



NOTE SUR LA QUALITÉ DES SOURCES DE RECRUTEMENT POUR LES ETUDES D'AUDIENCE

Ont participé à ce groupe :

- Membres du Comité Scientifique : Avner Bar-Hen, Jean Chiche, Anne-Marie Dussaix, Sébastien Lhote, Emmanuel Viennet et Thierry Vedel
- Collaborateurs du CESP : Noëlie Bonnard, Valérie Morrisson et Patricia Schultz

1. CONTEXTE

Aujourd'hui, les enquêtes de référence de mesure d'audience sont réalisées par téléphone ou par internet. C'est pourquoi le face à face, en tant que source de recrutement, n'est pas évoqué dans cette note. De même, certaines bases utilisées pour réaliser des enquêtes publiques (comme la base Fidéli) ne sont pas mentionnées car non accessibles aux études de mesure d'audience.

La diversification des sources de recrutement, c'est-à-dire le recours à d'autres sources que les bases de numéros de téléphone issues de l'annuaire ou générées aléatoirement, concerne aujourd'hui la plupart des études d'audience que le CESP contrôle. La baisse de la joignabilité pour les enquêtes par téléphone (avec un risque d'accélération avec la mise en place des numéros polyvalents par l'ARCEP au 1^{er} janvier 2023), la difficulté croissante à joindre certaines cibles (les jeunes et les CSP- par exemple) et le renchérissement des coûts que ces phénomènes entraînent, expliquent le recours à de nouvelles sources de recrutement par les instituts.

Ces nouvelles sources ont un point commun : elles sont qualifiées (par l'institut ou, le plus souvent, par un prestataire) et permettent de cibler des individus sur des critères socio-démographiques. Le recours à ces sources ne permet plus un recrutement en aléatoire ou sur des bases exhaustives.

Dans cette note, le CESP a souhaité :

- Lister les bonnes pratiques lorsque l'on introduit une nouvelle source de recrutement, en particulier quand il s'agit de recours à des bases qualifiées différentes d'un annuaire ou d'une base de numéros de téléphone générée aléatoirement.
- Donner une liste de critères permettant de valider la qualité de la source de recrutement.
- Préciser son avis sur l'opportunité d'utiliser chacune de ces sources.

Le recours à des sources qualifiées est un sujet qui concerne particulièrement le CESP car il est lié à la représentativité de l'échantillon recruté et interrogé. Les études d'audience ont pour objectif principal de fournir au marché publicitaire des indicateurs sur lesquels s'appuient l'achat et la vente des espaces publicitaires. Pour garantir une monnaie fiable et des échanges justes, il est nécessaire que la mesure d'audience ne soit pas biaisée, c'est-à-dire qu'elle soit représentative de l'ensemble des Français, quels que soient leurs comportements et leurs habitudes de consommation des médias.

En introduisant de nouvelles sources de recrutement, on peut améliorer la structure des échantillons mais on doit aussi s'assurer que cela n'introduit pas de nouveaux biais qui auraient un impact sur les audiences.

2. DÉFINITIONS

Il existe plusieurs sources de recrutement utilisées dans les études d'audience de référence.

Il est important de les définir pour les distinguer :

- Base aléatoire de numéros de téléphone : il y a plusieurs méthodes pour produire des bases de numéros de téléphone en aléatoire.
 - o Par déclinaison à partir de numéros tirés dans l'annuaire (cela ne concerne que les numéros géographiques)
 - o Par génération aléatoire à partir des préfixes attribués aux opérateurs de France métropolitaine par l'ARCEP. Cela concerne les numéros fixes ou mobiles. Les méthodes de génération ou de nettoyage des bases peuvent varier d'un institut à l'autre mais le principe reste le même.
- Vivier d'adresses propriétaires : il s'agit d'une base d'individus ayant été interrogés sur un dispositif et ayant accepté de répondre à d'autres enquêtes mises en œuvre par l'institut. Ces bases appartiennent à l'institut et sont gérées en interne. Il n'y a pas de système de gratification proposé. Cas particulier : l'institut peut se donner la possibilité de réinterroger des individus ayant déjà participé à l'étude d'audience.
- Panel et Access Panel : il s'agit d'individus qui déclarent qu'ils coopéreront pour une future collecte de données s'ils sont sélectionnés, généralement en échange d'une gratification. Il y a une relation qui s'inscrit dans le temps entre les membres du panel/access panel - qui comprennent l'engagement qui leur est demandé - et le prestataire en charge de cette base de données. Ces bases peuvent être recrutées via une ou plusieurs sources ou de manière aléatoire (la bonne représentativité du panel ou de l'access panel dépend de ces sources). Il faut souligner une différence importante entre un panel et un access panel : un panel est constitué de manière à être représentatif d'une population étudiée qui pourra être réinterrogée dans le temps (analyses longitudinales), tandis qu'un access panel est un vivier plus large, non représentatif, mais permettant de tirer des échantillons qui seront sollicités pour des enquêtes.
- Fournisseurs de fichiers : le prestataire est un agrégateur de fichiers ou de bases de contacts. Les individus listés dans ces fichiers n'ont pas de lien avec le fournisseur d'adresses (ils n'ont pas déclaré qu'ils souhaitaient coopérer avec ce fournisseur) et ils ne bénéficient pas de gratification pour leur participation à une enquête. Nous distinguons dans cette note deux types de fichiers : les fichiers « annuaires » qui couvrent une partie importante des numéros de téléphone actifs et les fichiers qualifiés qui permettent de cibler des profils spécifiques (jeunes, hauts revenus...).
- Parrainage : un individu recruté pour participer à une enquête est sollicité (et gratifié) pour fournir des contacts de personnes de son entourage susceptibles, elles aussi, de participer à l'enquête (entourage : au sein du foyer ou en dehors).
- « Intercepts » : il existe une autre catégorie de sources de recrutement non utilisée aujourd'hui dans les études d'audience de référence et qu'ESOMAR¹ désigne comme « Intercepts ». L'interception est une approche où les participants potentiels sont invités à répondre à une

¹ ESOMAR Global Research – March 2021 – Questions to help buyers of online samples.

enquête (pour une récompense ou non) pendant qu'ils sont engagés dans une autre activité, comme jouer à un jeu en ligne, lire des contenus, être actif sur un réseau social ou toute autre activité en ligne. Les participants « interceptés » peuvent être inconnus du fournisseur de l'échantillon ou avoir été pré-identifiés et qualifiés lors d'une expérience d'enquête antérieure.

3. BONNES PRATIQUES ET LISTE DES CRITÈRES

3.1. Bonnes pratiques

On peut lister des bonnes pratiques à suivre à plusieurs moments de l'intégration de ces nouvelles sources de recrutement :

Au moment de la sélection de ces sources en recueillant les informations nécessaires à l'évaluation de la qualité auprès du prestataire qui fournit les contacts :

- Définir de quels types de source il s'agit : vivier, panel, access panels, fournisseur de fichiers... pour comprendre le lien entre le prestataire et les individus susceptibles d'être contactés
- Obtenir des informations sur les sources qui ont permis la constitution du vivier, du panel, de l'access panel, du fichier d'adresses : l'objectif est de pouvoir déterminer s'il n'existe pas un défaut important de représentativité ou si une des sources de recrutement n'a pas un lien avec le sujet de l'enquête
- Demander les thématiques des études réalisées sur la source pour identifier s'il y a un lien entre le sujet de l'étude et les autres sujets traités le plus souvent par le prestataire.
- Connaître la méthode d'échantillonnage utilisée par l'institut ou le prestataire pour élaborer l'échantillon qui sera fourni et les règles de sélection (nombre de sélections possibles dans les tirages, nombre maximum de sollicitations autorisé...)
- Obtenir des informations pour pouvoir évaluer la méthode de gestion des contacts et leur exploitation. Quelles sont les règles de gestion : sélection des adresses et intégration dans les bases, recrutements des panélistes, nombre de sélections, de sollicitations / de participations, règles de nettoyage, système de relances et programme de gratification... ? L'objectif est de s'assurer que les contacts ne sont pas sur-sollicités, « experts » des sondages ou « chasseurs de prime », ce qui pourrait remettre en cause la qualité de leurs réponses.
- Enfin, décrire la méthode de qualification des profils et donner la récence et le rythme de mise à jour de ces informations

En réalisant des analyses au moment du test ou de l'intégration de la source dans le dispositif d'enquête :

- Mener des tests en amont, avant d'introduire la nouvelle source dans les bases de recrutement ou d'enquêtes.
- S'assurer de la déduplication de ces sources avec les bases d'enquêtes utilisées pour éviter les sollicitations multiples
- Garder, dans les fichiers, l'information « source de recrutement » pour pouvoir mener des analyses a posteriori (par exemple dans le recrutement d'un panel : y a-t-il un taux de churn différent selon la source ?)
- Mettre en place une validation des profils pour évaluer la qualité des sources sur ce critère (les jeunes sont-ils bien des jeunes ?)
- Mener des analyses pour estimer les impacts éventuels sur les niveaux de participation mais aussi sur les niveaux d'audience et partager ces analyses avec les commanditaires de l'enquête

- Intégrer une question sur la participation à des panels ou une variable d'intérêt

Dans l'utilisation de ces sources dans la durée :

- Limiter le poids des sources qualifiées (pour privilégier les sources aléatoires)
- Maintenir stable le poids de chaque source dans le temps sauf s'il est démontré qu'il n'y a pas d'effet sur les résultats en changeant le poids des sources.
- Garder une structure de l'échantillon proche d'une année sur l'autre pour limiter les impacts sur les niveaux d'audience liés à ces changements

Cas particulier des ré-interrogations

Pour éviter le risque de réponses de mauvaise qualité due à la connaissance du questionnaire :

- Limiter fortement le recours à cette méthode
- Établir des règles strictes : nombre de participations autorisées et durée minimum entre les ré interrogations
- Se donner les moyens de contrôler la participation de ces individus : durée de réponse au questionnaire, mise en place d'une durée minimum acceptable.

3.2. Première liste de critères

Dans cette partie, le CESP a élaboré une première liste de critères sur lesquels les bases peuvent être évaluées (certains critères ne s'appliquent pas à toutes les sources). Il n'y a pas d'indicateurs chiffrés pour les critères listés, car les analyses par source doivent être faites de manière ad hoc en prenant en compte l'ensemble des informations données.

Critères
Représentativité
Les éléments fournis par le prestataire permettent-ils de vérifier si la source utilisée vise à la représentativité de l'univers que l'on cherche à représenter ?
Le risque d'avoir une source liée à la mesure d'audience a-t-il été évalué ?
Quelle est la taille totale de la base utilisée ?
Quelle part de la cible visée (dans la base totale) représente l'échantillon fourni ?
Quelle est la méthode de tirage de l'échantillon au sein de la base ? Un certain aléa est-il préservé ?
Quels sont les critères de sélection utilisés par le prestataire pour tirer l'échantillon au sein de la base ?
Si la source est utilisée régulièrement : la méthode d'échantillonnage et les critères de sélection sont-ils stables dans le temps ?
Gestion des contacts
Combien de nouveaux contacts ont été ajoutés chaque mois au cours des 12 derniers mois ?
Quel est le taux de churn mensuel de chaque mois au cours des 12 derniers mois ?
Quelle est la durée maximum autorisée de présence du contact dans la base ?
Quel est le nombre maximum autorisé de sélections / de sollicitations sur 12 mois, sur un mois ?
Quel est le nombre moyen de sélections / de sollicitations observées au cours des 12 derniers mois pour l'échantillon fourni ?
Quel est le nombre moyen de sélections / de sollicitations observées au cours des 30 derniers jours pour l'échantillon fourni ?
Quels sont les critères pour exclure un contact ?
Quel est le protocole de gratification ?
Quel est le montant de l'incentive pour la participation à l'enquête ?
Quelles sont les règles de relances à la suite de la première sollicitation ?
Comment les données sont-elles mises à jour ?
Quelle est la fréquence de mise à jour des informations socio-démographiques ?
Quel est le nombre de participations pour chacun des 12 derniers mois pour l'échantillon fourni ?
Quel est le nombre de participations sur des enquêtes « médias » pour chacun des 12 derniers mois pour l'échantillon fourni ?
Quel est le taux de réponse moyen observé dans la base au cours des 12 derniers mois ? (Ne s'applique pas aux fournisseurs de fichiers)
Qualité du fichier fourni
Y a-t-il eu déduplication avec les autres bases d'enquêtes ?
Cohérence entre profils vendus et profils observés : quel est le pourcentage de profils cohérents ?
Quel est le taux de réponse observé sur l'échantillon fourni par rapport à l'ensemble de l'échantillon ?
Impact sur la qualité des réponses et sur les résultats
Quelle est la durée du questionnaire auprès de l'échantillon fourni par rapport à l'ensemble de l'échantillon ?
Des tests statistiques ou des analyses ont-ils été réalisés au cours des deux dernières années pour identifier s'il existe un lien entre la source et les niveaux d'audience ?

4. AVIS DU CESP

Le CESP donne dans le tableau suivant son avis quant à l'utilisation dans les études d'audience de référence de chacune des sources de recrutement.

Définitions :

- **Acceptable** (sous réserve du respect des bonnes pratiques) : l'institut peut envisager d'utiliser ce type de source en s'assurant que les critères de qualité sont respectés
- **À utiliser avec précaution** : le CESP attire l'attention sur les risques de perte de représentativité à utiliser cette source. Il recommande de s'assurer en amont de la qualité de la source avec le prestataire et de procéder à des analyses pour évaluer l'impact de l'introduction de cette source dans l'échantillon. Il recommande aussi de limiter le recours à ce type de sources à des cibles difficiles à recruter.
- **À proscrire**

Type de source qualifiée	Avis du CESP
Vivier pour de la ré interrogation : l'institut constitue une base avec des individus ayant déjà participé à l'enquête en vue d'être réinterrogé sur la même enquête	À utiliser avec précaution et selon les études en respectant un délai suffisant entre deux interrogations.
Vivier d'adresses propriétaires : l'institut constitue une base d'adresses avec des individus ayant participé à une ou plusieurs autres études menées par l'institut	Acceptable
Panel	Acceptable
Access panel	À utiliser avec précaution
Fichiers « annuaires » enrichis par un prestataire	Acceptable
Fichiers qualifiés	À utiliser avec précaution
Recrutements de plusieurs personnes dans le même foyer	À utiliser avec précaution
Recrutements par parrainage ou boule de neige	À proscrire
Intercepts	À proscrire

Voici quelques éléments complétant le tableau ci-dessus :

Pour les viviers d'adresses propriétaires et pour les panels, le CESP rappelle que leur utilisation, si elle est acceptable sous réserve du respect des bonnes pratiques, doit toutefois rester limitée. À noter que dans certains cas, les viviers sont issus d'études de référence auditées par le CESP. Dans ce cas, leur recours est acceptable sous réserve que l'audit soit satisfaisant.

Concernant le recrutement de plusieurs personnes dans un même foyer (dans le cadre d'enquêtes "individus"), il est nécessaire de prendre en compte l'effet de grappe dans l'analyse des résultats.

Les recrutements par parrainage sont des méthodes à exclure pour les études d'audience : l'effet « boule de neige » nuit gravement à la représentativité de l'échantillon. Le Comité Scientifique du CESP a d'ailleurs recommandé de supprimer des échantillons/panels les individus recrutés via parrainage.

En ce qui concerne les recrutements « Intercepts » ou à la volée, notamment sur les réseaux sociaux : il faudrait pouvoir faire des tests approfondis pour analyser les profils et les comportements des individus recrutés via ce mode de recrutement.

Enfin, le CESP tient à préciser que le recours à des sources qualifiées doit être limité : l'échantillon recruté / interrogé sur ces sources doit rester un complément de l'échantillon total dans le cadre d'une étude d'audience.