à **la santé** ne semblent pas non plus poser de problème particulier. Certes, la déclaration d'un handicap et le jugement porté sur son état de santé nécessitent, pour qu'ils soient comparables à la population générale, qu'on les mesure sur l'échantillon redressé des internautes (Tableau A 12, p.103). Mais tous les maux psychosomatiques (nervosité, maux de tête, insomnies, mal au dos) sont évoqués dans les mêmes proportions par les internautes et par l'ensemble de la population. Et les jugements portés sur le système de santé (sur le rôle des médecins ou sur le fait qu'on est mieux soigné si on a de l'argent et des relations) sont les mêmes chez les internautes et dans la population générale.

Les descripteurs du **logement**, dans près des trois-quarts des cas, ne montrent pas d'écarts entre les internautes et la population générale (Tableau A 7, p. 112). Le statut d'occupation du logement, le type de chauffage, la présence d'un jardin individuel... sont autant de critères sur lesquels les différences sont minces. Les opinions concernant le logement ne diffèrent pas non plus selon que l'on accède ou pas à Internet chez soi : la façon dont on comprend la crise actuelle du logement, le jugement sur son cadre de vie ou sur les dépenses de logement.. sont les mêmes chez les internautes et l'ensemble des Français.

Enfin, pour nombre d'indicateurs évoquant **la situation professionnelle** (durée de travail pour les actifs occupés, occurrence d'une période de chômage au cours de l'année passée, nature du contrat de travail ou type d'employeur ...), les internautes sont en tout point comparables à l'ensemble de la population (Tableau A 11, p.116). Mais on s'approche là des zones où des différences sensibles existent et où le redressement est nécessaire pour près d'un quart des variables.

## Conclusion

---

Ce travail ne vise pas à apporter une réponse définitive à la question de savoir si les enquêtes par Internet peuvent prétendre à la représentativité de la population générale. Il ne s'agit que d'une investigation permettant d'attirer l'attention sur les risques de mesure liés à un échantillonnage d'internautes.

Dans un grand nombre de cas, les internautes présentent des comportements et des opinions très similaires à ceux observés en population générale. Dans un certain nombre d'autres cas, ils sont clairement différents. Il est parfois possible de corriger les écarts, à travers notamment un calcul de pondération reposant sur le redressement et le calage par rapport à la structure sociodémographique de la population générale. Mais, dans certains cas, ce redressement est insuffisant : notamment en matière de nouvelles technologies, d'équipement des ménages, d'opinions en matière de mœurs, de pratiques culturelles et de loisirs, de moral économique des ménages ou de logement. Dans d'autres domaines, les erreurs de mesure semblent de moindre ampleur : situation professionnelle, perception de son état de santé et opinions sur le système de soin, opinions sur le fonctionnement de la société, jugements sur les politiques sociales et familiales, inquiétudes et sujets de préoccupations.

Ce type d'analyse pourrait être mené à travers d'autres enquêtes, dans lesquelles un certain nombre d'attitudes et de comportements sont enregistrés en même temps que l'on peut caractériser la personne interrogée selon qu'elle dispose ou pas d'un accès à Internet. Il serait encore plus intéressant de mener, parallèlement, des enquêtes strictement identiques dans leur questionnement, selon plusieurs modes de recueil de l'information : en face à face, au téléphone et par Internet.

Dans la littérature, les conclusions sont encore nuancées. Les enquêtes en ligne, comme les autres méthodes de collecte de données présentent des forces et des faiblesses. Une grande partie de la recherche au cours des dernières années a porté sur l'identification de ces forces et de ces faiblesses, pour surmonter les unes et exploiter les autres. Aujourd'hui, les chercheurs concluent qu'il est parfois possible d'utiliser les enquêtes en ligne comme mode de collecte des données pour les besoins de statistiques officielles mais que les erreurs sont encore importantes et qu'il est encore nécessaire d'approfondir la recherche pour établir comment procéder pour recueillir, dans ces conditions, des données de qualité. L'erreur due à l'autosélection est très forte. Elle est réduite par la repondération mais non résolue.

Cela ne doit pas occulter le fait que les modes traditionnels présentent eux aussi des faiblesses. Selon les sujets, les enquêtes en ligne peuvent être parfois préférables aux interviews téléphoniques. Les différences entre les deux modes sont parfois faibles et pas toujours à l'avantage de l'un ou l'autre. Les panels avec recrutement probabilistes, en équipant d'ordinateur et de connexion Internet les non-internautes, semblent être la meilleure solution, aujourd'hui, pour les études en ligne. Melanie Revilla et Willem Saris (2010)<sup>115</sup> ont procédé à la comparaison de deux enquêtes aux modes de passation différents. D'une part l'enquête ESS (European Social Survey) réalisée en face-à-face, d'autre part le Liss Panel avec une collecte par internet. Des questions de l'enquête ESS ont été posées aux panélistes du Liss Panel, ceci pendant la même période. Les structures des échantillons d'ESS et du LISS Panel apparaissent significativement différentes de la population générale. En revanche, ils n'observent pas de différences importantes entre les résultats apportés dans les deux enquêtes.

Le choix du mode d'enquête dépend en fait de la source d'erreur la plus importante à diminuer pour chaque investigation. Il dépend largement du thème étudié et de la méthode d'échantillonnage. Les panels avec recrutement probabiliste, par exemple, demandent un investissement coûteux et nécessitent l'existence, comme aux Pays-Bas, de registres d'adresses. De plus, les tests réalisés jusqu'à présent ont porté sur des enquêtes très spécifiques, et tous ne sont pas liés de la même manière au fait d'être équipé ou pas d'une connexion à Internet : à ce stade, il paraît encore délicat de généraliser les conclusions sans renouveler ce type d'analyse avec d'autres enquêtes. Aujourd'hui, il reste encore beaucoup à apprendre avant de pouvoir tirer des conclusions quant à l'efficacité des enquêtes en ligne pour obtenir des données de grande qualité.

---

<sup>115</sup> Melanie Revilla, Willem Saris, « A comparison of the quality of ESS questions in different data collection modes », Universitat Pompeu Fabra, February 2010



## Bibliographie

---

- ALIBERT David, BIGOT Régis et FOUCAUD David, *La dynamique des inégalités en matière de nouvelles technologies (Méthodes d'approche – Analyse évolutive)*, Cahier de recherche du CRÉDOC n°217, novembre 2005, <http://www.CRÉDOC.fr/pdf/Rech/C217.pdf>
- ARDILLY Pascal, *Les techniques de sondage*, Paris, Editions Technip, 2006, 675 p.
- BANNING Reinder, SCHOUTEN Barry, « A mixed-mode follow-up of panel refusers in the Dutch LFS », Discussion paper (09035), Statistics Netherlands, 2009
- BARTELS Larry M., « Panel Effects in the American National Election Studies », *Political Analysis*, 8, 1999, p. 1–20
- BAYART Caroline, *Le potentiel du web pour les enquêtes de mobilité urbaine*, thèse Université Lumière Lyon II, décembre 2009
- BETHLEHEM Jelke, « Reducing the bias of web survey based estimates », Document de travail 07001, Voorburg, Pays-Bas: Statistics Netherlands, 2007
- BETHLEHEM Jelke, STOOP Ineke, « Online panels. A paradigm theft? », in Trotman et al. (eds), *The Challenges of a Changing World*, Southampton, Association for Survey Computing, 2007, p. 113-131
- BETHLEHEM Jelke, « Peut-on établir des statistiques officielles à partir d'enquêtes en ligne reposant sur le principe de l'autosélection ? », Symposium 2008, Collecte des données : défis, réalisations et nouvelles orientations, La série des symposiums internationaux de Statistique Canada : recueil, 2009
- BETHLEHEM Jelke, « Selection Bias in Web Surveys », *International Statistical Review*, 78, 2, 2010, p. 161–188
- BIEMER Paul P., CHRIST Sharon L., « Weighting Survey Data », *In International Handbook of Survey Methodology*, eds. Edith de Leeuw, Joop, J. Hox, and Don, A. Dillman. New York, 2008, Taylor & Francis Group
- BIGOT Régis, CROUTTE Patricia, *La diffusion des technologies de l'information et de la communication dans la société française*, Collection des rapports, n° 269, CRÉDOC, décembre 2010, [http://www.cgiet.org/documents/2010\\_12\\_10\\_Rapport\\_CRÉDOC\\_2010.pdf](http://www.cgiet.org/documents/2010_12_10_Rapport_CRÉDOC_2010.pdf)
- BRANNEN Julia, « The Effects of Research on Participants: Findings from a Study of Mothers and Employment », *Sociological Review*, 41, 1993, p.328–46
- CALLEGARO Mario, DISOGRA Charles, « Computing response metrics for online panels », *Public Opinion Quarterly*, Volume 72, Issue 5, 2008, p. 1008-1032, <http://poq.oxfordjournals.org/content/72/5.toc>
- CHOQUETTE Keith A., HESSELBROCK Michie N., « Effects of Retesting with the Beck and Zung Depression Scales in Alcoholics », *Alcohol and Alcoholism*, 22, 1987, 277–83
- CHRISTIAN Leah M., DILLMAN Don A., « The Influence of Graphical and Symbolic Language Manipulations to Self-Administered Questions », *Public Opinion Quarterly*, 68, 2004, p. 57–80
- COEN Terrence, LORCH Jacqueline, PIEKARSKI Linda, « The Effects of Survey Frequency on Panelists' Responses », ESOMAR retrieved July 27, 2005, 2007, [www.websm.org](http://www.websm.org)
- COTE Anne-Marie, LAROCHE Danielle, « L'Internet : un nouveau mode de collecte au Recensement », Symposium 2008, Collecte des données : défis, réalisations et nouvelles orientations, Statistique Canada, 2009

- COUPER Mick P., « Web Survey : A Review of Issues and Approaches », *Public Opinion Quarterly*, Volume 64, N°4, décembre 2000, p. 464-494
- COUPER Mick P., TRAUGOTT Michael W., LAMIAS Mark J., « Web Survey Design and Administration », *Public Opinion Quarterly*, 65, 2001, p. 230-53
- COUPER Mick P., KAPTEYN Arie, SCHONLAU Matthias, WINTER Joachim, « Noncoverage and Nonresponse in an Internet Survey », *Social Science Research*, 36, 2007, p. 131-48
- COUPER Mick P., MILLER Peter V., « Web survey methods », Introduction, *Public Opinion Quarterly*, Volume 72, Issue 5, 2008, p. 831-835, <http://poq.oxfordjournals.org/content/72/5.toc>
- CUNNINGHAM John, SELBY Peter, KYPRI Kypros, HUMPHREYS Keith, « Access to the Internet among drinkers, smokers and illicit drug users: Is it a barrier to the provision of interventions on the World Wide Web? », *Informatics for Health and Social Care*, 31, 2006, p. 53-58
- CZAJA Ronald, BLAIR Johnny, *Designing surveys, A guide to decisions and procedures*, 2005, Thousand Oaks : Sage, [http://books.google.fr/books?id=GCWZVZXEF2gC&printsec=frontcover&dq=Ronald+Czaja+et+Johnny+Blair&source=bl&ots=dFqLqMSvIa&sig=e2y3NdDBmQwaNepWuB-47Q9s3B8&hl=fr&ei=7noLTc\\_5JcicOrWziIUJ&sa=X&oi=book\\_result&ct=result&resnum=1&sqi=2&ved=0CBgQ6AEwAA#v=onepage&q&f=false](http://books.google.fr/books?id=GCWZVZXEF2gC&printsec=frontcover&dq=Ronald+Czaja+et+Johnny+Blair&source=bl&ots=dFqLqMSvIa&sig=e2y3NdDBmQwaNepWuB-47Q9s3B8&hl=fr&ei=7noLTc_5JcicOrWziIUJ&sa=X&oi=book_result&ct=result&resnum=1&sqi=2&ved=0CBgQ6AEwAA#v=onepage&q&f=false)
- DE LEEUW Edith D., « Data quality in mail, telephone and face-to-face surveys », Amsterdam TT-Publikaties, 1992, <http://www.xs4all.nl/~edithl>
- DE LEEUW Edith D., DE HEER Wim, « Trends in household survey nonresponse : A longitudinal and international comparison », In Robert M. Groves, Don A. Dillman, John L. Eltinghe, Roderick .J.A. Little (Eds.), *Survey nonresponse*, New York: Wiley, 2002
- DE LEEUW Edith D., « To Mix or Not to Mix Data Collection Modes in Surveys », *Journal of Official Statistics*, Vol. 21, No. 2, 2005, p. 233-255
- DE LEEUW Edith D., « Choosing the Method of Data Collection », p. 113-135, in Edith D. de Leeuw, Joop J. Hox, Don A. Dillman, *International Handbook of Survey Methodology*, Taylor & Francis Group, 2008, 549 p.
- DILLMAN Don. A., *Mail and Internet Surveys. The Tailored Design Method*, 2007, Hoboken, NJ : Wiley
- DILLMAN Don A., SMYTH Jolene D., CHRISTIAN Leah Melani, *Internet, mail, and mixed-mode surveys, The tailored Design Method*, Third Edition, 2009, 499p.
- ELLIOTT Marc N., HAVILAND Amelia, « Utilisation d'un échantillon de convenance électronique comme complément à un échantillon probabiliste », *Techniques d'enquête*, Volume 33, N°2, 2007, p. 233-238, Statistique Canada, n°12-001-X au catalogue
- EYSENBACH Gunther, « Improving the Quality of Web Surveys: The Checklist for Reporting Results from Internet E-Surveys (Cherries) », *Journal of Medical Internet Research* 6(3) : e34, 2004
- FAAS Thorsten, « Online or Not Online?, A Comparison of Offline and Online Surveys Conducted in the Context of the 2002 German Federal Election », *Bulletin de méthodologie sociologique* n°82, 2004, p. 42-57, <http://bms.revues.org/index1086.html#ftn2>
- FAAS Thorsten et SCHOEN Harald, « Putting a questionnaire on the web is not enough : A comparison of online and offline surveys conducted in the context of the German federal election 2002 », *Journal of Official Statistics*, 22(2), 2006, p. 177-190
- FRICKER Scott, GALESIC Mirta, TOURANGEAU Roger, Ting YAN, « An experimental comparison of web and telephone surveys », *Public Opinion Quarterly*, 69, 2005, p.370-392

- FRIPPIAT Didier, MARQUIS Nicolas, « Les enquêtes par Internet en sciences sociales : un état des lieux », *Population*, Volume 65, n°2, 2010, p. 309-338
- GALESIC Mirta, TOURANGEAU Roger, COUPER Mick P., CONRAD Frederick G., « Eye-tracking data, New Insights on Response Order Effects and Other Cognitive Shortcuts in Survey Responding », *Public Opinion Quarterly*, Volume 72, Issue 5, 2008, 2008, p. 914-934, <http://poq.oxfordjournals.org/content/72/5.toc>
- GANASSALI Stéphane, MOSCAROLA Jean, « Protocoles d'enquête et efficacité des sondages par Internet », Journées E-Marketing AFM / AIM – Nantes, septembre 2002
- GOLOB Thomas F., « The Dynamics of Household Travel Time Expenditures and Car Ownership Decisions », *Transportation Research* 24A, 1990, p. 443-63
- GRÄF Lorenz, « Assessing Internet questionnaires: The online pretest lab. », In: Bernad Batinic, Ulf Dietrich Reips, Michael Bosnjak, Andreas Werner, eds, *Online social sciences*, Seattle, WA: Hogrefe & Huber, 2002
- GROVES Robert M., « Nonresponse rates and nonresponse bias in household surveys, *Public Opinion Quarterly*, Vol. 70, No. 5, Special Issue, 2006, p. 646-675
- GROVES Robert M., PEYTCHEVA Emilia, « The Impact of Nonresponse Rates on Nonresponse Bias. A Meta-Analysis » *Public Opinion Quarterly*, 72(2), 2008, p. 167-189
- HEERWEGH Dirk, Geert LOOSVELDT, « Face-to-face versus web surveying in a high-internet-coverage population, Differences in response quality », *Public Opinion Quarterly*, Vol. 72, No. 5 2008, p. 836-846
- HOOGENDOORN Adriaan, DAALMANS Jacco, « Nonresponse in the Recruitment of an Internet Panel Based on Probability Sampling », *Survey Research Methods*, Vol 3, No 2, 2009
- JÄCKLE Annette, Caroline ROBERTS, LYNN Peter, « Assessing the Effect of Data Collection Mode on Measurement », *ISER Working Paper Series*, No. 2008-08, February 2008
- KAPLOWITZ Michael D., HADLOCK Timothy D., and LEVINE Ralph, « A Comparison of Web and Mail Survey Response Rates », *Public Opinion Quarterly*, 68, 2004, p.98-101
- KASPRZYK Daniel, DUNCAN Greg, KALTON Graham, SINGH M.P., *Panel surveys*, John Wiley and Sons, New York, 1989
- KENYON Kristin, COUPER Mick P., TOURANGEAU Roger, « Experiments on Visual Effects in Web Surveys, AAPOR Annual Conference », Montreal, Canada, 2001.
- KNAPP Herschel, KIRK Stuart A., « Using Pencil and Paper, Internet and Touch-Tone Phones for Self-Administered Surveys: Does Methodology Matter? », *Computers in Human Behavior* 19, 2003, p. 117-34
- KREUTER Frauke, PRESSER Stanley, TOURANGEAU Roger, « Social desirability bias in CATI, IVR, and Web surveys, The effects of mode and question sensitivity », *Public Opinion Quarterly*, Vol. 72, No. 5, 2008, p. 847-865
- LEE Sunghee, « Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys », *Journal of Official Statistics* 22, 2006a, p. 329-49
- LEE Sunghee, « An Evaluation of Nonresponse and Coverage Errors in a Prerecruited Probability Web Panel Survey », *Social Science Computer Review*, 24, 2006b, p. 460-75.
- LEE Sunghee, VAILLANT Richard, « Estimation for volunteer panel web surveys using propensity score adjustment and calibration adjustment », *Sociological Methods Research*, 37(3), 2009, p. 319-343

- LE LAN Romuald, « Enquêtes ménages : vers la fin de la baisse des taux de réponse ? », *Courrier des Statistiques* n°128, INSEE, septembre-décembre 2009
- LENSVELDT-MULDERS Gerty J. L. M., LUGTIG Peter J., HUBREGTSE Marianne, « Separating Selection Bias and Non-coverage in Internet Panels using Propensity Matching », *Survey practice*, 2009
- LOOSVELDT Geert, SONCK Nathalie, « An evaluation of the weighting procedures for an online access panel survey », *Survey Research Methods*, 2(2), 2008, p. 93-105
- MALHOTRA Neil, 2008, « Completion time and response order effects in web surveys », *Public Opinion Quarterly*, Volume 72, Issue 5, p. 914-934, <http://poq.oxfordjournals.org/content/72/5.toc>
- MANFREDA Katja Lozar, VEHOVAR Vasja, « Internet Surveys », p. 264-284, in Edith D. de Leeuw, Joop J. Hox, Don A. Dillman, *International Handbook of Survey Methodology*, Taylor & Francis Group, 2008, 549 p.
- MATHIOWETZ Nancy A., LAIR Tamra J., « Getting Better? Changes or Errors in the Measurement of Functional Limitations », *Journal of Economic & Social Measurement*, 20, 1994, p. 237-62
- MATSUO Hisako, MCINTYRE Kevin P., TOMAZIC Terry, KATZ Barry, « The online survey : Its contributions and potential problems », *ASA Proceedings of the Section on Survey Research Methods*, 27, 2004, p. 3998-4000
- MEURS Henk, VANWISSEN Leo, VISSER Jacqueline, « Measurement Biases in Panel Data », *Transportation* 16, 1989, p. 175-94
- NAGELHOUT Gera E, WILLEMSSEN Marc C, THOMPSON Mary E, FONG Geoffrey T, VAN DEN PUTTE Bas, DE VRIES Hein, « Is web interviewing a good alternative to telephone interviewing? Findings from the International Tobacco Control (ITC) Netherlands Survey », *BMC Public Health*, 2010
- NEALON Jack, « The effects of male vs female telephone interviewers », *Proceedings of the survey research methods section of the american statistical association*, Alexandria, VA : American Statistical Association, 1983, p. 139-41
- O'MUIRCHEARTAIGH Colm, « Adapting to the Challenges of Household Survey Data Collection in a Multimode Environment », *Modernisation of Statistics Production (MSP2009)*, 2-4 November 2009
- PIAU Claire, *Quelques expériences sur la formulation des questions d'enquête à partir du matériau « Aspirations et Conditions de vie des Français »*, Cahier de recherche n°206, CRÉDOC, 2004, 65 p.
- PIERZCHALA Mark, « Expériences en matière d'enquêtes multimodales », *Symposium 2008, Collecte des données : défis, réalisations et nouvelles orientations*, La série des symposiums internationaux de Statistique Canada : recueil, 2008
- REVILLA Melanie, SARIS Willem, « A comparison of the quality of ESS questions in different data collection modes », *Universitat Pompeu Fabra*, February 2010
- ROOKEY Bryan D., HANWAY Steve, DILLMAN Don A., « Does a probability-based household panel benefit from assignment to postal response as an alternative to Internet-only? », *Public Opinion Quarterly*, Vol. 72, No. 5, 2008, p. 962-984
- ROSENBAUM Paul R., RUBIN Donald B., « The central role of the propensity score in observational studies for causal effects », *Oxford Journals, Mathematics & Physical Sciences, Biometrika*, Vol 70, Issue 1, 1983, p. 41-55

- SCHAEFER David R., DILLMANN Don A., « Development of a standard e-mail methodology : Results of an experiment », *Public Opinion Quarterly*, n°62, 1998, p.378-397
- SCHONLAU Matthias, VAN SOEST Arthur, KAPTEYN Arie, and COUPER Mick, « Selection Bias in Web Surveys and the Use of Propensity Scores », *Sociological Methods & Research*, 2009, 37, p. 291-318
- SCHONLAU Matthias, FRICKER Ronald D., ELLIOTT Marc N., « Conducting Research Surveys via E-mail and the Web », Rand Corporation, Santa Monica, CA, 2003,  
[http://books.google.fr/books?id=tvVeykDBU4cC&printsec=frontcover&dq=Schonlau+Fric ker+Elliott&source=bl&ots=h7TSHJZs2z&sig=ZC8bHXKekfnXBfD9yh9UJznXCd8&hl=fr&ei=cG4LTc\\_LI8Ow8gPt6amFDg&sa=X&oi=book\\_result&ct=result&resnum=4&ved=0CDsQ6AEwAw#v=onepage&q&f=false](http://books.google.fr/books?id=tvVeykDBU4cC&printsec=frontcover&dq=Schonlau+Fric ker+Elliott&source=bl&ots=h7TSHJZs2z&sig=ZC8bHXKekfnXBfD9yh9UJznXCd8&hl=fr&ei=cG4LTc_LI8Ow8gPt6amFDg&sa=X&oi=book_result&ct=result&resnum=4&ved=0CDsQ6AEwAw#v=onepage&q&f=false)
- SHARPE J. Patrick, GILBERT David G., « Effects of Repeated Administration of the Beck Depression Inventory and Other Measures of Negative Mood States », *Personal Individual Differences*, 24, 1998, p. 457-63
- SIKKEL Dirk, « Panel Surveys », p. 479-499, in Edith D. de Leeuw, Joop J. Hox, Don A. Dillman, *International Handbook of Survey Methodology*, Taylor & Francis Group, 2008, 549 p.
- STODDARD Jacqueline, AUGUSTSON Erik, « Smokers Who Use Internet and Smokers Who Don't, Data from the Health Information and National Trends Survey (HINTS) », *Nicotine & Tobacco Research*, 8(Suppl 1), 2006, p. 77-85.
- STURGIS Patrick, ALLUM Nick, BRUNTON-SMITH Ian, « Attitudes over Time: The Psychology of Panel Conditioning », in P. Lynn (édit.), *Methodology in Longitudinal Surveys*, 2007, p. 1-13, Chichester: Wiley
- TOEPEL Vera, DAS Marcel, VAN SOEST Arthur, « Effects of design in web surveys comparing trained and fresh respondents », *Public Opinion Quarterly*, Vol. 72, No. 5 2008, p. 985-1007
- TOURANGEAU Roger, COUPER Mick P., CONRAD Frederick, « Spacing, Position, and Order. Interpretive Heuristics for Visual Features of Survey Questions », *Public Opinion Quarterly*, 68, 2004, p. 368-93
- TOURANGEAU Roger, COUPER Mick P., CONRAD Frederick, « Color, Labels, and Interpretive Heuristics for Response Scales », *Public Opinion Quarterly*, 71, 2007, p. 91-112.
- VEHOVAR Vasja, BATAGELJ Zenel, MANFREDA Katja Lozar, ZATELEL Metka, « Nonresponse in web surveys », in Robert M. Groves, Don A. Dillman, John L. Eltinghe, Roderick .J.A. Little (Eds), *Survey Nonresponse*, New York: Wiley, 2002.
- VEHOVAR Vasja, MANFREDA Katja Lozar, « Web Surveys: Can Weighting Solve the Problem? », Proceedings of the Section on Survey Research Methods, American Statistical Association, 1999, p. 962-67
- WILLIAMS William H., « The Systematic Bias Effects of Incomplete Responses in Rotation SAMPLES », *Public Opinion Quarterly* 33, 1970, P. 593-602
- WILLIAMS William H., MALLOWS Colin L., « Systematic Biases in Panel Surveys Due to Differential Nonresponse », *Journal of the American Statistical Association*, 65, 1970, p. 1338-49

## Annexes

---

**Tableau A 1**  
**Taux d'accès à Internet à domicile en 1999-2001**

(en %)

		Est équipé d'Internet	N'est pas équipé d'Internet	Total
<b>Sexe</b>	. Homme .....	12	88	100
	. Femme .....	9	91	100
<b>Age</b>	. 18 - 24 ans .....	16	84	100
	. 25 - 39 ans .....	20	80	100
	. 40 - 59 ans .....	19	80	100
	. 60 - 69 ans .....	(5)	95	100
	. 70 ans et plus .....	(1)	99	100
<b>Enfant dans le foyer</b>	. Oui .....	14	86	100
	. Non .....	8	92	100
<b>Nombre de personnes dans le logement</b>	. Un .....	5	95	100
	. Deux .....	9	91	100
	. Trois .....	13	87	100
	. Quatre .....	16	84	100
	. Cinq et plus .....	13	87	100
<b>Diplôme</b>	. Aucun, Cep .....	2	98	100
	. Bepc .....	7	93	100
	. Bac .....	13	87	100
	. Diplôme du supérieur .....	26	74	100
<b>Profession</b>	. Indépendant .....	14	86	100
	. Cadre .....	37	62	100
	. Profession intermédiaire .....	20	80	100
	. Employé .....	10	90	100
	. Ouvrier .....	4	96	100
	. Reste au foyer .....	7	93	100
	. Retraité .....	2	98	100
. Etudiant .....	16	84	100	
<b>Revenus par unité de consommation</b>	. Catégories pauvres .....	7	92	100
	. Catégories modestes .....	4	96	100
	. Classes moyennes inférieures ...	6	93	100
	. Classes moyennes supérieures ..	12	88	100
	. Catégories aisées .....	18	82	100
	. Hauts revenus .....	27	73	100
<b>Lieu de résidence</b>	. Moins de 2 000 habitants .....	8	92	100
	. De 2 000 à 20 000 habitants ....	7	93	100
	. De 20 000 à 100 000 habitants .	7	93	100
	. Plus de 100 000 habitants .....	12	88	100
	. Paris et agglo. parisienne .....	17	83	100
<b>ZEAT</b>	. Region parisienne .....	17	82	100
	. Nord .....	8	91	100
	. Est .....	12	88	100
	. Bassin Parisien .....	6	94	100
	. Ouest .....	11	89	100
	. Sud Ouest .....	8	92	100
	. Centre Est .....	10	90	100
	. Mediterranee .....	8	92	100
<b>Total .....</b>	<b>10</b>	<b>90</b>	<b>100</b>	

Source : CRÉDOC, Enquêtes « Conditions de vie et Aspirations des Français ».

Les chiffres entre parenthèses reposent sur des effectifs faibles, ils sont à considérer avec précaution.

**Lecture** : sur la période 1999-2001, 37% des cadres disposent, à domicile, d'un accès à Internet. C'est le cas de 4% des ouvriers.

**Note** : la typologie de catégories de revenus repose sur la répartition de la population en déciles de niveau de vie (catégories pauvres : 1<sup>er</sup> décile ; catégories modestes : 2<sup>ème</sup> et 3<sup>ème</sup> déciles ; classes moyennes inférieures : 4<sup>ème</sup> à 6<sup>ème</sup> déciles ; classes moyennes supérieures : 7<sup>ème</sup> et 8<sup>ème</sup> déciles ; catégories aisées : 9<sup>ème</sup> décile ; hauts revenus : 10<sup>ème</sup> décile).

**Tableau A 2**  
**Structure de la population sur la période 1999-2001,**  
**selon qu'elle dispose ou pas d'Internet à domicile**

(en %)

		Est équipé d'Internet	N'est pas équipé d'Internet	Ensemble de la population
<b>Sexe</b>	. Homme .....	53	47	47
	. Femme .....	47	53	53
<b>Age</b>	. 18 - 24 ans .....	14	13	13
	. 25 - 39 ans .....	41	29	31
	. 40 - 59 ans .....	40	30	32
	. 60 - 69 ans .....	(4)	14	13
	. 70 ans et plus .....	(1)	14	12
<b>Enfant dans le foyer</b>	. Oui .....	50	34	36
	. Non .....	50	66	64
<b>Nombre de personnes dans le logement</b>	. Un .....	11	24	23
	. Deux .....	25	31	30
	. Trois .....	22	17	18
	. Quatre .....	27	16	17
	. Cinq et plus .....	15	12	12
<b>Diplôme</b>	. Aucun, Cep .....	5	29	27
	. Bepc .....	27	40	39
	. Bac .....	19	15	16
	. Diplôme du supérieur .....	48	16	19
<b>Profession</b>	. Indépendant .....	9	7	7
	. Cadre .....	23	4	6
	. Profession intermédiaire .....	22	10	11
	. Employé .....	16	16	16
	. Ouvrier .....	6	17	16
	. Reste au foyer .....	10	14	14
	. Retraité .....	4	25	22
. Etudiant .....	10	6	6	
<b>Revenus par unité de consommation</b>	. Catégories pauvres .....	6	9	9
	. Catégories modestes .....	7	19	18
	. Classes moyennes inférieures ...	16	27	26
	. Classes moyennes supérieures ..	20	17	18
	. Catégories aisées .....	15	8	9
	. Hauts revenus .....	22	7	9
<b>Lieu de résidence</b>	. Moins de 2 000 habitants .....	20	28	27
	. De 2 000 à 20 000 habitants ....	12	17	17
	. De 20 000 à 100 000 habitants ..	9	14	13
	. Plus de 100 000 habitants .....	33	27	28
	. Paris et aggl. parisienne .....	26	14	15
<b>ZEAT</b>	. Region parisienne .....	29	16	17
	. Nord .....	6	7	7
	. Est .....	10	9	9
	. Bassin Parisien .....	11	19	18
	. Ouest .....	14	14	14
	. Sud Ouest .....	9	11	11
	. Centre Est .....	11	12	12
	. Mediterranee .....	9	12	12
<b>Total .....</b>	<b>100</b>	<b>100</b>	<b>100</b>	

Source : CRÉDOC, Enquêtes « Conditions de vie et Aspirations des Français ».

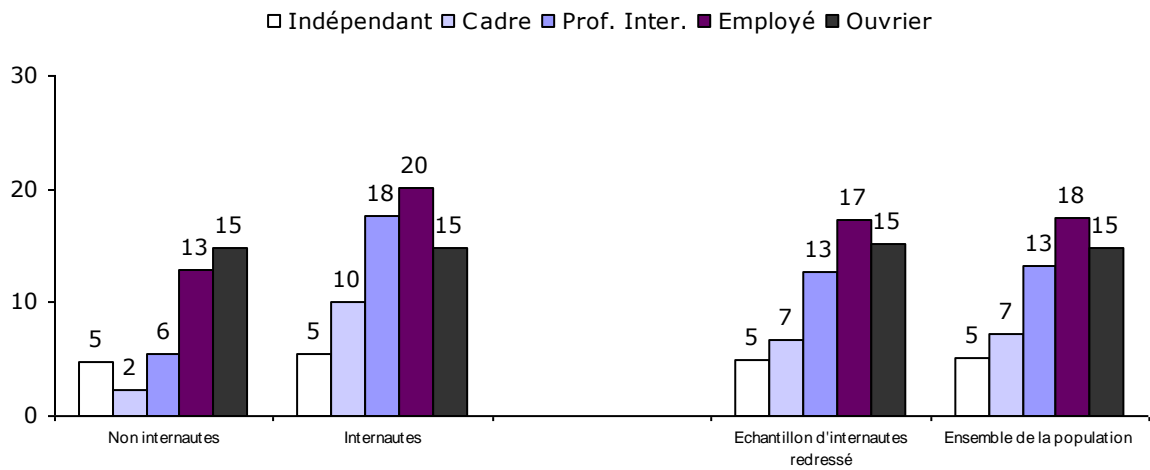
Les chiffres entre parenthèses reposent sur des effectifs faibles, ils sont à considérer avec précaution.

**Lecture** : sur la période 1999-2001, 48% de ceux qui ont, à domicile, un accès à Internet sont diplômés du supérieur, contre 19% de l'ensemble de la population en moyenne et 16% des personnes non équipées.

**Note** : la typologie de catégories de revenus repose sur la répartition de la population en déciles de niveau de vie (catégories pauvres : 1<sup>er</sup> décile ; catégories modestes : 2<sup>ème</sup> et 3<sup>ème</sup> déciles ; classes moyennes inférieures : 4<sup>ème</sup> à 6<sup>ème</sup> déciles ; classes moyennes supérieures : 7<sup>ème</sup> et 8<sup>ème</sup> déciles ; catégories aisées : 9<sup>ème</sup> décile ; hauts revenus : 10<sup>ème</sup> décile).

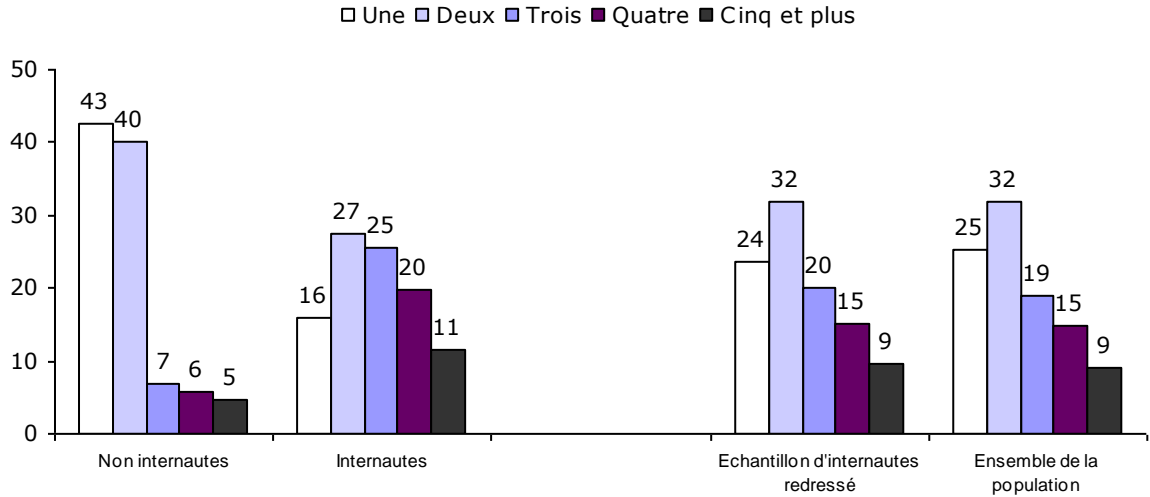


**Graphique A 1 : Effet du redressement sur la PCS, le cas des actifs**



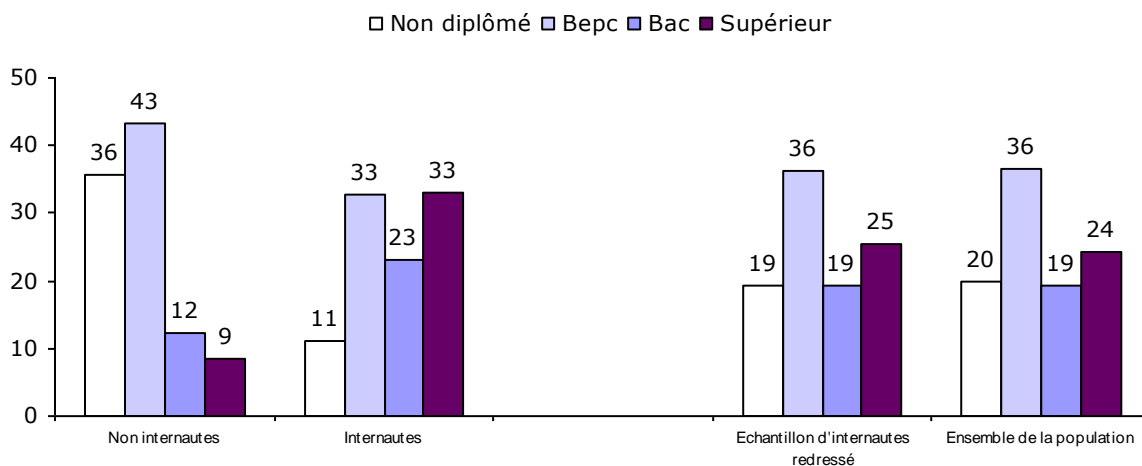
Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Graphique A 2 : Effet du redressement sur la taille du foyer**



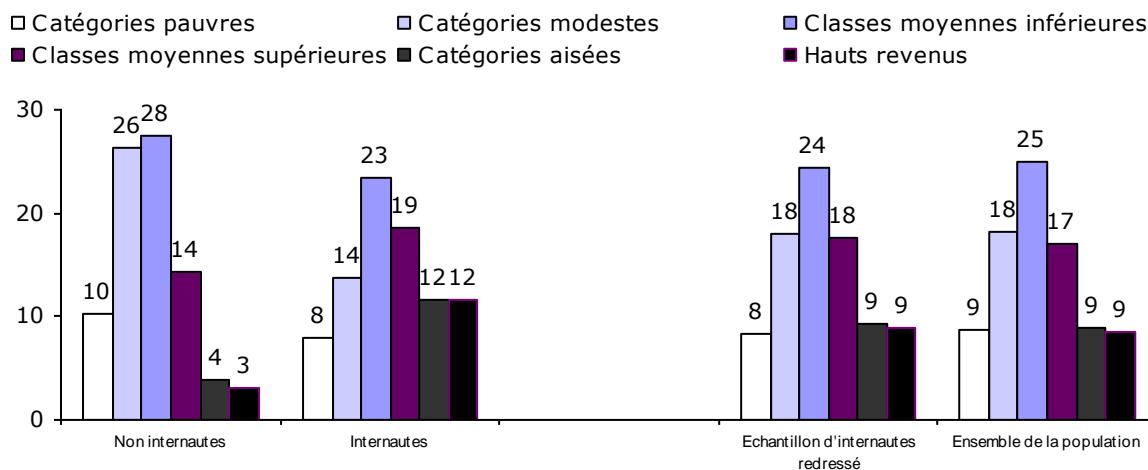
Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Graphique A 3 : Effet du redressement sur le niveau de diplôme**



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Graphique A 4 : Effet du redressement sur les revenus**



Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Note** : la typologie de catégories de revenus repose sur la répartition de la population en déciles de niveau de vie (catégories pauvres : 1<sup>er</sup> décile ; catégories modestes : 2<sup>ème</sup> et 3<sup>ème</sup> déciles ; classes moyennes inférieures : 4<sup>ème</sup> à 6<sup>ème</sup> déciles ; classes moyennes supérieures : 7<sup>ème</sup> et 8<sup>ème</sup> déciles ; catégories aisées : 9<sup>ème</sup> décile ; hauts revenus : 10<sup>ème</sup> décile).

## **GUIDE DE LECTURE POUR LES TABLEAUX SUIVANTS :**

L'analyse a porté sur un plus grand nombre de variables que ce qui est présenté dans les tableaux d'annexe ci-après. En effet, certaines questions sont financées par des souscripteurs au système d'enquête qui ont demandé à ce que leurs questions et leurs résultats restent confidentiels.

C'est le même souci de confidentialité qui nous a conduit, dans les tableaux, à ne faire apparaître que les écarts avec la population totale, et non les pourcentages en valeur absolue.

Dans la colonne de gauche des tableaux, on peut lire si la variable concernée est liée au fait d'avoir accès à Internet à domicile ou pas, toutes choses égales par ailleurs (c'est-à-dire à sexe, âge, diplôme, PCS, niveau de revenus, taille du foyer et taille d'agglomération comparables).

A droite, la couleur des rectangles indique si les écarts de pourcentages entre les échantillons sont significatifs ou pas (rouge : écart significatif ; blanc : écart non significatif).

**Tableau A 3 : l'équipement du ménage**

Influence intrinsèque d'Internet (*** : p<0.001; ** : p< 0.01; * : p < 0,05)		Ecart avec la population totale			Internaute /non internaute	Internaute /population totale	Echantillon d'internautes redressé/ population totale
		Non internaute	Internaute	Echantillon d'internautes redressé			
***	<b>Lave-vaisselle</b>	oui	-17	+ 9	+ 7		
		non	+ 17	-9	-7		
	<b>Téléviseur couleur</b>	oui	+ 1	0	0		
		non	-1	0	0		
***	<b>Magnétoscope</b>	oui	-8	+ 4	+ 4		
		non	+ 8	-4	-4		
***	<b>Congélateur</b>	oui	-6	+ 3	+ 3		
		non	+ 6	-3	-3		
***	<b>Lecteur DVD</b>	oui	-16	+ 9	+ 7		
		non	+ 16	-9	-7		
***	<b>Appareil photo numérique</b>	oui	-33	+ 18	+ 12		
		non	+ 33	-18	-12		
***	<b>Four micro-ondes</b>	oui	-8	+ 4	+ 4		
		non	+ 8	-4	-3		
***	<b>Micro-ordinateur à domicile</b>	oui, un seul	-32	+ 18	+ 23		
		oui, plusieurs	-18	+ 10	+ 4		
		non, aucun	+ 50	-28	-28		
***	<b>Téléphone fixe</b>	oui	-13	+ 7	+ 7		
		non	+ 13	-7	-7		
***	<b>Nb de postes téléphoniques</b>	aucun	+ 13	-7	-7		
		un seul	-2	+ 1	0		
		plusieurs	-12	+ 7	+ 7		
***	<b>Téléphone mobile</b>	oui	-19	+ 10	+ 9		
		non	+ 19	-10	-9		
***	<b>Voiture</b>	oui	-11	+ 6	+ 4		
		non	+ 11	-6	-4		
***	<b>A conduit</b>	oui	-14	+ 8	+ 4		
		non, mais a le permis	+ 7	-4	-2		
		n' a pas le permis	+ 6	-4	-2		
***	<b>Nb voitures</b>	une	+ 6	-3	0		
		plusieurs	-17	+ 10	+ 4		
		aucune	+ 11	-6	-4		
***	<b>Nb kilomètres parcourus</b>	moins de 3000 km	0	0	-1		
		3000 à 9000 km	+ 1	-1	-1		
		9000 à 19000 km	-5	+ 3	+ 3		
		19000 km et plus	-9	+ 5	+ 2		
		n.s.p., nr non concerné	-1	+ 1	0		
		+ 14	-8	-4			

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part des personnes ayant l'usage d'un lave-vaisselle est de 17 points inférieure à la population totale ; cette même part est supérieure de 9 points à la population totale chez les internautes et supérieure de 7 points dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes, entre les internautes et la population totale et entre l'échantillon d'internautes redressé et la population totale. Par ailleurs, l'accès à Internet à domicile est lié, toutes choses égales par ailleurs, à l'usage d'un lave-vaisselle (au seuil de 1%, voir colonne de gauche).

**Tableau A 4 : les attitudes vis-à-vis des nouvelles technologies et de l'informatique**

Influence intrinsèque d'Internet (*** : p<0.001; ** : p< 0.01; * : p < 0,05)	Ecart avec la population totale			Internaute /non internaute	Internaute /population totale	Echantillon d'internautes redressé/ population totale
	Non internaute	Internaute	Echantillon d'internautes redressé			
***	<b>Les découvertes scientifiques améliorent vie quotidienne</b>	un peu beaucoup pas du tout	+ 4 -9 + 5	-2 + 5 -3	-2 + 4 -2	
***	<b>La diffusion de l'informatique ...</b>	souhaitable inévitabile regrettable	-13 + 5 + 7	+ 7 -3 -4	+ 5 -2 -3	
***	<b>Est attiré par innovations technologiques</b>	très attiré assez attiré peu attiré pas du tout attiré	-7 -11 0 + 18	+ 4 + 6 0 -10	+ 3 + 4 + 1 -7	

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part des personnes pas du tout attirées par les produits comportant des innovations technologiques est de 18 points supérieure à la population totale ; cette même part est inférieure de 10 points à la population totale chez les internautes et inférieure de 7 points dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes, entre les internautes et la population totale et entre l'échantillon d'internautes redressé et la population totale. Par ailleurs, l'accès à Internet à domicile est lié, toutes choses égales par ailleurs, à cette attirance (au seuil de 1‰, voir colonne de gauche).

**Tableau A 5 : les pratiques culturelles, les loisirs**

Influence intrinsèque d'Internet (*** : p<0.001; ** : p< 0.01; * : p < 0,05)	Ecart avec la population totale			Internautes /non internautes	Internautes /population totale	Echantillon d'internautes redressé/ population totale
	Non internautes	Internautes	Echantillon d'internautes redressé			
***	<b>Est parti en vacances</b>	oui non	-19 + 19	+ 10 -10	+ 6 -6	
	<b>Rencontre régul. famille proche</b>	oui non	-1 + 1	0 0	0 0	
	<b>Reçoit amis ou relations</b>	tous les jours une fois par semaine une fois par mois plus rarement jamais	+ 1 -1 -8 + 5 + 3	-1 + 1 + 4 -3 -2	0 -2 + 2 + 1 -1	
	<b>Association sportive</b>	oui non	-6 + 6	+ 3 -3	0 0	
***	<b>Association culturelle</b>	oui non	-4 + 4	+ 2 -2	+ 3 -3	
	<b>Association confessionnelle</b>	oui non	-1 + 1	0 0	0 0	
	<b>Association de jeunes d'étudiants</b>	oui non	-3 + 3	+ 1 -1	0 0	
	<b>Association syndicale</b>	oui non	-3 + 3	+ 2 -2	+ 1 -1	
	<b>Association défense environnement</b>	oui non	-1 + 1	0 0	0 0	
	<b>Association parents élèves</b>	oui non	-3 + 3	+ 2 -2	0 0	
	<b>Association de consommateurs</b>	oui non	-1 + 1	0 0	+ 1 -1	
	<b>Membre d'un parti politique</b>	oui non	-1 + 1	+ 1 -1	+ 1 -1	
	<b>Autre association</b>	oui non	-2 + 2	+ 1 -1	0 0	
***	<b>Adhère à associations</b>	adhérent non adhérent	-12 + 12	+ 7 -7	+ 3 -3	
***	<b>Nb adhésions à associations</b>	aucune une deux ou plus	+ 12 -5 -7	-7 + 3 + 4	-3 + 1 + 2	
***	<b>Fréquente équipement sportif</b>	régulièrement exceptionnellement jamais	-8 -10 + 19	+ 5 + 6 -10	0 + 2 -2	
	<b>Fréquente bibliothèque</b>	régulièrement exceptionnellement jamais	-6 -4 + 10	+ 3 + 2 -6	0 0 0	
	<b>Fréquente lieu de culte</b>	régulièrement exceptionnellement jamais	+ 1 + 1 -2	-1 -1 + 1	0 -2 + 2	
***	<b>Fréquente cinéma</b>	régulièrement exceptionnellement jamais	-11 -10 + 22	+ 6 + 6 -12	+ 3 + 3 -6	
	<b>Regarde la télévision</b>	tous les jours assez souvent pas très souvent jamais	+ 7 -4 -2 -1	-4 + 2 + 1 0	-1 + 1 0 0	
***	<b>Abonné Canal +</b>	oui non	-7 + 7	+ 4 -4	+ 3 -3	

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part des personnes parties en vacances au cours des douze derniers mois est de 19 points inférieure à la population totale ; cette même part est supérieure de 10 points à la population totale chez les internautes et supérieure de 6 points dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes, entre les internautes et la population totale et entre l'échantillon d'internautes redressé et la population totale. Par ailleurs, l'accès à Internet à domicile est lié, toutes choses égales par ailleurs, au fait d'être parti en vacances au cours des douze derniers mois (au seuil de 1%, voir colonne de gauche).

**Tableau A 6 : les opinions sur la famille, les mœurs**

Influence intrinsèque d'Internet (*** : p<0.001; ** : p< 0.01; * : p < 0,05)	Ecart avec la population totale			Internautes /non internautes	Internautes /population totale	Echantillon d'internautes redressé/ population totale
	Non internautes	Internautes	Echantillon d'internautes redressé			
↓	<b>Famille seul endroit où on se sent bien</b>	oui	+ 8	-4	-1	
		non	-8	+ 4	+ 1	
	<b>Opinion sur le mariage</b>	indissoluble	+ 7	-4	-3	
		dissout si pb.grave	-1	+ 1	0	
		peut être dissout si accord	-6	+ 3	+ 3	
***	<b>Union homosexuelle</b>	plutôt favorable	-15	+ 8	+ 4	
		plutôt pas favorable	-2	+ 1	-1	
***	<b>Homoparentalité</b>	plutôt favorable	-13	+ 7	+ 5	
		plutôt pas favorable	-5	+ 3	1	
**	<b>Opinion à propos du travail des femmes</b>	jamais si enfant bas âge	+ 6	-3	-1	
		si besoin 2 salaires	+ 5	-3	-1	
		selon desir	-12	+ 7	+ 3	
		toujours travailler	0	0	0	
		jamais travailler	+ 1	0	0	

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part des personnes plutôt favorables à l'homoparentalité est de 13 points inférieure à la population totale ; cette même part est supérieure de 7 points à la population totale chez les internautes et supérieure de 5 points dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes, entre les internautes et la population totale et entre l'échantillon d'internautes redressé et la population totale. Par ailleurs, l'accès à Internet à domicile est lié, toutes choses égales par ailleurs, à cette opinion (au seuil de 1%, voir colonne de gauche).

**Tableau A 7 : le logement**

		Ecart avec la population totale			Internautes /non internautes	Internautes /population totale	Echantillon d'internautes redressé/ population totale
		Non internautes	Internautes	Echantillon d'internautes redressé			
Influence intrinsèque d'Internet (*** : p<0.001; ** : p< 0.01; * : p < 0,05)							
***	<b>Année achèvement du logement</b>	1914 et avant	+ 5	-3	-3		
		1915 - 1948	0	0	-1		
		1949 - 1961	+ 2	-1	-1		
		1962 - 1973	+ 4	-2	+ 1		
		1974 - 1982	0	0	0		
		1983 - 1992	-4	+ 2	+ 2		
		1993 - 2002	-3	+ 2	0		
		après 2002	-3	+ 2	0		
		nsp, nr	0	0	0		
***	<b>Nb pièces du logement</b>	une	+ 1	-1	0		
		deux	+ 4	-2	-2		
		trois	+ 2	-1	+ 1		
		quatre	+ 1	0	-1		
		cinq	-5	+ 3	+ 1		
		six et plus	-4	+ 2	0		
	<b>Est-ce suffisant ?</b>	oui	+ 3	-2	0		
		non	-3	+ 2	0		
	<b>Statut occupation du logement</b>	accédant	-7	+ 4	0		
		propriétaire	+ 6	-3	-1		
		locataire, logé gratuitement	+ 2	-1	0		
	<b>Le principal problème de la crise du logement</b>	manque logements	+ 1	-1	0		
		prix des logements	-1	0	-1		
		diffic. accession	-1	1	1		
		refus de louer	0	0	1		
*	<b>Zone où se trouve le logement</b>	ville, centre ville	0	0	+ 1		
		périphérie, banlieue	-2	+ 1	0		
		village, bourg	-1	0	+ 1		
		habitat dispersé	+ 3	-1	-1		
***	<b>Mode de chauffage principal du logement</b>	électricité	-2	+ 1	+ 1		
		gaz butane, propane	+ 1	0	0		
		gaz naturel (réseau)	-5	+ 3	+ 2		
		fioul, mazout	+ 4	-2	-1		
		charbon	0	0	0		
		bois	+ 3	-1	-2		
		chauffage urbain	+ 1	0	-1		
		nsp, nr	-1	+ 1	+ 1		
		autre	0	0	0		
*	<b>Type de chauffage</b>	individuel	0	0	-2		
		collectif	0	0	+ 2		
	<b>Opinion sur dépenses logement</b>	négligeables, petites	0	0	-1		
		lourdes ou très lourdes	0	0	+ 1		
	<b>Opinion sur cadre de vie</b>	plutôt satisfait	0	0	-2		
		plutôt pas satisfait	0	0	+ 2		
***	<b>Jardin individuel ?</b>	oui	-4	+ 2	-2		
		non	+ 4	-2	+ 2		
	<b>Résidence secondaire</b>	oui	-1	+ 1	+ 1		
		non	+ 1	-1	-1		

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part des personnes résidant dans des logements de cinq pièces et plus est de 9 points inférieure à la population totale ; cette même part est supérieure de 5 points à la population totale chez les internautes et supérieure de 1 point dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes ainsi qu'entre les internautes et la population totale. Par ailleurs, l'accès à Internet à domicile est lié, toutes choses égales par ailleurs, au nombre de pièces du logement (au seuil de 1%, voir colonne de gauche).



**Tableau A 8 : le moral économique**

		Ecart avec la population totale			Internaute/ non internautes	Internaute/ population totale	Echantillon d'internautes redressé/ population totale
		Non internautes	Internaute	Echantillon d'internautes redressé			
↓	<b>S'impose des restrictions budgétaires</b>	oui	0	0	+ 2		
		non	0	0	-2		
**	<b>Restrictions soins médicaux</b>	oui	+ 1	-1	+ 1		
		non	-1	+ 1	-1		
**	<b>Restrictions voiture</b>	oui	-6	+ 3	+ 4		
		non	+ 6	-3	-5		
**	<b>Restrictions équipement ménager</b>	oui	-1	+ 1	+ 2		
		non	+ 2	-1	-2		
**	<b>Restrictions alimentation</b>	oui	+ 1	-1	+ 2		
		non	-1	+ 1	-2		
**	<b>Restrictions vacances, loisirs</b>	oui	-2	+ 1	+ 3		
		non	+ 2	-1	-3		
**	<b>Restrictions habillement</b>	oui	0	0	+ 2		
		non	0	0	-2		
*	<b>Restrictions logement</b>	oui	+ 3	-2	-1		
		non	-3	+ 2	+ 1		
**	<b>Restrictions pour les enfants</b>	oui	-2	+ 1	+ 2		
		non	+ 2	-1	-2		
**	<b>Restrictions tabac, boissons</b>	oui	-3	+ 2	+ 1		
		non	+ 3	-2	-1		
**	<b>Restrictions téléphone</b>	oui	+ 4	-2	-1		
		non	-4	+ 2	+ 1		
**	<b>Que préférez-vous ? (actifs occupés)</b>	+ de pouvoir d'achat	+ 2	-1	+ 2		
		+ de temps libre	-2	+ 1	-2		
**	<b>Opinion niveau de vie perso depuis 10 ans</b>	mieux	-9	+ 5	0		
		pareil	+ 4	-2	-1		
		moins bien	+ 6	-3	+ 1		
**	<b>Opinion niveau de vie Français depuis 10 an</b>	mieux	0	0	0		
		pareil	+ 1	0	-1		
		moins bien	-1	+ 1	+ 1		
**	<b>Opinions conditions de vie dans 5 ans</b>	mieux	-6	+ 3	+ 1		
		pareil	+ 2	-1	-1		
		moins bien	+ 3	-2	0		
**	<b>Pense accroître son épargne</b>	plutôt oui	-8	+ 5	+ 1		
		plutôt non	+ 9	-5	-1		
**	<b>Pense accroître ses dépenses</b>	plutôt oui	-5	+ 3	-1		
		plutôt non	+ 6	-3	+ 1		
*	<b>Moment de faire achats importants ?</b>	oui	-3	+ 2	+ 1		
		non	+ 4	-2	-2		
		ni oui, ni non	-1	+ 1	+ 1		
**	<b>Le nombre de chômeurs va ...</b>	augmenter	+ 2	-1	-1		
		se stabiliser	-2	+ 1	0		
		diminuer	0	0	0		

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part des personnes disant se restreindre sur le poste budgétaire dédié à l'automobile est de 6 points inférieure à la population totale ; cette même part est supérieure de 3 points à la population totale chez les internautes et supérieure de 4 points dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes, entre les internautes et la population totale et entre l'échantillon d'internautes redressé et la population totale. Par ailleurs, l'accès à Internet à domicile est lié, toutes choses égales par ailleurs, à ce type de restriction (au seuil de 1%, voir colonne de gauche).

**Tableau A 9 : la situation familiale et personnelle**

		Ecart avec la population totale			Internaute /non internaute	Internaute /population totale	Echantillon d'internautes redressé/ population totale
		Non internaute	Internaute	Echantillon d'internautes redressé			
Influence intrinsèque d'Internet (*** : p<0.001; ** : p< 0.01; * : p < 0,05)							
Sexe	masculin	-3	+ 2	+ 1			
	féminin	+ 3	-2	-1			
*** Nombre de personnes dans le logement	une	+ 17	-9	-2			
	deux	+ 8	-4	0			
	trois	-12	+ 7	+ 1			
	quatre	-9	+ 5	0			
	cinq et plus	-4	+ 2	0			
Région	Région parisienne	-7	+ 4	+ 1			
	Nord	0	0	-1			
	Est	-1	+ 1	+ 2			
	Bassin Parisien	+ 6	-3	-3			
	Ouest	+ 1	0	0			
	Sud Ouest	+ 1	-1	0			
	Centre Est	+ 1	0	+ 1			
	Méditerranée	0	0	+ 1			
*** Statut matrimonial	célibataire	-1	+ 1	-1			
	marié(e)	-7	+ 4	+ 3			
	concubinage(yc PACS)	-6	+ 3	0			
	séparé(e) divorcé(e)	+ 4	-2	0			
	veuf(ve)	+ 10	-6	-2			
** Conjoint actif ?	oui, plein temps	-24	+ 9	0			
	oui, temps partiel	-3	+ 1	0			
	ne travaille pas	+ 27	-10	0			
*** Type de ménage	vit seul	+ 13	-7	-3			
	marié et cj. actif	-20	11	+ 2			
	marié et cj. inact.	+ 7	-4	+ 1			
*** Enqueté a eu des enfants	oui	+ 3	-2	+ 1			
	non	-3	2	-1			
*** Nombre enfants que l'enqueté a eu	aucun	-3	+ 2	-1			
	un	0	0	-2			
	deux	-2	+ 1	+ 1			
	trois	+ 2	-1	+ 1			
	quatre	+ 1	-1	0			
	cinq et plus	+ 2	-1	+ 1			
** Enfants de - 3 ans dans le foyer	oui	-5	+ 3	0			
	non	+ 5	-3	0			
*** Enfants de - 6 ans dans le foyer	oui	-8	+ 5	0			
	non	+ 8	-5	0			
*** Enfants de - 10 ans dans le foyer	oui	-10	+ 6	0			
	non	+ 10	-6	0			
*** Enfants de - 16 ans dans le foyer	oui	-15	+ 8	+ 1			
	non	+ 15	-8	-1			
Nombre idéal d'enfants	un	0	0	0			
	deux	-1	+ 1	0			
	trois	+ 1	-1	+ 1			
	quatre et plus	0	0	0			
	aucun	0	0	0			
	n.s.p., nr	0	0	0			
Age x sexe	moins de 25 ans, homme	-2	+ 1	-1			
	25 à 39 ans, homme	-5	+ 3	0			
	40 à 59 ans, homme	-3	+ 2	-1			
	60 ans et plus, homme	+ 7	-4	+ 2			
	moins de 25 ans, femme	-2	+ 1	+ 1			
	25 à 39 ans, femme	-7	+ 4	0			
	40 à 59 ans, femme	-2	+ 1	+ 2			
	60 ans et plus, femme	+ 14	-8	-4			
Sexe x activité	homme actif	-9	+ 5	-1			
	homme inactif	+ 6	-3	+ 2			
	femme active	-9	+ 5	0			
	femme inactif	+ 12	-7	-1			
*** Sexe x statut matrimonial	célibataire, homme	0	0	-2			
	marié, homme	-7	4	+ 2			
	divorcé, homme	+ 2	-1	0			
	veuf, homme	+ 2	-1	0			
	célibataire, femme	-1	1	1			
	mariée, femme	-6	3	0			
	divorcée, femme	+ 1	-1	0			
	veuve, femme	+ 9	-5	-2			
*** Région x âge	- 40 ans, province	-13	+ 7	+ 1			
	- 40 ans, Paris et agglo	-4	+ 2	-1			
	40 ans et + province	+ 20	-11	-2			
	40 ans et + Paris et agglo	-4	+ 2	+ 2			

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part des hommes est de 3 points inférieure à la population totale ; cette même part est supérieure de 2 points à la population totale chez les internautes et supérieure de 1 point dans l'échantillon d'internautes redressé. Il n'y a qu'entre internautes et non-internautes que la différence est significative. Par ailleurs, l'accès à Internet à domicile n'est pas lié, toutes choses égales par ailleurs, au genre (colonne de gauche).

**Tableau A 10 : le niveau de vie et le patrimoine**

Influence intrinsèque d'Internet (*** : p<0.001; ** : p< 0.01; * : p < 0,05)	Ecart avec la population totale			Internaute /non internaute	Internaute /population totale	Echantillon d'internautes redressé/ population totale
	Non internaute	Internaute	Echantillon d'internautes redressé			
**	<b>Allocataire de la CAF</b>	oui	-7	+ 4	+ 1	
		non	+ 8	-4	-1	
**	<b>Prestation logement (allocataires)</b>	oui	+ 22	-9	+ 1	
		non	-22	+ 9	-2	
***	<b>Minimum social (allocataires)</b>	oui	+ 8	-3	-2	
		non	-8	+ 3	+ 2	
***	<b>Possède valeurs mobilières</b>	oui	-5	+ 3	+ 2	
		non	+ 5	-3	-2	
**	<b>Possède biens immobiliers</b>	oui	-2	+ 1	+ 1	
		non	+ 2	-1	-1	
	<b>Possède biens fonciers</b>	oui	+ 1	-1	-1	
		non	-1	+ 1	+ 1	
**	<b>Possède produit épargne liquide</b>	oui	-6	+ 3	0	
		non	+ 6	-3	0	
***	<b>Possède produit assurance-vie</b>	oui	-8	+ 4	+ 3	
		non	+ 8	-4	-2	

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part des allocataires d'une prestation versée par la Caisse d'Allocations Familiales est de 7 points inférieure à la population totale ; cette même part est supérieure de 4 points à la population totale chez les internautes et supérieure de 1 point dans l'échantillon d'internautes redressé. La différence est significative aussi bien entre internautes et non-internautes qu'entre les internautes et la population totale. Par ailleurs, l'accès à Internet à domicile est lié, toutes choses égales par ailleurs, au fait de recevoir une prestation versée par la CAF (au seuil de 1%, voir colonne de gauche).

**Tableau A 11 : l'emploi, le chômage, la situation professionnelle**

Influence intrinsèque d'Internet  
(\*\*\* : p<0.001; \*\* : p< 0.01; \* : p < 0,05)

		Ecart avec la population totale			Internaute /non /internaute	Internaute /population totale	Echantillon d'internautes redressé/ population totale
		Non internaute	Internaute	Echantillon d'internautes redressé			
<b>Situation emploi</b>	actifs occupés	-18	+ 10	-1			
	inactifs	+ 19	-10	+ 1			
	chômeurs	0	0	0			
<b>Situation emploi détaillée</b>	actif	-18	+ 10	-1			
	étudiant (travaille)	0	0	0			
	chômeur (a travaillé)	0	0	0			
	au foyer	+ 3	-2	0			
	aide familial	0	0	0			
	retraité	+ 18	-10	+ 2			
	invalidé, malade	+ 2	-1	0			
	étudiant, élève	-4	+ 2	0			
recherche 1er emploi	-1	0	0				
<b>Situation (actifs occupés)</b>	salarié	-3	+ 1	0			
	à votre compte	+ 3	-1	0			
<b>Employeur (salariés)</b>	état, collectivités locales	-10	+ 3	-1			
	fonction publique	+ 1	0	0			
	secteur nationalisé	-1	0	+ 1			
	secteur privé	+ 10	-3	+ 1			
<b>Interim (salariés du privé)</b>	oui	+ 5	-2	-1			
	non	-5	+ 2	+ 1			
<b>Type contrat (salariés)</b>	à durée déterminée	+ 7	-2	-1			
	à durée indéterminée	-7	+ 2	+ 1			
<b>Durée travail (salariés sauf interim)</b>	temps plein	-5	+ 1	+ 1			
	temps partiel	+ 5	-1	-1			
<b>Durée travail (actifs occupés)</b>	moins de 35 heures	+ 4	-1	-1			
	35 heures	-3	+ 1	+ 1			
	36 à 39 heures	+ 1	0	0			
	40 à 45 h	-1	0	0			
	plus de 45 heures	-3	+ 1	+ 1			
ne sait pas	+ 1	0	0				
<b>Chômage an dernier (actifs occupés)</b>	oui	+ 4	-1	+ 1			
	non	-4	+ 1	-1			
<b>A déjà connu chômage</b>	oui	-7	+ 4	+ 4			
	non	+ 7	-4	-4			
	déjà au chômage	0	0	0			
<b>Menacé par chômage</b>	oui	-2	+ 1	0			
	non	-16	+ 9	-1			
	déjà au chômage	0	0	0			
	non concerné	+ 19	-10	+ 1			
<b>Cherche un emploi</b>	oui	-2	+ 1	0			
	non	+ 2	-1	0			

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part des actifs occupés est de 18 points inférieure à la population totale ; cette même part est supérieure de 10 points à la population totale chez les internautes et inférieure de 1 point dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes, ainsi qu'entre les internautes et la population totale. Par ailleurs, l'accès à Internet à domicile n'est pas lié, toutes choses égales par ailleurs, à la situation d'emploi (colonne de gauche).

**Tableau A 12 : la santé**

Influence intrinsèque d'Internet  
 (\*\*\*) : p<0.001; \*\* : p< 0.01; \* : p < 0,05)

		Ecart avec la population totale			Internautes	Internautes	Echantillon	Internautes	Internautes	Echantillon
		Non	Internautes	d'internautes	/non	/population	redressé/	/non	/population	redressé/
		internautes		redressé	internautes	totale	population	internautes	totale	population
							totale			totale
<b>Souffre de handicap, maladie chronique</b>	oui	+ 10	-5	+ 1						
	non	-9	+ 5	-1						
<b>A souffert de maux de tête</b>	oui	-3	+ 1	0						
	non	+ 3	-1	0						
<b>A souffert de mal au dos</b>	oui	+ 4	-2	0						
	non	-4	+ 2	0						
<b>A souffert de nervosité</b>	oui	0	0	0						
	non	0	0	0						
<b>A souffert d'un état dépressif</b>	oui	+ 1	-1	0						
	non	-1	0	-1						
<b>A souffert de insomnies</b>	oui	+ 5	-3	0						
	non	-5	+ 3	0						
<b>Etat de santé</b>	très satisfaisant	-7	+ 4	+ 1						
	satisfaisant	+ 2	-1	0						
	peu satisfaisant	+ 4	-2	0						
	pas du tout satisfaisant	+ 1	-1	0						
<b>Mieux soigné si on a argent et relations</b>	tout à fait d'accord	-3	+ 2	+ 2						
	assez d'accord	+ 1	-1	0						
	pas du tout d'accord	+ 2	-1	-2						
<b>La santé, c'est l'affaire des médecins</b>	tout à fait d'accord	+ 2	-1	0						
	assez d'accord	+ 3	-2	-1						
	pas très d'accord	-3	+ 1	0						
	pas du tout d'accord	-3	+ 2	0						
<b>Préoccupé par maladies graves</b>	oui	+ 4	-2	-1						
	non	-4	+ 2	+ 1						

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part de ceux qui souffrent d'un handicap ou d'une maladie chronique est de 10 points supérieure à la population totale ; cette même part est inférieure de 5 points à la population totale chez les internautes et supérieure de 1 point dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes, ainsi qu'entre internautes et population totale. Par ailleurs, l'accès à Internet à domicile n'est pas lié, toutes choses égales par ailleurs, à la déclaration d'un handicap ou d'une maladie chronique (colonne de gauche).

**Tableau A 13 : l'environnement**

Influence intrinsèque d'Internet  
 (\*\*\*) : p<0.001; \*\* : p< 0.01; \* : p < 0,05)

		Ecart avec la population totale			Internautes /non internautes	Internautes /population totale	Echantillon d'internautes redressé/ population totale
		Non internautes	Internautes	Echantillon d'internautes redressé			
↓	<b>Note de sensibilité à l'environnement de 1 à 7</b>	1	0	0			
		2	+ 1	0			
		3	0	0			
		4	+ 2	-1	-2		
		5	-2	+ 1	+ 1		
		6	0	0	-2		
		7	-2	+ 1	+ 2		
	<b>Action prioritaire à mener par l'Etat pour la protection de l'environnement</b>	protéger paysages	+ 1	-1	0		
		sauver flore, faune	+ 3	-2	-1		
		lutter contre le bruit	+ 3	-1	-1		
		réduire pollution air	+ 2	-1	-1		
		prévenir risques naturels	+ 1	0	0		
		lutter réchauf. Planète	-6	+ 4	+ 2		
		lutter pollution eau	0	0	-1		
		lutter risques nucléaires	-1	0	0		
		trier les déchets	0	0	0		
		utiliser techno propres	-3	+ 2	+ 1		
	<b>Action prioritaire à mener par vous pour la protection de l'environnement</b>	protéger paysages	0	0	0		
		sauver flore, faune	+ 3	-2	-1		
		lutter contre le bruit	+ 1	0	0		
		réduire pollution air	0	0	0		
		prévenir risques naturels	1	-1	-1		
		lutter réchauf. planète	-1	+ 1	+ 1		
		lutter pollution eau	+ 1	0	0		
		lutter risques nucléaires	0	0	0		
		trier les déchets	-3	+ 2	+ 1		
		utiliser techno propres	-1	0	0		
*	<b>Problème le plus préoccupant</b>	dégradation paysages	+ 4	-2	-2		
		disparition espèces	0	0	0		
		gêne liée au bruit	+ 2	-1	-1		
		pollution de l'air	+ 2	-1	0		
		catastrophes naturelles	-1	0	0		
		réchauffement planète	-6	+ 3	+ 1		
		pollution de l'eau	+ 1	-1	0		
		risques ind. nucléaire	0	0	0		
		déchets ménagers	0	0	0		
		technolog. pas propres	-2	+ 1	+ 1		
	<b>Prêt à payer taxes environnement</b>	oui	-3	+ 2	0		
		non	+ 3	-1	+ 1		
	<b>Asso. de défense environnement</b>	oui	-1	0	0		
		non	+ 1	0	0		
	<b>Se dit préoccupé par environnement</b>	oui	-4	+ 2	0		
		non	+ 4	-2	0		

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part de ceux qui se disent prêts à payer une taxe dédiée à l'environnement est de 4 points inférieure à la population totale ; cette même part est supérieure de 2 points à la population totale chez les internautes et égale à la population totale dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes. Par ailleurs, l'accès à Internet à domicile n'est pas lié, toutes choses égales par ailleurs, au fait de se déclarer prêt à payer une telle taxe (colonne de gauche).

**Tableau A 14 : les opinions sur la société**

Influence intrinsèque d'Internet  
(\*\*\* : p<0.001; \*\* : p< 0.01; \* : p < 0,05)

		Ecart avec la population totale			Internautes	Internautes	Echantillon	Internautes	Internautes	Echantillon
		Non	Internautes	d'internautes	/non	/population	d'internautes	/population	redressé/	Echantillon
		internautes		redressé	internautes	totale	redressé/	totale	population	totale
<b>Veut changer les institutions</b>	oui	-4	+ 2	+ 1						
	non	+ 3	-2	-1						
<b>Veut changer le système éducatif</b>	oui	-2	+ 1	+ 1						
	non	+ 2	-1	-1						
<b>Veut changer les services publics</b>	oui	-2	+ 1	+ 2						
	non	+ 2	-1	-2						
<b>Veut changer aide et protection sociale</b>	oui	0	0	0						
	non	0	0	0						
<b>La justice fonctionne</b>	très bien	+ 1	-1	-1						
	assez bien	0	0	0						
	assez mal	-4	+ 2	+ 1						
	très mal	+ 2	-1	0						
<b>La société a besoin de se transformer</b>	oui	+ 1	0	0						
	non	0	0	0						
<b>Par quel type de réformes</b>	progressives	-1	0	-2						
	radicales	0	0	+ 2						
<b>Pour mieux fonctionner la société a besoin de ...</b>	beaucoup plus d'ordre	+ 5	-3	0						
	un peu plus d ordre	0	0	0						
	un peu plus de liberté	-5	+ 3	0						
	beaucoup plus de liberté	0	0	0						

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part de ceux qui souhaitent changer le fonctionnement des institutions est de 4 points inférieure à la population totale ; cette même part est supérieure de 2 points à la population totale chez les internautes et supérieure de 1 point à la population totale dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes. Par ailleurs, l'accès à Internet à domicile n'est pas lié, toutes choses égales par ailleurs, à la volonté de changer le fonctionnement des institutions (colonne de gauche).

**Tableau A 15 : les opinions sur les politiques sociales**

Influence intrinsèque d'Internet  
(\* \*\* : p < 0.001; \* \* : p < 0.01; \* : p < 0,05)

	Ecart avec la population totale			Internaute/ non internaute	Internaute/ population totale	Echantillon d'internautes redressé/ population totale
	Non internaute	Internaute	Echantillon d'internautes redressé			
<b>Si chômage prolongé, s'adresse à ...</b>	parents, amis	-5	+3	-1		
	syndicat	0	0	0		
	CAF	+1	0	-1		
	aide sociale, assist sociale	+4	-2	0		
	organisme d'entraide	0	0	+1		
autre	-1	0	+1			
<b>Si pauvreté, c'est à cause de ...</b>	pas de chance	-2	+1	0		
	pas assez d'effort	+3	-2	0		
	nsp, nr	-1	0	0		
<b>Mesure à favoriser pour retour à l'emploi des RMIstes</b>	aide aux entreprises	+1	-1	-1		
	baisse des minima sociaux	+1	-1	-1		
	aide + salaire limité ds temps	-1	+1	+2		
	aide + salaire permanent	-1	0	0		
	prime retour emploi	0	0	0		
<b>Prendre en charge les plus démunis ...</b>	ça leur permet de vivre	-2	+1	-1		
	ça enlève sens responsa.	+1	0	+1		
	nsp, nr	+1	0	0		
<b>*** Minima sociaux concurrent salaires</b>	plutôt d'accord	-4	+2	+2		
	plutôt pas d'accord	+5	-3	-3		
<b>Bcp chômeurs pourraient retrouver emp.</b>	plutôt d'accord	+5	-3	0		
	plutôt pas d'accord	-5	+3	+1		
<b>Pour les plus démunis les pouvoirs publics ...</b>	font trop	0	0	0		
	font ce qu'il faut	+4	-2	-2		
	ne font pas assez	-4	+2	+2		
<b>Toucher le RMI ...</b>	n'incite pas à travailler	+4	-2	0		
	aide à s'en sortir	-5	3	0		
<b>Prestation à augmenter</b>	allocations familiales	-2	+1	-1		
	RMI	0	0	+1		
	prestations vieillesse	+7	-4	-1		
	prestations chômage	0	0	0		
	remboursements maladie	-1	+1	+1		
	aides au logement	-3	+2	+1		
	pensions invalidité	0	0	0		
nsp, nr	+1	0	-1			
<b>Prestation à diminuer</b>	allocations familiales	0	0	-1		
	RMI	+2	-1	0		
	prestations vieillesse	0	0	-1		
	prestations chômage	-1	0	+1		
	remboursements maladie	-3	+1	0		
	aides au logement	0	0	0		
	pensions invalidité	0	0	0		
nsp, nr	2	-1	0			
<b>RMIste peut retrouver emploi ?</b>	jamais	+2	-1	-1		
	oui, si aide	-3	+2	+1		
<b>Connait bénéficiaire RMI ?</b>	oui	-3	+2	+3		
	non	3	-2	-3		
<b>* Minimas sociaux doivent permettre de</b>	disposer de revenus	+2	-1	-3		
	travailler	-3	+2	+3		
	nsp, nr	+1	-1	0		

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part de ceux qui pensent qu'il faut augmenter en priorité les prestations vieillesse est de 7 points supérieure à la population totale ; cette même part est inférieure de 4 points à la population totale chez les internautes et inférieure de 1 point à la population totale dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes ainsi qu'entre les internautes et la population totale. Par ailleurs, l'accès à Internet à domicile n'est pas lié, toutes choses égales par ailleurs, au choix des prestations vieillesse comme prestation sociale à augmenter (colonne de gauche).



**Tableau A 16 : les inquiétudes, les préoccupations**

Influence intrinsèque d'Internet  
(\*\*\* : p<0.001; \*\* : p< 0.01; \* : p < 0,05)

		Ecart avec la population totale			Internaute /non /population internaute	Internaute /population totale	Echantillon d'internautes redressé/ population totale
		Non internaute	Internaute	Echantillon d'internautes redressé			
<b>Préoccupé par le chômage</b>	oui	+ 1	0	+ 1			
	non	-1	0	-1			
<b>Préoccupé par les maladies graves</b>	oui	+ 4	-2	-1			
	non	-4	+ 2	1			
<b>Préoccupé par l'environnement</b>	oui	-4	+ 2	0			
	non	+ 4	-2	0			
<b>Préoccupé par la pauvreté France</b>	oui	0	0	0			
	non	0	0	0			
<b>Préoccupé par la pauvreté monde</b>	oui	-4	+ 2	+ 1			
	non	+ 4	-2	-1			
<b>Préoccupé par la drogue</b>	oui	+ 2	-1	0			
	non	-2	+ 1	0			
<b>Préoccupé par la violence, insécurité</b>	oui	+ 3	-2	0			
	non	-3	+ 2	0			
<b>Préoccupé par les tensions internat.</b>	oui	-2	+ 1	0			
	non	+ 2	-1	0			
<b>Préoccupé par les conflits sociaux</b>	oui	-2	+ 1	+ 1			
	non	+ 2	-1	-1			
<b>Préoccupé par l'immigration</b>	oui	+ 2	-1	0			
	non	-2	+ 1	0			
<b>Préoccupé par l'Europe</b>	oui	-1	0	0			
	non	+ 1	0	0			
<b>Inquiet des maladies graves</b>	inquiet	+ 1	-1	0			
	pas inquiet	-1	+ 1	0			
<b>Inquiet du risque d'agression dans rue</b>	inquiet	+ 4	-2	+ 1			
	pas inquiet	-4	+ 2	-1			
<b>** Inquiet d'un accident de la route</b>	inquiet	-1	+ 1	+ 1			
	pas inquiet	+ 1	-1	-2			
<b>Inquiet du risque de chômage</b>	inquiet	+ 1	-1	+ 2			
	pas inquiet	-1	+ 1	-2			
<b>Inquiet du risque de guerre</b>	inquiet	+ 5	-3	-1			
	pas inquiet	-5	+ 3	+ 1			
<b>Inquiet accident centrale nucléaire</b>	inquiet	+ 7	-4	-2			
	pas inquiet	-8	+ 4	+ 2			
<b>Inquiet risques conso. prod. alimentaires</b>	inquiet	+ 1	-1	-1			
	pas inquiet	-1	+ 1	+ 1			
<b>Sentiment de sécurité dans vie quotidienne</b>	tout à fait en sécurité	-3	+ 1	0			
	assez en sécurité	+ 1	-1	-1			
	peu en sécurité	0	0	+ 1			
	pas du tout en sécurité	+ 1	-1	0			

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part de ceux qui se disent préoccupés par les maladies graves est de 4 points supérieure à la population totale ; cette même part est inférieure de 2 points à la population totale chez les internautes et inférieure de 1 point dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes. Par ailleurs, l'accès à Internet à domicile n'est pas lié, toutes choses égales par ailleurs, au fait d'être préoccupé par les maladies graves (colonne de gauche).

**Tableau A 17 : les opinions sur les politiques familiales**

		Ecart avec la population totale			Internaute/ non internautes	Internaute/ population totale	Echantillon d'internautes redressé/ population totale
		Non internautes	Internaute	Echantillon d'internautes redressé			
Influence intrinsèque d'Internet (*** : p<0.001; ** : p< 0.01; * : p < 0,05)							
<b>Préférence durée congé parental</b>	3 ans, montant forfaitaire	+ 5	-2	-1			
	+ court, montant + élevé nsp, nr	-4 0	+2 0	1 0			
<b>Mode de garde préféré pour enfant - 1 an</b>	crèche collective	-2	+1	0			
	nourrice agréée	+1	0	0			
	nourrice non agréée	0	0	0			
	crèche familiale	-1	+1	0			
	garde à domicile	-1	+1	+2			
	grands parents	+3	-2	-2			
	halte garderie autre	0 0	0 0	0 0			
<b>Mode de garde préféré pour enfant 1-3 ans</b>	crèche collective	-3	+2	0			
	nourrice agréée	+2	-1	+1			
	nourrice non agréée	0	0	0			
	crèche familiale	-1	0	-1			
	garde à domicile	-1	+1	+1			
	grands parents	+4	-2	-2			
	halte garderie autre	-1 0	0 0	+1 0			
<b>Mesure préférée pour aider familles, garde enfant - 3 ans</b>	aides réduction activité + de places en crèche	-1 -1	+1 +1	+1 +1			
	créer jardins enfants	+2	-1	-1			
	aides pr garde indiv	+1	0	-1			
	mieux inform. parents	+1	0	0			
	mieux répartir géogr.	-1	+1	+1			
<b>Mesure préférée pour aider familles, enfant 12-17 ans</b>	relever montant P.F.	0	0	+1			
	dyper activités pr enfants	-2	+1	0			
	réduire temps travail	+1	0	-1			
	plus d'enseignants	-2	+1	+1			
	salaire au foyer échanges prof. enf.	+3 0	-2 0	-1 0			
<b>* La famille doit aider jeune en difficulté</b>	oui	-2	+1	+1			
	non	+2	-1	-1			
<b>L'état doit aider jeune en difficulté</b>	oui	+1	0	-1			
	non	0	0	0			
<b>** PF seules (allocataires)</b>	oui	-2	0	-1			
	non	+2	0	+1			
<b>Opinion sur transmission des données de la DGI</b>	mesure utile	-3	+2	0			
	mesure injustifiée nsp, nr	+3 0	-2 0	-1 0			
<b>Allocations logement devraient être attribuées ...</b>	à un plus grand nombre	-3	+2	-1			
	aux plus démunis nsp, nr	+3 0	-2 0	+1 0			
<b>Les aides aux familles sont ...</b>	suffisantes	+1	-1	-1			
	insuffisantes	-3	+2	+1			
<b>Mieux vaut, pour aider les familles ...</b>	aides en espèces	+2	-1	-2			
	aides équipements, serv.	-3	+2	+2			
<b>Modulation Allocation Rentrée Scolaire ...</b>	une bonne chose	-1	0	0			
	une mauvaise chose nsp, nr	+1 0	-1 0	-1 +1			
<b>Les montants Allocation Rentrée Scolaire ...</b>	tout a fait adaptés	+3	-2	-1			
	plutôt adaptés	+1	0	0			
	plutôt pas adaptés	-2	+1	0			
	pas du tout adaptés nsp, nr	-2 0	+1 0	+1 0			

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part de ceux qui jugent les montants de l'Allocation de Rentrée Scolaire plutôt adaptés est de 4 points supérieure à la population totale ; cette même part est inférieure de 2 points à la population totale chez les internautes et inférieure de 1 point dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes. Par ailleurs, l'accès à Internet à domicile n'est pas lié, toutes choses égales par ailleurs, aux jugements portés sur les montants de l'ARS (colonne de gauche).

**Tableau A 17 - suite : les variables concernant les opinions sur les politiques familiales**

Influence intrinsèque d'Internet (*** : p<0.001; ** : p< 0.01; * : p < 0,05)	Ecart avec la population totale			Internautes /non internautes	Internautes /population totale	Echantillon d'internautes redressé/ population totale
	Non internautes	Internautes	Echantillon d'internautes redressé			
↓	<b>Préférence pour système qui favorise</b>	enfant de 6 à 10 ans	-1	+ 1	0	
		enfant de 11 à 14 ans	-1	+ 1	-1	
		enfant de 15 à 18 ans	+ 2	-1	0	
		âge indifférent	-1	0	0	
		nsp, nr	+ 1	0	0	
**	<b>Pour aider familles, mieux vaut donner ...</b>	les mêmes pour tous	+ 1	-1	-1	
		plus pour les démunis	+ 4	-2	0	
		selon coût réel enfant	-6	+ 3	+ 2	
***	<b>Action la plus utile pour jeunes adultes</b>	aides au logement	+ 1	0	0	
		alloc. au-delà 20 ans	+ 2	-1	0	
		bourses étudiants	-7	+ 4	+ 3	
		aides à l'insertion	+ 3	-1	-2	
		déductions fisc. parents	+ 1	-1	-1	
		RMI avant 25 ans	0	0	0	
	<b>Action la moins utile pour jeunes adultes</b>	aides au logement	0	0	-1	
		alloc. au-delà 20 ans	0	0	0	
		bourses étudiants	+ 1	0	0	
		aides à l'insertion	+ 1	-1	0	
		déductions fisc. parents	-1	+ 1	0	
		RMI avant 25 ans	0	0	+ 1	
	<b>Objectif principal actuel des politiques familiales</b>	soutenir la natalité	-1	0	-1	
		lutter contre la pauvreté	+ 5	-3	-1	
		aider les familles	-3	+ 2	+ 2	
		aider fam. jeunes enfants	0	0	-1	
		aider fam. grands enfants	+ 1	-1	0	
		lutter contre inégalités	0	0	0	
		aider à bien se loger	-1	+ 1	+ 1	
		aider si pb. familiaux	0	0	0	
		nsp, nr	0	0	0	
		*	<b>Objectif principal souhaité à l'avenir pour les politiques familiales</b>	soutenir la natalité	+ 1	0
lutter contre la pauvreté	-2			1	1	
aider les familles	0			0	0	
aider fam. jeunes enfants	+ 1			0	0	
aider fam. grands enfants	+ 1			-1	-1	
lutter contre inégalités	-1			+ 1	+ 1	
aider à bien se loger	-1			+ 1	-1	
aider si pb. familiaux	+ 1	-1	0			

Source : CRÉDOC, Enquête « Conditions de vie et Aspirations des Français », début 2009.

**Lecture** : chez les non-internautes, la part de ceux qui pensent que pour mieux aider les familles il faut verser des aides qui correspondent au coût réel de l'enfant est de 6 points inférieure à la population totale ; cette même part est supérieure de 3 points à la population totale chez les internautes et supérieure de 2 point dans l'échantillon d'internautes redressé. La différence est significative entre internautes et non-internautes. Par ailleurs, l'accès à Internet à domicile est lié, toutes choses égales par ailleurs, à l'aide qu'on juge préférable pour mieux aider les familles (au seuil de 1%, voir colonne de gauche).