

Journée d'échange Association Française du Marketing – Syntec Etudes Marketing & Opinion

L'échantillonnage en ligne,
Ou comment surmonter ses péchés originaux
Van Terradot, Dirigeante
Novatest

Attrait et défiance vis-à-vis de l' échantillonnage en ligne

- Des **avantages évidents**
 - ✓ Accessibilité de cibles nombreuses et variées
 - ✓ Recrutement possible des profils rares à faible taux d'incidence
 - ✓ Rapidité de la collecte
 - ✓ Informations préalables et pré qualification des répondants
 - ✓ Sincérité et richesse des réponses
 - ✓ Coût parfois plus intéressant
- Mais aussi des **reproches récurrents** :
 - ✓ Non représentativité
 - ✓ Echantillonnage non probabiliste
 - ✓ Risque de « professionnalisation » des répondants
 - ✓ Mode de collecte : un internaute peut –il représenter un non internaute ?
 - ✓ Projectabilité des résultats (validité externe)

Ces reproches tiennent à la nature même des échantillons online, et font figure de péchés originaux. La question est alors de savoir si on peut-on les surmonter, et si oui, comment ?

1. Non représentativité : de quelle représentativité parlons nous ?

• Représentativité nationale ?

✓Le profil des Internautes n'est toujours pas exactement le même que celui de la population française : 53 % d'hommes contre 48.2 % dans la population française, 34.4 % de CSP + contre 20.8 %, 22 % en Ile de France contre 18.5 %...)

Mais avec un taux de pénétration moyen de 60 %, les différences tendent à disparaître. Si on a 72 % de taux de pénétration d'internet chez les 12-17 ans, il n'est que de 40 % chez les 60-69 ans, mais ceux-ci ont la plus forte progression, passant en 1 an de 21 % à 40 %.

• Mais faut-il rechercher une représentativité nationale ?

✓La plupart des études marketing ne recherchent pas une représentativité nationale (sauf les sondages d'opinion sur des sujets généraux, les sondages politiques, les omnibus...)

✓Mais la représentativité d'une population cible particulière : les conducteurs de voitures, les adolescents de 15 à 25 ans, les femmes, les médecins généralistes...

• Faut-il faire un échantillonnage en ligne ou pas ?

✓C'est une question de pertinence : peu pertinent s'il s'agit d'interroger un échantillon de personnes de + 65 ans.

✓C'est une question de moyens : a t -on les ressources nécessaires pour y recruter l'échantillon voulu (liste d'emails, access panel assez grand...)

✓C'est une question qui va perdre de son importance au fur et à mesure que le taux de pénétration d'Internet augmente.

2. Echantillonnage non probabiliste : Pas de salut en dehors du sondage aléatoire?

- **Les 2 erreurs qui ont marqué l'histoire (américaine) des méthodes d'échantillonnage:**
 - ✓ **1936** : où The Literary Digest avec ses 2 millions de votes de lecteurs s'était trompé en annonçant l'élection d'Alfred Landon, alors que George Gallup, avec 4 000 interviews selon la méthode des quotas avait prédit juste, la victoire de Franklin Delano Roosevelt.
Même si The Literary Digest avait prédit juste lors de 5 élections précédentes, l'histoire n'a retenu que cet échec cuisant, et ouvert l'ère des sondages scientifiques
 - ✓ **1948** : L'échec retentissant des prévisions électorales de George Gallup, annonçant la victoire de Thomas Dewey, alors que ce fut Harry Truman qui l'emporta, a jeté le discrédit sur la méthode des quotas qu'il utilisait (on oublie souvent de dire que les autres sondages probabilistes s'étaient aussi trompés !)
- **Depuis, le monde des études est partagé** entre les tenants les échantillonnages probabilistes, pour qui c'est la seule méthode scientifique, la seule qui permette le calcul des marges d'erreur, et les autres, des pragmatiques, essentiellement les défenseurs de la méthode des quotas.

2. Pourquoi l'échantillonnage en ligne ne peut pas être probabiliste

- **Il n'existe pas d'annuaire exhaustif des emails**
- **On ne peut pas générer de manière automatique des emails** (comme c'est le cas pour la numérotation automatique téléphonique, où un logiciel peut faire des séries de 10 chiffres)
- Même s'il existe des listes et des bases d'emails, les utiliser pour des études peut enfreindre les **lois anti-spams** dans certains pays (en France, les invitations pour les études ne sont pas considérées comme des spams)

Mais le véritable échantillonnage probabiliste existe –t-il ?

- **Le face à face**, qui fait souvent figure de terrain le plus « qualitatif » et le plus cher est en fait rarement probabiliste :
 - ✓ Pour des questions de coût : pour limiter les frais de déplacement, les enquêtes sont souvent cantonnées aux abords des grandes villes
 - ✓ Les enquêtes en rue sont souvent limitées aux grands artères passantes
 - ✓ Les digicodes rendent l'accès aux immeubles et les random routes de plus en plus difficiles
 - ✓ Certaines banlieues difficiles ne sont jamais visitées (de même que les quartiers trop huppés)
 - ✓ De même que certaines régions entières où il y a rarement des enquêteurs (le Centre de la France, et les zones rurales en règle générale).

2. Le vrai échantillonnage probabiliste n'existe pas

- **Le téléphone** : est en théorie le plus aléatoire des modes de collecte, grâce à la numérotation automatique, mais ne l'est pas en pratique car :
 - ✓ Pas toujours utilisé : une étude auprès de 83 sociétés dans 17 pays européens, américains et asiatiques (Humphrey Taylor 1997), a montré que 4 pays n'y recourent jamais, 4 le pratiquent presque toujours, et 9 qui ne le pratiquent que de manière partielle (dont la France)
 - ✓ Non représentatif des abonnés au seul téléphone portable (16 % en 2007), souvent plus jeunes et de CSP plus modestes, des VOIP et des téléphones via ordinateur de type Skype, ces derniers en forte croissance
 - ✓ Le plus souvent du faux aléatoire car contrôlés par des quotas (régions, sexe, âge, taille de la commune, CSP...)
 - ✓ Filtré par les messageries vocales, l'affichage des numéros, les listes rouges ou listes « Do not call »
- **Online** :
 - ✓ Sauf pour les enquêtes sur sites web, nécessite la constitution de bases préalables
 - ✓ Recourt le plus souvent aux online access panels
 - ✓ En utilisant la méthode des quotas
 - ✓ Et le redressement pour compenser les différences entre la structure de l'échantillon et la structure de la population visée
 - ✓ Mise en place d'études parallèles de contrôle

3. Access panel : quel impact sur la qualité des réponses ?

- Les échantillonnages en ligne peuvent recourir à des pop ups sur des sites, des fichiers clients ou à des base d'emails.
Excluons d'emblée ces pseudo-études que l'internaute peut venir choisir sur des sites qui les proposent, à cause des risques évidents liés à l'autosélection;
- Mais le plus souvent, les études online font appel à des access panels online, qui sont des réservoirs de répondants potentiels qui se sont inscrits pour répondre de manière répétée à des enquêtes en ligne.
- C'est souvent le **recours aux access panel online** qui est mis en cause derrière les discussions sur l'échantillonnage en ligne.
- Les **questions** que cela recouvre concernent :
 - ✓ **l'origine du recrutement**, online ou aussi offline (diversité du sourcing)
 - ✓ Le fait d'**incentiver** les répondants pour les motiver à répondre
 - ✓ La **duplication** ou le fait d'appartenir à plusieurs panels
 - ✓ Le **mode de collecte** en lui-même

3. Sourcing : y a-t'il une différence de qualité selon l'origine du recrutement ?

- Il est couramment admis qu'un recrutement diversifié est un gage de qualité du panel, et que c'est mieux de recruter par différents modes (online et offline)
- Une étude menée en 2007 auprès de 4 427 panélistes extraits de manière aléatoire du panel XL online (Kristof De Wulf et Sam Berteloot) a analysé les différences de réponses selon le moyen de recrutement (téléphone, face à face, emailing, formulaire posté sur des sites, campagne de pub sur Internet). Les caractéristiques socio-démographiques ont été redressées pour apparier les échantillons, afin que les différences ne soient dûes qu'à la différence de l'origine du sourcing.

Les mesures portent sur :

- ✓ le taux de participation à l'étude
 - ✓ le temps de remplissage du questionnaire
 - ✓ La cohérence des réponses
 - ✓ La richesse des réponses aux questions ouvertes (en nombre de mots)
 - ✓ Le style de réponse (monotone ou varié)
 - ✓ Le niveau de connaissance
 - ✓ Le profil de consommation
- Les résultats montrent :
 - ✓ Pas de différence dans les comportements de consommation (alimentation, boissons, équipement, loisirs, activités politiques...)
 - ✓ Pas de différence dans la qualité des réponses (cohérence, richesse), sauf chez les recrutés par emailing (plus de sans réponse), et par internet (taux de clicks plus élevé)
 - ✓ Pas de différence dans les motivations à participer à des études

3. Incentives : Risque de « professionnalisation » des répondants ?

- Répondre à des enquêtes, est-ce une profession ? Il y a déjà 7 ans, les Américains disaient que bientôt, quand vous voudrez interviewer un consommateur, il vous dira : « Adressez-vous à mon impresario ! »
- Le fait de recourir le plus souvent à un système d'incentive (tirage au sort, loterie, points monnayables, cadeaux...) fait planer le doute quant au profil de ces répondants, que certains n'hésitent pas à qualifier de « chasseurs de primes ».
- Est-ce que c'est le gain qui motive l'inscription à un panel ? De nombreuses études semblent montrer que l'attrait du gain ne constitue pas la motivation première, même si les résultats peuvent varier en fonction du système de rémunération et du mode de recrutement.

Une étude auprès de 2 084 panélistes du panel XL online qui utilise comme incentives des tirages au sort et des versements à des œuvres caritatives donne comme motivation (Annelies Verhaghe 2007) :

- ✓ Je veux exprimer mon opinion : 32 %
- ✓ Je veux aider le chercheur en participant à l'étude : 27 %
- ✓ Je veux avoir une chance de gagner un prix : 20 %
- ✓ J'aime répondre à des questionnaires : 10 %
- ✓ C'est une façon de soutenir les œuvres caritatives : 8 %
- ✓ J'aime apprendre : 2 %
- ✓ Autres : 1 %

3. Duplication : Quel impact sur la qualité des réponses ?

- On sait qu'il existe une proportion non négligeable de panélistes qui sont **inscrits dans plusieurs panels**. L'ampleur du phénomène varie selon les sources :
 - ✓ Comley (UK, 2005) estime à 75 % des panélistes qui sont inscrits à au moins 3 panels différents, et à 68 % ceux qui aimeraient répondre à plusieurs enquêtes par semaine
 - ✓ Comscore Networks (USA, Casro 2006) estime que 1 % des 10 plus gros panels américains représente 34 % des questionnaires remplis.
 - ✓ Mais Kristof de Wulf et Sam Berteloot (Belgique, 2007) n'ont trouvé dans le panel XL online que 28.6 % qui appartiennent à au moins un autre panel...
- Derrière la question de la « professionnalisation » des répondants se posent la question de l'impact sur la qualité des réponses du fait des multiples participations à des enquêtes. De nombreuses études ont été menées (Fulgoni, 2005, Smith et Hofma Brown 2005, Noyce 2006) sans qu'on puisse montrer que les échantillons qui contiennent plus d'hyperactifs donnent des résultats moins valides ou moins fiables.
- 2 études en Angleterre menées par Harris Interactive (2006) auprès de 1983 panélistes dont 133 multipanélistes, et de 2 071 personnes dont 166 de multipanélistes) a montré que :
 - ✓ Pas de différence dans l'âge ou le sexe
 - ✓ Légèrement plus de chômeurs chez les multipanélistes (6.7 % contre 5.2 %)
 - ✓ Ils répondent plus, mais abandonnent aussi plus, d'où un taux de réponse identique
 - ✓ Mais la qualité de leurs réponses est légèrement supérieure (cohérence, richesse)
 - ✓ Ne sont pas plus intéressés par les incentives
 - ✓ Montrent par contre un plus grand désir d'apprendre des choses nouvelles

3. Les différents types de non qualité

•Smith, Holland et Hofma Brown (2006) proposent de distinguer:

•Les **hyperactifs**, qui répondent beaucoup et sont multipanélisés : ils sont plus souvent des femmes, travaillant dans l'industrie, moins actifs politiquement, lisent moins les journaux et écoutent moins la radio. Ils répondent plus, abandonnent moins, et donnent des réponses aussi riches aux questions ouvertes, pas plus « monotones » que la moyenne. Ils ne répondent pas plus vite que les autres, et accordent la même importance au prix dans les critères d'achat.

•Les **fraudeurs**, identifiés grâce à leur adresse IP et aux cookies, et à leurs réponses atypiques par rapport à la loi de Benford sur la fréquence relative d'apparition des valeurs numériques : on estime leur proportion entre 5 et 6 %, pas plus élevée que dans d'autres modes de collecte.

•Les **inattentifs** : ont des durées de réponses plus courtes, des taux de « Sans réponse » plus élevés, davantage de réponses illogiques, donnent un nombre de réponses aux questions à réponses multiples plus faible. Certains auteurs suggèrent de les éliminer au bout de 3 indices d'inattention. Mais il ne faut pas oublier qu'une proportion d'inattentifs plus élevée peut aussi être dû à un design de questionnaire déficient (trop long, trop fastidieux, grilles de réponses trop ardues)

•Les **conditionnés**, dont les réponses peuvent changer suite à une exposition à des informations provenant d'autres études.

• Les auteurs recommandent une approche multi-critères pour instaurer des règles de qualité, mais sans sur-réagir, et faire plus de mal que de bien par rapport à la situation avant les corrections.



4. Mode de collecte : des internautes peuvent-ils représenter des non internautes ?

- A profil égal, le fait d'interroger des internautes donne t-il les mêmes réponses que le même échantillon de non internautes ?
- Une étude comparative menée en Belgique auprès de 1015 personnes (Niels Schillewaert, Annelies Verhaghe, Kristof de Wulf, Bert Weijters, 2007) avec le même questionnaire passé auprès de 4 sous échantillons correspondant aux 4 principaux modes de collecte :
 - postal par un échantillon aléatoire
 - face à face par la méthode de la route au hasard
 - téléphone, par échantillon probabiliste (par numérotation automatique ou RDD)
 - par access panel avec la méthode des quotas

Les 4 échantillons ont été appariés sur le plan de l'âge et du sexe, afin que les différences ne soient dues qu'au mode de collecte.

Les résultats ont montré que :

- ✓ Pas de différence au niveau socio-démographique, sauf pour le postal qui est plus rural
- ✓ Pas de différence dans les comportements de consommation (produits courants, proportion d'early adopters, sensibilité aux marques), sauf pour le téléphone qui consomme moins de produits courants, mais plus de voyages et de produits financiers, et le online qui est plus consommateur d'appareils photo numériques, et plus early adopter

4. Mode de collecte : les résultats changent-ils selon le mode de collecte ?

- Peu de différences dans les attitudes, mesurées par l'échelle de Short, Schwartz Value System ou SSVS (qui distinguent 10 valeurs de vie : Universalisme, Engagement, Réussite, Autonomie, Stimulation, Sécurité, Intégration, Tradition, Hédonisme), sauf :
 - ✓ Le téléphone moins attaché à la réussite, l'engagement social, l'intégration
 - ✓ L'access panel accorde moins d'importance à la sécurité
 - ✓ Le postal qui y accorde plus d'importance.
- Au global, c'est le téléphone qui présente le plus d'écart par rapport aux autres modes de collecte.
- Harris Interactive a conduit plus de 200 études comparatives en parallèle entre téléphone et online sur 2 ans, et n'a pas trouvé de résultats différents sur la plupart des sujets.
- Intersurvey a comparé les réponses sur 1 an des panélistes qui avaient internet avant d'être recrutés pour son panel, à celles des panélistes qui ne l'avaient pas, et est arrivé à la même conclusion : le fait d'avoir Internet ne change pas les réponses.

5. Projectabilité des résultats online : Peut-on généraliser les résultats online à des populations offline ?

- Si l'origine de l'échantillon est **très liée à la marque du client** (recrutement online sur le site web du client, recrutement sur la base des clients, ou des inscrits à une newsletter du client), il faut faire attention aux **risques de biais** possibles, car l'échantillon ne sera pas représentatif ni de l'ensemble des internautes, et encore moins de la population globale (comme pour tout fichier ayant un profil de recrutement spécifique) :
 - ✓ Eviter d'utiliser ce type de recrutement pour des questions de notoriété, ou de consommation des produits liés à la marque.
- Sinon, le **problème est le même que pour tout autre mode de collecte** :
 - ✓ Il faut disposer de données de cadrage pour connaître la structure de la population cible (par des sources secondaires, par des questions dans un omnibus, ou par des études parallèles de contrôle)
 - ✓ Et effectuer des redressements pour compenser les écarts éventuels entre la structure de l'échantillon et la structure de la population globale.

Ex : Post test de campagne multimédias auprès des propriétaires d'automobiles de 25-49 ans. L'étude principale a eu lieu online, mais en parallèle des questions dans un omnibus CAPI pour avoir le profil en termes d'âge, de profession du chef de famille et de l'habitat servant de base au redressement. Des questions de contrôle sur l'impact spontané, assisté et la mémorisation de la campagne ont été posées. Pas de différence globalement entre online et offline sur ces dernières questions.

6. Conclusions

- Discuter la validité de l'échantillonnage en ligne suggère implicitement qu'il existe des terrains parfaits, à la représentativité absolue. Or aucune méthode de recueil n'est parfaite, chacune a ses avantages et ses biais. Ce n'est pas parce que certaines sont anciennes qu'il faut les oublier. Si les access panels online (pas les échantillons !) sont souvent plus urbains, plus jeunes, mieux éduqués, et plus riches que la population globale, le postal est souvent plus vieux, plus ouvrier, moins éduqué, et le téléphone, plus inactif, excluant les abonnés exclusifs de mobiles qui sont plus jeunes, et ont de plus bas revenus.
- De fait, l'échantillonnage parfait n'existe pas, et les frontières entre aléatoire et non aléatoire est souvent mince. Il existe un continuum de méthodes plus ou moins aléatoires, et des échantillons plus ou moins représentatifs.
- Au-delà de l'échantillonnage, c'est la qualité des résultats qui est en jeu, et leur projectabilité à la population de référence qui en est l'enjeu. A la lumière des recherches connues à ce jour, rien ne permet de remettre en cause la validité externe des résultats des échantillons online, même si il s'agit de constats plus empiriques que théoriques.

Van Terradot
01 43 55 95 64
vterradot@novatest.fr

6. Conclusions

- A partir du moment où les croyances, les attitudes et les comportements des internautes ne diffèrent pas de ceux des non internautes, rien ne s'oppose à ce que l'on extrapole les résultats à la population cible toute entière, et à prédire les comportements des non internautes à partir de ceux des internautes, en prenant les mêmes précautions que pour les autres modes de collecte.
- Ainsi que l'a déclaré un client : « *Je faisais partie des sceptiques au départ, et force m'a été de constater, aussi surprenant que cela puisse paraître, que les résultats que nous avons obtenus en ligne ont complètement corroboré les résultats de l'étude de contrôle en face à face.* »