



**EFFET DU PASSAGE EN LIGNE DU QUESTIONNAIRE
COMPORTEMENTS ET CONSOMMATIONS ALIMENTAIRES EN
FRANCE**

Juliette GAMBARETTI, Patricia CROUTE et Pascale HEBEL

CAHIER DE RECHERCHE N°338

DECEMBRE 2017

Pôle Consommation et Entreprise

Dirigé par Pascale HEBEL

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

SOMMAIRE

RESUME	4
INTRODUCTION	5
1 ETAT DE L'ART	6
2 METHODOLOGIE	8
1. Collecte des données.....	8
Les questionnaires.....	8
L'enquête en face-à-face	8
L'enquête en ligne	9
2. Redressement classique des deux échantillons sur les variables de quotas	10
3. Comparaison des réponses entre les deux échantillons	11
Regroupement des variables d'intérêt par blocs.....	11
Création d'indicateurs résumés des écarts entre les deux échantillons.....	13
Questions sujettes à un biais de désirabilité sociale	13
4. Tests de redressement de l'échantillon des panélistes.....	14
Redressements par calage sur marge.....	14
Redressements par pondération en utilisant l'inverse du score de propension	16
Combinaison des deux méthodes.....	17
3 RESULTATS.....	18
Des différences observées sur les questions auxiliaires.....	18
Des différences sociodémographiques entre les deux échantillons malgré les quotas.....	18
Des différences sur les autres questions auxiliaires	19
2. Des différences observées sur les questions d'intérêt	20
Des écarts dans la structure des réponses qui sont liés à la façon de remplir le questionnaire	20
Des écarts de réponses qui reflètent des différences d'opinions et de comportements entre les deux échantillons interrogés.....	21
Des écarts peu visibles sur les questions sujettes à un biais de désirabilité sociale.....	31
3. Effets des redressements de l'échantillon des panélistes	32
Redressements par calages sur marges.....	32
Redressements par pondération en utilisant l'inverse du score de propension	36
Combinaison des deux méthodes de redressement	39
4 DISCUSSION.....	40
1. Des différences observées sur les questions auxiliaires	40
Des différences sociodémographiques entre les deux échantillons malgré les quotas.....	40
Des différences sur les autres questions auxiliaires	41
2. Des différences observées sur les questions d'intérêt	42
Des écarts dans la structure des réponses qui sont liés à la façon de remplir le questionnaire	42
Des écarts de réponses qui reflètent des différences d'opinions et de comportements entre les deux échantillons interrogés.....	43
3. Effet des redressements supplémentaires de l'échantillon interrogé en ligne sur les écarts de réponses aux questions d'intérêt	46
Méthodes de redressement utilisées.....	46
Variables de redressement utilisées	47
5 CONCLUSION	49
BIBLIOGRAPHIE	51

RESUME

Les méthodes de collectes de données quantitatives se font de plus en plus avec de nouveaux modes d'interrogation. Après le téléphone, la diffusion d'internet a permis de réaliser des enquêtes auprès d'échantillons de volontaires. Ces évolutions posent la question de la représentativité des échantillons ainsi constitués ainsi que celle de la sensibilité des réponses aux modes d'interrogation, notamment avec le moindre effet de désirabilité sociale dans les modes auto-administrés.

En analysant l'enquête Comportement Alimentaire en France du CREDOC, nous comparons deux sous-échantillons réalisés sur la même période, l'un est interrogé en Face à Face et l'autre par Internet. Nous distinguons les différences dues aux caractéristiques des répondants et celles imputables aux conditions de collecte. Comme l'ont fait auparavant Thompson et al (2013) et Croute (2015), nous proposons une méthode pour produire une estimation des résultats en mode Online que l'on aurait obtenus avec une méthodologie en Face en Face afin d'avoir une continuité barométrique entre un système en Face à Face réalisé depuis 1988 et un nouveau système en Online à partir de 2016.

Mots clés : access panel ; score de propension ; désirabilité sociale ; calage sur marge ; multimode.

INTRODUCTION

Depuis 1988, le CRÉDOC réalise tous les trois ans l'enquête « Comportements et Consommations Alimentaires en France (CCAF) », qui a pour objectif d'analyser le rapport à l'alimentation des Français sous toutes ses formes. Actuellement menée en face-à-face, cette enquête comporte deux volets :

- Un volet « Comportements et Attitudes », qui étudie les opinions, les comportements et les attitudes des Français dans le domaine alimentaire, en soumettant un questionnaire à un échantillon représentatif de la population française.
- Un volet « Consommation Alimentaire », qui étudie qualitativement et quantitativement la consommation alimentaire des Français en recueillant sous forme de carnets de consommation la consommation alimentaire directe d'un échantillon représentatif de la population française.

En raison des nombreux avantages pratiques et économiques que présentent les enquêtes en ligne par rapport aux enquêtes en face-à-face, le CRÉDOC souhaite réaliser à l'avenir le volet « Comportements et Attitudes » en posant le questionnaire en ligne plutôt qu'en face-à-face comme cela a pu être réalisé pour son enquête Conditions de Vie et Aspirations (Bigot et al (2010) et Croutte et al (2015)). Cependant, il est aujourd'hui reconnu que les enquêtes en ligne présentent un problème de représentativité de la population générale. Il est donc nécessaire d'étudier préalablement dans quelle mesure le passage en ligne du questionnaire impacte les réponses recueillies. Pour cela, en parallèle à la réalisation en face-à-face de l'enquête CCAF de 2016, le CRÉDOC a soumis le même questionnaire à un panel d'internautes. Le premier objectif de cette étude est donc d'identifier quels sont précisément les effets du passage en ligne de l'enquête en étudiant les écarts entre les réponses obtenues en ligne et en face-à-face. Le deuxième objectif est de rechercher des méthodes statistiques pour diminuer ces écarts, dans le but de gommer les effets liés au mode de collecte et d'augmenter la représentativité de l'enquête en ligne.

1 ETAT DE L'ART

Depuis quelques années, la méthodologie d'enquête a été amenée à évoluer considérablement. D'après Dillman et al. (2006) les enquêtes en face-à-face dominaient largement dans les années 1960. Dans les années 1970, la modernisation, entre autres du téléphone et des photocopies, a permis aux enquêtes téléphoniques et par courrier de se développer considérablement. Les enquêtes par courrier qui présentaient l'avantage d'être auto-administrées, étaient alors utilisées lorsque les enquêteurs disposaient de listes d'adresses adéquates. Les enquêtes téléphoniques étaient quant à elles intéressantes car elles permettaient de recueillir des données sur une grande échelle géographique. De plus elles présentaient un taux de participation généralement assez élevé (Croutte, 2015; Dillman, 2006).

Cependant, depuis les années 1990, plusieurs obstacles se sont mis sur la voie des enquêtes par téléphone. D'abord, la qualité des bases de sondages diminue : les numéros sur listes rouges augmentent, les opérateurs se multiplient et les téléphones fixes sont de plus en plus substitués aux téléphones mobiles (Croutte, 2015). En parallèle, une forme de lassitude se fait sentir chez les personnes, de plus en plus sollicitées ; et le nombre de personnes qui filtrent les appels ou raccrochent leurs téléphones augmente considérablement. Par conséquent, ces enquêtes deviennent de moins en moins représentatives de la population générale.

Dans le même temps, la révolution du numérique s'est mise en place et les enquêtes sur Internet ont commencé à connaître un véritable engouement. Ces dernières présentent en effet plusieurs avantages. Premièrement, elles peuvent permettre de diminuer considérablement le coût et la durée de la collecte, en partie puisque ceux-ci ne sont plus liés à la taille de l'échantillon. Elles abolissent également les frontières dans le recueil de données et permettent de couvrir des aires géographiques importantes (Dillman, 2006; Evans & Mathur, 2005)

De plus, un autre avantage majeur de ces enquêtes vient du fait qu'elles soient auto-administrées. L'enquêteur est absent et ne peut donc pas influencer le participant, que ce soit de façon négative ou positive. Certes il ne peut pas le motiver à répondre ni apporter des précisions. Mais il ne peut pas non plus influencer les réponses du participant via un processus de désirabilité sociale : seul devant son écran, l'enquêté ne cherchera pas forcément à se montrer sous son meilleur jour. (Duffy, Terhanian, Bremer, & Smith, 2005; Evans & Mathur, 2005; Kreuter, Presser, & Tourangeau, 2008)

Enfin, la simplicité d'utilisation facilite grandement la passation du questionnaire. Les répondants peuvent répondre à leur rythme et aux plages horaires qui les arrangent (tard le soir, le weekend...). De plus, les filtres sont gérés automatiquement, les différents items peuvent être proposés en rotation aléatoire. Différents outils peuvent également être facilement introduits pour faciliter la réponse : aides et instructions, images, sons, vidéos... (Croutte, 2015; Evans & Mathur, 2005)

Cependant, les enquêtes en ligne présentent un inconvénient majeur lorsque l'on cherche à généraliser les résultats à la population cible : leur représentativité (Couper, 2011; Duffy et al., 2005). Ceci est notamment dû au fait que la couverture est encore insuffisante. En effet, en France, toute la population ne dispose pas d'un ordinateur, même si, d'après le baromètre du numérique réalisé par

le CRÉDOC en 2015, 80% de la population dispose d'au moins un ordinateur à son domicile (Brice, Croutte, Jauneau-Cottet, & Lautié, 2015). Ils ont également recensé 84% d'internautes en France et 68% d'internautes qui se connectent tous les jours. Toute la population n'est donc pas connectée et des différences marquées existent entre les personnes connectées et les non-connectées, induisant un biais de sélection lorsque les participants d'une enquête sont recrutés sur internet. Plus précisément, plus une personne est jeune et plus elle a de chances d'être internaute. Par exemple, parmi les plus de 70 ans, seuls 38% sont des internautes. Le lien avec le diplôme est également très net (la proportion d'internautes variant presque du simple au double en fonction du niveau de diplôme). Une personne seule est également moins souvent internaute que celle qui évolue dans un foyer de trois personnes ou plus ; et une personne qui dispose de bas revenus a également de moindres chances d'être internaute. Enfin, les hommes sont surreprésentés chez les personnes connectées (Brice et al., 2015).

Différents travaux ont été réalisés depuis le début des années 1980 pour améliorer la représentativité des enquêtes menées sur Internet. Deux principales méthodes sont aujourd'hui utilisées pour redresser les populations d'enquêtés en ligne. La première méthode est la post-stratification par calage sur marge, qui consiste à caler la distribution de certaines variables de l'échantillon en ligne pour qu'elles aient la même distribution que dans un échantillon de référence (Bayart & Bonnel, 2010; Blasius & Brandt, 2010; Croutte, 2015). Une autre méthode consiste à utiliser des scores de propension pour pondérer l'échantillon des panélistes (Austin, 2011; Couper, 2011; Croutte, 2015; Lee, 2006; Rosenbaum, 1987; Rosenbaum & Rubin, 1984; Schonlau, 2007).

2 METHODOLOGIE

1. Collecte des données

Les questionnaires

Le volet « Comportements et Attitudes » de l'enquête CCAF s'intéresse aux opinions et aux comportements en termes d'achats, de cuisine, etc. Il se décompose en trois questionnaires différents qui sont posés à trois groupes de population :

- Un questionnaire MENAGES auquel répond le ou la responsable du ménage en matière alimentaire, c'est-à-dire la personne qui fait le plus souvent les courses et/ou la cuisine.
- Un questionnaire ADULTES posé à tous les individus de 15 ans et plus de chaque ménage.
- Un questionnaire ENFANTS posé à tous les enfants de 3 à 14 ans de chaque ménage.

Pour étudier les effets d'un passage en ligne du questionnaire, ce cahier de recherche s'intéresse aux questionnaires MENAGES et ADULTES. Ces derniers abordent deux types de questions : des questions d'intérêt et des questions auxiliaires. Les questions d'intérêt sont celles qui ont pour but l'étude des opinions, des attitudes et des comportements alimentaires en France au niveau des foyers ou des individus.

- Les questions d'intérêt du questionnaire MENAGES portent sur les achats, la préparation des repas, la prise des repas et les types de produits consommés.
- Les questions d'intérêt du questionnaire ADULTES portent sur les pratiques et les opinions de l'alimentation, la restauration hors foyer, la consommation hors repas, la santé et la nutrition et les risques alimentaires.

En plus de ces questions d'intérêt, des questions auxiliaires permettent d'obtenir des informations générales sur les personnes interrogées.

- Pour le questionnaire adressé aux ménages, des questions sociodémographiques sont posées, ainsi que des questions sur l'équipement du ménage et sur l'utilisation des écrans et d'internet et des services de livraison à domicile.
- Pour le questionnaire des adultes, en plus des questions sociodémographiques, des questions sont posées concernant le poids des individus, leur taille, leur niveau de sédentarité et d'activité physique, les sports pratiqués et leur consommation de tabac.

L'enquête en face-à-face

L'enquête en face-à-face correspond à la réalisation classique en face-à-face du volet « Comportements et attitudes » de l'enquête CCAF 2016. Elle a été réalisée par les enquêteurs professionnels de la société ISL (groupe GfK) d'octobre 2015 à juillet 2016.

Pour s'assurer que la population de l'enquête CCAF 2016 soit représentative de la population française, la méthode des quotas a été appliquée par la société ISL pour sélectionner l'échantillon des ménages interrogés. Cette méthode d'échantillonnage consiste à s'assurer de la représentativité d'un échantillon en lui affectant une structure similaire à celle de la population de base sur certains critères sociodémographiques. En respectant cette structure, les ménages sont sélectionnés de manière aléatoire. Les critères sociodémographiques qui ont été utilisés sont les suivants : l'âge de la personne responsable du ménage, sa catégorie socio-professionnelle, la région d'habitation du ménage, la taille d'agglomération, le nombre de personnes vivant dans le foyer et le type de famille (personne seule, couple sans enfant, couple avec enfant, famille monoparentale ou autre type de famille). Les cinq premiers critères sont utilisés depuis 1988 pour réaliser l'enquête CCAF, alors que le dernier critère a été rajouté pour la dernière version de l'enquête en 2016. D'après l'INSEE, la personne de référence du ménage (anciennement chef du ménage) est le plus souvent de la personne de référence de la famille quand il y en a une, ou de l'homme le plus âgé, en donnant priorité à l'actif le plus âgé. Mais depuis 2004, la définition a été modifiée pour ne pas tenir compte du sexe des personnes ; pour prendre uniquement en compte l'âge, le fait d'apporter ou non des ressources au ménage et d'être actif ou non (« Site Internet de l'INSEE - Définition personne de référence du ménage », 2017). Les quotas ont été fixés à partir d'une enquête de référence de l'INSEE, l'enquête « emploi 2014 » (« Enquête Emploi 2014 ; Activité, chômage, emploi et inactivité en 2014 - Insee Résultats - 173 », 2017).

Conformément à la méthodologie utilisée depuis 1988, les enquêteurs se sont rendus au domicile des ménages sélectionnés où ils ont interrogé les personnes de références en matière alimentaire avec le questionnaire MENAGES. Au total, ce premier questionnaire a été renseigné par 1519 ménages. Cette taille d'échantillon a été volontairement augmentée par rapport aux précédentes enquêtes CCAF (1000 ménages en 2003, 1200 en 2007, 2010 et 2013). Le questionnaire ADULTES a ensuite été posé, cette fois à tous les membres du foyer âgés de 15 ans et plus (y compris le responsable en matière alimentaire ayant déjà rempli le questionnaire MENAGES). Au total, 2456 adultes ont été interrogés en 2016. Le questionnaire ENFANTS a été posé à toutes les personnes entre 3 et 15 ans. Les carnets de consommations alimentaires ont également été distribués lors de cette visite, avant d'être vérifiés puis récupérés lors d'une seconde visite de l'enquêteur. Mais ces derniers ne font pas l'objet de cette étude.

L'enquête en ligne

Pour réaliser l'enquête en ligne, exactement le même questionnaire « Comportements et attitudes » a été posé à un panel de consommateurs en ligne. L'enquête a été menée du 16 juin au 5 juillet 2016. Contrairement à l'enquête en face-à-face, l'enquête en ligne a pu être menée sur une fenêtre temporelle plus réduite, pour des raisons de praticité et puisque la partie concernant les carnets alimentaires, très sensible aux effets de saisonnalité, n'a pas été réalisée en ligne.

La sélection des interrogés a été confiée au prestataire TOLUNA, en charge du terrain de l'enquête en ligne. Contrairement à l'enquête en face-à-face, les ménages n'ont pas pu être sélectionnés selon la méthode des quotas, le prestataire ne disposant pas d'informations sociodémographiques directes par ménages, mais uniquement par individus. La méthode de sélection par quotas a donc cette fois été appliquée aux individus, également d'après l'enquête « emploi 2014 ». Pour l'échantillon des

adultes, les quotas ont été appliqués sur les variables suivantes : l'âge de l'interviewé, son sexe, sa catégorie socio-professionnelle, la taille de l'unité urbaine et la région d'habitation.

Le questionnaire ADULTES a été rempli sur internet par 1692 individus de plus de 15 ans. A la fin du questionnaire, une question a été ajoutée pour demander à l'interviewé s'il était ou non la personne de référence en matière alimentaire. Si l'individu répondait oui, il était alors invité à répondre en plus au questionnaire MENAGES. Finalement 1217 personnes ont rempli le questionnaire MENAGES par internet. Au total, le passage du questionnaire durait entre trente minutes et une heure, suivant si l'interviewé était responsable en matière alimentaire ou non.

2. Redressement classique des deux échantillons sur les variables de quotas

Toujours dans un souci de représentativité de la population générale, les échantillons ainsi sélectionnés sont redressés sur les variables de quotas afin de s'assurer qu'ils respectent bien les quotas imposés. Ce redressement est très important ici, puisque les recrutements se sont fait différemment selon le mode de collecte : la méthode des quotas ayant été appliquée aux ménages pour l'enquête en face-à-face, et aux adultes pour l'enquête en ligne.

Ainsi, les échantillons des ménages (obtenus en ligne et en face-à-face) ont été redressés par la méthode du calage sur marge sur les variables de quotas suivantes :

- **L'âge de la personne responsable du ménage** en 6 modalités : « 15-24 ans », « 25-34 ans », « 35-44 ans », « 45-54 ans », « 55-64 ans » et « 65 ans ou plus ».
- **La Profession et Catégorie Socioprofessionnelle (PCS) du responsable du ménage** en 7 modalités : « Agriculteur exploitant », « Artisan, commerçant, chef d'entreprise », « Cadre, profession intellectuelle supérieure », « Profession intermédiaire », « Employé », « Ouvrier », « Non concerné / inactif ».
- **La région d'habitation du ménage**, en 9 modalités, suivant les Zones d'Etudes et d'Aménagement du Territoire (ZEAT) de l'INSEE : « Région parisienne », « Bassin parisien », « Ouest », « Est », « Nord », « Sud-Ouest », « Centre Est », « Méditerranée » (« Site Internet de l'INSEE - Définition ZEAT », 2017).
- **La taille de l'unité urbaine où vit le ménage**, en 5 modalités : « Moins de 2 000 habitants », « 2 000 à 20 000 habitants », « 20 000 à 100 000 habitants », « Plus de 100 000 habitants hors agglomération parisienne » et « Agglomération parisienne ».
- **Le nombre de personnes vivant dans le foyer**, en 5 modalités : « 1 personne », « 2 personnes », « 3 personnes », « 4 personnes » et « Plus de 5 personnes ».
- **Le type de famille** en 5 modalités : « Personne seule », « Couple sans enfant », « Couple avec enfants », « Famille monoparentale », « Autre type de famille ».

De la même façon, les échantillons des adultes (obtenus en ligne et en face-à-face) ont été redressés par la méthode du calage sur marge sur les variables de quotas suivantes :

- **L'âge de l'individu interrogé** en 6 modalités : « 15-24 ans », « 25-34 ans », « 35-44 ans », « 45-54 ans », « 55-64 ans » et « 65 ans ou plus ».
- **Le sexe de l'individu interrogé** : « Homme » ou « Femme ».

- **La Profession et Catégorie Socioprofessionnelle (PCS) de l'individu** en 7 modalités : « Agriculteur exploitant », « Artisan, commerçant, chef d'entreprise », « Cadre, profession intellectuelle supérieure », « Profession intermédiaire », « Employé », « Ouvrier », « Non concerné / inactif ».
- **La taille de l'unité urbaine où vit l'individu**, en 5 modalités : « Moins de 2 000 habitants », « 2 000 à 20 000 habitants », « 20 000 à 100 000 habitants », « Plus de 100 000 habitants hors agglomération parisienne » et « Agglomération parisienne ».
- **La région d'habitation du ménage**, en 9 modalités, suivant les Zones d'Etudes et d'Aménagement du Territoire (ZEAT) de l'INSEE : « Région parisienne », « Bassin parisien », « Ouest », « Est », « Nord », « Sud-Ouest », « Centre Est », « Méditerranée ».

Dans les deux cas, les marges appliquées sont fixées par l'enquête « Emploi 2014 » de l'INSEE (« Enquête Emploi 2014 ; Activité, chômage, emploi et inactivité en 2014 - Insee Résultats - 173 », 2017). Les échantillons recueillis en ligne et en face-à-face présentent donc la même répartition vis-à-vis de ces variables.

3. Comparaison des réponses entre les deux échantillons

Le premier objectif est d'étudier comment sont affectées les réponses au questionnaire « Comportement et attitudes alimentaires » lorsque les données sont collectées en ligne plutôt qu'en face-à-face. Le deuxième but est de proposer des méthodes pour diminuer les écarts observés. Pour cela, des indicateurs ont été mis au point pour évaluer ces écarts de réponses entre les deux modes de collecte.

Regroupement des variables d'intérêt par blocs

Les différentes variables (correspondant à 288 questions pour le questionnaire adressé aux ménages, et 230 pour celui posé aux adultes) ont d'abord été regroupées en blocs. Les questions portant sur les habitudes et les attitudes ont été séparées de celles qui portent sur les opinions, les représentations, ou les raisons d'achat. Le Tableau 1 et le Tableau 2 ci-dessous listent les différents groupes de questions abordées, avec pour chaque bloc le nombre de questions et le nombre de modalités totales de réponse correspondant.

Tableau 1 : Regroupement des variables d'intérêt du questionnaire MENAGES

TYPE DE QUESTION – Questionnaire « MENAGES »	NOMBRE DE QUESTIONS	NOMBRE DE MODALITES
Types de produits alimentaires achetés	60	152
Produits frais non préparés	7	21
Produits prêts à l'emploi	7	21
Produits bio	14	42
Produits régionaux	8	17
Produits AOC	8	17
Produits issus du commerce équitable	2	6
Consommation de cuisine exotique	14	28
Raisons d'achat de produits alimentaires	69	138
Prix	3	6
Mode de production	14	28
Sécurité et effets sur la santé	8	16
Saveur	4	8
Aspect du produit et informations en magasin	14	28
Imaginaire derrière les produits	5	10
Autre	9	18
Acceptation de payer plus cher pour des signes officiels de qualité	12	24
Lecture des étiquetages alimentaires	24	50
Lecture d'informations autre que le prix et la DLC	14	30
Lecture des DLC	10	20
Préparation des repas	33	73
Organisation pour la préparation des repas	10	26
Intégration des exigences de santé dans la préparation des repas	23	47
Caractéristiques des repas	15	48
Structure des repas	9	31
Invités, apéritifs	6	17
Approvisionnement en produits alimentaires	21	71
Circuits de distribution fréquentés	8	32
Enseignes de grande distribution fréquentées	9	27
Organisation pour faire les courses	4	12
Situation du foyer vis-à-vis de la sécurité alimentaire	10	22
Attitudes vis-à-vis du gaspillage alimentaire	59	119
Types de produits jetés ou congelés	9	25
Produits jetés lorsqu'ils sont abimés, ont mauvais gout, mauvaise odeur	7	14
Produits jetés car le ménage n'a pas de congélateur (ou il est trop petit)	7	14
Produits jetés lorsque la DLC est dépassée	5	10
Produits jetés lorsqu'ils ne sont pas bons	7	14
Produits jetés lorsqu'ils sont achetés en trop grande quantité	7	14
Produits jetés lorsque les individus ne savent pas comment consommer les restes	7	14
Produits jetés pour d'autres raisons	7	14

Source : CRÉDOC, enquête CCAF 2016

Tableau 2 : Regroupement des variables d'intérêt du questionnaire ADULTES

TYPE DE QUESTION – Questionnaire « ADULTES »	NOMBRE DE QUESTIONS	NOMBRE DE MODALITES
Structure des prises alimentaires	49	114
Rythme des repas	7	21
Lieux des repas	17	41
Circonstances de consommations hors repas	25	52
Raisons d'achat et critères de qualités	27	71
Raisons d'achat	16	48
Critères de qualité des produits alimentaires	11	23
Opinion sur les liens entre alimentation et plaisir	26	62
Gout pour la cuisine et ouverture culinaire	11	32
Gouts préférés	5	10
Importance du plaisir et de la convivialité pour bien manger	10	20
Opinion sur les liens entre alimentation et santé	35	70
Éléments importants dans la prévention en matière de santé	7	14
Importance de la santé et du mode de production pour bien manger	17	34
Attentes de l'alimentation pour être en bonne santé	11	22
Attitudes vis-à-vis des liens entre alimentation et santé	27	57
Aliments-santé consommés	5	11
Éléments surveillés dans l'alimentation	9	18
Régimes suivis	13	28
Perception des risques alimentaires	36	89
Types de risques alimentaires	19	38
Types de produits alimentaires à risques	17	51
Opinion sur le système alimentaire	30	77
Confiance envers les acteurs de l'alimentation	15	30
Opinion sur la réglementation en matière alimentaire	5	15
Opinion sur les OGM	8	24
Opinion sur l'évolution de l'alimentation	2	8

Source : CRÉDOC, enquête CCAF 2016

Création d'indicateurs résumés des écarts entre les deux échantillons

Pour évaluer les écarts de réponses entre les deux modes de collecte, le pourcentage de réponses obtenues en face-à-face et en ligne a été relevé pour toutes les modalités de chaque question. Puis le nombre de points d'écart a été déterminé par la valeur absolue de la différence entre ces deux pourcentages. L'enquête CCAF étant la propriété du CRÉDOC et confidentielle, les seuls résultats qui sont présentés dans ce cahier de recherche sont les différences de points et non les pourcentages de réponse directs. L'écart moyen de points par blocs a ensuite été déterminé en calculant la moyenne des écarts pour toutes les modalités de chaque bloc. De la même façon, l'écart moyen pour l'ensemble du questionnaire a été calculé, pour chaque questionnaire.

Les différences obtenues ont été testées statistiquement pour chaque question au moyen d'un test de khi deux, pour déterminer si la distribution des réponses à une question est indépendante du mode de collecte, en ligne ou en face-à-face, pour chaque question. Les tests de khi deux ont été réalisés sous SAS 9.4 avec un seuil de risque de 5%.

Questions sujettes à un biais de désirabilité sociale

De nombreux auteurs ont signalé un biais introduit par la présence ou l'absence de l'enquêteur lors du remplissage du questionnaire. La plupart du temps, les études ont montré qu'en face-à-face, les interrogés avaient tendance à se reporter plus ou moins consciemment à la norme sociale établie,

en citant plus facilement des réponses attendues par cette norme. Au contraire, avec des questionnaires auto-administrés, les personnes ont tendance à être plus honnêtes et plus franches. (Duffy et al., 2005). Une attention particulière a donc été portée aux questions qui pouvaient présenter un biais de désirabilité et des blocs supplémentaires ont donc été créés regroupant les questions qui peuvent être concernées.

- Pour l'échantillon des ménages, les questions qui ont été considérées comme les plus concernées par ces biais sont celles portant sur le gaspillage alimentaire et les produits alimentaires jetés, et celles sur les inquiétudes vis-à-vis de la sécurité alimentaire.
- Dans le cas de l'échantillon des adultes, les questions suivantes ont été regroupées : la fréquence de consommation en dehors des repas, l'utilisation des distributeurs automatiques de snacks et de boissons, la fréquentation des enseignes de restauration rapide, le temps passé devant les écrans, le nombre de cigarettes fumées par jour, l'image qu'ont les individus de leur poids et concernant leur activité physique et enfin la mesure avec laquelle les individus surveillent leur poids et font attention à leur alimentation.

4. Tests de redressement de l'échantillon des panélistes

Le deuxième objectif de cette étude est de trouver des méthodes de redressement permettant d'augmenter la représentativité de l'enquête en ligne, pour que l'enquête CCAF puisse à l'avenir être réalisée sur Internet. Ceci est nécessaire pour que les résultats des futures enquêtes conduites en ligne puissent être comparés avec les résultats obtenus les années précédentes en face-à-face. L'expertise du CRÉDOC réside en effet dans sa capacité à étudier des évolutions au cours du temps et à prédire les tendances futures.

Considérons que l'enquête en face-à-face est représentative de la population. On cherche donc à retrouver par l'enquête en ligne les résultats obtenus en face-à-face. L'objectif est donc de trouver des moyens statistiques pour modifier l'échantillon obtenu en ligne, et lui seul, pour que les résultats obtenus soient semblables à ceux obtenus en face-à-face. L'efficacité des moyens testés sera donc évaluée par leur capacité à diminuer les écarts de réponse entre les deux modes de collecte.

Redressements par calage sur marge

Comme l'a fait précédemment dans son étude Patricia Croutte au CRÉDOC (Croutte, 2015), la première hypothèse qui a été testée ici pour tenter de diminuer les différences entre les modes de collectes est de rajouter des variables supplémentaires dans le calage sur marges aux variables de quotas utilisées de façon « classique » pour l'enquête CCAF (voir partie 2. Redressement « classique » des deux échantillons).

L'hypothèse 0 correspond au redressement classique avec les variables de quotas habituelles. Pour les hypothèses suivantes, différentes variables auxiliaires disponibles ont été testées parmi celles qui peuvent expliquer qu'un individu ait une plus grande probabilité d'être un panéliste ou non, en suivant les caractéristiques des panélistes citées dans la littérature : le fait qu'ils soit plus diplômés, plus urbains, plus modernes, plus utilisateurs d'internet... (Brice et al. 2015). Des variables permettant de refléter la modernité des répondants ont aussi été utilisées (Varedian & Forsman, 2003). En

incluant ces variables dans le redressement, les caractéristiques de l'échantillon en ligne sont rendues identiques à celles de l'échantillon en face-à-face du point de vue de ces variables auxiliaires.

Si ces caractéristiques ont un impact sur les comportements alimentaires, leur insertion dans le redressement devrait diminuer les différences observées sur les variables d'intérêt. Les variables auxiliaires suivantes ont été rajoutées aux variables de quotas initiales:

Echantillon des ménages :

- Hypothèse 1 : le niveau de diplôme le plus élevé du chef de famille
- Hypothèse 2 : le sexe du responsable des achats en matière alimentaire
- Hypothèse 3 : la fréquence à laquelle les membres du ménage prennent leurs différents repas (petit déjeuner, déjeuner et dîner) devant la télévision
- Hypothèse 4 : l'équipement électroménager de la cuisine du foyer (lave-vaisselle, plaques à induction ou vitrocéramiques ou plaques au gaz), qui reflète la modernité du ménage
- Hypothèse 5 : la fréquence d'utilisation d'Internet pour faire les courses
- Hypothèse 6 : la fréquence d'utilisation des services de livraison à domicile

Echantillon des adultes :

- Hypothèse 1 : le niveau de diplôme le plus élevé de l'individu interrogé
- Hypothèse 2 : la situation de l'habitation (en zone urbaine dense, en banlieue/périurbain pavillonnaire, dans un village/bourg, dans un hameau ou dans une zone isolée)
- Hypothèse 3 : la fréquence à laquelle l'individu pratique une activité physique
- Hypothèse 4 : le nombre de cigarettes fumées par jour
- Hypothèse 5 : la situation professionnelle (en poste ou non) et maritale de l'individu
- Hypothèse 6 : le temps passé devant les écrans par semaine (télévision, ordinateur et jeux-vidéo)
- Hypothèse 7 : l'ensemble des variables précédentes

Toutes ces variables sont des variables catégorielles, le redressement par calage sur marge consiste donc à modifier les pondérations initiales observées dans un échantillon en remplaçant la distribution de variables auxiliaires par la distribution de ces variables dans une population de référence (Bayart & Bonnel, 2010). Pour cela, le calage sur marge attribue à chaque individu un nouveau poids : le poids de calage. Cette méthode a été utilisée pour imposer dans l'échantillon en ligne une distribution identique sur ces variables à celle mesurée en face-à-face. La méthode du « raking ratio » a été utilisée, méthode itérative qui consiste à minimiser la somme des distances entre les poids initiaux et les poids de calage pondérée par les poids initiaux, sous la contrainte du calage (Bayart & Bonnel, 2010). Les calages sur marges ont été réalisés sur le logiciel SAS 9.4. en utilisant la macro « CALMAR », développée par l'INSEE (Sautory, 1993).

Il est important de noter que contrairement au redressement « classique », où les marges utilisées viennent d'une étude externe (l'enquête emploi 2014), les marges supplémentaires utilisées pour redresser l'échantillon en ligne par rapport à l'échantillon en face-à-face sont cette fois les marges observées sur l'échantillon en face-à-face pour la variable correspondante.

Pour juger de l'efficacité des différents redressements testés, leurs impacts sur les écarts moyens entre les réponses ont été analysés pour chaque bloc de question, ainsi que leurs impacts sur l'écart

moyen global pour la totalité du questionnaire concerné. Les méthodes de redressement les plus efficaces étant celles qui permettent le plus de diminuer les différences entre les deux modes de collecte. Cependant un arbitrage doit être réalisé entre l'efficacité du redressement et les modifications de l'échantillon qui sont nécessaires pour y parvenir. Pour cela, les caractéristiques des poids attribués aux individus pour chaque redressement doivent être également étudiées. Un redressement qui distribue des poids trop hétérogènes ne peut être validé, puisqu'il donne à un petit nombre d'individus un poids très important alors que certains individus ne sont presque pas représentés. Ce problème a été soulevé par Paul R. Rosenbaum dès 1987 : le redressement direct est certes attractif mais lorsqu'il se base sur de nombreuses sous-catégories de populations, il peut donner des poids très élevés à des petites quantités de données, souvent d'une façon erratique (Rosenbaum, 1987). Les caractéristiques des poids obtenus pour chaque redressement ont donc été étudiées pour chaque hypothèse. Etant donné les modes de sélection différents entre les deux échantillons, et puisque le but est d'étudier l'effet des redressements sans les appliquer pour analyser des résultats, il sera pour la suite considéré qu'une répartition des poids (minimum/maximum) inférieure à 200 est acceptable dans le cadre de cette étude méthodologique.

Redressements par pondération en utilisant l'inverse du score de propension

Inspiré également du précédent travail de Patricia Crouette (Crouette, 2015), un autre type de redressement couramment utilisé dans ce genre de problématiques a été testé. Cette méthode permet de créer un nouvel échantillon « artificiel » où la distribution de variables auxiliaire est à priori indépendante du mode de collecte des données (Austin, 2011). Autrement dit, cette méthode permet de simuler une sélection aléatoire des individus dans les deux modes de collecte, vis-à-vis de ces variables.

Pour cela, un score de propension a été estimé pour chaque individu du panel en ligne. Ce score de propension représente la probabilité, pour chaque individu, d'appartenir effectivement au panel en ligne (plutôt qu'au groupe interrogé en face-à-face). L'hypothèse est que, en donnant plus de poids aux individus ayant une faible probabilité d'appartenir au panel online (et donc une plus forte probabilité d'appartenir aux interviewés en face-à-face), et moins de poids à ceux ayant une forte probabilité d'appartenir au panel online, l'écart entre les deux échantillons peut être diminué. Pour cela, différentes utilisations du score de propension sont possibles. La plus intuitive a été choisie ici : chaque individu de la base online a reçu un poids égal à l'inverse de son score de propension d'appartenir à l'enquête en ligne.

Les scores de propension ont été estimés par une régression logistique. Le modèle a été réalisé avec le logiciel SAS 9.4. La variable binaire de réponse du modèle était le type d'enquête à laquelle a participé l'individu : en face-à-face ou en ligne, et les variables explicatives ont été choisies parmi les variables auxiliaires qui pouvaient expliquer qu'un individu ait plus de chance d'appartenir à l'échantillon recueilli en ligne ou non. L'option de sélection automatique des variables significatives du modèle par la méthode *stepwise* a été utilisée.

Echantillon des ménages

Les variables prises en compte dans le modèle de régression logistique et sélectionnées par la méthode *stepwise* pour l'échantillon des ménages sont les suivantes :

- Les variables de quotas
- D'autres variables sociodémographiques (le diplôme le plus élevé obtenu par le responsable du ménage, la situation de l'habitation, le type d'habitation : maison ou appartement ; et le statut marital et professionnel du responsable du ménage)
- L'âge et le sexe du responsable des achats en matière alimentaire
- La prise des trois repas de la journée devant la télévision par les membres du ménage
- L'utilisation par le responsable des achats d'Internet pour faire les courses et d'une tablette pour suivre des recettes de cuisine
- L'équipement du ménage en plaques de cuisine (à gaz, électriques et à induction), en congélateur et en lave-vaisselle
- L'utilisation par le ménage de services de livraison à domicile pour les courses ou les repas.

Echantillon des adultes

Pour l'échantillon des adultes, les variables sélectionnées sont les suivantes :

- Les variables de quotas
- D'autres variables sociodémographiques concernant l'individu : le diplôme le plus élevé obtenu par l'individu, la situation de l'habitation, le type d'habitation : maison ou appartement, le statut marital et professionnel du responsable du ménage, la nationalité et l'origine (étrangère ou non) de l'individu
- L'âge et la catégorie socioprofessionnelle du chef de famille
- L'indice de masse corporelle de l'individu (IMC)
- Le temps passé devant les écrans par l'individu par semaine (ordinateur, télévision et jeux-vidéos)
- La fréquence de pratique d'une activité physique
- Le nombre de cigarettes fumées par l'individu par jour

Combinaison des deux méthodes

Enfin, une dernière hypothèse a été étudiée : la combinaison des deux méthodes précédentes pour voir si cela permet d'améliorer la qualité du redressement. Les calages sur marges les plus efficaces (c'est-à-dire ceux qui ont donné une meilleur qualité de redressement, sans présenter une trop grande hétérogénéité des poids) ont alors été réalisés à nouveau mais en utilisant cette fois comme poids initial, non pas un poids identique égal à 1 pour chaque individu, mais l'inverse du score de propension.

3 RESULTATS

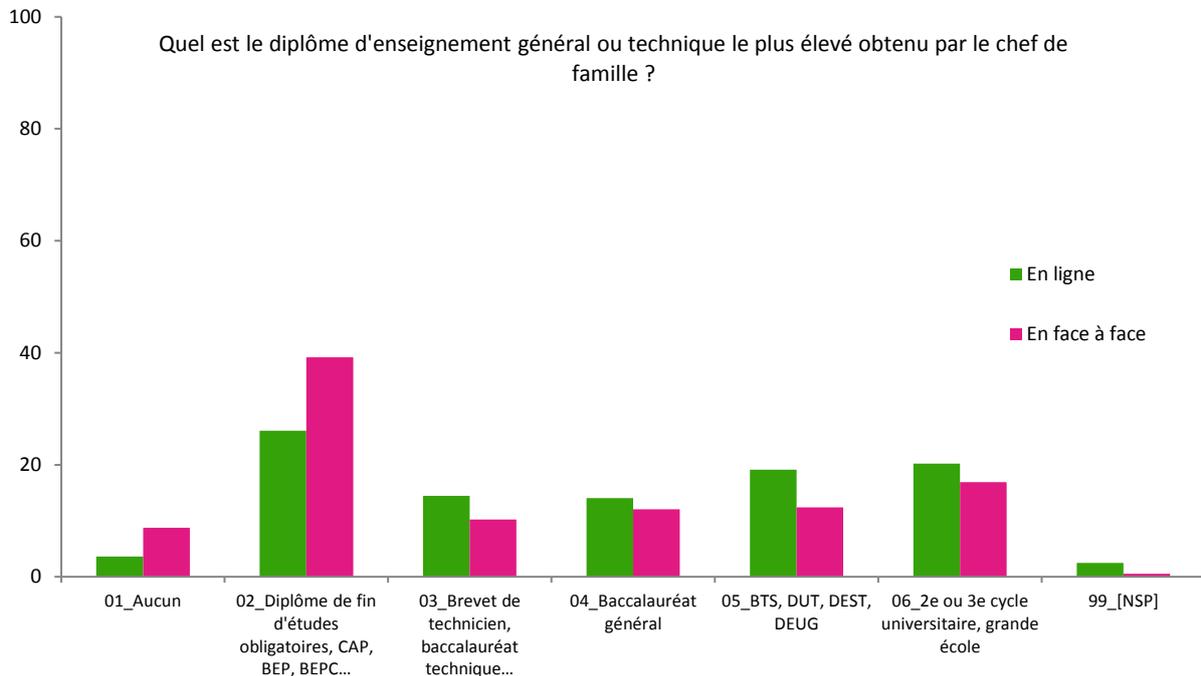
Des différences observées sur les questions auxiliaires

Des différences sociodémographiques entre les deux échantillons malgré les quotas

Une fois redressés suivant les variables de quotas, les échantillons obtenus en ligne et en face-à-face ont les mêmes caractéristiques vis-à-vis de ces variables, et devraient donc avoir des caractéristiques proches sur les autres paramètres sociodémographiques. Quelques différences ressortent cependant sur des variables sociodémographiques autres que les variables de quotas.

Pour l'échantillon des adultes, les différences entre les modes de collecte sont peu importantes, avec seulement 7 modalités de réponse sur 124 présentant plus de 5 points de différence entre les modes de collecte. Cette différence se situe sur le niveau de diplôme du chef de famille, avec davantage de niveaux de diplômes plus élevés chez les panélistes et davantage de niveaux de diplôme plus faibles dans l'échantillon en face-à-face (Graphique 1). Les panélistes vivent également un peu moins en zone urbaine ou périurbaine que les personnes interrogées en face-à-face (-6 points).

Graphique 1 : Niveau de diplôme du chef de famille selon le mode de collecte pour le questionnaire ADULTES



Source : CREDOC, enquête CCAF 2016

Concernant l'échantillon des ménages, les différences entre les deux modes de collecte sont un peu plus importantes. La même tendance peut être observée dans les niveaux de diplôme du chef de famille, et les différences entre les habitations rurales ou urbaines sont accentuées, avec 57% des panélistes qui vivent en zone urbaine ou périurbaine contre 67% des interrogés en face-à-face. On

retrouve également plus de logements sous forme de maisons individuelles en face-à-face et plus d'appartements en ligne (+3 points).

Enfin, une différence très marquée est à noter concernant le sexe du responsable des achats en matière alimentaire et donc la personne ayant répondu au questionnaire des ménages. En face-à-face, 76% des responsables des achats en matière alimentaire sont des femmes, contre seulement 60% en ligne.

Des différences sur les autres questions auxiliaires

Au-delà de ces quelques différences sociodémographiques, le passage en ligne du questionnaire crée des écarts sur d'autres questions auxiliaires.

Dans le cas de l'échantillon des ménages, des écarts importants s'observent lorsque les foyers sont interrogés sur l'équipement électroménager de leur cuisine. Des taux identiques de possession de frigos, fours, fours à micro-ondes, cuisinières à bois et congélateurs sont visibles (pas plus de 3 points de différence). Cependant, les panélistes ont moins souvent des cuisinières butane ou propane que les interrogés en face-à-face (-17 points), et plus souvent des plaques électriques (+8 points), des plaques à induction ou vitrocéramiques (+6 points). Ils sont également plus nombreux à posséder des lave-vaisselle (+8 points).

D'autres écarts témoignent également de comportements et d'équipements plus modernes chez les panélistes. Les panélistes sont plus nombreux à se faire livrer leurs courses alimentaires chez eux (+7 points), et à se faire livrer des repas tout prêts, le weekend (+20 points) et en semaine (+18 points). Les panélistes font également beaucoup plus souvent leurs courses en ligne (+24 points). Enfin, les panélistes ont moins tendance à manger leurs repas devant la télévision, quel que soit le repas : petit-déjeuner (-13 points), déjeuner (-9 points) et dîner (-9 points).

Dans le cas de l'échantillon des adultes, les questions auxiliaires portent sur l'anthropométrie et l'hygiène de vie des interrogés. Si aucune différence anthropométrique n'est visible entre les deux modes de collectes (tailles, poids et IMC des individus comparables), des écarts vis-à-vis de l'hygiène de vie des individus peuvent en revanche être notés. En effet, les panélistes font du sport un peu plus souvent que les interrogés en face-à-face (10 points de plus à en faire au moins d'une fois par mois). Pour ce qui est des sports pratiqués, peu d'écarts s'observent entre les modes de collecte, avec un écart moyen de 1,7 point pour les différents sports. Des questions ont également été posées concernant l'image qu'avaient les individus d'eux même : *plus de panélistes se trouvent « un peu trop gros » (+13 points), alors que les IMC sont semblables, et « plutôt pas actifs » ou « pas actif du tout » (+13 points)*. Les panélistes sont également moins souvent fumeurs (-8 points), même si parmi les fumeurs, les nombres moyens de cigarettes fumées par jour sont semblables entre les deux modes de collecte.

De façon surprenante, les interrogés en face-à-face sont beaucoup plus sédentaires que les panélistes : ils passent plus d'heures par semaine à regarder la télévision (« 20h ou plus par semaine » : +16 points) et à jouer aux jeux vidéo (« plus de 5h » : +3 points). En revanche les panélistes utilisent plus souvent des ordinateurs (« 7h ou plus » : +9 points)

2. Des différences observées sur les questions d'intérêt

Au-delà de ces différences observables sur les questions auxiliaires, des différences peuvent être observées sur les questions d'intérêt des questionnaires, qui portent sur les comportements et les opinions dans le domaine alimentaire.

Des écarts dans la structure des réponses qui sont liés à la façon de remplir le questionnaire

Le premier type de différences à noter ne porte pas sur le contenu, mais sur la forme des réponses, qui s'explique par la façon de remplir le questionnaire qui est différente en ligne et en face-à-face.

Une utilisation plus fréquente de la modalité « Ne Sait Pas » par les panélistes

Les panélistes ont plus souvent recours à la modalité « Ne sait pas » (NSP) lorsqu'elle est proposée. Pour le questionnaire « ménages », en ligne, les réponses « NSP » représentent en moyenne 1,5% des réponses aux questions pour lesquelles cette proposition est disponible, contre 0,7% en face-à-face. Pour le questionnaire « adultes », elles représentent respectivement 6,6% et 3,7%.

Des différences importantes sont par ailleurs à noter dans les types de questions qui génèrent ces réponses. Dans le cas du questionnaire ADULTES, les réponses « NSP » sont les plus retrouvées pour les questions portant sur les produits issus d'Organismes Génétiquement Modifiés (OGM). Pour ces questions, le recours à cette modalité représente de 21% à 39% des réponses par question pour l'échantillon en ligne contre 12% à 23% des réponses pour l'échantillon en face-à-face, selon la question. Viennent ensuite les questions portant sur l'opinion des interrogés sur les différentes réglementations en vigueur dans le domaine alimentaire (13 à 17% pour l'échantillon en ligne contre 6 à 8% des réponses par question pour l'échantillon en face-à-face).

Dans le cas du questionnaire adressé aux ménages, le recours à la modalité « NSP » est plus faible puisque la plupart des questions qui proposent cette alternative recueillent moins de 5 points d'écart entre les modes de collecte pour cette modalité. Pour quelques questions, les panélistes ont tout de même beaucoup utilisé cette modalité. C'est le cas des questions portant sur l'achat de produits du commerce équitable (12% des réponses à cette question pour les panélistes contre 1% des réponses pour les interrogés en face-à-face), sur l'achat de produits disposant d'une Appellation d'Origine Contrôlée (AOC) (11% contre 3%) et sur la consommation de produits de fabrication artisanale (8% contre 1%).

Un test a été réalisé pour voir si, en supprimant *a posteriori* les réponses « ne sait pas », les écarts observés sur les autres modalités entre les deux modes de collecte ne pouvaient pas être réduits. Les réponses « NSP » ont alors été remplacées par des valeurs manquantes. Or, cette manipulation n'a pas permis de diminuer les écarts sur les autres modalités pour chaque question, les écarts ont même parfois été augmentés. Le recours à la modalité « NSP » ne semble donc pas toujours être indépendante de la « vraie » réponse recherchée, correspondant réellement au comportement ou à l'opinion de la personne interrogée. Pour la suite de cette étude, la modalité « NSP » a donc été conservée telle qu'elle pour toutes les questions qui la proposaient.

Une moindre utilisation des modalités extrêmes de la part des panélistes

Une autre différence est notable dans la structure des réponses entre les modes de collecte. Les panélistes ont tendance à utiliser beaucoup moins souvent les modalités extrêmes, en particulier les modalités « jamais », « pas du tout », « pas du tout d'accord », que les interrogés en face-à-face. Ils utilisent en particulier beaucoup moins les modalités situées à la fin de la liste des modalités proposées. Ce phénomène est visible pour les deux questionnaires, pour la plupart des questions ayant plusieurs modalités de réponses échelonnées. Ceci entraîne des écarts importants lorsque l'on mesure les écarts entre les deux échantillons, qui ne sont pas forcément dus à des différences de comportement ou d'opinion, simplement à la façon dont sont posées les questions (à l'oral ou par écrit).

Une hypothèse a alors été testée pour réduire cet effet du mode de collecte sur la structure des réponses. Pour cela les modalités proches ont été regroupées : les modalités, « rarement » et « jamais » ont par exemple été regroupées, ainsi que les modalités « souvent » et « parfois ». Cela a permis de diminuer les écarts sur les questions concernées (passage de 5,8 à 4,3 points d'écart pour le questionnaire des adultes, et de 6,3 à 4,4 pour celui des ménages). Cette transformation des données permet donc de limiter cet effet dû au mode de collecte, et les écarts observés en valeur absolue sur les différentes modalités des questions sont en théorie uniquement dus aux différences d'opinions ou de comportement. Dans la suite de l'étude, les chiffres évoqués sont ceux obtenus avec les modalités regroupées.

Des écarts de réponses qui reflètent des différences d'opinions et de comportements entre les deux échantillons interrogés

En plus de ces différences dans la forme des réponses, des écarts sont visibles sur le fond également, qui traduisent cette fois des différences à la fois dans les comportements et dans les opinions entre les deux échantillons étudiés.

Au total sur toutes les modalités, un écart moyen de 5,31 points entre les réponses obtenues en ligne et obtenues en face-à-face a été calculé pour le questionnaire MENAGES (Tableau 3 ci-dessous). Pour 85% des modalités, les personnes interrogées ont répondu avec moins de 10 points de différences entre les deux modes de collecte et 58% des modalités présentent des différences de moins de 5 points. Le Tableau 3 montre l'écart moyen global entre les deux modes de collectes et les écarts moyens par blocs de questions pour le questionnaire auprès des ménages, ainsi que les écarts minimums et maximums obtenus. Se trouve également le nombre de questions dont la distribution des réponses est statistiquement indépendante de la façon dont le questionnaire a été passé : en ligne ou en face-à-face (c'est-à-dire les questions présentant une p-value du test du χ^2 supérieur à 5%) sur le nombre total de questions du bloc. Pour le questionnaire des ménages, près de 37% des questions recueillent des réponses indépendantes du mode de collecte (107 questions sur 288).

Tableau 3 : Ecart moyens, minimums et maximums par blocs de questions pour le questionnaire des MENAGES et nombre de questions dont la distribution des réponses est indépendantes du mode de collecte

GRUPE DE QUESTIONS	Ecart moyen	Ecart min	Ecart max	Questions indépendantes du mode de collecte
GENERAL	5,31	0,0	26,5	107/288
Types de produits alimentaires achetés	6,0			
Produits frais non préparés	1,8	0,0	5,4	3/7
Produits prêts à l'emploi	3,8	0,0	14,2	3/7
Produits bio	3,6	0,0	13,5	8/14
Produits régionaux	6,8	1,2	14,6	1/8
Produits AOC	8,8	1,7	24,1	1/8
Produits issus du commerce équitable	12,7	3,0	24,9	0/2
Consommation de cuisine exotique	10,9	0,5	25,1	3/14
Raisons d'achat de produits alimentaires	4,5			
Prix	3,1	1,8	5,1	2/3
Mode de production	4,3	0,6	6,5	7/14
Sécurité et effets sur la santé	5,3	0,2	7,9	3/8
Saveur	10,6	5,6	20,9	0/4
Aspect du produit et informations en magasin	5,3	0,9	20,0	3/14
Imaginaire derrière les produits	2,2	0,3	4,2	4/5
Autre/Ne sait pas	1,8	0,3	4,7	5/9
Acceptation de payer plus cher pour des signes officiels de qualité	4,6	0,8	11,9	6/12
Lecture des étiquetages alimentaires	6,4			
Lecture d'informations autre que le prix et la DLC	5,9	0,1	26,5	5/14
Lecture des DLC	7,2	0,9	17,3	2/10
Préparation des repas	5,8			
Organisation pour la préparation des repas	7,1	0,8	16,9	1/10
Intégration des exigences de santé dans la préparation	5,1	0,3	15,8	9/23
Caractéristiques des repas	5,0			
Structure des repas	5,1	0,1	15,8	1/9
Invités, apéritifs	4,7	0,2	12,7	3/6
Approvisionnement en produits alimentaires	4,3			
Circuits de distribution fréquentés	2,1	0,1	15,1	5/8
Enseignes de grande distribution fréquentées	5,0	0,2	22,3	1/9
Organisation pour faire les courses	8,7	0,1	22,7	0/4
Situation du foyer vis-à-vis de la sécurité alimentaire	6,4	0,4	20,7	3/10
Attitudes vis-à-vis du gaspillage alimentaire	5,2			
Types de produits jetés ou congelés	4,2	0,1	18,7	1/9
Produits jetés lorsqu'ils sont abimés, ont un mauvais goût, une mauvaise odeur	9,1	3,3	18,7	4/7
Produits jetés car le ménage n'a pas de congélateur (ou il est trop petit)	4,6	0,4	14,8	3/7
Produits jetés lorsque la DLC est dépassée	6,6	1,5	11,5	3/5
Produits jetés lorsqu'ils ne sont pas bons	4,0	0,4	7,9	4/7
Produits jetés lorsqu'ils sont achetés en trop grande quantité	6,9	2,0	11,8	4/7
Produits jetés lorsque les individus ne savent pas comment consommer les restes	4,5	0,7	11,5	5/7
Produits jetés pour d'autres raisons	2,3	0,0	5,6	4/7

Source : CRÉDOC, enquête CCAF 2016

Pour le questionnaire ADULTES, le Tableau 4 ci-dessous montre une différence moyenne de 4,32 points par modalité sur tout le questionnaire, avec 65% de modalités à moins de 5 points de différences, et 88% avec moins de 10 points. Mais ici seulement 28% des questions recueillent des réponses indépendantes du mode de collecte (64 questions sur 229).

Tableau 4 : Ecart moyens, minimums et maximums par blocs de questions pour le questionnaire des ADULTES et nombre de questions dont la distribution des réponses est indépendante du mode de collecte

GROUPE DE QUESTIONS	Ecart moyen	Ecart min	Ecart max	Questions indépendantes du mode de collecte
GENERAL	4,32	0,0	25,7	64/230
Structure des prises alimentaires	3,3			
Rythme des repas	3,2	0,0	12,4	2/7
Lieux des repas	2,6	0,0	11,7	7/17
Circonstances de consommations hors repas	4,0	0,1	12,0	9/25
Raisons d'achat et critères de qualités	5,0			
Raisons d'achat	4,9	0,0	16,1	1/16
Critères de qualité des produits alimentaires	5,4	0,1	21,3	5/11
Opinion sur les liens entre alimentation et plaisir	3,8			
Gout pour la cuisine et ouverture culinaire	4,3	0,3	14,7	0/11
Gouts préférés	4,4	1,7	7,3	1/5
Importance du plaisir et de la convivialité pour bien man-	2,6	0,2	6,5	4/10
Opinion sur les liens entre alimentation et santé	4,2			
Eléments importants dans la prévention en matière de santé	3,3	0,3	10,4	3/7
Importance de la santé et du mode de production pour bien manger	4,5	0,1	18,2	5/17
Attentes de l'alimentation pour être en bonne santé	4,4	0,1	14,7	6/11
Attitudes vis-à-vis des liens entre alimentation et	4,9			
Aliments-santé consommés	7,6	0,8	11,1	0/5
Eléments surveillés dans l'alimentation	3,0	0,2	7,6	4/9
Régimes suivis	5,0	0,2	15,0	3/13
Perception des risques alimentaires	3,5			
Types de risques alimentaires	2,7	0,1	10,3	9/19
Types de produits alimentaires à risques	4,2	0,1	13,7	0/17
Opinion sur le système alimentaire	6,2			
Confiance envers les acteurs de l'alimentation	3,0	0,1	11,0	5/15
Opinion sur la réglementation en matière alimentaire	7,0	0,5	12,8	0/5
Opinion sur les OGM	7,6	0,2	15,5	0/8
Opinion sur l'évolution de la qualité de l'alimentation	12,2	1,1	25,7	0/2

Source : CRÉDOC, enquête CCAF 2016

Questions portant sur les habitudes et les comportements

Comportements d'achat

Concernant les comportements d'achats des ménages, le passage du questionnaire en ligne donne des résultats semblables pour ce qui est des circuits de distribution fréquentés (hypermarchés, supermarchés, superettes, hard-discount, sur internet, épiceries, marchés...). Dans le cas des circuits de grande distribution, les écarts entre les modes de collectes sont acceptables, avec 5 points de différence en moyenne dans la fréquentation des différentes enseignes (Leclerc, Carrefour, Auchan...). Mais des différences notables apparaissent à propos de l'organisation pour faire les courses.

En effet, les courses sont beaucoup moins souvent réalisées par le conjoint (-10 points) ou à plusieurs (-15 points).

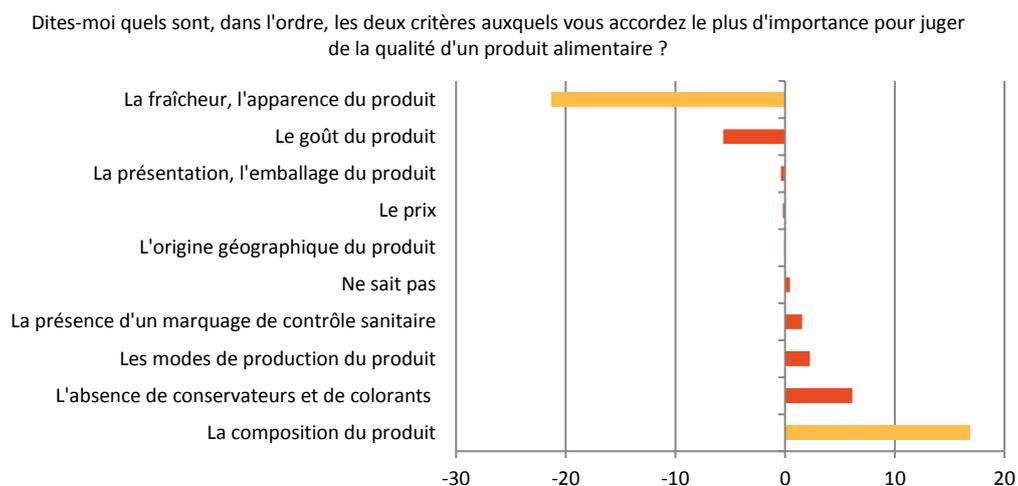
Pour ce qui est des produits achetés, le mode d'enquête donne des résultats semblables pour l'achat de produits « classiques » comme les produits frais non préparés (1,8 points de différence) ou les produits prêts-à-l'emploi (3,8 points), ou même pour les produits issus de l'agriculture biologique (3,6 points). Les panélistes achètent en revanche plus de produits régionaux (+6,8 points), de produits AOC (+8,8 points) et de produits issus du commerce équitable (+12,7 points). De la même façon, les panélistes consomment plus de types différents de cuisines exotiques que les interrogés en face-à-face (+10,9 points).

Raisons d'achat

Quel que soit le mode de collecte, les responsables des achats semblent accorder une importance semblable aux différentes raisons d'achat de produits alimentaires. Peu d'écarts s'observent en effet sur la plupart des questions portant sur les raisons d'achat, notamment le prix (3,1 points de différence), l'imaginaire associé au produit (2,2 points) et le mode de production (4,3 points). A noter tout de même que dans le cas du mode de production biologique, presque deux fois plus de panélistes que d'interrogés en face-à-face ont répondu que ce mode de production était une raison importante d'achat. Des différences assez faibles apparaissent également pour les raisons d'achat portant sur la sécurité et les conséquences sur la santé (5,3 points) et sur les informations sur le produit en magasin (2,9 points). Les panélistes et les interrogés en face-à-face ont en revanche des avis très différents lorsqu'il s'agit de considérer ou non la saveur et l'aspect des produits. Ces deux paramètres semblent en effet beaucoup moins importants pour les panélistes que pour les personnes interrogées en face-à-face (-19,8 points d'écart pour l'aspect des fruits et des légumes et -10,6 points pour la saveur de différents produits)

De la même façon, lorsque les adultes sont interrogés sur les critères importants de qualité des produits alimentaires, peu de différences apparaissent entre les modes de collecte pour la plupart des questions. Des écarts apparaissent sur seulement deux critères de qualité : la composition est beaucoup plus citée par les panélistes (+17 points), alors que l'apparence du produit est beaucoup plus importante pour les interrogés en face-à-face (+21 points) (Voir Graphique 2).

Graphique 2 : Ecart entre les réponses obtenues en ligne par rapport à la référence des réponses obtenues en face-à-face pour les questions portant sur les critères de qualité des produits alimentaires



Source : CREDOC, enquête CCAF 2016

Lecture du graphique : L'histogramme ci-dessus représente le nombre de points de différence entre les réponses recueillies en ligne par rapport aux réponses recueillies en face-à-face. Par exemple, pour la question portant sur la composition du produit, les panélistes ont plus souvent répondu que ce critère était important par rapport aux personnes interrogées en face-à-face, à hauteur de 17 points de plus. Au contraire, ils ont répondu avec 21 points de moins qu'ils trouvaient la fraîcheur et l'apparence du produit importants. Les écarts importants, supérieurs à 10 points sont représentés en rouge pour montrer que ces questions recueillent des différences marquées selon le mode de collecte

Dans ce sens, il faut aussi noter que les panélistes lisent beaucoup plus la composition sur les étiquettes des produits (+22 points), et un peu plus la teneur en matières grasses (+8 points) et en sucre (+7 points), alors que pour les autres informations présentes sur les étiquettes, les deux échantillons déclarent les lire de façon comparable.

Concernant l'importance accordée aux signes de qualités, peu de différences sont visibles sur les questions demandant si les responsables des ménages sont prêts à payer plus cher pour la présence de signes officiels de qualités (4,6 points de différence sur le bloc). Les panélistes sont cependant moins prêts à payer plus cher dans le cas de la viande de boucherie (-12 points).

Circonstances des repas et des consommations alimentaires hors repas

Pour ce qui est des rythmes et des lieux de consommations alimentaires, les deux modes de collectes donnent des résultats semblables. En effet, les blocs de question portant sur les lieux des repas (2,6 points d'écart), sur les rythmes des repas (3,2 points) et sur les circonstances de consommations hors repas sont parmi ceux qui présentent les différences les plus faibles. Apparaît juste une différence dans la fréquentation des enseignes de restauration rapide. Si autant de consommateurs réguliers sont relevés dans les deux populations, les panélistes sont plus souvent des consommateurs occasionnels (+10 points), et moins souvent des non-consommateurs (-12 points).

Préparation des repas

Pour ce qui est de la préparation des repas, les deux populations ont des habitudes différentes et les responsables alimentaires semblent un peu plus impliqués que les interrogés en face-à-face dans la préparation des repas. Les panélistes passent en effet un peu plus de temps à préparer les repas : ils sont par exemple plus nombreux à mettre plus d'une heure à préparer le repas du soir le weekend (+15 points) ou en semaine (+15 points). De plus, les panélistes font un peu plus de pâtisserie (+10 points).

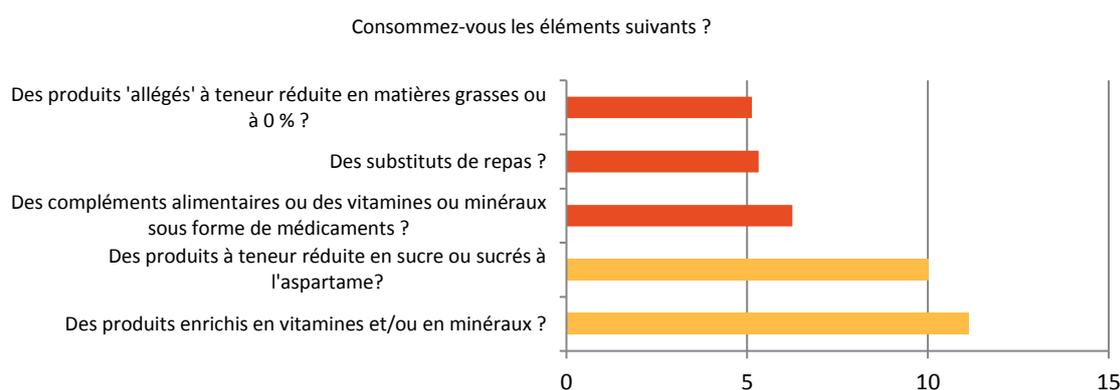
On note également quelques différences dans les comportements relatifs au gaspillage alimentaire. Alors que les deux populations présentent peu de différences vis-à-vis des produits qu'ils jettent (2,1 points d'écart), les panélistes ont beaucoup plus facilement tendance à congeler des produits alimentaires (+15,5 points).

Comportements en lien avec la santé

Les questions abordant les comportements alimentaires en lien avec la santé donnent des résultats globalement semblables avec les deux modes de collecte. Les individus surveillent les différents éléments de leur alimentation de façon comparable (vitamines, calcium, protéines, sucres lents, magnésium, fer, sucre), même si les panélistes font un peu plus attention aux fibres (+8 points). Les panélistes sont cependant plus nombreux à suivre des régimes alimentaires (+15 points), et davantage pour rester en forme (+10 points) et pour maigrir (+8 points) que les interrogés en face-à-face. Ils sont également plus nombreux à avoir diminué leur consommation de viande en 2015 (+12 points). Sinon, les deux populations suivent des régimes végétariens, sans lactose et sans gluten dans des proportions similaires.

Des écarts plus importants apparaissent aussi à propos de la consommation d'aliments-santé (voir Graphique 3). En effet, les panélistes consomment plus de produits allégés en sucre (+10 points) et en matières grasses (+5 points), plus de produits enrichis en vitamines ou en minéraux (+11 points), de compléments alimentaires (+10 points) et de substituts de repas (+5 points).

Graphique 3 : Ecarts entre les réponses obtenues en ligne par rapport à la référence des réponses obtenues en face-à-face pour les questions portant sur les « aliments-santé » consommés.



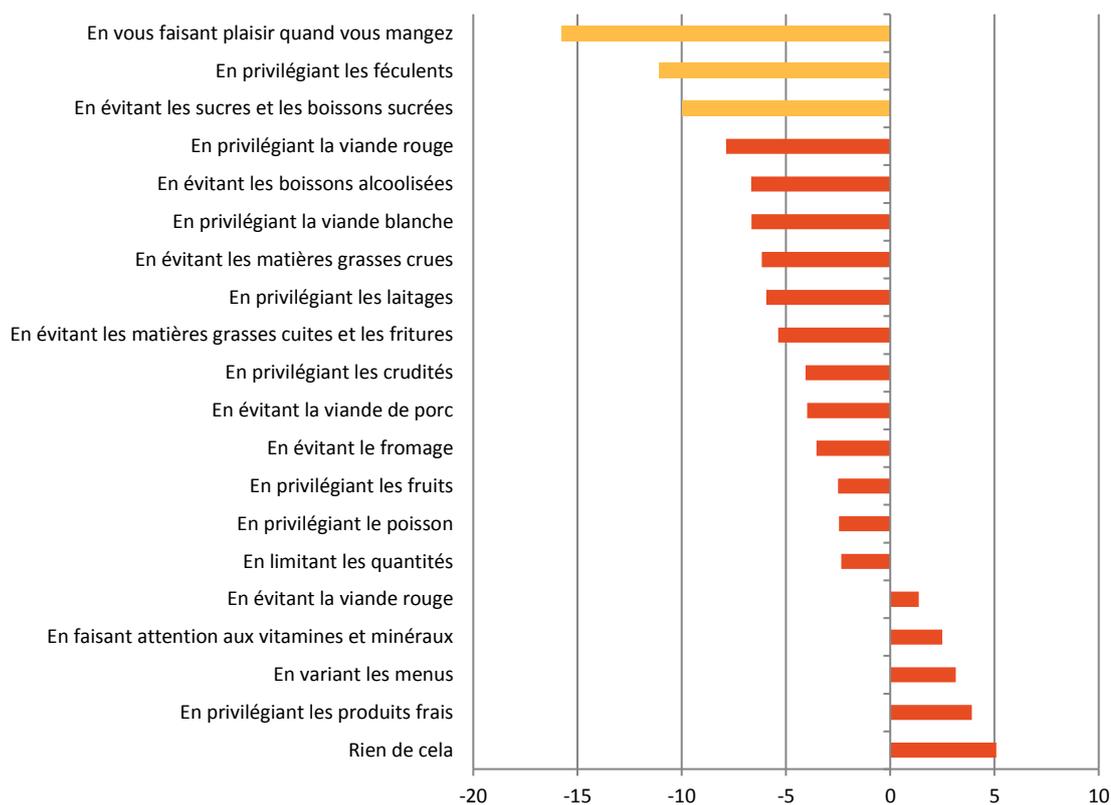
Source : CREDOC, enquête CCAF 2016

En revanche, peu de différences s'observent dans la manière qu'ont les responsables des ménages de tenir compte des exigences de santé dans leur alimentation et celle de leur foyer (voir Graphique 4). Les écarts sont en effets faibles pour la plupart des attitudes citées dans le questionnaire : « en faisant manger de tout, en équilibrant les repas », « en privilégiant les légumes » ou « les fruits », « en évitant les matières grasses cuites » ou « crues », « en privilégiant les produits frais » ... etc.

Sur ces 23 questions, seules 4 présentent des écarts notables. Les interrogés en ligne ont en effet moins souvent répondu « oui » aux affirmations suivantes : « en évitant les sucres » (-10 points), « en privilégiant les féculents » (-11 points), et surtout beaucoup moins souvent répondu « en vous faisant plaisir » (-16 points) que les interrogés en face-à-face.

Graphique 4 : Ecarts entre les réponses obtenues en ligne par rapport à la référence des réponses obtenues en face-à-face pour les questions portant sur les exigences de santé liées à l'alimentation.

Comment tenez-vous compte des exigences de santé dans votre alimentation et celle de votre foyer ?



Source : CREDOC, enquête CCAF 2016

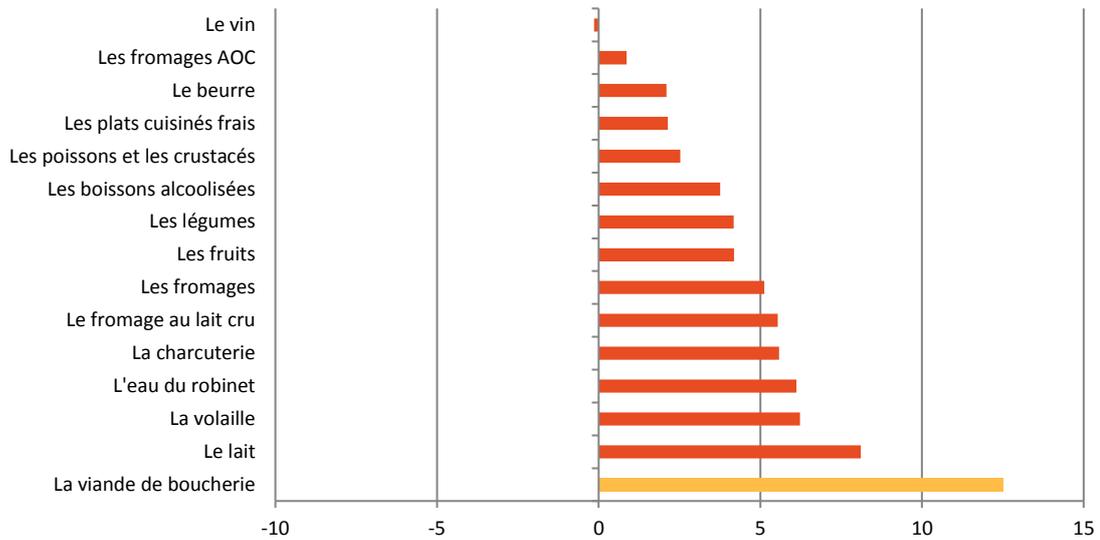
Questions portant sur les opinions et les représentations

Perception des risques (adultes)

Peu de différences sont visibles entre les deux modes de collecte vis-à-vis des produits considérés comme présentant des risques pour la santé (4,2 points de différence), sauf dans le cas de la viande de boucherie (Graphique 5). Les panélistes la considèrent en effet beaucoup plus comme un produit à risques important (+13 points).

Graphique 5 : Ecarts entre les réponses obtenues en ligne par rapport à la référence des réponses obtenues en face-à-face pour les questions portant sur les produits perçus comme à risque.

Selon vous les produits suivants présentent-ils aujourd'hui des risques pour la santé ?

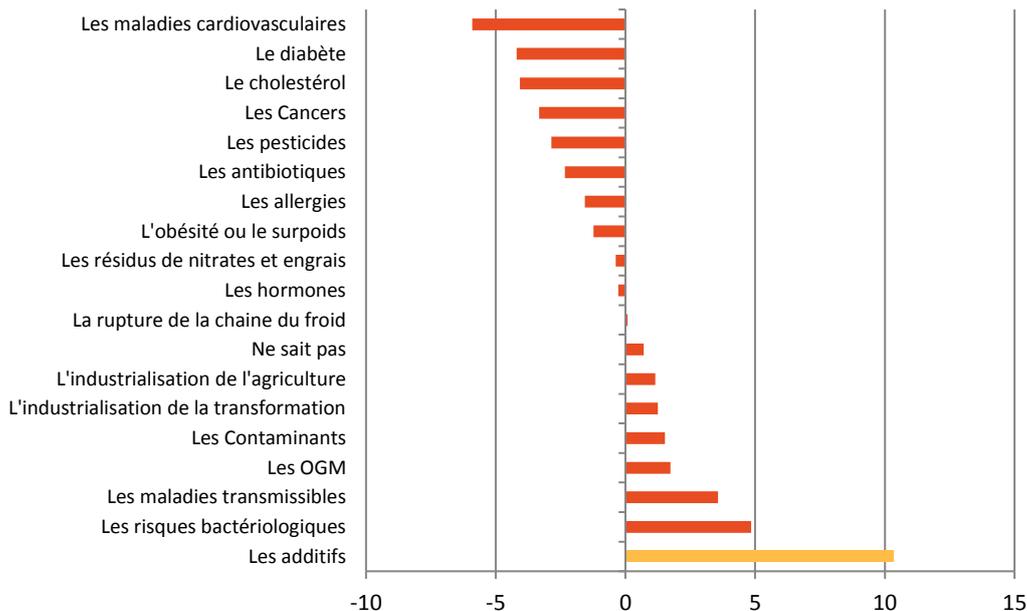


Source : CREDOC, enquête CCAF 2016

Concernant les types de risques qui inquiètent le plus les populations, peu de différences se détachent (2,7 points de différence en moyenne sur le bloc). On note cependant une caractéristique particulière des panélistes : ils sont beaucoup plus inquiets à propos des risques liés aux additifs (+10 points) (voir Graphique 6 ci-dessous).

Graphique 6 : Ecarts entre les réponses obtenues en ligne par rapport à la référence des réponses obtenues en face-à-face pour les questions portant sur les types de risques

Parmi les risques éventuels liés à l'alimentation, quel est celui qui vous inquiète le plus ?



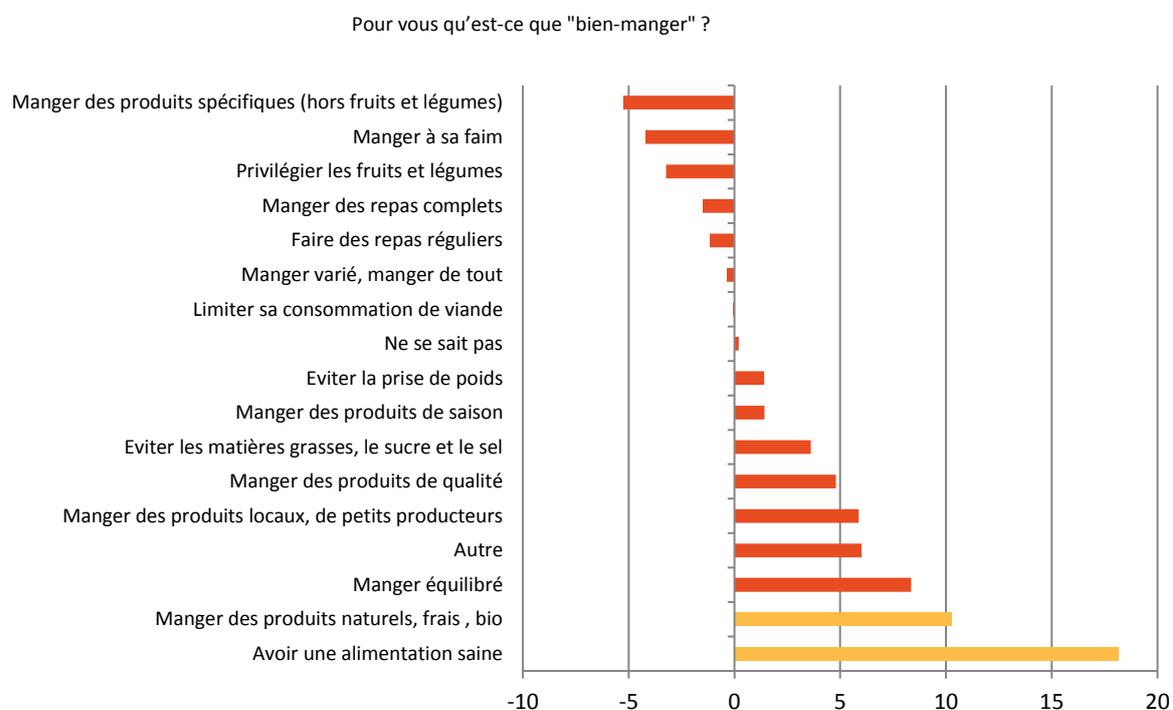
Source : CREDOC, enquête CCAF 2016

Opinion sur les liens entre alimentation et santé

La question « pour vous, qu'est-ce que bien manger » est posée dans le questionnaire des adultes. Cette question est une question ouverte ; et les réponses ont été codées *a posteriori* sous différentes catégories. Les réponses ont été séparées en deux blocs : les réponses de types « santé et mode de production » et les réponses de type « plaisir et convivialité » (qui sont abordées dans la sous-partie suivante).

Pour les réponses de type « santé et mode de production », les avis sont assez comparables entre les deux populations pour la plupart des réponses. Cependant, les panélistes accordent beaucoup plus d'importance au fait d'avoir une alimentation saine que les interrogés en face-à-face (+18 points) et au fait de manger des produits bio, naturels et frais (+10 points) et au fait de manger équilibré (+8 points) (Graphique 7 ci-dessous).

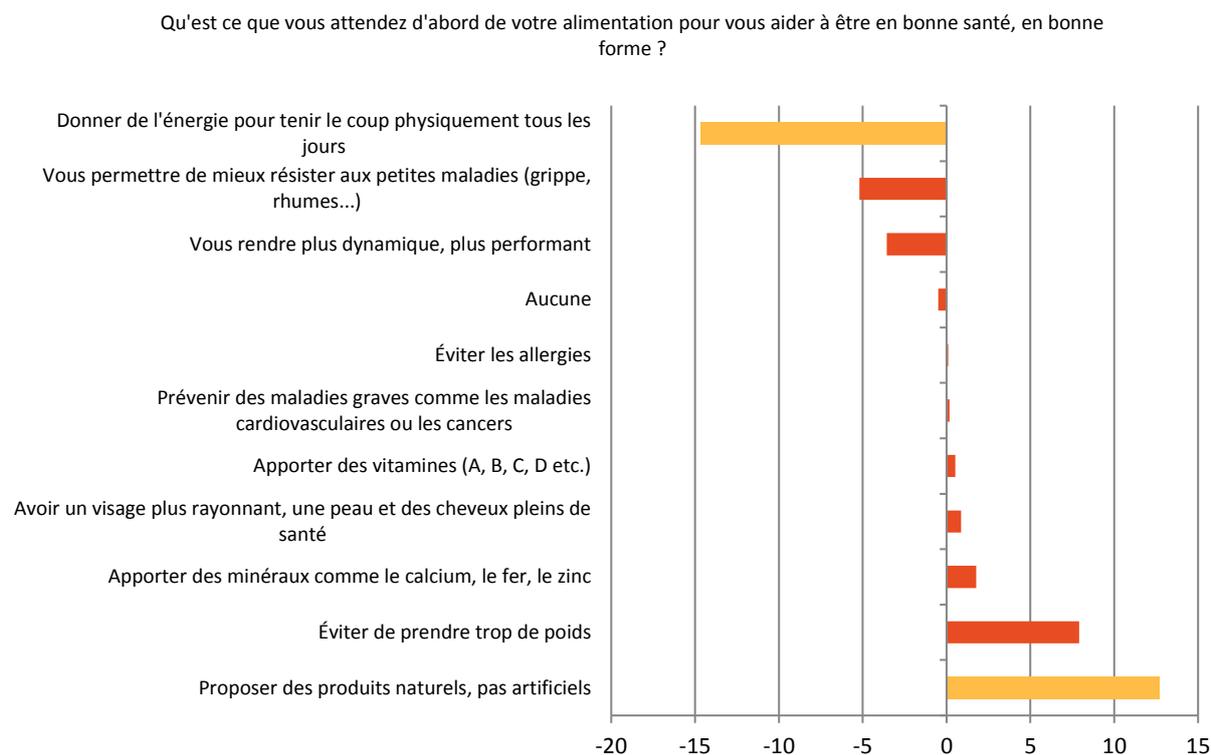
Graphique 7 : Ecart entre les réponses obtenues en ligne par rapport à la référence des réponses obtenues en face-à-face pour les réponses de type « santé et mode de production » à la question « Pour vous, qu'est-ce que « bien manger » ? »



Source : CREDOC, enquête CCAF 2016

De façon plus ciblée, la question fermée « Qu'est-ce que vous attendez d'abord de votre alimentation pour vous aider à être en bonne santé, en bonne forme ? » est posée dans ce questionnaire. Ici aussi, les réponses sont semblables pour la plupart mais quelques réponses présentent des écarts importants (voir Graphique 8 ci-dessous). Une fois de plus, les panélistes témoignent d'une plus grande attente vis-à-vis des produits naturels (+13 points) et attendent davantage que leur alimentation leur permette d'éviter la prise de poids (+8 points). Au contraire, les interrogés en face-à-face attendent, beaucoup plus que les panélistes, que leur alimentation leur fournisse l'énergie nécessaire pour tenir le coup physiquement (+15 points).

Graphique 8 : Ecarts entre les réponses obtenues en ligne par rapport à la référence des réponses obtenues en face-à-face pour les questions portant sur les attentes de l'alimentation en matière de santé.



Source : CREDOC, enquête CCAF 2016

Opinion sur les liens alimentation et plaisir

Peu de différences sont visibles à propos des représentations des liens entre alimentation et plaisir entre les deux échantillons. Le bloc regroupant les réponses de type « hédonisme et convivialité » à la question « Pour vous, qu'est-ce que « bien manger » ? » présente très peu de différences entre les deux modes de collectes (2,6 points). Quant aux questions sur l'hédonisme et l'ouverture culinaire en général, elles obtiennent également des réponses semblables pour les deux échantillons. Cependant, les panélistes semblent plus apprécier « cuisiner au quotidien » (+10 points) et « changer, essayer de nouveaux produits, de nouvelles marques » (+10 points).

Opinion sur le système alimentaire

Dans le questionnaire réservé aux adultes, plusieurs questions sont posées dans le but d'étudier l'opinion des Français sur le système alimentaire. Les questions sur la confiance accordées par les individus aux différents acteurs de l'alimentation fournissent des réponses semblables avec les deux modes de collecte (3,0 points de différence pour le bloc).

En revanche, les opinions sur la réglementation et les contrôles en vigueur en matière alimentaire divergent clairement entre les deux échantillons. Les interrogés en face-à-face sont en effet plus nombreux que les panélistes à trouver « plutôt suffisant[s] » les réglementations et les contrôles mis en place dans le domaine de l'alimentation. C'est le cas pour la réglementation française en matière

alimentaire (+13 points), la réglementation européenne en matière alimentaire (+9 points), les contrôles effectués en France sur la fabrication des produits alimentaires (+12 points), les contrôles effectués sur les produits alimentaires importés de l'étranger (+5 points) et les contrôles effectués en France sur les lieux de vente des produits alimentaires (+9 points). Cependant, ces différences doivent être interprétées avec précaution puisque pour ces questions, la modalité « NSP » a été très utilisée par les personnes interrogées, et deux fois plus utilisée par les panélistes que par les personnes ayant répondu en face-à-face.

Les adultes sont également interrogés sur leur opinion à propos des produits alimentaires issus d'Organismes Génétiquement Modifiés (OGM). Les panélistes sont plus nombreux à répondre qu'ils trouvent les produits contenant des OGM de plus en plus présents dans les aliments (-6 points). Ils sont moins nombreux à répondre qu'ils les trouvent dangereux (-7 points), mais sont aussi moins nombreux à leur trouver des points positifs. Cependant, il faut ici aussi tenir compte des très nombreuses réponses « ne sait pas » : cette modalité atteint en effet jusqu'à 39% des réponses pour ces questions.

Concernant l'avis sur l'évolution de l'alimentation, de très grandes différences apparaissent. Les panélistes sont beaucoup plus nombreux à trouver que la qualité et le goût des aliments se sont détériorés ces dix dernières années (+17 points pour le goût et +19 points pour la qualité) et moins nombreux à trouver qu'elle n'a pas changé (-23 points pour le goût et -26 points pour la qualité).

Inquiétudes vis-à-vis de la sécurité alimentaire

Une dernière différence est à noter entre les deux échantillons à propos des inquiétudes des ménages vis-à-vis de la sécurité alimentaire. Les panélistes sont plus nombreux à répondre qu'il leur arrive « souvent » ou « parfois » d'être inquiets à l'idée de manquer d'aliments (+8 points) et plus nombreux à répondre qu'il leur arrive « souvent » ou « parfois » de ne pas pouvoir manger du poisson, de la volaille ou de la viande tous les deux jours par manque de moyens financiers (+12 points). Ils sont aussi moins nombreux à répondre qu'ils peuvent manger tous les aliments qu'ils veulent (+21 points).

Des écarts peu visibles sur les questions sujettes à un biais de désirabilité sociale

Pour ce qui est des questions présentant un possible biais de désirabilité sociale, aucune tendance claire ne se dégage qui permettrait de conclure que le biais lié à l'absence ou la présence de l'enquêteur est important sur la totalité des questions.

Dans le cas du questionnaire des ménages, peu de différences sont observables entre les deux modes de collecte à propos des produits jetés. Pour ce qui est des inquiétudes en revanche, de grandes différences sont visibles, avec des panélistes qui déclarent être plus inquiets vis-à-vis de la sécurité alimentaire, peut-être parce qu'ils n'avaient pas honte de le déclarer, seuls devant leur ordinateur. La présence de l'enquêteur a donc peut-être introduit un biais de désirabilité sociale à ce niveau dans le questionnaire adressé aux ménages.

Pour le questionnaire des adultes : les panélistes ont tendance à se déclarer un peu moins actifs, plus consommateurs de restauration rapide, et plus utilisateurs de distributeurs d'aliments, ce qui

pourrait témoigner d'un biais de désirabilité sociale. Cependant, les panélistes se déclarent aussi moins fumeurs, moins sédentaires, et plus attentifs à leur poids et leur alimentation.

Il semble donc que pour le questionnaire « Comportements et Attitudes » de l'enquête CCAF, le biais de désirabilité sociale constitué par la présence de l'enquêteur soit peu présent, sauf peut-être pour les questions se reportant aux inquiétudes vis-à-vis de la sécurité alimentaire.

3. Effets des redressements de l'échantillon des panélistes

Suite à cette étude des divergences de réponses aux questions d'intérêt observées sur les deux questionnaires (adultes et ménages) selon le mode de collecte (en ligne ou en face-à-face), cette partie a pour but d'étudier les effets des redressements testés pour redresser l'échantillon recueilli en ligne par rapport à celui recueilli en face-à-face.

Redressements par calages sur marges

Echantillon des ménages : Pour le questionnaire adressé aux ménages, les différentes variables entrées successivement dans le calage sur marge ont un impact sur l'indicateur global des différences entre les deux échantillons. En effet, ces hypothèses permettent de passer d'un écart moyen du nombre de points entre les pourcentages de réponses aux différentes modalités de 5,31 à 4,80 dans le meilleur des cas, et à 5,68 dans le pire des cas (Tableau 5-A ci-dessous.)

- Les hypothèses 1 et 4 augmentent l'écart global entre les deux échantillons. En redressant l'échantillon avec le niveau de diplôme du chef de famille, l'écart global passe de 5,31 à 5,42. En le redressant avec l'équipement en électroménager du foyer, l'écart augmente également : de 5,31 à 5,68 (Tableau 5-A.). Les différences de niveau de diplôme et de modernité de l'équipement de la cuisine des ménages des panélistes ne semblent donc pas être liées aux différences de comportements alimentaires de ces panélistes par rapports aux comportements des personnes interrogés en face-à-face.
- En revanche, les autres hypothèses réduisent les écarts de points, les hypothèses 3, 5 et 6 permettant même de réduire les écarts pour tous les blocs de questions sauf un. Le fait de regarder la télévision pendant les repas, d'utiliser internet pour faire ses courses et de faire appel aux services de livraison à domicile semblent donc être des comportements qui expliquent en partie les différences de comportements alimentaires observées chez les panélistes (Tableau 5-A.).

Tableau 5 : Ecart moyen global et écarts par blocs entre les deux modes de collecte obtenus avec les différentes hypothèses de redressement pour l'échantillon des MENAGES

	A. Redressement par calages sur marges avec variables supplémentaires						
	Hypothèse 0 : Quotas	Hypothèse 1 : Niveau de diplôme du chef de famille	Hypothèse 2 : Sexe du responsable des achats	Hypothèse 3 : Prise des repas devant la télé	Hypothèse 4 : Equipement de la cuisine	Hypothèse 5 : Utilisation d'internet en alimentaire	Hypothèse 6 : Utilisation des services de livraison à domicile
Approvisionnement en produits alimentaires	4,3	4,0	4,7	4,1	4,5	4,2	4,1
Raisons d'achat de produits alimentaires	4,5	5,2	4,7	4,4	4,9	4,4	4,4
Caractéristiques des repas	5,0	4,5	4,9	5,2	5,4	4,6	4,9
Attitudes vis-à-vis du gaspillage alimentaire	5,2	5,0	5,0	5,0	5,2	4,8	4,8
Préparation des repas	5,8	6,3	5,2	5,7	6,2	5,3	5,8
Types de produits alimentaires achetés	6,0	6,0	6,1	5,7	6,6	5,5	5,4
Situation du foyer vis-à-vis de la sécurité alimentaire	6,4	5,9	6,1	6,1	6,4	7,0	6,4
Lecture des étiquetages alimentaires	6,4	6,5	6,2	6,2	6,9	6,0	6,3
GENERAL	5,31	5,42	5,28	5,16	5,68	5,00	5,04

	B. Pondération par l'inverse du score de propension	C. Combinaison des deux méthodes
	Hypothèse 7 : 1/ps	Hypothèse 8 : 1/ps et Quotas
Approvisionnement en produits alimentaires	4,2	4,4
Raisons d'achat de produits alimentaires	3,5	4,6
Caractéristiques des repas	5,3	4,6
Attitudes vis-à-vis du gaspillage alimentaire	4,6	4,8
Préparation des repas	5,0	4,9
Types de produits alimentaires achetés	5,6	5,4
Situation du foyer vis-à-vis de la sécurité alimentaire	6,0	6,5
Lecture des étiquetages alimentaires	6,0	6,4
GENERAL	4,80	5,03

 diminution de l'écart par rapport à l'hypothèse 0

Source : CREDOC, enquête CCAF 2016

Lecture du tableau : Le tableau ci-dessous résume les écarts moyens obtenus entre les deux modes de collecte par blocs de question, et globalement sur l'ensemble du questionnaire. On voit par exemple que le redressement correspondant au redressement classique auquel le niveau de diplôme du chef de famille a été ajouté (hypothèse 1) augmente l'écart moyen global sur l'ensemble du questionnaire en le faisant passer de 5,31 à 5,42 points. En revanche, l'écart moyen sur le bloc regroupant les questions relatives à l'approvisionnement en produits alimentaire est diminué grâce à cette hypothèse puisqu'il passe de 4,3 à 4,0.

Cependant, toutes ces hypothèses doivent être rejetées car pour chaque hypothèse, pour caler l'échantillon en ligne par rapport à celui en face-à-face sur les nouvelles variables, le modèle attribue aux individus des poids présentant une répartition très importante, avec des poids minimum/maximum largement supérieurs à 200 (Tableau 6-A.). Cela signifie qu'il attribue des poids très importants à certains individus surreprésentés, alors que certains ne sont quasiment pas pris en compte. Ces hypothèses doivent donc être rejetées.

Comme pour l'échantillon des adultes, une hypothèse a été testée pour tenter d'ajouter simultanément toutes les variables testées séparément. La répartition des poids était encore plus élevées (min/max supérieur à 100 000 000), les résultats ne sont donc pas présentés.

Tableau 6 : Caractéristiques des poids attribués aux individus lors des différentes hypothèses de redressement pour l'échantillon des MENAGES

		Poids initial	Calage sur les marges suivantes		Caractéristiques des poids obtenus				Ecart moyen	
					Min	Max	Max/Min	Médiane		
Echantillon en face à face	Redressement classique	1	Variables de quotas classiques : Catégorie d'agglomération - Type de ménage - Age de la personne de référence - Nombre de personnes dans le foyer - ZEAT - PCS du chef de famille	-	0,11	6,48	58,43	0,73	-	
Echantillon en ligne	Hypothèse 0 : Redressement classique	1	Variables de quotas classiques	-	0,01	34,43	6805,12	0,41	5,32	
	A. Redressements par calages sur marges avec des variables supplémentaires	Hypothèse 1	1	Variables de quotas classiques	+ Niveau de diplôme du chef de famille	0,00	62,33	63634,80	0,30	5,43
		Hypothèse 2	1	Variables de quotas classiques	+ Sexe du responsable des achats alimentaires	0,00	41,60	8933,08	0,40	5,30
		Hypothèse 3	1	Variables de quotas classiques	+ Fait que les membres du foyer mangent devant la télé	0,00	29,48	8199,00	0,40	5,17
		Hypothèse 4	1	Variables de quotas classiques	+ Equipement de la cuisine	0,00	54,56	14301,97	0,40	5,69
		Hypothèse 5	1	Variables de quotas classiques	+ Recours aux commandes sur internet pour les courses ou les repas	0,00	38,46	274854,04	0,33	5,02
		Hypothèse 6	1	Variables de quotas classiques	+ Recours au services de livraison à domicile	0,00	37,04	39774,02	0,39	5,05
	B. Pondération par l'inverse du score de propension	Hypothèse 7	1 / ps	-	-	0,51	10,02	19,58	0,84	4,81
	C. Combinaison des deux méthodes	Hypothèse 8	1 / ps	Variables de quotas classiques	-	0,00	43,31	41058,12	0,37	5,04

Source : CREDOC, enquête CCAF 2016

Echantillon des adultes : Pour l'échantillon des adultes, le redressement par calage sur marges avec des variables supplémentaires permet également de diminuer les écarts entre les modes de collecte. Ces hypothèses permettent de passer d'un écart moyen de 4,32 à 4,07 pour les plus efficaces (Tableau 7 ci-dessous).

Tableau 7 : Ecart moyen global et écarts par blocs entre les deux modes de collecte obtenus avec les différentes hypothèses de redressement pour l'échantillon des ADULTES

	A. Redressement par calages sur marges avec variables supplémentaires							
	Hypothèse 0 : Quotas	Hypothèse 1 : Niveau de diplôme de l'individu	Hypothèse 2 : Situation de l'habitation	Hypothèse 3 : Fréquence de pratique d'une activité physique	Hypothèse 4 : Nombre de cigarettes fumées par jour	Hypothèse 5 : Situation professionnelle et maritale	Hypothèse 6 : Temps passé devant les écrans par semaine	Hypothèse 7 : Ensemble des variables testées
Structure des prises alimentaires	3,3	3,2	3,4	3,3	3,4	3,2	3,0	3,1
Perception des risques alimentaires	3,5	3,7	3,5	3,7	3,5	3,6	3,2	3,7
Opinion sur les liens entre alimentation et plaisir	3,8	3,8	3,7	3,9	3,7	3,7	3,4	3,6
Opinion sur les liens entre alimentation et santé	4,2	4,2	4,3	4,1	4,3	4,3	4,0	4,2
Attitudes vis-à-vis des liens entre alimentation et santé	4,9	4,5	4,9	4,5	4,8	4,8	4,7	4,8
Raisons d'achat et critères de qualités	5,0	5,0	5,0	5,0	5,0	5,1	4,9	5,0
Opinion sur le système alimentaire	6,2	6,4	6,3	6,2	6,3	6,1	6,1	6,4
GENERAL	4,32	4,29	4,35	4,30	4,33	4,28	4,07	4,30

	B. Pondération par l'inverse du score de propension		C. Combinaison des deux méthodes	
	Hypothèse 8 : 1/ps	Hypothèse 9 : 1/ps et quotas	Hypothèse 10 : 1/ps quotas et temps passé devant les écrans	
Structure des prises alimentaires	3,5	3,2	3,4	
Perception des risques alimentaires	3,5	3,4	3,1	
Opinion sur les liens entre alimentation et plaisir	3,8	3,7	3,4	
Opinion sur les liens entre alimentation et santé	4,2	4,2	4,0	
Attitudes vis-à-vis des liens entre alimentation et santé	4,8	4,6	4,6	
Raisons d'achat et critères de qualités	5,0	4,8	4,7	
Opinion sur le système alimentaire	6,4	6,4	6,6	
GENERAL	4,37	4,23	4,19	

 diminution de l'écart par rapport à l'hypothèse 0

Source : CREDOC, enquête CCAF 2016

- Contrairement à l'échantillon des ménages, l'hypothèse 1, qui ajoute le diplôme de l'individu dans le redressement, améliore un peu la situation puisqu'elle diminue l'écart moyen (de 4,32 à 4,29 points) et diminue les écarts pour tous les blocs sauf deux (Tableau 7-A.). Cependant, la diminution de l'écart total est assez limitée.
- De la même façon, presque toutes les autres variables testées ont un impact assez faible sur l'écart global entre les deux modes de collecte. C'est le cas pour la situation de l'habitation, la fréquence de pratique d'une activité physique, le nombre de cigarettes fumées, les situations maritales et professionnelles, et même pour la combinaison de toutes les variables testées (Tableau 7-A.).
- L'utilisation des variables relatives au temps passé devant les différents écrans a en revanche un impact notable sur l'écart global entre les deux échantillons. L'hypothèse 6 permet en effet de faire passer l'écart global de 4,32 à 4,07 en diminuant l'écart pour tous les blocs de questions (Tableau 7-A.). Cette hypothèse augmente également légèrement la proportion de modalités présentant moins de 5 points d'écart (le pourcentage passe de 66% à 70%) et moins de 10 points d'écart (le pourcentage passe de 89% à 91%). Cette hypothèse permet aussi d'augmenter le nombre de questions dont la distribution des réponses est indépendante du mode de collecte de 64 à 77 questions sur 229.

Ici, contrairement à ce qui peut être observé pour l'échantillon des adultes, ces redressements n'entraînent pas des poids trop dispersés (Tableau 8-A. ci-dessous). Pour celles qui améliorent la situation, ces hypothèses peuvent donc être validées.

Tableau 8 : Caractéristiques des poids attribués aux individus lors des différentes hypothèses de redressement pour l'échantillon des ADULTES

		Poids initial	Calage sur les marges suivantes		Caractéristiques des poids obtenus				Ecart moyen	
					Min	Max	Max/Min	Médiane		
Echantillon en face à face	Redressement classique	1	Variables de quotas classiques : Catégorie d'agglomération - Age - Sexe - PCS de l'individu - ZEAT	-	0,42	3,21	7,59	0,97	-	
Echantillon en ligne	Hypothèse 0 : Redressement classique	1	Variables de quotas classiques	-	0,48	4,76	9,91	0,94	4,32	
	A. Redressements par calages sur marges avec des variables supplémentaires	Hypothèse 1	1	Variables de quotas classiques	+ Niveau de diplôme de l'individu	0,17	7,60	45,84	0,77	4,29
		Hypothèse 2	1	Variables de quotas classiques	+ Situation de l'habitation	0,18	5,61	30,87	0,96	4,35
		Hypothèse 3	1	Variables de quotas classiques	+ fréquence à laquelle l'individu fait du sport	0,20	7,82	38,18	0,83	4,30
		Hypothèse 4	1	Variables de quotas classiques	+ Nombre de cigarettes fumées par jour	0,11	5,06	48,12	0,91	4,33
		Hypothèse 5	1	Variables de quotas classiques	+ situation professionnelle et familiale	0,21	6,09	29,13	0,93	4,28
		Hypothèse 6	1	Variables de quotas classiques	+ temps passé devant les écrans	0,10	9,09	92,09	0,82	4,07
		Hypothèse 7	1	Variables de quotas classiques	+ Ensemble des variables auxiliaires testées	0,02	16,78	846,70	0,60	4,30
	B. Pondération par l'inverse du score de propension	Hypothèse 8	1/ps	-	-	0,52	14,84	28,60	0,92	4,37
	C. Combinaison des deux méthodes	Hypothèse 9	1/ps	Variables de quotas classiques	-	0,22	9,31	41,94	0,77	4,23
Hypothèse 10		1/ps	Variables de quotas classiques	+ temps passé devant les écrans	0,11	14,33	125,69	0,68	4,19	

Source : CREDOC, enquête CCAF 2016

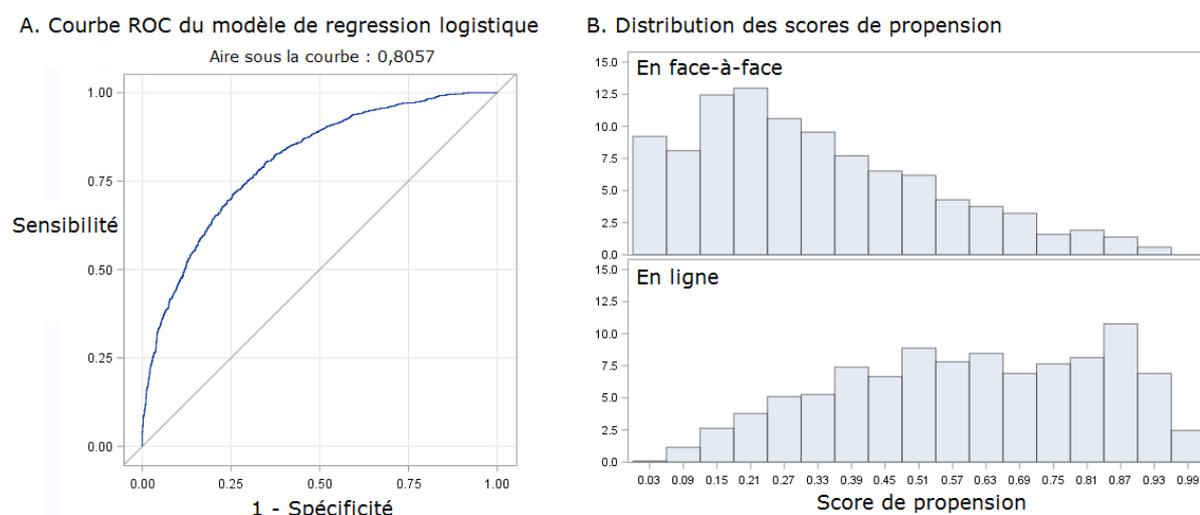
Redressements par pondération en utilisant l'inverse du score de propension

Efficacité des régressions logistiques utilisées pour l'estimation des scores de propension

Echantillon des ménages :

Pour l'échantillon des ménages, la courbe ROC du modèle de régression logistique utilisé pour estimer les scores de propension est caractérisée par une aire sous la courbe de 0,806. Cela signifie que le modèle classe correctement les individus avec une probabilité de 80,6%, il permet donc de classer assez efficacement un individu selon qu'il ait été interrogé en ligne ou en face-à-face (Graphique 9 ci-dessous). En effet il est bien visible que le modèle attribue des scores de propension plus élevés aux individus qui appartiennent en effet à l'échantillon en ligne, et des scores de propension plus faibles aux individus interrogés en face-à-face, le score de propension étant la probabilité d'appartenir à l'échantillon en ligne.

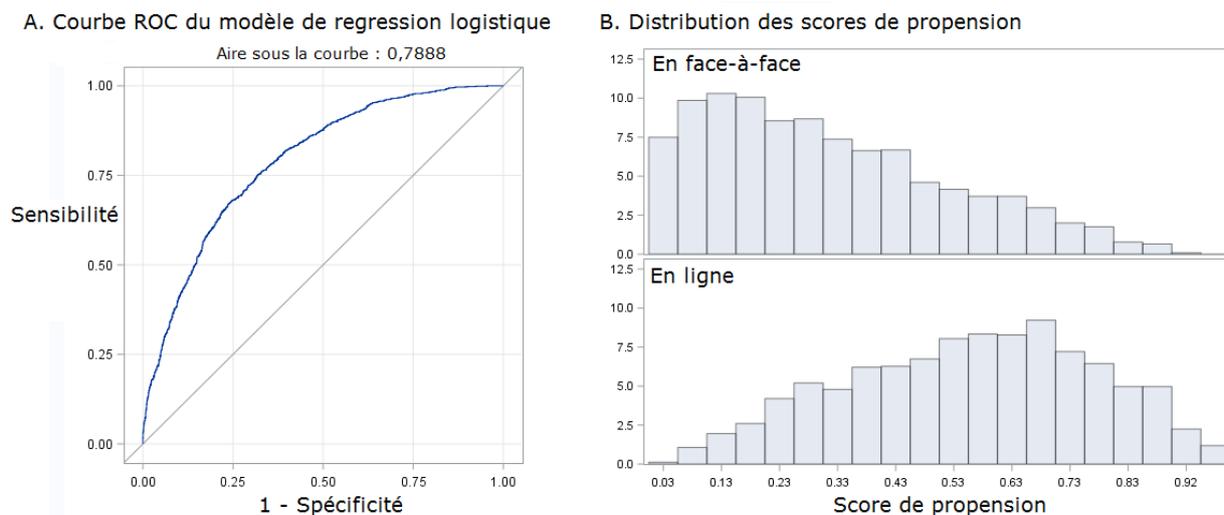
Graphique 9 : Courbe ROC du modèle de régression logistique et distribution des scores de propension obtenus pour l'échantillon des MENAGES



Echantillon des adultes :

Pour l'échantillon des ménages, le modèle de régression logistique aboutit cette fois à 78,9% d'affectation correctes. L'introduction des variables dans le modèle permet donc ici aussi de classer assez efficacement un individu selon qu'il ait été interrogé en ligne ou en face-à-face. Cependant, en observant la répartition des scores de propensions affectés aux individus de l'échantillon en ligne (Graphique 10), il est net que peu d'individus se voient affecter des scores de propension proches de 1. Le modèle ne permet donc pas de détecter très efficacement les individus qui ont des caractéristiques très « typiques » des panélistes. En pondérant l'échantillon des panélistes par l'inverse du score de propensions estimés par ce modèle, il y a un risque de ne pas pouvoir affecter des faibles poids à ces individus.

Graphique 10 : Courbe ROC du modèle de régression logistique et distribution des scores de propension obtenus pour l'échantillon des ADULTES



Source : CREDOC, enquête CCAF 2016

Efficacité des redressements par pondération avec l'inverse des scores de propension

Dans le cas de l'échantillon des ménages, la pondération par l'inverse du score de propension est la méthode qui permet le plus efficacement de diminuer les écarts observés entre les modes de collecte. L'hypothèse 7 permet en effet de faire passer l'écart moyen de 5,31 à 4,80 points, en diminuant la différence pour tous les blocs sauf un (Tableau 4-B.). Elle augmente également légèrement la proportion de modalités présentant moins de 5 points d'écart (le pourcentage passe de 58% à 64%) et moins de 10 points d'écart (le pourcentage passe de 85% à 86%). Elle augmente également le nombre de questions dont la distribution des réponses est indépendante du mode de collecte, qui passe de 107 à 121 sur 288. De plus, cette méthode est très encourageante puisque, contrairement aux méthodes de calage sur marge, elle n'entraîne pas une répartition des nouveaux poids trop dispersée (Tableau 6-B.).

Pour l'échantillon des adultes en revanche, la pondération par l'inverse des scores de propension ne permet pas d'améliorer la situation. L'hypothèse 8 augmente en effet légèrement l'écart total et ne diminue les écarts pour la moitié des blocs de questions d'intérêt (Tableau 5-B.). Ceci est probablement dû à la remarque faite précédemment : les variables introduites dans modèle de régression logistique ne permettent pas de détecter efficacement les individus qui ont des caractéristiques extrêmes des panélistes (Graphique 10-B.). En pondérant l'échantillon des panélistes par l'inverse du score de propension, il est possible que peu d'individus de ce type se soient vu affecter des poids faibles. Ce qui expliquerait pourquoi cette hypothèse n'a pas permis de rapprocher l'échantillon recueilli en ligne de celui recueilli en face-à-face.

Combinaison des deux méthodes de redressement

Les dernières hypothèses consistent à supposer que de combiner les deux méthodes précédentes pourraient permettre de diminuer plus efficacement les écarts entre les modes de collecte. Pour les deux questionnaires, la méthode de pondération par le score de propension a donc été combinée avec les calages sur marge les plus efficaces testés précédemment.

Pour le questionnaire adressé aux ménages, l'hypothèse 8 correspond à la méthode du score de propension combiné au calage sur marge avec les variables classiques. Ceci afin de tester si le fait d'utiliser l'inverse des scores de propension comme poids initiaux permettait d'obtenir une dispersion finale des poids plus acceptables. Cela n'est pas le cas puisqu'un rapport minimum/maximum supérieur à 40 000 est obtenu (Tableau 6-C.), on ne peut donc pas accepter cette hypothèse. Par contre, malgré cette répartition trop importante des poids, cette méthode permet de diminuer l'écart global entre les modes de collecte, même si elle ne réduit les écarts que pour la moitié des blocs de questions (Tableau 5-C.).

Pour le questionnaire adressé aux adultes, la méthode du score de propension a été combinée au calage sur marge avec les variables de quotas (hypothèse 9). Ici, la combinaison des deux méthodes donne de meilleurs résultats que leur utilisation seule. L'hypothèse 9 diminue en effet légèrement l'écart global et diminue l'écart pour tous les blocs sauf un (Tableau 7-C.). Dans une dernière hypothèse (hypothèse 10), la méthode du score de propension a été combinée à l'hypothèse de calage sur marge qui donnait le redressement le plus efficace : l'introduction du temps passé devant les différents écrans. La combinaison des deux méthodes n'améliore pas cette fois le résultat obtenu avec l'hypothèse 6. Dans les deux cas, la répartition des scores obtenus est correcte (Tableau 8-C.)

4 DISCUSSION

1. Des différences observées sur les questions auxiliaires

Des différences sociodémographiques entre les deux échantillons malgré les quotas

Le mode de recrutement est fondamentalement différent entre les enquêtes en ligne et en face-à-face. En face-à-face, les interrogés peuvent être recueillis dans des conditions proches d'un tirage aléatoire, malgré un biais de non-réponse. Les enquêtes en ligne sont en revanche des enquêtes non-probabilistes. Les personnes interrogées sont des volontaires qui ont dû s'inscrire préalablement sur des sites internet, il existe donc un biais de sélection. De plus, une partie de la population (celle qui n'a pas accès à Internet et qui ne s'inscrit jamais que les sites d'enquêtes) est inaccessible par les enquêteurs, ce qui ajoute un biais de couverture. Or, comme cette partie de la population a des caractéristiques particulières, une enquête en ligne est beaucoup moins représentative de la population générale. Ceci est confirmé par de nombreux auteurs ayant comparé la représentativité d'enquêtes en ligne à des enquêtes en face-à-face ou au téléphone (Beck, Yan, & Wang, 2009; Chang & Krosnick, 2009; Yeager et al., 2011). Le problème est qu'il n'existe pas aujourd'hui de méthode précise pour sélectionner un échantillon représentatif de la population générale à partir d'un panel d'individus connectés à Internet. (Taylor, Nelson, Grandjean, Anatchkova, & Aadland, 2009). De plus, à ces biais de sélection et de couverture spécifiques des enquêtes en ligne, s'ajoutent des biais de non-réponse qui sont plus importants pour les enquêtes en ligne.

C'est pourquoi, comme c'est souvent le cas dans ce genre d'étude (Couper, 2011; Lee, 2006; Rosenbaum & Rubin, 1984; Schonlau, 2007), les panélistes se retrouvent avec des caractéristiques différentes de la population interrogée en face-à-face, davantage représentative de la population générale. Les panélistes présentent ici surtout des niveaux de diplômes plus élevés, malgré un redressement sur la catégorie socioprofessionnelle.

En plus de cette limitation, un autre problème se pose dans le cas de cette étude. Il s'agit de la différence des méthodes de sélection des personnes interrogées en ligne et en face-à-face. Le prestataire TOLUNA n'était pas en mesure de sélectionner les participants par la méthode des quotas sur les ménages, puisque les panels de volontaires sont répertoriés par individus, alors que pour l'enquête CCAF habituelle en face-à-face, la sélection se fait par la méthode des quotas appliquée aux ménages. La sélection s'est donc faite sur les individus pour l'enquête en ligne, alors qu'elle s'est faite sur les ménages en face-à-face. Ceci est surtout problématique pour la base des ménages, puisque le redressement « classique » donne des répartitions de poids déjà beaucoup trop hétérogènes. L'ajout de variables supplémentaires augmente donc encore plus cette répartition. Il n'est donc pas possible de conclure clairement sur l'utilité ou non des redressements testés. A l'avenir, il sera donc indispensable de trouver un moyen pour sélectionner les panélistes en appliquant la méthode des quotas directement aux ménages plutôt qu'aux adultes. Il faudra pour cela travailler en partenariat avec le prestataire TOLUNA ou trouver un prestataire capable de fournir ce service.

De plus, il est intéressant de s'interroger sur la pertinence des variables de quotas choisis. La plupart des variables de quotas utilisées le sont depuis les premières enquêtes CCAF, ce qui permet d'étudier l'évolution des réponses aux différentes questions au cours du temps. Elles ont été choisies en tenant compte de l'hypothèse suivante : si l'échantillon choisi est représentatif sur les grandes variables sociodémographiques, il sera représentatif sur les autres variables que l'on veut étudier. Les variables choisies sont donc celles dont la distribution dans la population est connue (notamment parce qu'elles sont aisées à obtenir) et parce qu'elles sont corrélées avec les comportements à étudier. Cependant, ces variables sont-elles les plus pertinentes aujourd'hui ? D'après l'expérience du qu'a le CRÉDOC des enquêtes alimentaires, il apparaît que les comportements alimentaires sont surtout dépendants du niveau de diplôme. Il pourrait donc être intéressant d'ajouter directement cette variable aux quotas. Les revenus des ménages et des individus pourraient également être intéressants à prendre en compte, d'autant plus que la distribution de ces deux variables dans la population générale est également assez facile à obtenir.

Au-delà des variables de quotas elles-mêmes, on peut s'interroger sur les marges utilisées pour les variables de quotas, qui représentent la distribution des variables dans la population générale. Il faut noter ici que les marges sont obtenues à partir d'une autre enquête (l'enquête de l'INSEE « emploi 2014 »). D'éventuels biais présents sur cette enquête se répercuteraient donc sur l'enquête CCAF. De plus, il existe toujours un décalage temporel, les distributions utilisées pour redresser l'échantillon de l'enquête CCAF sont celles mesurées quelques années précédentes. Cependant, étant donné la méthodologie et la fiabilité de l'enquête « emploi » et la vitesse d'évolution des caractéristiques sociodémographiques, on peut considérer que les marges utilisées ici sont bien représentatives de la population française.

Des différences sur les autres questions auxiliaires

Des écarts sont visibles sur d'autres variables auxiliaires que les variables sociodémographiques. S'observe notamment une grande différence dans le sexe de la personne responsable des achats alimentaires: en ligne, ce rôle est beaucoup plus assuré par des hommes qu'en face-à-face. On peut supposer que ce décalage est dû à un problème de définition claire du responsable des achats dans le questionnaire rempli en ligne et donc de compréhension de la part du répondant, plus qu'à un vrai écart. Les femmes sont en effet encore clairement aujourd'hui les principales acheteuses des ménages (Mathé & Hébel, 2013).

En face-à-face, l'enquêteur est physiquement présent au côté de tous les membres du foyer, il peut donc apporter des précisions sur la définition du responsable des achats et discuter avec les différents adultes pour identifier qui endosse vraiment ce rôle. En ligne, la seule indication est « êtes-vous le responsable des achats en matière alimentaire, c'est-à-dire celui qui fait le plus souvent les courses ou la cuisine ? ». Si la personne répond « Oui », il est automatiquement dirigé vers le questionnaire des ménages. Un panéliste peut donc répondre « Oui » sans trop réfléchir à cette question et être considéré comme la personne de référence du foyer en matière alimentaire sans l'être vraiment.

Ceci est problématique puisque les hommes et les femmes ont des perceptions très différentes de la consommation et des façons différentes d'appréhender les achats (Mathé & Hébel, 2013). Pour la

suite, il est donc indispensable que l'enquête en ligne permette d'identifier de façon sûre si la personne interrogée est réellement le responsable des achats. Il faudrait peut-être pour cela étoffer la définition et préciser que si la personne est bien responsable, elle devra alors répondre à un questionnaire sur les achats et les habitudes précises du foyer.

2. Des différences observées sur les questions d'intérêt

Des écarts dans la structure des réponses qui sont liés à la façon de remplir le questionnaire

Une utilisation plus fréquente de la modalité « NSP » par les panélistes

Le fait que les panélistes utilisent beaucoup plus souvent la modalité « NSP » que les personnes interrogées en face-à-face est assez problématique puisque cela entraîne une perte d'information. Ce phénomène a été observé par plusieurs auteurs dans le passé (Heerwegh & Loosveldt, 2008). En effet, les panélistes ont tendance à utiliser facilement cette modalité lorsqu'ils ne sont pas motivés par la présence d'un enquêteur. Ici, ce phénomène est d'autant plus visible sur les questions qui demandent réflexion, comme les questions portant sur les opinions personnelles ou celles pouvant présenter une forme de désirabilité sociale.

Par la suite, il sera donc nécessaire de limiter le plus possible le recours à cette modalités. Pour cela, il est possible de supprimer cette modalité pour certaines questions, notamment les questions factuelles, afin de forcer le panéliste à réfléchir à la réponse qui correspond le plus à sa situation. Pour les questions d'opinion, où l'individu peut réellement ne pas du tout savoir la réponse, il peut être intéressant de rajouter une précision sous l'énoncé de la question, pour rappeler l'importance d'une réponse précise pour l'interprétation du questionnaire. Il est également possible de mettre en place des systèmes de boucle, pour que les questions auquel l'interrogé n'a pas répondu réapparaissent plusieurs fois afin de les inciter à choisir une autre modalité. Enfin, des méthodes d'imputation, comme la méthode « hotdeck », permettent de recoder cette modalité si elle est trouvée en faibles proportions. (Andridge & Little, 2010)

Une moindre utilisation des modalités extrêmes de la part des panélistes

Les panélistes ont tendance à utiliser beaucoup moins souvent les modalités extrêmes que les interrogés en face-à-face. On peut tenter de proposer une explication. Lorsqu'un individu répond à un questionnaire en ligne, il est seul devant son écran et personne ne l'incite à répondre aux questions de façon attentive. Il peut donc lire rapidement les questions et les listes de réponses proposées, souvent listées de haut en bas. Ainsi, souvent, les panélistes liraient de moins en moins attentivement les modalités proposées de haut en bas, et répondent moins facilement par les modalités extrêmes en bas de liste. De nombreux auteurs ont observé ce phénomène. (Croutte, 2015; Heerwegh & Loosveldt, 2008)

La théorie utilisée pour expliquer le manque de précision dans les réponses des panélistes est celle du « satisficing » ou « théorie du seuil de satisfaction de l'individu » (Krosnick, 2002). Elle affirme que le fait de répondre à une question de façon optimale demande un travail cognitif important. La personne interrogée doit d'abord interpréter correctement le sens de la question et comprendre la réponse attendue, faire appel aux informations nécessaires pour répondre par un travail de mémoire,

intégrer ces informations de façon synthétique, et les reporter en prenant en compte les alternatives proposées (Tourangeau, Rips, & Rasinski, 2000). Sans le concours d'un enquêteur extérieur, les panélistes peuvent plus facilement chercher à prendre des raccourcis dans ce chemin mental pour fournir une réponse moins précise mais qui leur paraît suffisante car elle limite les efforts fournis.

Ceci est problématique puisque, même s'il n'est pas à l'origine de perte directe d'information, il est à l'origine de perte de précision. En effet, ce problème ne peut être réglé qu'en regroupant les réponses en catégories plus larges : en regroupant par exemple les réponses « rarement » et « jamais ». Or, la nuance entre ces réponses peut être importante pour l'interprétation de l'enquête. Par la suite, la réalisation éventuelle de l'enquête CCAF en ligne devra donc intégrer cette perte de précision dans les quelques questions proposant de nombreuses modalités. Cependant, ces questions ne concernent qu'une partie du questionnaire et leurs réponses, même en regroupant les modalités, permettent tout de même d'étudier les tendances en matière de comportements et attitudes alimentaires.

Des écarts de réponses qui reflètent des différences d'opinions et de comportements entre les deux échantillons interrogés

De nombreuses questions présentent des écarts faibles entre les deux modes de collecte

Pour la plupart des questions, les écarts de réponses obtenus entre le questionnaire en ligne et le questionnaire en face-à-face sont faibles. Ceci est très positif puisque, pour ce type de questions, un changement de mode de passage du questionnaire n'aura pas d'impact sur les réponses obtenues.

Pour le questionnaire posé aux adultes, les questions concernant la fréquence des repas « sautés » donnent des résultats comparables en ligne et en face-à-face. Les deux modes de collecte donnent également les mêmes résultats concernant les lieux où mangent les individus (à la maison, au restaurant...), sauf pour ce qui est de la restauration rapide, davantage fréquentée par les panélistes. Pour ce qui est des critères de qualité énoncés pour juger les produits alimentaires, les enquêtés déclarent les mêmes raisons quel que soit le mode de collecte, sauf en ce qui concerne la fraîcheur, l'apparence du produit et sa composition.

Les réponses aux questions portant sur les goûts préférés (sucré, salé...) ne dépendent pas non plus du mode de collecte, de la même façon que les réponses de type « hédonisme et convivialité » à la questions « Pour vous, qu'est-ce que « bien manger » ? ». L'enquête en ligne donne également des résultats semblables à l'enquête en face-à-face concernant les éléments nutritionnels surveillés dans l'alimentation, à part pour les fibres qui récoltent un peu plus d'attention chez les panélistes. Elle donne également des résultats semblables à l'enquête en face-à-face concernant les types de risques alimentaires, sauf pour le cas des additifs, qui inquiètent beaucoup plus les panélistes. Enfin, le questionnaire proposé aux adultes fournit les mêmes résultats concernant la confiance qu'accordent les adultes aux différents acteurs de l'alimentation avec les deux modes de collecte.

Le questionnaire soumis au responsable des achats des ménages comporte également plusieurs types de questions où peu d'écarts notables sont visibles entre les deux modes de collecte. Ceci est vrai pour les questions concernant les fréquences d'achat des différents produits frais non préparés et des produits prêts à l'emploi (même si les panélistes consomment un peu plus de légumes et de

fruits prêts à l'emploi). Les deux modes de collecte recueillent également des résultats semblables concernant les circuits de distribution fréquentés (marchés, épiceries, supermarchés, hypermarchés...).

Plusieurs raisons d'achat sont également semblables entre les deux modes de collecte : c'est le cas du prix, des informations sur les produits disponibles en magasin, et de l'imaginaire derrière les produits. Il apparaît également que les panélistes et les personnes interrogées en face-à-face acceptent dans les mêmes proportions de payer plus cher pour des produits s'ils présentent des signes officiels de qualité, sauf pour la viande de boucherie. Enfin, le questionnaire des ménages met en évidence les mêmes habitudes relatives aux produits alimentaires jetés avec les deux modes de collecte.

Des questions qui mettent en évidence un profil particulier des panélistes

Cependant, plusieurs questions recueillent des réponses différentes selon le mode de collecte, mettant en évidence des opinions et des attitudes différentes entre les deux populations interrogées.

Dans le domaine de l'alimentation, les panélistes semblent accorder plus d'importance et à la convivialité et au plaisir dans la préparation et la consommation des repas, et recherchent davantage l'originalité en cuisine. Ils ont en effet plus tendance à répondre qu'ils aiment cuisiner, découvrir de nouveaux produits et qu'ils n'achètent pas toujours les mêmes marques et les mêmes produits. Le moment du repas semble être plus important pour les panélistes puisqu'ils ont tendance à passer un peu plus de temps à préparer, ainsi qu'à consommer leur repas. Cette importance du « fait maison » est une tendance en forte augmentation depuis 2013 chez les Français. Elle est le signe d'un retour du plaisir dans l'alimentation, tant du point de vue de la préparation que du point de vue du goût attendu (Mathé, Beldame, & Hébel, 2014). Ce phénomène montre la volonté d'aller vers une cuisine faite soi-même, à partir de produits bruts, pour des raisons de goût et de qualité. Cuisiner soi-même permet d'allier le plaisir de préparer et de déguster et apporte de la confiance dans son alimentation (Mathé & Hébel, 2015).

Mais ce phénomène illustre aussi indirectement un déficit de confiance vis-à-vis des produits industriels tels que les produits transformés ou les plats préparés, faisant percevoir le « fait maison » comme garantie d'un « bien manger ». Ceci peut notamment être liée à la crise de la viande de cheval qui a contribué à créer un climat de défiance vis-à-vis des produits industriels (Mathé et al., 2014). Cette défiance risque d'ailleurs d'être renforcée par la crise actuelle des œufs contaminés au Fipronil. Cette tendance est exacerbée chez les panélistes. Les panélistes accordent en effet plus d'importance aux produits naturels, non transformés, produits de façon artisanale et qui respectent l'environnement. Cela se manifeste entre autres par le fait que les panélistes ont plus peur des additifs que les personnes interrogées à leur domicile, alors que l'apparence des produits est beaucoup moins importante pour eux. Par conséquent, ils lisent beaucoup plus les étiquetages et particulièrement la composition des produits.

On remarque également que la préservation de l'environnement est une problématique importante pour les panélistes. Ils sont également plus sensibles au soutien aux terroirs et aux petits producteurs. Ceci se manifeste par des achats plus fréquents de produits issus de l'agriculture biologique ou du commerce équitable, de produits disposant d'une AOC, ou de produits achetés directement

chez les producteurs. Ces préoccupations sont également assez récentes, et davantage retrouvées chez les personnes éduquées et ayant des revenus plus élevés pouvant davantage accéder à des produits respectueux de l'environnement, des éleveurs et des conditions d'élevage, sont souvent plus chers.

Autre point qui sépare les panélistes des interrogés en face-à-face : les panélistes surveillent plus leur alimentation pour éviter de grossir. Pour des IMC équivalents à ceux des interrogés en face-à-face, ils sont plus nombreux à se trouver un peu trop gros et pas assez actifs. Ils semblent donc être un peu plus sujets à la pression de la société d'idéal de minceur. Ils font plus attention à la composition et à la teneur en sucre et en matière grasse écrites sur l'étiquetage des produits. Pourtant, ils déclarent moins de comportements restrictifs que les personnes interrogées en face-à-face. Ils sont en effet moins nombreux à déclarer éviter les matières grasses, les sucres, le fromage, l'alcool... Ces comportements sont en effet des comportements en diminution depuis 2017. Avant cette période, les Français, qui avaient alors été l'objet d'une diffusion massive de messages nutritionnels à partir du premier PNNS en 2001, accordaient de plus en plus d'importance aux liens entre alimentation et santé et se trouvaient tiraillés entre la tradition culinaire nationale et l'augmentation des pathologies du comportement alimentaire (Mathé et al., 2014). Mais la crise économique a mis un terme à cette progression comme le montrent les résultats de l'enquête précédente, CCAF 2013 (Hébel, Lehuédé, Mathé, & Siounandan, 2013). Depuis, les préoccupations se tournent vers des produits plus naturels, et vers le « fait maison ».

Malgré ces comportements restrictifs moins déclarés par les panélistes, ces derniers consomment en moyenne beaucoup plus d'aliments allégés en sucre ou en matière grasse. Il est possible que, plutôt que de se priver de certaines catégories d'aliments, les panélistes préfèrent faire appel aux innovations en matière de santé en consommant des produits allégés pour concilier plaisir et maintien ou perte de poids.

Les panélistes ont donc tendance à avoir des préoccupations plus modernes que les personnes interrogées en face-à-face. D'autres éléments étayent cette théorie. Les panélistes utilisent par exemple plus facilement les distributeurs de nourriture ou de boisson, fréquentent plus souvent les enseignes de restauration rapide et font plus souvent appel aux services de livraison de courses ou de repas à domicile. Par ailleurs, ils prennent également moins souvent des petits déjeuners le matin, les sauts de petits déjeuners ayant tendance à se développer dans la population française (Hébel, 2013).

Enfin, les panélistes ont des avis plus tranchés et plus négatifs vis-à-vis du système alimentaire. Si ils font autant confiance aux différents acteurs de l'alimentaire que les personnes interrogées en face-à-face, ils sont en moyenne beaucoup moins satisfaits par tous les types de contrôles et de réglementation en vigueur dans le domaine alimentaire, en France et à l'étranger. Ils ont également un point de vue assez négatif sur l'évolution de l'alimentation au cours du temps puisqu'ils sont beaucoup plus nombreux à trouver qu'à la fois le goût et la qualité de l'alimentation s'est détériorée ces dernières années. Ceci reflète la tendance des Français à faire de moins en moins confiance en la qualité et dans les marques de l'offre alimentaire (Hébel, 2008).

En conclusion, il apparaît que les panélistes ont des comportements plus modernes et plus associés aux comportements des classes aisées, même si l'échantillon des panélistes a été redressé par rapport à l'âge et à la catégorie socioprofessionnelle. On retrouve en effet plus chez eux les grandes tendances liées aux niveaux de vie plus élevés étudiées depuis les dernières années dans le domaine alimentaire : importance des produits naturels et respectueux de l'environnement, volonté de manger sainement pour ne pas grossir, utilisation des services connectés De plus, on retrouve chez eux une plus grande importance accordée à la convivialité, au « fait-maison » et à la découverte culinaire.

Le fait que l'enquête en ligne surestime de cette façon les comportements plus moderne peut être intéressant pour certains clients du CRÉDOC puisque cela permettrait à l'avenir de détecter plus rapidement les dernières tendances en matière alimentaire et d'anticiper les nouveaux comportements, même si une quantification précise des phénomènes sera difficile.

Des biais de désirabilité sociale liés à la présence ou l'absence de l'enquêteur peu visibles

Ici, le biais de désirabilité sociale semble peu présent, sûrement dû au fait que les questions sur les comportements alimentaires sont moins sujets à controverse et donc à ce biais de désirabilité sociale. Cependant, le fait que les panélistes interrogés se déclarent plus inquiets vis-à-vis de la sécurité alimentaire et plus négatifs vis-à-vis du système alimentaire peut être dû à ce phénomène. On peut alors se demander si, pour ce genre de question, le passage du questionnaire en ligne ne donne pas des résultats plus « proches de la réalité » que le questionnaire en face-à-face. Mais cela peut aussi être dû au biais lié à la rémunération des panelistes (voir ci-dessus).

De manière générale, cela pousse à s'interroger sur la pertinence de garder l'enquête réalisée en face-à-face comme référence. Plusieurs auteurs montrent en effet que lorsqu'il s'agit de questions sensibles, les niveaux de déclaration sont plus élevés en ligne (Kreuter et al., 2008). Cependant, ce type d'enquête ne pourra pas être considéré comme une référence tant qu'il ne pourra pas atteindre toutes les catégories d'individus de la population.

3. Effet des redressements supplémentaires de l'échantillon interrogé en ligne sur les écarts de réponses aux questions d'intérêt

Méthodes de redressement utilisées

Les résultats des redressements de l'échantillon des ménages sont difficilement exploitables à causes d'une répartition des poids déjà très hétérogènes dans l'hypothèse 0. L'ajout de variables supplémentaires a donc encore plus augmenté le risque de distribuer des poids aux individus de façon erratique. Il sera donc indispensable pour la suite de pouvoir sélectionner directement les ménages interrogés en ligne par la méthode des quotas. Il semble tout de même que le fait d'ajouter aux quotas la fréquence de repas pris devant la télé, l'utilisation d'internet pour faire ses courses ou l'utilisation des services de livraison à domicile pourrait permettre de diminuer efficacement les écarts entre les deux modes de collecte.

En revanche, la pondération par l'inverse du score de propension donne des résultats concluants pour le questionnaire des ménages. Cette méthode de redressement ne distribue pas des poids trop hétérogènes aux individus, même si de nombreuses variables sont intégrées dans le modèle de calcul des scores de propension. De plus, cette méthode permet de diminuer efficacement l'écart moyen et les écarts sur les différents blocs de questions de l'enquête.

Les résultats de l'enquête auprès des adultes sont plus facilement interprétables puisque les deux échantillons sont, dès le début, assez semblables sur les caractéristiques sociodémographiques de base. Les tests de redressement par calage sur marges ne donnent donc pas des poids trop hétérogènes. Pour la plupart des hypothèses qui ont été testées, l'effet sur la diminution des écarts entre les deux échantillons sont assez limitées, sauf pour l'hypothèse 6. Ajouter aux variables de quotas le temps passé devant la télévision et l'ordinateur ainsi que le temps passé à jouer aux jeux-vidéos permet en effet de diminuer efficacement les écarts entre les modes de collecte.

Contrairement à ce que l'on a vu pour l'enquête des ménages, la pondération par l'inverse du score de propension n'est pas plus efficace ici que les méthodes de calage sur marge, elle augmente même les différences entre les deux échantillons. Cela est dû au fait que le modèle de régression logistique permettant de calculer les scores de propension ne détecte pas assez efficacement les individus ayant une forte probabilité d'appartenir à la population des panélistes. Ici, le score de propension n'a pas permis de rapprocher la population des panélistes de celle des interrogés en face-à-face.

La pondération par l'inverse du score de propension semble donc pouvoir être une méthode efficace, puisqu'elle permet de redresser une population en prenant en compte de nombreuses variables sans entraîner une répartition très élevée des poids attribués à chaque individu. Mais elle est efficace uniquement si les variables introduites dans le modèle de régression logistique sont pertinentes et permettent de discriminer efficacement les deux populations. Cela est dû au fait que ce type de redressement est moins « fort » que le calage sur marge. En effet, le calage sur marge impose une distribution identique entre les deux populations sur les variables entrées dans le modèle de calage, alors que l'utilisation du score de propension permet juste de « rapprocher » les deux échantillons vis-à-vis de ces variables. En effet, cette méthode permet juste de « mimer » un design d'étude qui aurait sélectionné les deux populations de façon aléatoire vis-à-vis de ces variables (Austin, 2011). Ceci est d'autant plus marquant ici puisque la pondération n'est appliquée qu'aux individus interrogés en ligne, pour ne pas modifier la population de référence des interrogés en face-à-face...

Dans tous les cas, l'effet des différents redressements restent assez limités, même pour les plus efficaces. La littérature suggère en effet que ces différentes méthodes permettent la plupart du temps de diminuer les biais induits par le mode de collecte, mais ne les éliminent généralement pas. (Lee, 2006; Schonlau, 2007)

Variables de redressement utilisées

L'efficacité contrastée des tentatives de redressement montre l'importance des variables auxiliaires dont dispose le statisticien pour redresser les échantillons. Pour redresser un échantillon obtenu en ligne, le statisticien doit disposer de variables qui peuvent expliquer qu'un individu ou un ménage ait plus ou moins de chance d'être interrogé en ligne.

Il aurait été pertinent d'ajouter des variables supplémentaires au questionnaire avant de réaliser les deux enquêtes en parallèles. Notamment, concernant les variables sociodémographiques, il aurait ici été très intéressant de pouvoir calculer les revenus des ménages par UC. Les panélistes ayant tendance à avoir des revenus plus élevés que la population générale, introduire cette variable dans les redressements aurait pu permettre d'ajuster nos deux populations. D'autant plus que les revenus ont un impact très fort sur les comportements alimentaires (Hébel, 2012). Cependant, pour l'enquête en ligne, tous les individus de chaque ménage n'ayant pas été interrogés, il n'a pas été possible d'avoir accès au nombre d'enfants et d'adultes de chaque ménage. Il serait utile par la suite d'intégrer ces variables au questionnaire en ligne.

D'autres variables mériteraient également d'être étudiées. Ce sont toutes les variables qui cristallisent les différences entre les panélistes et la population générale et dont l'introduction dans les redressements permettraient de redresser la population des panélistes pour qu'elle ressemble le plus possible à la population générale (Croutte, 2015). Il s'agit notamment des variables dites « enquêtographiques » de type « A quelle fréquence participez-vous à des enquêtes en ligne ? », « au téléphone ? », « en face-à-face ? ». Ces dernières permettent clairement de repérer les individus qui ont une plus grande probabilité d'être atteints par une enquête en ligne.

D'autres variables relatives à l'équipement informatique et connecté seraient également nécessaires. Elles permettent de mesurer les équipements informatiques des ménages ou individus et donc de refléter s'ils ont des chances d'être atteints par une enquête en ligne. Ces questions peuvent porter sur la possession d'ordinateurs, de smartphones, de tablettes, d'une connexion à internet... Ces variables pourraient être très pertinentes surtout que le temps passé sur les différents écrans correspond aux variables qui permettent de diminuer le plus efficacement les écarts entre les deux modes de collecte.

Enfin, il pourrait être intéressant d'étudier l'effet de variables dites « attitudinales ». Ces dernières ont été utilisées dans différentes études pour mesurer la présence ou l'absence d'un biais de désirabilité sociale (Croutte, 2015; Lee, 2006; Schonlau, 2007). Ce sont des questions qui ne concernent pas le thème de l'étude et peuvent paraître étranges à l'interrogé, mais elles permettent de mesurer si l'individu est sujet à une pression de désirabilité sociale ou non lors du remplissage du questionnaire. Par exemple, S. Lee (Lee, 2006) a utilisé des questions portant sur la classe sociale dans laquelle l'interrogé s'estimait personnellement, le fait de voter régulièrement aux élections, le fait d'avoir une religion, et l'opinion vis-à-vis des minorités. M. Schonlau (Schonlau, 2007) a utilisé des questions portant sur les sentiments des personnes interrogées vis-à-vis des personnes homosexuelles, le fait d'avoir voyagé récemment, d'avoir lu un livre récemment, d'être gêné par les fouilles à l'aéroport... Cependant, il est clair que ce type de variable ne pourra pas permettre d'effacer tous les biais liés aux enquêtes en ligne. Mais il est important de continuer les recherches vers cette voie car il serait très intéressant de disposer de variables de ce type qui seraient efficaces dans la plupart des situations (Schonlau, 2007).

Cependant, il faut garder à l'esprit qu'il est très difficile de trouver un redressement qui serait vraiment efficace dans tous les cas. Notamment parce que les caractéristiques de la population cible vis-à-vis de tous ces paramètres ne sont pas forcément faciles à recueillir (particulièrement dans le cas des attitudes et des comportements). Si le biais de couverture de la population en ligne peut être

partiellement gommé par des redressements, il est nécessaire de connaître *a posteriori* la distribution dans la population générale des variables à introduire, ce qui est rarement le cas (Andrews, Nonnecke, & Preece, 2003).

5 CONCLUSION

Cette étude comparative est très encourageante puisque pour de nombreuses questions du questionnaire « Attitudes et Comportements » de l'enquête CCAF, le passage en ligne du questionnaire fournit des réponses semblables à l'enquête de référence réalisée en face-à-face. Cependant, pour certaines questions, des différences notables apparaissent, témoignant du manque de représentativité dont peuvent faire preuve les enquêtes en ligne. Pour cause, le mode de recrutement des personnes interrogées en ligne qui ne peut pas être probabiliste à cause du recours à un panel de volontaire et à cause de la couverture de connexion à internet qui n'est pas encore totale en France. Ainsi, malgré des quotas précis fixés sur plusieurs variables sociodémographiques, les panélistes qui témoignent parfois nettement de comportements différant de ceux de la population interrogée en face-à-face. Ils ont des comportements et des représentations plus modernes vis-à-vis de leur alimentation et on retrouve plus nettement chez eux les dernières tendances observées en matière d'alimentation. Cela est dû au fait que les personnes connectées sont en moyenne plus jeunes, plus aisées et plus éduquées que la population générale. En effet, un panel ne permet pas d'atteindre toutes les catégories de la population, notamment les personnes plus âgées, moins éduquées et dans des situations plus précaires

Plusieurs hypothèses ont donc testé plusieurs hypothèses pour tenter de réduire ces biais de sélection et de couverture. Pour cela, deux types de méthodes statistiques ont été utilisées pour redresser l'échantillon recueilli en ligne afin de le rendre artificiellement plus proche de la population recueillie en face-à-face. Nos résultats sont encourageants puisque plusieurs de nos tests ont permis d'ame-nuïser les différences de comportements alimentaires entre les deux populations. Cependant, plusieurs limites se sont posées. Tout d'abord, l'impossibilité de sélectionner des ménages sur Internet par la méthode des quotas est un problème qui devra être réglé pour pouvoir recueillir en ligne une population de ménages plus proche de la population générale. De plus, quelques variables qui auraient été très pertinentes pour redresser l'échantillon des panélistes qui n'étaient pas disponibles pour travail méthodologique. Cependant, les résultats encourageants obtenus laissent présager que l'introduction d'autres variables dans le questionnaire pourrait permettre efficacement de redresser efficacement l'échantillon de panélistes pour le rendre plus représentatif de la population générale.

Enfin, le recours à une enquête en ligne présente des points positifs. Le fait de remplir le questionnaire seul devant son ordinateur peut limiter les biais liés à la présence de l'enquêteur et permettre à l'interrogé de répondre de façon plus honnête, sans chercher à apparaître sous son meilleur jour. De plus, le fait qu'une enquête réalisée en ligne surestime les comportements plus moderne peut être intéressant puisque cela permettrait à l'avenir de détecter plus rapidement les dernières tendances en matière alimentaire, même si une quantification précise des phénomènes sera difficile.

Finalement, ce travail de méthodologie pourrait donc permettre par la suite de passer la partie « Attitudes et Comportements » de l'enquête CCAF en ligne, moyennant quelques manipulations de l'échantillon. Ce passage en ligne permettra de faciliter considérablement l'étape de collecte des données et de supprimer l'étape de saisie des questionnaires, étape longue et fastidieuse qui peut être source d'erreurs. Le prix de réalisation de l'enquête sera également considérablement diminué, puisqu'il pourrait être réduit à 5000 euros contre 40 000 euros nécessaires actuellement pour réaliser l'enquête en face-à-face. De plus, un nombre plus élevé de participants pourra être interrogé, rendant l'enquête plus attractive pour les clients. Ceci va permettre au CRÉDOC de rester concurrentiel et de continuer à vendre son étude CCAF aux nombreux organismes et entreprises déjà clients et de démarcher de nouveaux clients, après les avoir clairement informé de la méthodologie de l'enquête. Cependant, il est encore difficile aujourd'hui de concevoir un passage en ligne du questionnaire destiné aux enfants, ainsi que du volet « Consommation » basé sur des carnets alimentaire. Ces parties de l'enquête devront donc continuer à être menées en face-à-face.

BIBLIOGRAPHIE

- Andrews, D., Nonnecke, B., & Preece, J. (2003). Conducting research on the internet: online survey design, development and implementation guidelines. *International Journal of Human-Computer Interaction.*, 16(2), 185- 210.
- Andridge, R. R., & Little, R. J. A. (2010). A review of Hot Deck imputation for survey non-response. *International statistical review - Revue internationale de statistique*, 78(1), 40- 64. <https://doi.org/10.1111/j.1751-5823.2010.00103.x>
- Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, 46(3), 399- 424. <https://doi.org/10.1080/00273171.2011.568786>
- Bayart, C., & Bonnel, P. (2010). Enquête déplacements web - face-à-face : quelle comparabilité ? *Les Cahiers scientifiques du transport*, (57), 141- 167.
- Beck, K. H., Yan, A. F., & Wang, M. Q. (2009). A comparison of web-based and telephone surveys for assessing traffic safety concerns, beliefs, and behaviors. *Journal of Safety Research*, 40(5), 377- 381.
<https://doi.org/10.1016/j.jsr.2009.07.007>
- Bigot R., C. P. (2010). Enquête en ligne : peut-on extrapoler les comportements et les opinions des internautes à la population générale ? CREDOC, Cahier de recherche n°273.
- Blasius, J., & Brandt, M. (2010). Representativeness in Online Surveys through Stratified Samples. *Bulletin of Sociological Methodology*, 107(1), 5- 21. <https://doi.org/10.1177/0759106310369964>
- Brice, L., Croutte, P., Jauneau-Cottet, P., & Lautié, S. (2015). Baromètre du numérique, édition 2015 - Pour le Conseil Général de l'Economie et l'Autorité de Régulation des Communications Electroniques et des Postes. CREDOC.
- Chang, L., & Krosnick, J. A. (2009). National Surveys Via Rdd Telephone Interviewing Versus the Internet Comparing Sample Representativeness and Response Quality. *Public Opinion Quarterly*, 73(4), 641- 678.
<https://doi.org/10.1093/poq/nfp075>
- Couper, M. P. (2011). The Future of Modes of Data Collection. *Public Opinion Quarterly*, 75(5), 889- 908.
<https://doi.org/10.1093/poq/nfr046>
- Croutte, P. (2015). Une approche de l'effet du passage sur Internet d'une enquête en population générale. CREDOC - Cahier de recherche N°C323.
- Dillman, D. A. (2006). *Mail and Internet Surveys: The Tailored Design Method -- 2007 Update with New Internet, Visual, and Mixed-Mode Guide* (2 edition). Hoboken, N.J: Wiley.
- Duffy, B., Terhanian, G., Bremer, J., & Smith, K. (2005). Comparing data from online and face-to-face surveys. *International Journal of Market Research*, 47(6), 615- 639.

- Enquête Emploi 2014 ; Activité, chômage, emploi et inactivité en 2014 - Insee Résultats - 173. (2017). Consulté 25 mars 2017, à l'adresse <https://www.insee.fr/fr/statistiques/1406356>
- Evans, J. R., & Mathur, A. (2005). The value of online surveys. *Internet Research*, 15(2), 195 - 219. <https://doi.org/10.1108/10662240510590360>
- Hébel, P. (2008). Synthèse des baromètres sur l'alimentation. CREDOC - publication N°S3424.
- Hébel, P. (2012). In *Comportements et consommations alimentaires en France* (Tech&Doc 2012).
- Hébel, P. (2013). Le petit-déjeuner en perte de vitesse. CREDOC - publication N°259, ISSN 0295-9976.
- Hébel, P., Lehuédé, F., Mathé, T., & Siounandan, N. (2013). *Comportements et consommations alimentaires en France 2013 - Rapport général*.
- Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus web surveying in a high-Internet-coverage population - differences in response quality. *Public Opinion Quarterly*, 72(5), 836 - 846. <https://doi.org/10.1093/poq/nfn045>
- Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social Desirability Bias in CATI, IVR, and Web Surveys The Effects of Mode and Question Sensitivity. *Public Opinion Quarterly*, 72(5), 847 - 865. <https://doi.org/10.1093/poq/nfn063>
- Krosnick, J. A. (2002). The causes of no-opinion responses to attitude measures in surveys: they are rarely what they appear to be. The Ohio State University.
- Lee, S. (2006). Propensity score adjustment as a weighting scheme for volunteer panel web surveys. *Journal of Official Statistics*, 22(2), 329 - 349.
- Mathé, T., Beldame, D., & Hébel, P. (2014). Evolution des représentations sociales du bien manger. CREDOC - cahier de recherche N°316.
- Mathé, T., & Hébel, P. (2013). Comment consomment les hommes et les femmes ? CREDOC - Cahier de recherche N°309.
- Mathé, T., & Hébel, P. (2015). Le plaisir du cuisiné maison : pour le goût et la qualité. CREDOC - publication N°275, ISSN 0295-9976.
- Rosenbaum, P. R. (1987). Model-Based Direct Adjustment. *Journal of the American Statistical Association*, 82(398), 387 - 394. <https://doi.org/10.2307/2289440>
- Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association*, 79(387), 516 - 524. <https://doi.org/10.1080/01621459.1984.10478078>
- Sautory, O. (1993). La macro CALMAR - Redressement d'un échantillon par calage sur marges. INSEE. Consulté à l'adresse <http://www.insee.fr/fr/methodes/outils/calmar/doccalmar.pdf>
- Schonlau, M. (2007). Are 'Webographic' or attitudinal questions useful for adjusting estimates from Web surveys using propensity scoring? *Survey Research Methods*, 1(3), 155 - 163. <https://doi.org/10.18148/srm/2007.v1i3.70>
-

Effet du passage en ligne du questionnaire comportements et consommations alimentaires en France

Site Internet de l'INSEE - Définition personne de référence du ménage. (2017). Consulté 10 avril 2017, à l'adresse

<https://www.insee.fr/fr/metadonnees/definition/c1192>

Site Internet de l'INSEE - Définition ZEAT. (2017). Consulté 28 mars 2017, à l'adresse <https://www.insee.fr/fr/metadonnees/definition/c1910>

Taylor, P. A., Nelson, N. M., Grandjean, B. D., Anatchkova, B., & Aadland, D. (2009). Mode Effects and Other Potential Biases

in Panel-based Internet Surveys: Final Report. Mode Effects in Internet Panels, USEPA 83359101.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge University Press.

Varebian, M., & Forsman, G. (2003). Comparing propensity score weighting with other weighting methods: a case study on

Web data. The American Association for Public Opinion Research (AAPOR) 58th annual conference.

Yeager, D. S., Krosnick, J. A., Chang, L., Javitz, H. S., Levendusky, M. S., Simpser, A., & Wang, R. (2011). Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples.

Public Opinion Quarterly, 75(4), 709- 747. <https://doi.org/10.1093/poq/nfr020>
