
EFFICACITE D'UN PROTOCOLE MULTIMODE (TELEPHONE ET INTERNET)

Christophe BARRET, Mady CISSE, Émilie GAUBERT, Zora MAZARI, Manon OLARIA

() Céreq, Équipe ingénierie et gestion d'enquêtes,*

christophe.barret@cereq.fr, mady.cisse@cereq.fr, emilie.gaubert@cereq.fr,
zora.mazari@cereq.fr, manon.olaria@cereq.fr

Mots-clés : Protocole multimode, collecte, relance, téléphone, internet, sms

1. Résumé

Cette étude porte sur l'évaluation des protocoles de collecte multimode mis en œuvre dans le cadre des expérimentations en parallèle des enquêtes Générations du Céreq. Ce dispositif d'enquêtes s'intéresse à l'insertion professionnelle et au cheminement des sortants du système éducatif, tous niveaux de formation confondus, lors de leurs premières années de vie active.

L'enquête Génération est historiquement administrée avec une collecte monomode téléphone. Dans une démarche de rénovation du dispositif d'enquêtes, des expérimentations ont été réalisées en 2015 et 2016 pour envisager la possibilité d'introduire une collecte multimode (téléphone et internet). Au préalable, le dispositif historique des enquêtes Génération sera présenté afin d'introduire le protocole d'appel.

Le protocole des enquêtes multimodes, menées en parallèle de l'enquête téléphonique, a été construit autour de plusieurs phases de collecte. Dans un premier temps, l'objectif est d'obtenir des questionnaires par internet uniquement avec l'envoi de lettre-avis et de relances mails. À l'issue de cette première phase, les relances téléphoniques démarrent avec une incitation à répondre par internet tout en continuant les relances par mail. Enfin, une dernière phase dite de « rattrapage » vise à contacter tous les individus n'ayant pas encore répondu pour réaliser l'enquête par téléphone (en cas de refus, la possibilité de répondre par internet était toujours offerte). Pendant toute la collecte, les enquêtés ont la possibilité de reprendre un questionnaire commencé en changeant ou non de mode de réponse.

Une description détaillée des protocoles expérimentaux appliqués sera réalisée au travers d'un bilan de collecte des différentes phases. L'étude tentera de montrer l'intérêt de mixer un protocole multimode séquentiel avec un protocole multimode concurrentiel. De plus, une évaluation des modes de contact sera menée. Une attention particulière sera, tout d'abord, portée sur les individus ayant changé de mode de collecte en cours d'enquête afin de les quantifier et de déterminer leur exposition au risque d'abandon. Il s'agira, ensuite, d'analyser les effets sur la probabilité de contact et de réponse des lettres-avis, papier ou électronique, ainsi que des relances téléphoniques, mails et sms.

L'objectif de ce travail est de définir le futur protocole de l'enquête Génération qui permettrait d'optimiser les taux de réponse, notamment sur internet, tout en limitant les éventuelles nuisances liées à la multiplication des relances multi-supports (papier, mail, téléphone et sms) auprès des individus.

2. Abstract

The Generation surveys main objective is to facilitate a regular evaluation of the education-to-work transition over the first three years of the young people's working lives. Experiments were conducted on the surveys of the 2010 and 2013 cohorts as part of the system's overhaul programme. The aim of these experiments was to investigate the impact of adding a self-administered, web-based data gathering tool on the quality of the responses. This article discusses the effects of a three-step data collection process on the response rate. This process switched from a fully web-based data collection phase during the first weeks to a competitive second phase, where people could choose either web or phone. A phone-only catch-up phase was introduced in the last weeks to convince non-respondents to take part in the survey.

Introduction

Le Centre d'études et de recherches sur les qualifications (Céreq) mène depuis à présent vingt ans le dispositif des enquêtes Générations qui s'intéresse à l'insertion professionnelle et au cheminement des sortants du système éducatif, tous niveaux de formation confondus, lors de leurs premières années de vie active. Ces enquêtes sont historiquement administrées avec une collecte monomode téléphone.

Depuis 2012, le Céreq est dans une démarche de réflexion pour rénover son dispositif d'enquêtes Génération. Cette initiative résulte, d'un travail de refonte à caractère scientifique et d'une adaptation technologique visant à moderniser le questionnaire. L'implémentation d'internet, dans la collecte historiquement téléphonique, a, par ailleurs, d'autres avantages. Il permet notamment de toucher une génération de jeunes connectés, peu joignables par téléphone, ou résidant à l'étranger, non couverts actuellement par l'enquête. Il offre également la possibilité de réduire le budget du dispositif, puisque le coût marginal d'une enquête réalisée sur internet est moindre comparé à une enquête réalisée par téléphone.

Des expérimentations ont donc été lancées en parallèle des enquêtes Génération afin d'appréhender l'impact de l'introduction d'internet comme deuxième mode sur la collecte et de définir à terme un protocole d'enquête multimode (notamment de contact, de relances, ...) adapté en favorisant la participation des jeunes. L'objet de l'étude est d'analyser spécifiquement l'efficacité des protocoles mis en œuvre au cours des différentes expérimentations produites, notamment, en mettant en évidence l'intérêt de mixer un protocole multimode séquentiel avec un protocole multimode concurrentiel.

En préambule, l'expertise du protocole de contact de l'enquête téléphonique sera présentée notamment la qualité des numéros de téléphone issus de la base de sondage et de la phase d'enrichissement des coordonnées. Les effets du protocole d'appel sur la probabilité de contact et de réponse seront également mis en évidence en particulier en termes de nombre de tentatives d'appel pour chaque numéro disponible.

Par la suite, les protocoles multimodes mis en place et leurs variantes seront présentés puis feront l'objet d'une évaluation quant à leur efficacité. Ainsi, des modèles seront mis en œuvre pour tester l'effet des lettres-avis et des différentes formes de relances sur la probabilité de contact et de réponse. Enfin une analyse, inhérente à la spécificité d'une collecte multimode, sur le comportement des individus ayant changé de mode de réponse sera également proposée.

1. Le Céreq et le dispositif des enquêtes Génération monomode téléphone

1.1. Le Centre d'études et de recherches sur les qualifications

Créé en 1971, le Céreq est un établissement public sous la double tutelle du ministère de l'Éducation nationale et du ministère du travail. Il est un expert majeur dans le domaine de la relation formation-emploi depuis près de 50 ans. Le Céreq s'interroge sur des questions liées au rôle de la formation initiale et continue dans les parcours professionnels, à l'insertion professionnelle des jeunes, à l'évolution des métiers, du travail, des compétences et des certifications. Il a également pour objectif de développer la connaissance des liens entre formations, travail et emploi ainsi que d'accompagner les acteurs de la formation et du travail en les outillant dans la mise en place de dispositifs d'évaluation ou d'observation. Pour cela trois thèmes d'études principaux sont abordés en combinant plusieurs approches méthodologiques : la transition de l'école à l'emploi, la formation tout au long de la vie et l'évolution des métiers. Le Céreq s'interroge également sur les conditions de l'acquisition des compétences par la formation initiale et continue, par l'exercice d'une activité professionnelle. Ses recherches se tournent autour des conditions d'accès aux emplois, des conditions de la mobilité professionnelle et sociale, de l'évolution des qualifications liées aux transformations des technologies et de l'organisation du travail. Aussi, il évalue les formations, les dispositifs et politiques publiques mis en œuvre afin d'éclairer les décideurs et acteurs en formulant des avis et propositions en matière de politiques de formation et d'enseignement. Pour cela, le Céreq travaille en partenariat avec douze centres associés implantés dans des laboratoires de recherche universitaire ou du CNRS, lui permettant d'approfondir les analyses sur les dynamiques formation-travail-emploi et de développer des recherches.

1.2. Le dispositif des enquêtes Génération

En 1997, le Céreq met en place un dispositif d'enquêtes original permettant d'étudier l'accès à l'emploi des jeunes à l'issue de leur formation initiale. Tous les trois ans, une nouvelle enquête est réalisée auprès de jeunes sortis du système éducatif une même année donnée quel que soit le niveau ou le domaine de formation atteint, d'où la notion de "Génération". Il s'agit d'une enquête téléphonique (CATI : Computer-Assisted Telephone Interviewing). L'enquête permet de reconstituer les parcours des jeunes au cours de leurs trois premières années de vie active et d'analyser ces parcours au regard notamment du parcours scolaire et des diplômes obtenus.

L'enquête Génération peut comprendre plusieurs interrogations. La première interrogation, trois ans après la sortie du système éducatif, est principalement tournée vers la production d'indicateurs et des usages institutionnels. Les interrogations suivantes, à cinq et sept ans, sont plutôt centrées sur les usages analytiques (notamment sur la question des parcours, des mobilités sur moyen terme). Ces ré interrogations permettent aussi d'approfondir certains constats issus de l'exploitation de la première interrogation (par exemple sur les insertions les plus problématiques).

L'enquête repose principalement sur un calendrier qui décrit mois par mois la situation des jeunes et sur la collecte d'informations plus précises concernant le premier emploi et l'emploi occupé au bout de trois années passées sur le marché du travail. Ce dispositif permet non seulement d'analyser les trajectoires d'entrée dans la vie active mais aussi de distinguer, d'une génération à l'autre, les aspects structurels et conjoncturels de l'insertion.

1.3. Protocole de contact et d'appel

Constituer une base d'élèves nécessite une collecte de fichiers auprès des établissements de formations initiales dispensées en France entière (y compris les départements d'outre-mer). Or, il n'existe pas de base d'élèves nominative centralisée couvrant l'ensemble des formations. C'est pourquoi, la base de sondage de chaque enquête Génération est constituée à partir de différentes

sources telles que des bases rectorales, la base du système d'information sur la formation des apprentis, de bases universitaires, et de bases récupérées directement auprès des établissements. Ainsi, les coordonnées, indispensables pour joindre les individus (mail(s), adresse postale, téléphone(s)), dépendent des données recueillies par les établissements scolaires lors de l'inscription des étudiants. Au moment de l'enquête, ces données, datant de trois ans, peuvent être obsolètes (ex : coordonnées du logement étudiant, des parents). C'est pourquoi, une phase d'enrichissement des coordonnées des individus composant l'échantillon est nécessaire pour les actualiser.

Tout au long de l'enquête, les individus ont la possibilité de mettre à jour leurs coordonnées en se connectant sur le site dédié à l'enquête (via le lien et avec identifiants fournis dans la lettre-avis) ou en contactant la hotline. De plus, lorsqu'un tiers répond à l'appel téléphonique, il est possible de collecter un voire deux nouveau(x) numéro(s) de l'individu recherché (un portable et/ou un fixe, *le cas échéant*). Le protocole d'enrichissement est reproduit à l'identique pour les enquêtes expérimentales multimode et le sera également pour la généralisation en 2020.

Une analyse de l'efficacité de cette phase est disponible en Annexe 1.1.

Le protocole d'appel mis en place sur l'enquête Génération impose au prestataire de brasser l'ensemble du fichier au cours du premier mois, c'est-à-dire que l'ensemble des individus sont appelés au moins une fois au cours de ce mois. Le protocole d'appel implique également que vingt tentatives d'appels soient réalisées sur chaque numéro de téléphone tant que l'individu n'a pas répondu à l'enquête ou n'a pas exprimé un refus définitif de participer. Cependant, si le numéro est détecté comme faux (message de l'opérateur indiquant que le numéro n'existe pas) ou abouti à un hors cible, les tentatives d'appels cessent sur ce numéro et on passe automatiquement au suivant. Une analyse du protocole d'appel est disponible en Annexe 1.2.

Ce protocole inclut également la possibilité de laisser un message sur répondeur. Il est déposé lorsqu'un numéro est appelé au moins six fois sans succès (répondeur, occupé ou ne répond pas). L'objet du message explicite rapidement l'importance de la participation à l'enquête et indique le numéro vert pour prendre rendez-vous ou réaliser l'enquête. Les règles de rappels sont disponibles en Annexe 1.2.

L'alternance d'appels en numéro masqué et démasqué (soit une semaine sur deux, soit un jour sur deux) avec, en cas de rappel de l'enquêté, un contact direct avec le plateau téléphonique via le numéro vert dédié à l'enquête, a été intégré depuis peu au protocole. L'objectif est de toucher des publics aux comportements différents face à un appel d'un inconnu. Certains individus ne répondront pas à un numéro masqué mais répondront à un numéro affiché inconnu (d'autant plus si celui-ci est régionalisé) et inversement. En revanche, il est très difficile d'évaluer l'efficacité de cette alternance, en particulier en raison de difficulté à obtenir les parodonnées nécessaires.

Ce protocole d'appel téléphonique est appliqué quasiment à l'identique aux expérimentations multimodes, ce qui permettrait en théorie de déduire un effet « pur » de l'introduction d'internet comme mode de contact et de collecte supplémentaire. En pratique, c'est plus compliqué que cela. En effet, malgré les efforts des prestataires pour respecter le protocole imposé, dans le cadre des enquêtes expérimentales, le niveau d'insistance des appels n'atteint pas le même niveau que pour une enquête principale jugée prioritaire. Pour la généralisation, ce problème ne se posera plus et le niveau d'insistance des appels téléphoniques sera rigoureusement identique au standard des enquêtes Génération.

2. Description des protocoles de collecte multimode

Des expérimentations ont été développées en parallèle des enquêtes principales monomode téléphonique afin d'introduire une collecte alternative par internet, jugée plus moderne et moins coûteuse. Ce mode offre des solutions ergonomiques, qui permet de compenser au moins en partie l'absence d'un enquêteur, car il offre une certaine interactivité et un soutien à l'enquêté via des liens

hypertexte, des info-bulles, des argumentaires ciblés, des foires aux questions (FAQ)... Il s'avère par ailleurs être, pour l'enquêté, un mode peu intrusif qui lui permet de répondre au moment où il le souhaite et peut aussi limiter les biais de désirabilité sociale. Cependant, une partie de la population n'a pas accès ou est peu encline à répondre par internet, certains peuvent présenter des difficultés en lecture difficilement compatibles avec ce mode de collecte autoadministré. Ces populations possèdent de plus des caractéristiques socioéconomiques particulières (Bibliographie). Le maintien de la collecte par téléphone semble donc essentiel pour couvrir l'ensemble du public des enquêtes Génération.

Après une première expérimentation de collecte monomode internet, deux expérimentations multimodes ont suivi, sur une première interrogation et sur une ré interrogation.

2.1. Expérimentation monomode internet

Une première enquête expérimentale monomode internet, avec un protocole 100 % web tant sur le mode de contact que sur le mode de collecte, a été initiée en 2013 pour observer le comportement des jeunes vis-à-vis d'un questionnaire sur internet. Conscients des difficultés énoncées dans la littérature, cette approche permettait de faire un état des lieux des enjeux à venir dans la réflexion de l'écriture du questionnaire d'une enquête Génération en première interrogation, du développement de l'outil informatique de collecte multimode en visant particulièrement l'ergonomie du CAWI et de la définition des protocoles de contact et de relances.

Le questionnaire de l'enquête monomode internet était identique à celui de l'enquête usuelle. Seuls les protocoles de contact et de relances ont été adaptés. Un premier mail avis informatif pour annoncer le démarrage de l'enquête est envoyé aux individus de l'échantillon disposant d'un mail. L'adresse email est soit issue de la base de sondage (seulement un quart des individus de l'échantillon disposait d'un mail) soit retrouvée via une phase d'enrichissement. En parallèle, ont été intégrés dans cette enquête expérimentale, les jeunes de l'enquête principale repérés comme résidant à l'étranger.

Un protocole de dix relances par mail, à un rythme hebdomadaire en variant les jours et les heures d'envoi, est mis en place afin d'inciter les individus à répondre, ou, pour ceux ayant commencé le questionnaire sans le finir, à inviter à le terminer. Si, en répondant à l'enquête, l'individu ressentait un besoin d'information sur l'enquête ou d'aide au remplissage du questionnaire, deux canaux de contact lui étaient offerts : une adresse email technique et une page Facebook dédiée à l'enquête. L'adresse email technique a été sollicitée à de multiples reprises pour une aide ponctuelle à la compréhension des questions, mais également pour formuler des remarques à propos de l'ergonomie du questionnaire. En conséquence, pour l'assistance en ligne, un nombre plus important d'enquêteurs que prévu a été mobilisé.

Les résultats de cette expérimentation montrent un faible taux de réponse. Les hypothèses sont multiples pour expliquer la non-réponse (Bibliographie). Le premier facteur explicatif est le premier contact réalisé sur seulement un tiers de l'échantillon, du fait de l'indisponibilité du mail sur une très grande majorité de l'échantillon et de l'échec de la phase d'enrichissement des mails qui se sont avérés de très mauvaise qualité. Au-delà des problèmes techniques, la mauvaise ergonomie du questionnaire a par ailleurs entraîné un nombre important d'abandons en cours d'enquête. Cependant, les relances mails impactent positivement le taux de réponse y compris après la cinquième relance.

L'ensemble des répondants de cette première expérimentation ont été ré interrogés dans le cadre de l'enquête à cinq ans après la sortie du système éducatif. Le questionnaire a été adapté, l'outil de collecte CAWI a été développé selon les standards du web et s'est révélé nettement plus ergonomique. Le protocole de collecte, lui, est resté inchangé : enquête monomode internet, envoi d'un mail avis personnalisé et dix relances mails sur la période d'observation. À la différence de la

précédente interrogation, le protocole intègre la possibilité de joindre le Céreq via un numéro vert pour une aide ponctuelle ou pour répondre à des questions autour de l'enquête en complément du mail de contact et de la page Facebook.

L'objectif principal de cette réédition était d'observer l'attrition dans le cadre d'un protocole 100 % web. Bien que la participation soit nettement plus affirmée que lors de la première interrogation, l'attrition demeure plus importante que pour une enquête monomode téléphone classique. En revanche, le taux d'abandon en cours d'enquête a été très nettement réduit. Cela étant sans nul doute imputable à l'amélioration sensible de l'ergonomie du questionnaire.

Le protocole de contact centré uniquement sur les relances mails ne permet pas d'aboutir à un taux de réponse satisfaisant et ne répond pas aux exigences d'une enquête Génération portant sur l'ensemble des jeunes quel que soit leur niveau de formation. En effet, les répondants aux enquêtes par internet sont classiquement issus de l'enseignement supérieur ce qui entraîne un fort biais de sélection en l'absence de mode alternatif. Ces conclusions amènent le Céreq à s'orienter vers un protocole multimode (téléphone et internet).

2.2. Expérimentations multimodes téléphone/internet

De nouvelles expérimentations dans une version multimode sont réalisées dans le cadre d'une première et d'une ré interrogation. Le protocole a été mis en place pour la première fois sur une ré interrogation. Il a ensuite été amélioré pour être testé lors d'une première interrogation. L'objectif est d'obtenir un taux de réponse proche de celui de l'enquête monomode téléphone historiquement mise en place, tout en tentant de maximiser le taux de réponse par internet. La construction de ce protocole découle donc de cet objectif. Le protocole définit une collecte multimode classique incitant à répondre en priorité par internet avec un rattrapage par téléphone.

L'ensemble des expérimentations ont été réalisées en parallèle de l'enquête monomode téléphonique. Les enquêtes multimodes intègrent une nouveauté par rapport à l'enquête monomode internet. Des individus considérés historiquement hors champ de l'enquête téléphonique ont la possibilité de répondre à l'enquête par internet uniquement, à titre exploratoire et expérimental. Ils sont ainsi basculés dans l'expérimentation multimode une fois leur acceptation de répondre par internet et à la condition d'une adresse mail valide pour réceptionner le lien de connexion au questionnaire. Les individus concernés sont résidant hors France (métropolitaine et DOM) au moment de l'interrogation. Jusqu'ici non intégrés dans le champ, le Céreq souhaite les inclure dans la prochaine enquête rénovée car ils sont considérés comme des sortants d'un établissement scolaire français. Dans le cadre de la dernière expérimentation multimode, la proposition de répondre par internet a été faite également aux post-initiaux docteurs. L'ensemble de ces individus, quel que soit le moment où ils sont basculés dans l'enquête expérimentale, ne prennent pas part aux différentes phases ; ils ne peuvent répondre à l'enquête que par internet et, à ce titre, sont relancés uniquement par mail.

Les objectifs principaux de ces expérimentations sont de mesurer l'efficacité d'un protocole de contact amélioré, d'évaluer l'adaptation du questionnaire multimode pour réduire les effets de mesure et des plans de dialogue ajustés facilitant les bascules d'un mode de collecte à un autre (cf. Annexe 2).

Le premier protocole d'enquête construit autour de trois phases distinctes dans le temps privilégie une collecte au préalable internet. Il a été mis en place dans le cadre d'une ré interrogation.

Phase 1 - Incitation à répondre par internet : une première phase de trois semaines comprend l'envoi de la lettre-avis destinée à informer du démarrage de l'enquête intégrant un lien personnalisé pour participer directement par internet.

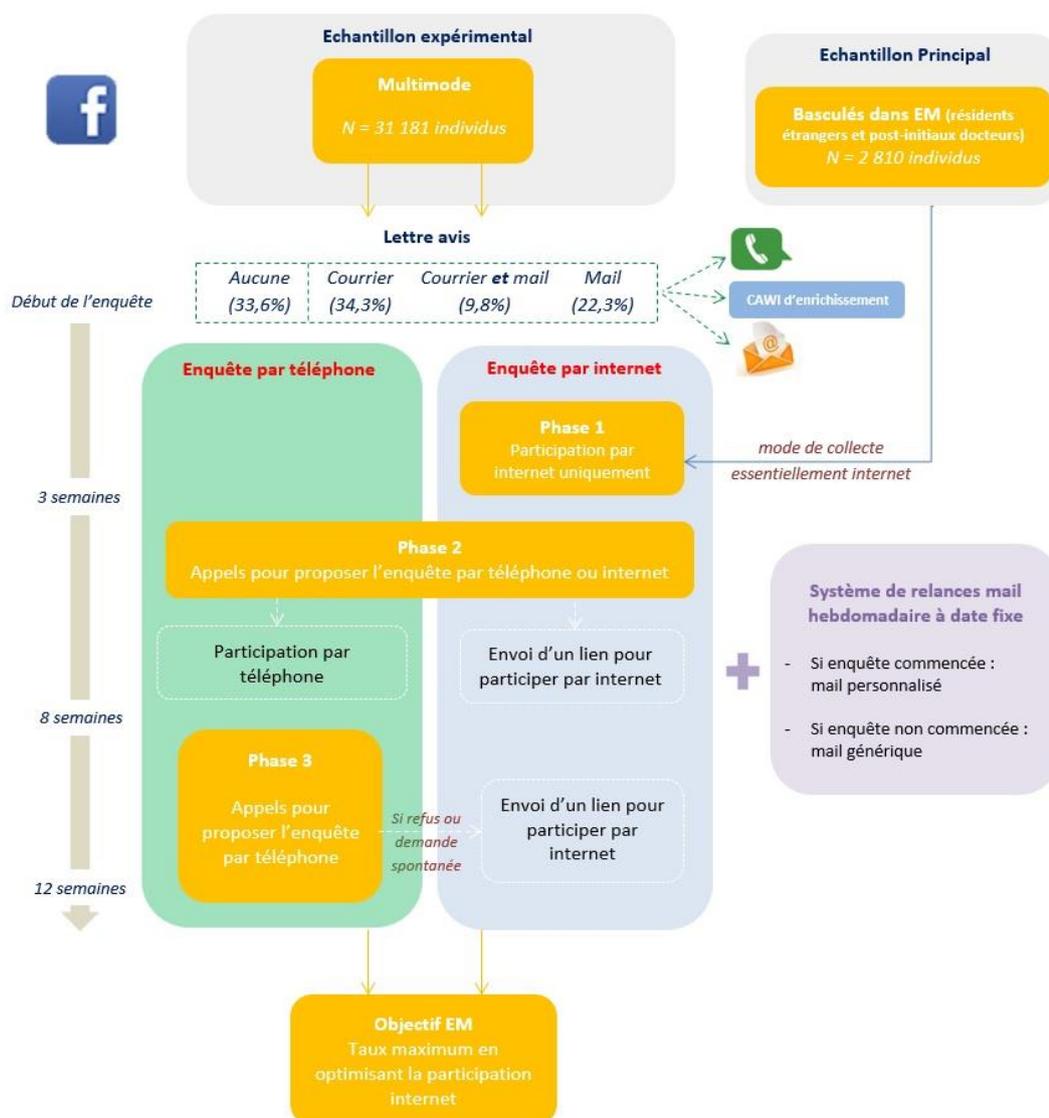
Phase 2 - Choix de répondre par internet ou téléphone : une seconde phase de cinq semaines où le plateau téléphonique est lancé. Le choix est donné aux enquêtés contactés par téléphone de participer soit par internet soit par téléphone selon leur préférence.

Phase 3 – Rattrapage pour répondre par téléphone : une dernière phase de collecte d'une durée de quatre semaines pour proposer aux enquêtés de participer par téléphone.

S'il n'y avait que la phase 1 et une phase 3 « pur » téléphone, on serait dans le cadre d'un protocole multimode séquentiel. L'introduction de la phase 2, où le choix est laissé à l'internaute, et le fait que lors de la phase 3 il soit toujours possible de répondre par internet (suite à une relance mail ou suite à un refus de répondre par téléphone), le protocole multimode retenu est de type concurrentiel. L'objectif est de maximiser les réponses sur internet, mode de collecte le moins coûteux. Le protocole séquentiel ne conduirait pas forcément à un taux de réponse général dégradé mais la part des répondants par internet serait plus réduite.

Ce protocole a servi de base pour la construction du protocole de la deuxième expérimentation multimode et a été adapté pour répondre aux caractéristiques d'une première interrogation. 33 991 individus composent l'échantillon multimode ; il inclut 2 810 individus hors champ de l'enquête monomode téléphone historique qui participent à l'enquête multimode à titre expérimental. Ces individus particuliers mis à part, l'échantillon multimode est soumis aux différentes étapes du protocole détaillées ci-dessous.

Figure 1 - Logigramme du protocole de l'enquête multimode¹



Source : Enquête expérimentale multimode menée en 2016 auprès de la Génération 2013, Céreq

¹ Les statistiques de la lettre-avis ne concernent que les individus de l'échantillon expérimental. Les individus de l'échantillon principal ont reçu un mail pour les inviter à participer à l'enquête multimode avec le lien de connexion.

2.2.1. Première phase de collecte essentiellement internet

L'enjeu de cette phase est de tenter d'informer l'ensemble de l'échantillon et les inviter à répondre au questionnaire de l'enquête par internet. Cette phase permet d'obtenir les réponses spontanées sur internet sans introduire d'animation téléphonique (hormis la hotline). L'utilisation de la lettre-avis sous plusieurs formes permet à la fois de contacter des individus disposant d'une adresse mail et ceux dont seule l'adresse postale est connue. Pour ces derniers un flash-code sur la lettre-avis permet de se connecter directement au questionnaire.

La lettre-avis a une double utilité : inviter l'individu à participer directement à l'enquête (identifiants de connexion fournis) ou à corriger ses coordonnées téléphoniques avec la possibilité de prendre un rendez-vous pour être contacté.

Chronologiquement, une lettre-avis courrier, est envoyée aux individus disposant d'une adresse postale mais pas d'adresse mail dans le fichier d'import. Un mail avis est ensuite envoyé à tous les individus présentant une adresse mail dans l'échantillon. Les envois de mails sont échelonnés dans le temps et espacés de quelques jours (deux lots d'envoi) afin de réguler d'une part, l'éventuelle charge de sollicitation sur le numéro vert qui est actif dès le début d'envoi des lettres-avis et d'autre part, d'éviter que l'envoi de mails, massif compte tenu du volume, soit blacklisté par certains opérateurs. L'espace personnel de l'individu est quant à lui actif dès le lendemain de l'envoi de la lettre-avis papier afin que l'individu puisse se connecter dès réception du courrier et/ou mail.

À la différence des précédentes expérimentations, une voire deux lettres-avis sont transmises lorsqu'une adresse mail et une adresse postale sont disponibles ; avec l'hypothèse d'un meilleur effet sur la probabilité de contact (cf. Figure 1). 27,7 % de l'échantillon expérimental disposent des deux informations, mais le choix a été fait de n'envoyer les deux formes de lettres-avis qu'à 42,3 % d'entre eux selon une règle aléatoire.

Dans la suite du document, les statistiques englobent l'ensemble des individus de l'échantillon multimode.

Après transmission de la lettre-avis, 24,4 % des individus de l'échantillon (expérimental et principal) ont reçu uniquement un mail avis, 33,3 % uniquement un courrier avis et 10,2 % les deux versions de la lettre. Parmi les envois de la lettre-avis courrier, une partie non quantifiable pour l'enquête multimode² peut être imputée à des NPAI (N'habite Pas à l'Adresse Indiquée). Au total, ce premier contact n'a pas abouti pour un jeune sur trois pour des raisons d'indisponibilité de l'adresse mail et/ou postale et d'invalidité de l'adresse mail (badmails) représentant pour ce dernier cas 18 % d'entre eux.

Près de la moitié des contactés (45,5 %)³ se sont connectés une première fois au questionnaire par internet. Parmi eux, 74 % remplissent le questionnaire en intégralité (sans relance téléphonique). En raison des difficultés à identifier la période de connexion au questionnaire sur internet, ce chiffre englobe l'ensemble de la durée d'enquête, des individus ayant pu se connecter sur internet en phase 2 à un moment où ils étaient injoignables par téléphone.

Au cours de cette première phase, les appels ne sont pas passés pour les individus disposant d'une adresse mail valide afin de leur laisser un temps suffisant pour réaliser une réponse spontanée. Pour les autres, ceux dont l'adresse mail est invalide ou indisponible, des tentatives d'appels sont réalisées dès cette phase. Plus précisément, ces individus démarrent la seconde phase de manière anticipée (une semaine plus tôt que les autres individus).

² Les NPAI représentent 13,8% des envois de courrier pour l'enquête principale (monomode téléphone)

³ Un contacté est un individu qu'il a été possible de joindre au téléphone et/ou qui s'est connecté sur internet.

2.2.2. Seconde phase de collecte multimode

Le protocole de cette seconde phase correspond à l'ouverture du plateau téléphonique (hors hotline ouverte depuis le début) pour l'ensemble des individus et donc introduit clairement le mode de collecte téléphonique en parallèle d'internet. 26,5 % des contactés sont pour la première fois appelés par téléphone dans la phase 2.

L'expérimentation monomode internet a montré ses limites en termes de taux de réponse lorsqu'internet seul est utilisé pour la collecte, et de biais de sélection par la présence du mail dans la base de sondage, largement majoritaire pour les diplômés de l'enseignement supérieur. L'introduction du téléphone permet donc de contacter des publics non atteignables lors d'une enquête monomode internet et de les inciter à répondre par internet.

Ainsi, le mode de collecte par internet est proposé à tout individu contacté par téléphone pour la première fois. L'incitation à répondre à internet peut être abordée de différentes manières que le Céreq a testées dans ses expérimentations. Pour la première expérimentation multimode, internet a été proposé directement à l'individu et d'une manière assez imposée. À l'écoute, ce plan de dialogue est apparu trop brusque, il a donc été rapidement assoupli, mais internet était toujours proposé dès la phase de contact. La formulation améliorée permet de s'assurer du consentement de l'enquêté à répondre par internet pour éviter au maximum que ce mode de réponse puisse être une forme de refus déguisé (seuls 15 % des individus choisissant le mode de collecte par internet ont finalement participé lors de cette expérimentation). Pour la seconde expérimentation internet – téléphone, l'idée était de construire un plan de dialogue sur le même principe que celui reformulé lors de la première enquête : une alternative entre internet et téléphone avec insistance modérée d'internet. La particularité de la seconde expérimentation est d'être une première interrogation, avec comme premier module à l'issue de la phase de contact, le module visant à déterminer si l'individu est éligible à l'enquête. Deux options pour le protocole étaient possibles ; soit proposer à l'individu de réaliser l'enquête sur internet directement lors de la phase de contact, comme ce fut le cas pour la première interrogation multimode ; soit déterminer si l'individu est dans le champ, puis si c'est le cas, lui proposer de poursuivre l'enquête sur internet.

Les deux options présentées dans le cadre d'une première interrogation présentent chacune leurs avantages et inconvénients. La première option peut, d'un côté, s'avérer moins coûteuse puisque l'individu peut réaliser l'ensemble du questionnaire sur internet, y compris la partie filtre qui élimine les individus hors champ de l'enquête. Or, pour une première interrogation, l'adhésion à l'enquête lors de la phase de contact (très courte) semble ne pas être suffisante, se traduisant par un refus de répondre sur internet ou par un refus déguisé pour les individus acceptant de répondre sur internet. La phase de relance et de rattrapage téléphonique sera plus importante et plus difficile, et donc plus coûteuse.

C'est pourquoi la seconde option a été mise en place dans le cadre de l'enquête multimode de 2016. Elle permet de renforcer l'adhésion de l'individu à l'enquête par le remplissage des premières questions (qui font référence à son dernier établissement scolaire) et de le convaincre de poursuivre. La proposition de poursuivre sur internet seulement aux individus qui sont dans le champ évite de relancer inutilement – par mail et par téléphone – les individus qui n'auraient pas répondu par internet (contrairement au choix émis lors de la conversation téléphonique) et qui ne sont pas dans le champ de l'enquête.

Plus précisément, le protocole mis en place en phase 2 est le suivant. À chaque appel téléphonique, l'individu répond en premier lieu à quelques questions pour identifier si l'individu est éligible. S'il est qualifié (i.e., dans le champ), il peut soit poursuivre l'entretien par téléphone, soit participer par internet. Dans ce dernier cas, l'enquêteur valide ou saisit une nouvelle adresse mail et transmet en direct un mail d'invitation contenant un lien personnalisé pour répondre à l'enquête sur internet.

Un plateau visant à brasser au maximum les numéros de téléphone disponibles est mis en place au début de cette phase. Pour rappel, chaque numéro de téléphone est appelé jusqu'à vingt fois dans le but de maximiser les chances de contact et de réponse.

Au sein de la phase 2, 61,3 % des individus identifiés dans le champ souhaitent poursuivre l'enquête par internet. Parmi eux, seuls 14,5 % terminent effectivement l'enquête sur internet ; 20,3 % ont dû être rappelés car ne se connectant pas ou ne finissant pas l'enquête sur internet et ont réalisé l'enquête par téléphone.

2.2.3. Troisième phase de collecte essentiellement téléphone

Cette dernière phase de collecte se concentre sur des appels téléphoniques pour inviter les individus à participer à l'enquête par téléphone. Quel que soit le nombre et le type de relances reçues ou le stade du remplissage du questionnaire, chaque individu de l'échantillon est contacté pour commencer ou terminer l'enquête. 28 % des contactés sont concernés par cette dernière phase du protocole multimode.

Une adaptation du plan de dialogue est réalisée pour poursuivre l'enquête directement par téléphone avec l'individu, autrement dit, le mode internet n'est plus mentionné. Cela étant, la possibilité de répondre par internet est toujours possible, soit sur demande spontanée de l'individu, soit proposé en ultime recours par l'enquêteur en cas de refus de l'individu de répondre par téléphone. Comme sur toute la période de la collecte, l'enquêteur peut à tout moment transmettre un mail d'invitation à participer à l'enquête par internet. Cela a été le cas pour 132 individus de la phase 3 avec seulement 3 répondants.

2.3. Système de relances mails et SMS

En parallèle des relances téléphoniques propres à chaque phase, des relances mails hebdomadaires sont programmées ; les mails sont soit génériques dans le cas de questionnaire non commencé, soit personnalisés dans le cas de questionnaire commencé. La relance peut également se faire ponctuellement par SMS.

En complément de la lettre-avis, un système de relances est organisé autour du statut de l'enquêté une semaine après la réception de celle-ci. Pour les individus n'ayant pas encore répondu à l'enquête et disposant d'un mail valide, une relance mail hebdomadaire est envoyée variant le jour et l'heure d'envoi. Sinon, pour ceux dont l'adresse mail est invalide ou indisponible, des tentatives d'appels sont réalisées. L'objectif est de vérifier l'identité de la personne recherchée et son appartenance au champ de l'enquête. Un mail de connexion au questionnaire est immédiatement transmis après avoir collecté l'adresse mail. Dans le cas où l'individu ne dispose pas de mail ou refuse de participer à l'enquête par internet, il est possible de répondre directement par téléphone. 38,7 % des individus dans le champ n'ont pas souhaité poursuivre sur internet.

Pour les individus ayant commencé leur questionnaire par internet sans le terminer, des relances mail personnalisées pour terminer l'enquête sont transmises. Lors de la première expérimentation, ces relances personnalisées s'effectuaient à $J + 4$ et $J + 8$, J étant la dernière date de connexion. À cela s'ajoutait une relance téléphonique à $J + 12$ pour terminer le questionnaire par téléphone, sauf si l'enquêté demandait à terminer le questionnaire par internet. Lors de la deuxième expérimentation, ce protocole jugé complexe a été simplifié. Le principe des mails personnalisés est maintenu mais ils sont envoyés à un rythme hebdomadaire (le même jour que les mails de relance génériques). Pour information, pour ceux l'ayant commencé par téléphone, et indépendamment des phases, l'enquêteur suit le protocole de l'enquête téléphonique, à savoir fixer un rendez-vous pour terminer l'enquête.

Pour maximiser les chances de réussite, une ultime relance a été programmée par SMS sur une population cible de l'enquête n'ayant pas encore répondu. Le contenu du SMS présente en quelques mots le sujet de l'enquête et propose le numéro vert pour une future participation.

2.4. Expérimentation multimode téléphone/internet : échantillon embarqué

Un protocole alternatif a également été expérimenté, à visée principalement méthodologique. La mise en place d'un échantillon embarqué, de taille réduite, permet de comparer les réponses de deux populations aux caractéristiques sociodémographiques similaires. L'objectif est de disposer de deux populations identiques, sans déformation due aux préférences des individus pour un mode de réponse afin de supprimer l'effet de sélection, pour pouvoir apprécier des différences de réponse des individus entre le mode, le cas échéant.

Sur cet échantillon, le mode de collecte des individus a été déterminé de manière aléatoire afin de leur être imposé. Un individu qui répond spontanément par téléphone, alors que le mode imposé est internet, est invité à participer à l'enquête par internet et inversement.

L'affectation du mode de collecte se faisant après l'acceptation de répondre au questionnaire, ce protocole permet de gommer les effets de sélection liés au mode de contact. Toutefois, le taux d'abandon en cours est plus important que celui de l'enquête multimode classique, dû à l'imposition de changement de mode de collecte (pour ceux qui ne sont pas spontanément sur le bon mode).

3. Efficacité du protocole

Les prochains résultats sont issus de l'exploitation de la dernière enquête multimode de 2016. L'enquête multimode de 2015 est également utilisée lorsque des données sont indisponibles ou partielles pour réaliser certaines analyses. À noter que les deux protocoles sont très proches, cependant ils ont été appliqués à une première interrogation (2016) et une ré interrogation (2015).

L'objet de l'étude est d'étudier principalement l'efficacité du protocole de collecte en trois phases. Quel est le comportement des individus face au possible changement de mode ? Quel est l'impact sur la probabilité de réponse ? Quel est l'effet du type de lettre-avis, courrier ou numérique, sur la probabilité de contact ? Quel est l'effet des relances et spécifiquement de l'envoi de SMS sur le taux de participation ? L'exploitation du fichier historique d'appels et de connexions, à l'aide de modélisation, est la source de données principale permettant d'établir les effets positifs (ou négatifs) sur la probabilité de contact et de réponse à l'enquête.

3.1. Étude des changements de mode de collecte

L'enquête multimode propose aux individus de répondre par internet, par téléphone, ou en utilisant les deux modes. Plusieurs cas de figure sont envisageables du point de vue du mode de connexion au questionnaire. Soit l'individu commence et termine son enquête sur le même mode. Soit l'individu commence l'enquête sur un mode, puis au moins une connexion suivante se fait sur un autre mode. Il peut s'agir d'un individu qui a commencé à répondre aux premières questions de l'enquête sur internet, mais l'interrompt en cours ; dans ce cas, des tentatives de relance par téléphone sont réalisées afin de terminer l'enquête avec un enquêteur. Il peut également s'agir d'un tiers indiquant que celui-ci n'est pas disponible. Après collecte des coordonnées mails de l'individu recherché, un mail lui est envoyé pour que l'enquête soit réalisée par internet. Dans cette partie consacrée à l'étude de l'utilisation de plusieurs modes de collecte, une enquête sera dite « commencée » à partir du moment où, par téléphone l'appel sera décroché, et par internet, l'individu se sera connecté au questionnaire.

Il est donc possible que les individus aient commencé leur enquête sur un mode et le termine sur un autre, voire même cumulent plusieurs changements de mode tout au long de l'enquête. À titre d'exemple, un individu a changé jusqu'à 7 fois de modes de collecte, en commençant par internet et basculant entre internet et téléphone pour finir sur internet.

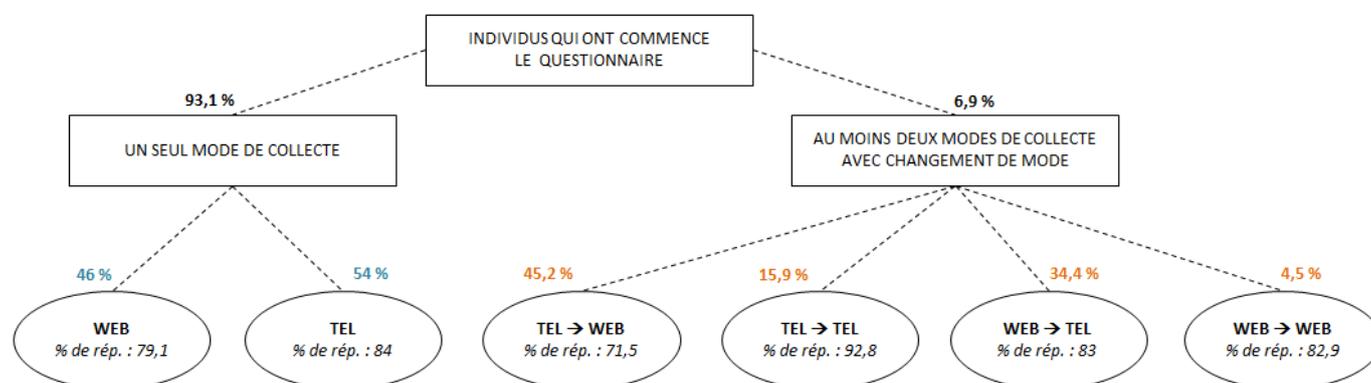


Figure 2 - Répartition des individus ayant commencé le questionnaire selon le mode de collecte initial et final

Source : Enquête expérimentale multimode menée en 2016 auprès de la Génération 2013, Céreq

Parmi l'ensemble des individus qui ont commencé à répondre au questionnaire, 93,1 % ont participé à l'enquête (partiellement ou totalement) sur un seul mode⁴ et 6,9 % par les deux modes de collecte.

79,3 % des individus dont la participation à l'enquête croise internet – téléphone ont terminé l'enquête (individus dans le champ et questionnaire terminé, et individus hors champ).

3.1.1. Bascule d'internet vers téléphone

La bascule entre internet et téléphone est souvent involontaire et correspond à des cas de rattrapage d'enquêtes commencées et non terminées. 38,9 % des individus qui ont changé de mode de collecte sont concernés par une bascule d'internet en mode initial vers le téléphone, à un moment donné de la collecte.

Plus de 11,5 % des individus se connectent pour la dernière fois au questionnaire sur internet. Il s'agit dans la moitié des cas d'individus à qui un mail est envoyé dès la phase de contact pour répondre sur internet (hors champ). Il peut également s'agir d'individus qui en cours d'enquête, décident de ne plus poursuivre par téléphone.

Parmi les individus ayant basculé d'internet vers le téléphone, 83 % ont terminé l'enquête, qu'ils soient dans le champ ou non.

Les individus se connectent en moyenne 3,7 fois au questionnaire mais 44,3 % des individus ne réalisent que deux connexions au questionnaire, soit internet puis téléphone. Les autres individus (55,7 %) enchaînent plusieurs connexions consécutives sur le même mode et/ou basculent à nouveau d'un mode à l'autre.

⁴ À titre indicatif, 54 % participent à l'enquête par internet et 46 % par téléphone. 85 % des individus participant à l'enquête avec un seul mode le font sur le mode en une seule fois, par exemple l'individu se connecte sur internet et termine l'enquête directement.

3.1.2. Bascule de téléphone vers internet

Dans cette situation, 61,1 % des individus qui ont changé de mode de collecte, ont commencé par téléphone et ont par la suite effectué un changement vers internet. Dans 91,4 % des cas, la bascule s'effectue juste après avoir été appelé. Les 8,6 % restants sont essentiellement des individus qui acceptent de réaliser ou poursuivre l'enquête sur internet, mais qui se révèlent être des refus déguisés pour limiter l'intervention de l'enquêteur. Le délai dépassé, ils sont rappelés pour réaliser l'enquête par téléphone, sauf indication immédiate de l'individu de l'effectuer sur internet.

La dernière participation à l'enquête se fait finalement par internet pour 74 % d'entre eux, dont 1 % qui bascule plusieurs fois entre internet et téléphone pour finir sur internet. 26 % des individus sont rattrapés sur le mode initial téléphone, enchaînant dans 94 % des cas *a minima* téléphone – internet – téléphone.

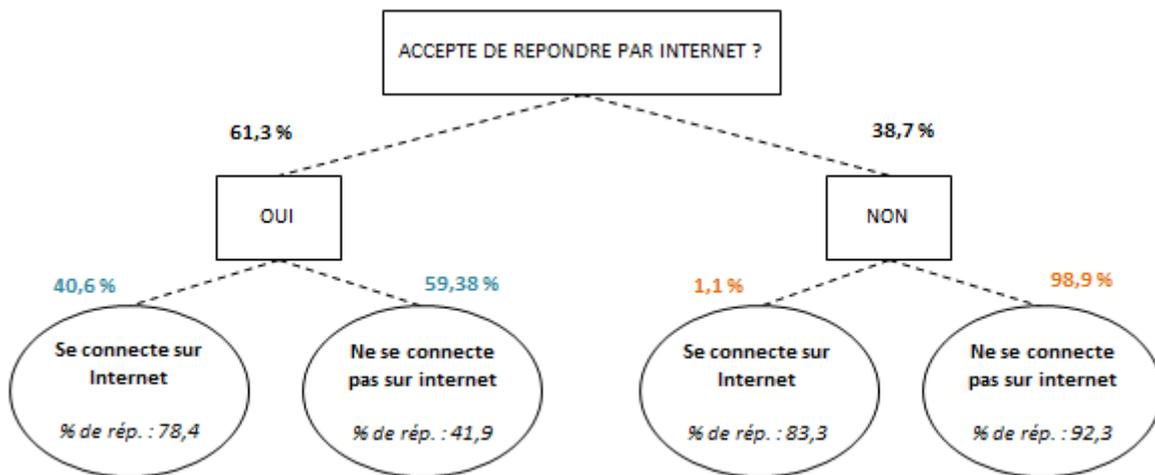
Par ailleurs, 77 % des basculés de téléphone vers internet terminent l'enquête et ils le font essentiellement sans être à nouveau sollicité par un enquêteur.

Les individus utilisent en moyenne 3,2 fois les modes internet et téléphone pendant la durée de l'enquête.

3.1.3. Étude de la mise en concurrence des deux modes de collecte, avec insistance sur internet

En phase 2 de l'enquête, chaque individu contacté par téléphone se voit proposer le choix de répondre de répondre par téléphone ou internet, tout en priorisant le mode internet.

Figure 3 - Accepte de répondre par internet ?



Source : Enquête expérimentale multimode menée en 2016 auprès de la Génération 2013, Céreq

Ainsi, 61,3 % disent accepter de basculer sur internet, et pourtant, seuls 40,6 % d'entre eux se sont réellement connectés au questionnaire. Cette proportion, non négligeable, met en lumière une distinction nette entre les intentions de bascules et les individus qui changent effectivement de mode de collecte. Cet écart entre le basculement énoncé et le basculement réel traduit un phénomène de refus déguisé de répondre à l'enquête. Parmi les individus qui ne se sont pas connectés par internet alors qu'ils ont manifesté leur intention de le faire, 41,9 % ont finalement, après appel d'un enquêteur, terminé l'enquête. Par contre, 78,4 % de ceux qui ont respecté leur intention de basculer sur internet ont participé à l'enquête que ce soit par téléphone ou par internet.

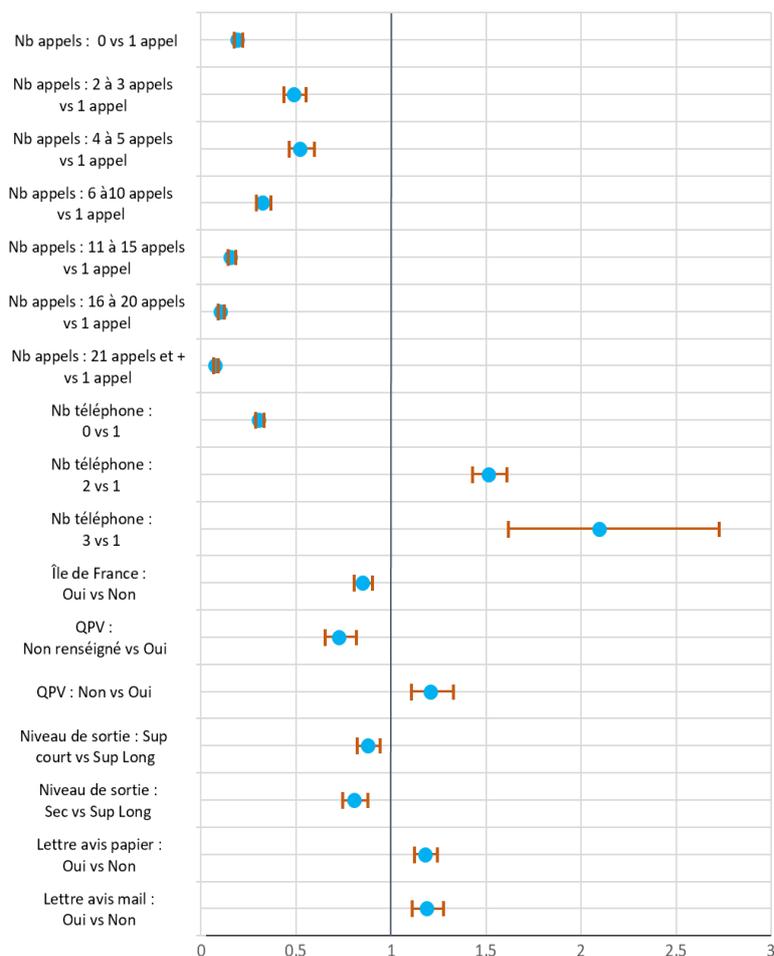
3.2. Analyse de la probabilité de contact

Le protocole de contact téléphonique mis en œuvre dans l'enquête multimode est quasi identique à celui d'une enquête monomode historique. À cela s'ajoute l'envoi des lettres avis transmises sous différentes formes à tous les individus disposant d'un mail ou d'une adresse postale valides. Quels sont les effets positifs (ou négatifs) de ce protocole sur la probabilité de contact ?

Pour définir le premier contact dans une enquête multimode, des conditions spécifiques sont déterminées selon le mode de collecte. Par téléphone, un contact est la réponse à un appel avec identification de l'individu ou d'un tiers (proche). Par internet, un contact est matérialisé par la connexion au questionnaire individuel avec des identifiants fournis dans la lettre-avis.

Afin d'estimer et expliquer la probabilité de contact, un modèle de régression logistique a été utilisé. Des variables individuelles issues de la base de sondage ont été introduites telles que le sexe de l'enquêté, le niveau de sortie, la région de l'établissement de formation ainsi que le lieu de résidence (niveau quartier). À cela ont été ajoutés les indicateurs de contact suivants : la réception d'une lettre-avis mail, l'envoi d'une lettre-avis courrier (sans certitude de réception), l'addition des deux types de lettres-avis, le nombre de numéro de téléphone disponibles et le nombre total d'appels effectués. Le fait de recevoir un courrier avis et un mail avis semble ne pas être significatif ainsi que le sexe de l'individu et sont ainsi retirés du modèle. D'autres indicateurs de type tracking (paradonnées) tels qu'ouvrir le mail avis ou cliquer sur le lien de connexion, n'ont pas été utilisés car l'information disponible est partielle dans l'historique d'appels.

Figure 4 – Représentation graphique des odds ratios du modèle logistique sur la probabilité de contact sur l'ensemble de l'échantillon



Source : Enquête expérimentale multimode menée en 2016 auprès de la Génération 2013, Céreq

Le premier facteur explicatif et qui influe le plus sur la probabilité de contact est le nombre de numéros de téléphone disponible. Le choix du nombre de numéros envoyés en production est limité à 3 à la différence de l'enquête monomode qui est de 8. L'objectif étant de parvenir à respecter un protocole d'appel, intégrant le mode de collecte par internet avec alternance de relances mails, sur les trois mois de collecte. Toutefois de nouveaux numéros peuvent être enrichis en cours d'enquête. Ainsi, la probabilité de contact augmente suivant le nombre de numéros de téléphone disponible. Un individu, avec au maximum trois numéros de téléphone, a un « risque » 2,1 fois plus élevé d'être contacté que celui qui n'en dispose qu'un seul, toutes choses égales par ailleurs. Le positionnement du numéro de téléphone qui a permis le premier contact est également un critère qui peut être significatif. Néanmoins, la construction de cet indicateur de qualité du numéro de téléphone est difficile par manque de précision dans l'historique d'appels de cette enquête.

Le protocole d'enquête prévoit de contacter les individus de l'échantillon par tous les moyens (papier, mail, tel et sms) pour limiter les biais de sélection liés au mode de contact. S'agissant du nombre de tentatives d'appels, plus il augmente plus les chances de contact diminuent. Cet effet est mécanique mais interroge tout de même sur la faiblesse des chances d'aboutir à un contact au-delà de la dixième tentative d'appel (tous numéros confondus).

Recevoir un mail ou un courrier avis influe positivement et identiquement (oddsratio=1,2) sur la probabilité d'être contacté. Par ailleurs, l'impact de la lettre-avis sur le mode de collecte varie selon le type de lettre reçue comme l'indique le tableau suivant.

Tableau 1 - Premier contact au questionnaire selon le type de lettre-avis reçue

Mode de premier contact parmi les individus contactés	Aucune	Mail	Papier	Mail + Papier
Internet	21,0	75,3	23,0	76,1
Téléphone	79,0	24,7	77,0	23,9
Total	100	100	100	100

Source : Enquête expérimentale multimode menée en 2016 auprès de la Génération 2013, Céreq

Ainsi, parmi les individus contactés, le premier contact est dans trois cas sur quatre réalisé sur internet suite à la réception d'un mail avis. À l'inverse, lorsqu'il reçoivent uniquement une lettre-avis papier ou aucune lettre-avis (faute de coordonnées disponibles) alors le premier contact s'opère, dans près de 80 % des cas, par téléphone. Lorsque les individus reçoivent les deux types de lettre-avis (mail et papier) le comportement est proche de ceux qui reçoivent uniquement un mail avis.

Le fait d'être contacté dépend également du lieu de résidence. Résider en région parisienne ou encore plus dans un quartier prioritaire de la politique de la ville (QPV) amenuisent les chances de contact. Les probabilités de contact sont également plus élevées pour les sortants des formations du supérieur par rapport aux sortants du secondaire.

Pour améliorer les chances de contact, la lettre-avis est un élément indispensable pour annoncer le lancement de l'enquête. La réception simultanée au format mail et courrier à caractère officiel doit avoir un effet rassurant pour déclencher la participation. En amont de l'enquête, la procédure d'enrichissement des numéros et leur classement, en fonction de la qualité, sont une condition *sine qua non* à la réussite de l'enquête.

En définitive, 35 % de l'échantillon a été approché la première fois par mail, 33 % par courrier, le complément par téléphone. La part des jeunes ayant effectivement débuté le questionnaire par téléphone est de 53,7 %, le complémentaire par internet. Ce résultat montre un certain équilibre dans le choix du mode de collecte dès le premier contact.

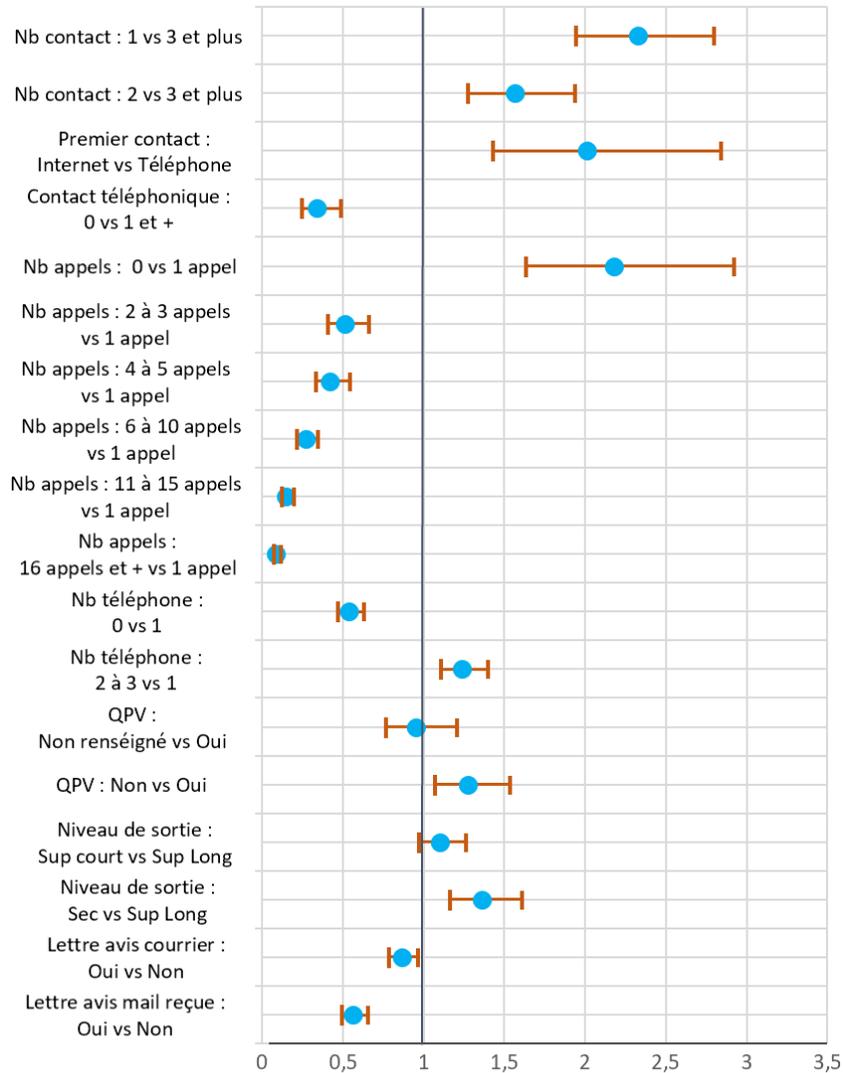
3.3. Analyse de la probabilité de réponse

Le protocole d'enquête s'appuie sur différents leviers pour optimiser le taux de réponse. L'objectif, dans le cadre de l'enquête expérimentale de 2016, est de parvenir, avec un protocole multimode, à un taux de réponse au moins équivalent à celui d'une enquête monomode téléphonique. L'intégration du mode de collecte internet implique la mise en place d'un système de relances spécifique combinant téléphone et mail. De plus, des relances de type SMS sont également prévus. Quels sont les effets de la multiplication des relances sous différentes formes sur la probabilité de réponse ?

Pour mesurer l'efficacité de ce protocole, un modèle de régression logistique a été utilisé sur la base des individus contactés au moins une fois. Les variables introduites sont de même nature que celles du modèle précédent lié à la probabilité de contact. D'autres indicateurs ont été intégrés tels que : le nombre de contact par téléphone ou par mail, le mode de premier contact, le mode de dernier contact, avoir eu au moins un contact par téléphone, avoir eu au moins un contact par internet, avoir reçu un message laissé sur répondeur et avoir pris un rendez-vous par téléphone ou par mail. Ce modèle permettant de mesurer que partiellement l'effet des relances mails et en l'absence

d'information sur les relances sms, une évaluation est réalisée sur l'expérimentation multimode de la ré interrogation de 2015 (cf. Figure 6 - Effet des relances sur la réponse en fonction du mode).

Figure 5 – Représentation graphique des odds ratio du modèle logistique sur la probabilité de répondre parmi les contactés



Source : Enquête expérimentale multimode menée en 2016 auprès de la Génération 2013, Céreq

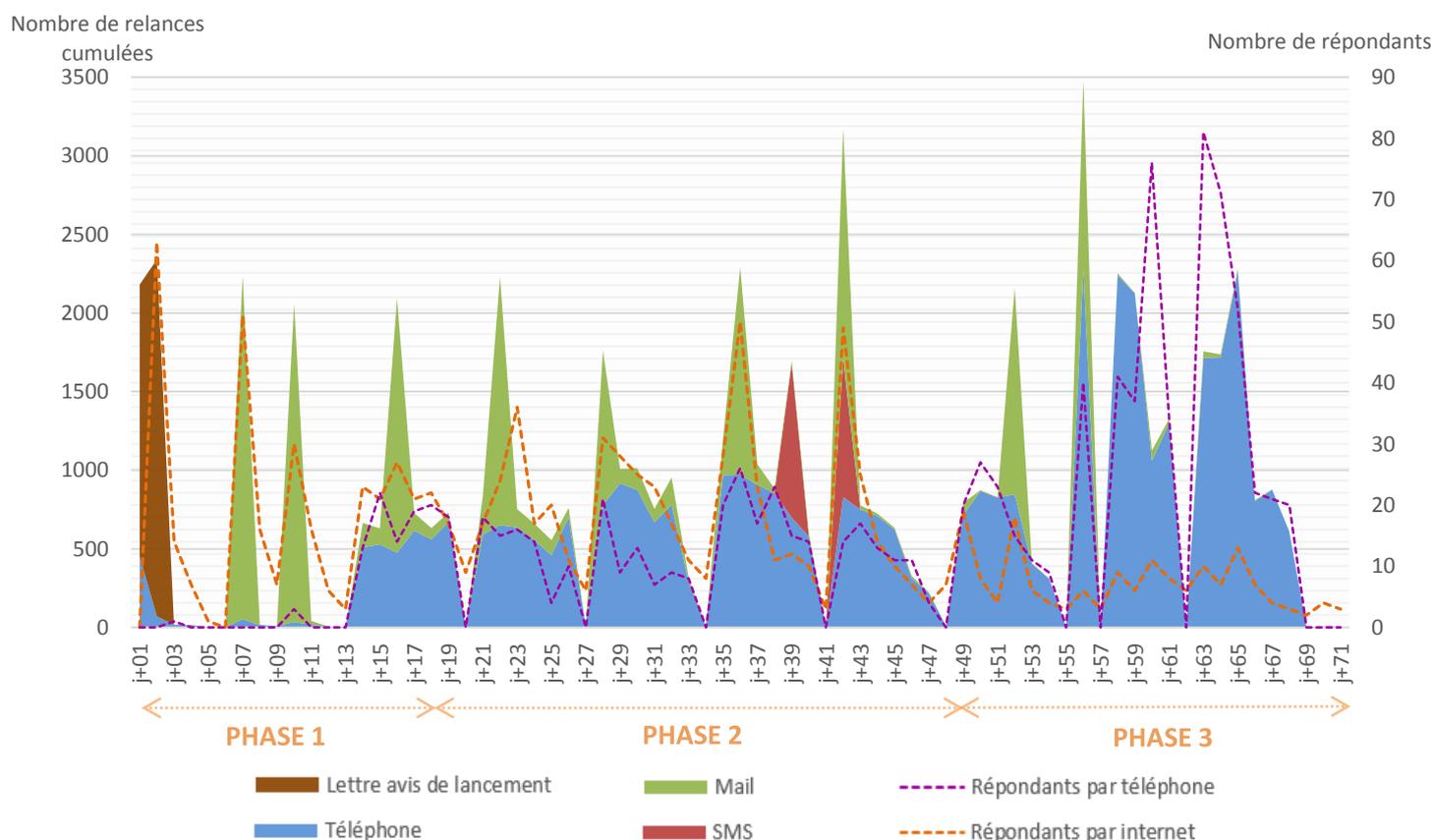
Le nombre de numéros de téléphone disponible joue favorablement sur la probabilité de réponse mais l'effet est moins prononcé que pour la probabilité de premier contact. Tout comme le modèle précédent, plus le nombre de tentatives d'appels est élevé, plus la probabilité de répondre diminue. Toutefois, le fait de ne pas être appelé double les « chances » de répondre par rapport à un individu appelé une seule fois. Ce dernier résultat traduit simplement le fait que les individus qui n'ont pas été appelés et qui ont été contactés sont ceux qui se sont connectés au questionnaire internet. De plus, ces individus vont plus souvent concrétiser leur connexion par une réponse complète au questionnaire que ceux contactés par téléphone.

Parmi les critères significatifs, recevoir un mail ou un courrier avis, qui jouait positivement sur le premier contact, semble ici influencer négativement sur la probabilité de réponse. Il faut cependant relativiser ce constat avec le résultat précédent puisqu'une partie de l'effet des lettres-avis est captée par la modalité « 0 appel » qui joue quant à elle très favorablement sur la probabilité de réponse par rapport aux individus appelés.

Concernant les caractéristiques des individus, les sortants du niveau secondaire contactés répondraient plus souvent que leurs homologues du supérieur. Ce résultat contre-intuitif est lié au modèle de réponse sur les individus contactés. Il peut s'expliquer par l'absence d'adresse mail dans la base de sondage pour cette catégorie d'individus. Par conséquent, étudiants de la première phase de la collecte uniquement internet, ils sont appelés plus fréquemment sur le reste de la période d'observation. L'appel téléphonique ou plutôt l'effet enquêteur reste un levier efficace pour la réalisation d'une enquête. Il faut relativiser ce paradoxe en rappelant que les sortants du supérieur sont beaucoup plus facilement contactés que ceux du secondaire. Ainsi, en faisant un modèle de réponse en une étape sur tout l'échantillon (contactés et non contactés), les sortants de l'enseignement supérieur ont effectivement, toutes choses égales par ailleurs, plus de chances de répondre à l'enquête que leurs homologues du secondaire.

La probabilité de réponse est également expliquée par les relances multimodales. Ces effets sont estimés à partir des données de l'expérimentation multimode de 2015. Il s'agit ici de mettre en évidence l'intérêt de mixer un protocole multimode séquentiel avec un protocole multimode concurrentiel.

Figure 6 - Effet des relances sur la réponse en fonction du mode



Source : Enquête expérimentale multimode menée en 2015 auprès de la Génération 2010, Céreq

Le graphique, ci-dessus, fait apparaître les différents types de relances utilisés en cours de collecte. Les données représentées sont cumulées et journalières. Pour exemple, le jour 16, environ 2 000 relances sont lancées dont 1 500 mails transmis et 500 appels téléphoniques réalisées. Les courbes

affichent le nombre de répondants selon le mode de collecte. Un axe secondaire a été ajouté pour en simplifier leur lecture.

Des relances spécifiques sont prévus selon la phase de collecte variant les jours et les heures. Au démarrage de l'enquête, les premiers courriers avis sont envoyés, s'en suivent les mails avis en plusieurs lots espacés dans le temps (heures) afin de ne pas saturer le numéro vert. L'effet de la lettre-avis est très rapidement visible par un pic de connexions spontanées au questionnaire par internet. Une petite activité téléphonique existe, liée aux appels entrants sur le numéro vert, pour rassurer et informer les individus sur l'enquête.

Lors de la phase 1, des relances mails hebdomadaires sont mises en place pour inciter à répondre par internet. À chaque vague de mails envoyé, de nouveaux questionnaires par internet sont remplis et validés. Toutefois, le nombre de répondants diminue au cours des trois premières semaines d'enquête, l'effet des relances mails s'essouffle, conformément au constat déjà observé sur l'expérimentation monomode internet de 2015.

Pour favoriser la réponse par ce mode de collecte, des relances téléphoniques sont introduites pour produire une forme d'« animation » faite par le téléenquêteur incitant à poursuivre ou participer par internet ; bien que lui soit proposé le choix du mode de collecte. Cela génère ainsi un regain de répondants internet dont le nombre se stabilise durant la phase 2.

Pour les individus ne disposant d'aucun mail, des appels téléphoniques démarrent dès la dernière semaine de phase 1. Le nombre de répondants par téléphone s'accroît au même moment. L'alternance des appels et des relances mails, au cours de la phase 2, permet de collecter des questionnaires par téléphone et par internet de manière relativement équilibrée. Mais surtout, offre la possibilité de contacter des individus aux caractéristiques différentes quel que soit la propension à répondre à l'un ou l'autre des modes de collecte.

Pour booster le nombre de répondants et offrir une ultime chance de répondre par internet, une tentative de relance par sms est faite, en fin de phase 2, sur une partie des individus non répondants, ayant commencé à répondre ou non à l'enquête auparavant. Environ 2 000 individus sont sélectionnés et contactés sur deux jours ouvrés (lundi et vendredi). Un modèle sur cette expérimentation a également été construit pour tester l'effet du SMS de relance, il semble indiquer un effet bénéfique sur la probabilité de répondre mais paradoxalement cet effet n'est pas immédiatement perceptible mais visible dans les jours et semaines qui suivent. En fin de collecte, un répondant sur trois a reçu un sms additionnel au système de relances en vigueur.

En phase 3, les relances mails sont arrêtées. Le protocole prévoit un rattrapage massif par téléphone. Ainsi, très peu de questionnaires sont réalisés par internet. En parallèle, le volume des appels téléphoniques double en conséquence de l'augmentation des effectifs d'enquêteurs suite au basculement des enquêteurs rattachés à l'enquête principale monomode téléphonique qui se termine. L'impact de cet accroissement s'observe également sur le nombre de questionnaires réalisés. Étant donné le nombre de répondants la dernière semaine de plateau, il est vraisemblable que le protocole n'a pas atteint son maximum et qu'il y avait encore un potentiel de répondants à l'issue du plateau. Autrement dit, si le plateau téléphonique avait été prolongé de quelques semaines, le taux de réponse final aurait été plus important. Les trois semaines de la phase 1 empiètent sur l'efficacité du protocole d'appel. L'insistance téléphonique a été moindre que pour l'enquête monomode téléphone (20 tentatives par numéro disponible). Pour maximiser le taux de réponse, la période d'observation pourrait être prolongée d'au moins trois semaines pour respecter le niveau d'insistance standard des enquêtes Génération lors de la généralisation en 2020 (ou à durée constante prévoir des effectifs d'enquêteurs plus importants). Le degré d'insistance des appels conjugué aux nombreuses relances mails doivent cependant être dosés au mieux pour éviter un sentiment d'« overdose » qui serait contre-productif.

Tableau 2 – Répartition des répondants selon le mode de collecte et les phases du protocole (en %)

Phase	Internet	Téléphone
1	31,4	9,6
2	54,7	35,1
3	13,9	55,4
Total	50	50

Source : Enquête expérimentale multimode menée en 2015 auprès de la Génération 2010, Céreq

Le protocole mis en place a permis d'obtenir autant de répondants sur internet que par téléphone (répartition 50/50). En faisant l'hypothèse, comme évoqué plus haut, d'un prolongement du plateau téléphonique en phase 3, le taux de réponse aurait été proche de celui de l'enquête principale monomode téléphonique de 2015 (60 %). Cette hypothèse semble réaliste et cohérente avec la littérature (Bibliographie) : le taux de réponse multimode est analogue à celui de l'enquête monomode. Avec cette hypothèse, la répartition entre les deux modes aurait été légèrement différente (40 % par internet et 60 % par téléphone).

Si l'on observe la répartition par phase, on constate tout l'intérêt de la phase 2 qui permet d'obtenir plus de la moitié des répondants internet (54,7 %).

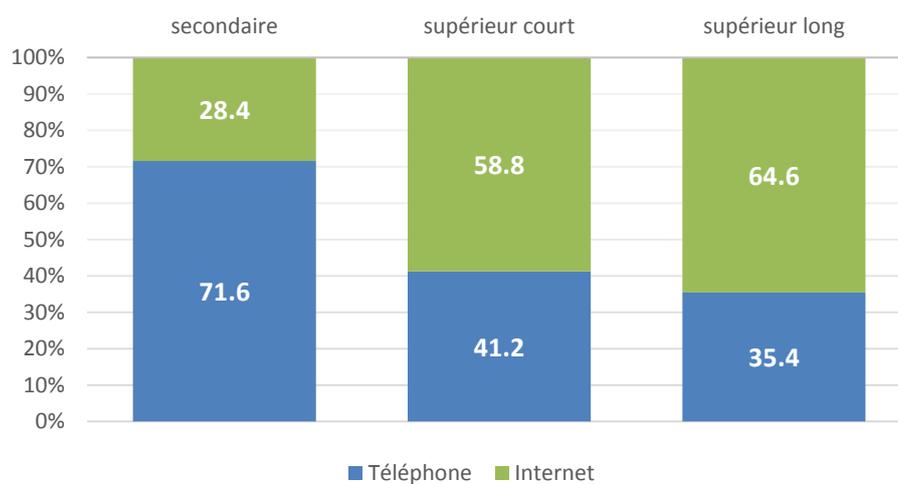
3.4. Un effet de sélection important lié au mode de collecte

Dans une collecte multimode telle que présentée ici, les individus ont la possibilité et le choix de répondre sur leur mode préférentiel. Ainsi les profils des répondants sur un mode peuvent sensiblement varier du profil des répondants sur le mode alternatif. C'est l'effet de sélection, brièvement abordé dans cette partie (pour plus de détails, se reporter à Bibliographie).

Sur la base des répondants à la dernière enquête expérimentale multimode, quelles sont les spécificités des individus faisant le choix de participer à un mode plutôt qu'à l'autre ? Pour estimer la probabilité de répondre à l'enquête par internet plutôt que par téléphone, un modèle de régression logistique a été utilisé. Plusieurs variables individuelles ont été introduites afin de caractériser le profil des répondants internet versus téléphone : le sexe de l'enquêté, l'âge à la sortie des études, le niveau de sortie, la région de l'établissement de formation. À cela, ont également été intégrés des indicateurs sociodémographiques tels que le mode de cohabitation, le lieu de résidence (niveau quartier) et la situation professionnelle à la date de l'enquête.

Les répondants par internet se caractérisent principalement par leur niveau de sortie de formation initiale. Les sortants du supérieur répondent plus souvent par internet que ceux du secondaire (cf. Figure 7 - Représentation graphique du mode de collecte des répondants selon le niveau de sortie). Pour les sortants du supérieur, ce sont les plus jeunes qui répondent par internet alors que pour ceux issus de l'enseignement secondaire, ce sont les plus âgés. Les femmes sont également plus disposées à répondre sur internet que les hommes à caractéristiques identiques. Le lieu de résidence du répondant influe sur son mode de réponse. En effet, si l'individu réside dans un quartier prioritaire de la politique de la ville, la tendance est plutôt de participer par téléphone. Le mode de cohabitation de l'individu impacte également la probabilité de répondre par internet. Un individu se déclarant vivre seul répond plus souvent par internet que celui vivant chez ses parents.

Figure 7 - Représentation graphique du mode de collecte des répondants selon le niveau de sortie



Source : Enquête expérimentale multimode menée en 2016 auprès de la Génération 2013, Céreq

Conclusion

Le Céreq, dans le cadre de la rénovation des enquêtes Génération, a décidé de basculer en 2020 vers une collecte multimode internet-téléphone. L'introduction d'internet comme mode de collecte supplémentaire présente plusieurs intérêts : un gage de modernité, un mode particulièrement adapté à la cible de l'enquête, les jeunes, une augmentation potentielle de la couverture en touchant des individus difficiles à joindre sur le mode historique (téléphone), une source d'économie puisqu'une enquête sur internet est moins coûteuse qu'une enquête passée avec un enquêteur par téléphone.

Cela étant ce mode de collecte impose un certain nombre de contraintes pour éviter de dégrader la qualité de l'enquête. Tout d'abord, il est nécessaire de développer une bonne ergonomie de questionnaire et pas uniquement pour l'aspect esthétique. En effet, cela permet de limiter les abandons en cours et également de limiter en partie les effets de mesure avec le téléphone (gestion des modalités « ne sait pas », des questions à choix multiples, des consignes, des définitions-précisions, des argumentaires, ...).

Le multimode, dans le sens où deux modes de collecte sont disponibles, impose également un certain nombre de contraintes. En particulier, il est important de prévoir un protocole de collecte adéquat, qui garantit à la fois de maximiser le taux de réponse global tout en optimisant le nombre de répondants sur le mode le moins onéreux, à savoir internet. Ainsi, il est nécessaire d'assurer une bonne gestion des passages d'un mode à l'autre à n'importe quel moment du questionnaire et ce autant de fois que possible. Cela nécessite d'organiser les bascules avec des plans de dialogue adaptés pour reprendre le questionnaire en cours en fonction des différentes situations possibles. Même si la très grande majorité des répondants n'utilise qu'un seul mode de collecte, il ne faut pas négliger ce dispositif qui permet de limiter la non-réponse, de surcroît sur des profils potentiellement atypiques.

Grâce aux trois expérimentations menées en parallèle des enquêtes principales, le Céreq a pu tester cinq protocoles de collecte différents, allant du monomode internet à l'affectation aléatoire d'un mode de collecte en passant par des protocoles multimodes dit « classiques ». Ces derniers combinent un protocole multimode séquentiel, intégrant plusieurs phases de collecte, avec un multimode concurrentiel, proposant le choix du mode de réponse. Les enseignements sont multiples. Une enquête 100 % web pour une enquête de cette ampleur ne permet pas d'atteindre des taux de réponse satisfaisant et conduit à une population trop sélectionnée (Razafindranovona, 2015). Il faut donc *a minima* introduire une « animation » téléphonique pour inciter à répondre sur internet. Par ailleurs, les plans de dialogue qui imposent (obligatoirement) de basculer de mode (notamment dans le sens téléphone vers internet) sont jugés trop abrupts par les enquêtés et conduisent à une forte non-réponse sur ces individus.

Le multimode classique apparaît donc comme le meilleur compromis. Il permet de garantir des taux de réponse analogues aux standards des enquêtes Génération tout en optimisant le nombre de réponse sur internet. Il s'agit d'un protocole en trois phases dont la première permet d'obtenir les réponses spontanées sur internet ; la seconde correspond au démarrage du plateau téléphonique avec un choix laissé aux enquêtés de répondre sur internet (avec envoi immédiat d'un mail personnalisé) ou de poursuivre par téléphone ; et enfin la troisième peut être qualifiée de rattrapage téléphonique puisque le choix du mode n'est plus proposé spontanément (mais uniquement en cas de refus ou de demande de l'enquêté). Tout au long de l'enquête et quelles que soient les phases, tous les supports de communication sont utilisés pour provoquer le contact avec l'individu (lettres et mails avis, mails de relances, SMS et appels téléphoniques).

La deuxième phase est essentielle pour obtenir un nombre satisfaisant de questionnaires internet. Pour autant, elle fait peser un risque sur le taux de réponse général. En effet, certains individus voient la proposition de répondre sur internet comme une échappatoire à l'enquête téléphonique et de ce fait demandent à basculer sur ce mode pour déguiser le refus de poursuivre le questionnaire. Si

ces individus sont « rattrapés » par la suite grâce aux relances mails personnalisés ou les rappels téléphoniques pour finir l'enquête, cela n'est pas dommageable. En revanche, si ces relances échouent, il s'agit d'une perte pour l'enquête car sur un protocole uniquement téléphone, la proportion d'individus qui seraient parvenus au terme du questionnaire aurait été plus importante grâce au rôle de l'enquêteur chargé de mener l'individu à renseigner le questionnaire jusqu'à son terme. Cet effet négatif sur le taux de réponse global est compensé par l'augmentation potentielle de la couverture citée plus haut sans que l'on puisse dire quel effet domine. Ainsi, un ajustement sur la durée des phases, la deuxième pouvant être raccourcie et la troisième rallongée, permettrait d'aboutir au respect complet du protocole d'insistance des appels téléphoniques.

Un autre point de vigilance concernant les enquêtes multimodes internet et téléphone concerne le besoin accru en données. En effet, pour construire des indicateurs de suivi et d'analyse de qualité de l'enquête, des données supplémentaires sont nécessaires et doivent être anticipées au moment de la rédaction du cahier des charges concernant le développement du questionnaire (CATI et CAWI). En particulier, il est indispensable de savoir le mode de réponse utilisé pour chaque question puisque l'enquêté peut basculer d'un mode à l'autre en cours de remplissage de questionnaire. Par ailleurs, il est important de disposer d'un fichier historique des événements datés, qui intègre toutes les interactions entre le gestionnaire d'enquête et l'enquêté dans les deux sens (d'une part, envoi de courriers, mails, sms, appels téléphoniques sortants et d'autre part les ouvertures de mail, les connexions au questionnaire individuel, les envois d'un mail ou appel vers la hotline, ...). De telles données sont utiles pour améliorer sensiblement le traitement de la non-réponse et par ailleurs permettent d'évaluer au mieux le protocole mis en place.

Bibliographie

- [1] De Peretti G., Razafindranovona T., « Les enquêtes multimode : multi-solution ou multi-problème? », 45^{ème} Journées de Statistique de la SFDS, 2013
- [2] Barret C., Dzikowski C., « La collecte par internet est-elle l'avenir des enquêtes Génération du Céreq? », 12^{ème} Journées de Méthodologie Statistique, 2015
- [3] Dillman D.A., Smyth J.D., Melani Christian L., « Internet, mail, and mixed-mode surveys - The tailored design method », *Third Edition*, 2009.
- [4] Barret C, Cissé M., « Agrégation de données multimode : Impact sur la modélisation des variables présentant un effet de mesure », 13^{ème} Journées de Méthodologie Statistique, 2018
- [5] Dillman D.A., Phelps G., Tortora R., Swift K., Kohrell J., Berck J., Messer B.L., « Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet », *Elsevier -Social Science Research*, may 2008.
- [6] Fripiat D., Marquis N., « Les enquêtes par Internet en sciences sociales : un état des lieux », *I.N.E.D | Population*, vol 65, pp 309-338, 2010.
- [7] Barret C., Dzikowski C., « Protocoles multimodes effet de mesure et agrégation », Colloque francophone sur les sondages, 2016.
- [8] Razafindranovona T., « La collecte multimode et le paradigme de l'erreur d'enquête totale », *Document de travail Méthodologie statistique*, n°2015/01, Insee, 2015.

Annexes

1. Expertise du protocole de contact de l'enquête téléphonique monomode

L'ensemble des informations dont on dispose dans la base de sondage sont des données issues directement des établissements de formation. Ainsi, la présence des coordonnées dont on a besoin pour joindre les individus (mail(s), adresse postale, téléphone(s)) dépend des données recueillies par chaque établissement. De plus, lorsque ces coordonnées sont disponibles, elles sont anciennes de trois ans et peuvent ainsi ne plus être correctes (ex : coordonnées du logement étudiant, des parents). C'est pourquoi, une phase d'enrichissement des coordonnées est nécessaire pour les actualiser.

Bien qu'un enrichissement de mails soit possible, le Céreq ne réalise pas cette procédure car trop onéreuse pour le peu de résultats qu'elle fournit.

C'est pourquoi, nous réalisons uniquement l'enrichissement des coordonnées postales et téléphoniques. Le premier consiste, via un prestataire spécialisé, à une restructuration, normalisation et validation postale (RNVP) et à une actualisation des adresses des individus ayant déménagés. Il est nécessaire en amont du second afin d'obtenir des résultats optimaux. Le second se fait en deux étapes. Premièrement, en amont de l'enquête via un prestataire spécialisé. Deuxièmement, nous laissons la possibilité aux enquêtés d'actualiser eux même leurs coordonnées téléphoniques une semaine avant et pendant la durée de l'enquête via leur espace personnel, disponible sur le site dédié à l'enquête, à l'aide de leurs identifiants fournis dans les lettres-avis.

1.1. Évaluation de la qualité des numéros de téléphone

L'enrichissement des coordonnées téléphoniques a pour but de maximiser les chances de contacts et ainsi de favoriser un taux de réponse maximum. L'objectif principal est de collecter des numéros de téléphone pour les individus n'en disposant pas et compléter avec de nouveaux numéros les coordonnées téléphoniques déjà présentes dans la base.

Un protocole de recherche spécifique par étapes successives est défini pour récupérer un maximum de numéros de téléphone avec un taux de fiabilité acceptable. La recherche peut aboutir à la collecte d'un ou plusieurs numéros supplémentaires qui peuvent être liés soit à l'individu, soit à un membre de sa famille. Les deux sources d'informations utilisées par le prestataire sélectionné est l'annuaire de France Télécom et une base de données contenant des numéros de téléphones mobiles et des numéros spéciaux en lien avec les lignes dégroupées (nommée base partenaire).

Etape 1 - Recherche de numéro fixe : une première recherche est réalisée, dans l'annuaire France Télécom, sur la base des individus pour lesquels les critères nom, prénom et adresse complète aboutissent à un seul écho. Pour pallier aux éventuelles imprécisions des données individuelles, une entorse aux critères est faite en intégrant de légers écarts d'orthographe sur le nom, le prénom ou l'adresse, voire en l'absence du prénom ou de l'adresse.

Etape 2 - Recherche de numéro portable : une seconde recherche est lancée, dans la base partenaire, sur la base des individus avec les mêmes critères de sélection que la phase précédente.

Etape 3 - Recherches élargies : dans la dernière recherche, seuls les individus exclus des deux phases précédentes sont considérés. Ils sont recherchés dans les deux sources de données. Dans un premier temps, le critère de la localisation géographique est levé pour augmenter le nombre d'échos et par conséquent les possibilités de recueil de nouveaux numéros. La recherche se fait donc sur nom de famille et prénom au sein de la commune, puis est élargie au département, puis à la région, puis à la France entière. Parmi les individus sélectionnés, la présence d'homonyme est donc possible. Dans un second temps, la recherche est effectuée en conservant uniquement le

nom de famille et la commune puis est élargie au département. Ainsi parmi les individus sélectionnés, la présence de membres de la famille est possible.

L'échantillon, constitué pour la dernière enquête téléphonique monomode, renseignait au moins un numéro de téléphone pour 71 % des individus. Après enrichissement, 89,5 % des individus de l'échantillon disposent d'au moins un numéro de téléphone soit 18,5 % de plus qu'au départ. De plus, l'enrichissement a permis de retrouver des numéros de téléphone dont nous disposions initialement dans la base de sondage. Ainsi, 22,1 % des numéros de téléphone présents dans la base de sondage ont été confirmés. Dans ce dernier cas, pour l'analyse qui suit, l'origine du numéro reste base de sondage.

Le Céreq détermine l'ordonnancement des numéros de téléphone, par ordre de fiabilité, en fonction des critères de qualité, afin d'augmenter la probabilité de contact de chaque individu. La logique de classement est de prioriser le numéro de téléphone portable :

1. Sélection du numéro de téléphone portable issu de la base de sondage
2. Sinon sélection du numéro de téléphone portable issu de la phase d'enrichissement en numéro de portable
3. Sinon sélection du numéro de téléphone fixe issu de la phase d'enrichissement en numéro fixe
4. Sinon sélection du numéro téléphone fixe issu de la base de sondage
5. Sinon sélection du numéro de téléphone (fixe ou portable) issu de la recherche élargie

Bien que pendant plusieurs années le numéro de fixe était préféré au portable dans le but de privilégier un endroit calme et isolé pour répondre dans de meilleures conditions à l'enquête, le portable est à présent priorisé au fixe. En effet, il est considéré comme un outil permettant d'obtenir plus facilement et surtout plus rapidement une enquête. Ainsi, au lancement de l'enquête, les numéros de téléphone disponibles pour les individus de l'échantillon se répartissaient de la manière suivante :

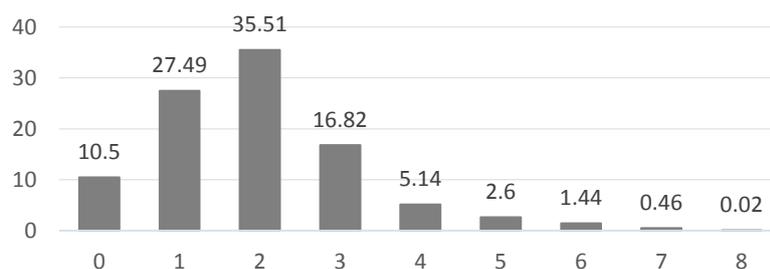
Tableau 3 - Répartitions des numéros de téléphones disponibles dans la base parmi les individus en possédant au moins 1

Source des numéros de téléphone	Nombre d'occurrence possible par individu	% d'individus en possédant au moins
Base de sondage	1 à 3	71
Numéro fixe issu de l'étape 1 de la recherche	1	14
Numéro de portable issu de l'étape 2 de la recherche	1	22
Recherches élargies issues de l'étape 3	1 à 7	34

Source : Enquête téléphonique menée en 2016 auprès de la Génération 2013, Céreq

Chaque individu pouvait avoir un maximum de 8 numéros comme le montre le graphique ci-après.

Figure 8 - Répartition du nombre de numéros disponible par individu (en %)



Source : Enquête téléphonique menée en 2016 auprès de la Génération 2013, Céreq

En moyenne, 1,9 numéro était disponible par individu.

Au cours de l'enquête, les individus ont la possibilité de se connecter sur le site du Céreq pour mettre à jour leurs coordonnées téléphoniques ; information indiquée sur la lettre-avis. De plus, lorsqu'un tiers réponds à l'appel, il est possible de collecter un voire deux nouveau(x) numéro(s) de l'individu recherché (un portable et/ou un fixe, *le cas échéant*). Ainsi, 11,8 % des individus de l'échantillon ont eu leurs coordonnées enrichis entraînant une augmentation des numéros de téléphone de 6,3 %. Le nombre maximum de numéros de téléphone disponibles est alors passé à 9. 20 % des individus disposaient alors uniquement d'au moins un numéro fixe, 17 % d'uniquement au moins un numéro portable et 52 % d'au moins un des deux. Le numéro enrichi était automatiquement appelé en priorité.

La totalité des individus ayant au moins un numéro de téléphone ont été appelés. En effet, notre protocole d'appel impose au prestataire, un brassage complet du fichier au cours du premier mois de l'enquête. L'objectif a été atteint. Les individus ne disposant pas de numéro de téléphone, sont plus difficile à contacter, notre seul moyen de contact étant la lettre-avis. En effet, la lettre-avis leur donne accès comme précisé plus haut, au CAWI d'enrichissement des coordonnées mais également à un numéro vert (gratuit), sur lequel ils peuvent nous joindre pour réaliser l'enquête.

Chaque numéro de téléphone est censé être appelé 20 fois tant que l'individu n'a pas répondu à l'enquête. Au total, sur l'ensemble des numéros de téléphone disponibles (de 1 à 9 selon les individus), 75 % des numéros de téléphone ont été appelé au moins une fois. À savoir, que si un individu, possédant 7 numéros, répond sur le second numéro alors les 5 autres sont considérés comme non appelés et comptabilisés dans les 25 % restants. En moyenne, 1,6 numéro a été appelé par individu.

Parmi les répondants, 62 % ont réalisé l'enquête sur un numéro issu de la base de sondage, 21 % sur un numéro enrichi en cours d'enquête, 10 % sur un numéro de portable issu de l'étape 2, 4 % sur un numéro issu des recherches élargie de l'étape 3 et 3 % sur un numéro fixe issu de l'étape 1. Ainsi, les enrichissements effectués en aval de l'enquête ont permis de récupérer 17 % des répondants, ce qui laisse penser qu'il est primordial de tous les réaliser.

En s'intéressant à l'efficacité des numéros de téléphone selon leur source d'enrichissement, on constate sans surprise que les numéros enrichis en cours d'enquête (par un tiers ou par l'individu lui-même) permettent d'aboutir à une enquête dans plus d'un cas sur deux. Ensuite, ce sont les numéros de la base de sondage, *a fortiori* s'ils ont été confirmés pendant la phase de recherche, qui produisent les meilleures chances de contact et de réponse. Concernant les numéros enrichis en amont de l'enquête, ce sont les portables qui produisent les meilleurs résultats suivis par les téléphones fixes. En revanche, la recherche élargie (où l'on desserre certains critères comme le prénom ou la localisation géographique) conduit naturellement plus fréquemment à des classements « hors-cible », à des faux numéros et surtout à la non-réponse. Leur efficacité est donc très limitée.

Ce bilan d'efficacité de l'enrichissement permet donc de juger de la qualité de ce protocole qui a également été mis en place sur l'expérimentation multimode. Les conclusions de ce bilan permettront d'ajuster au mieux ce protocole pour le rendre le plus optimal possible dans le but de l'intégrer au nouveau dispositif multimode des enquêtes Génération.

1.2. Évaluation du protocole d'appel

Dans les enquêtes Générations, le protocole d'appel prévoit vingt tentatives maximum à réaliser sur chaque numéro, à des heures et des jours différents (non réponse au téléphone ou répondeur), suivant des règles de rappels identiques pour tous avant de passer au numéro de téléphone suivant.

Tableau 4 - Règles de rappel globales

	Rappel	Fiche vierge	Ne répond pas ou Occupé	Répondeur
Règles de répétition	Date et heure choisies par l'enquêteur dans le calendrier	Présentation automatique après les rappels et avant les autres catégories	Rappel automatique à « J + 1 heure » ; si la 2 nd tentative est infructueuse rappel automatique à « J + 1 jour » avec décalage de 1 heure Présentation obligatoire sur un samedi	Rappel automatique à « J + 15 minutes » ; avec un maximum de 2 tentatives sur la même journée Présentation obligatoire sur un samedi
Priorité	1	2	3	3

Source : Enquête téléphonique menée en 2016 auprès de la Génération 2013, Céreq

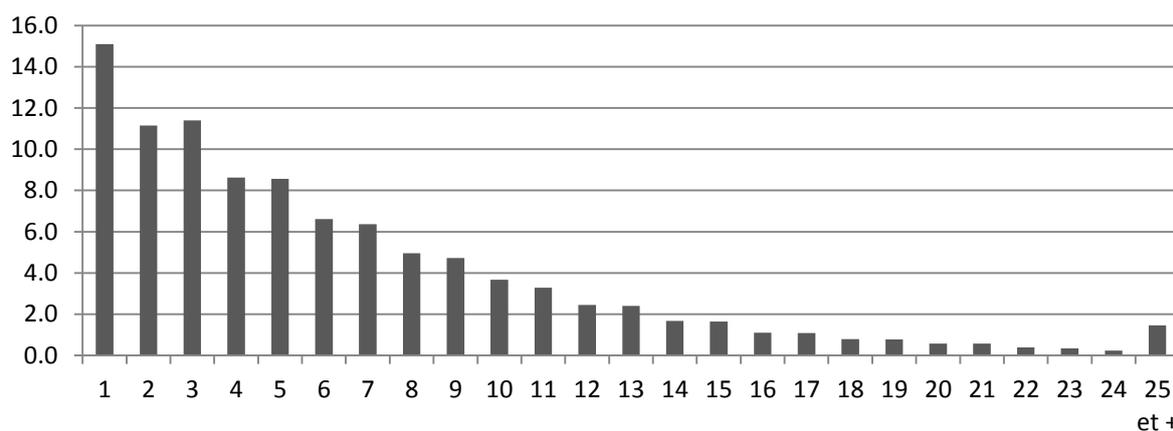
Tous les numéros de téléphone sont exploités jusqu'au classement définitif de la fiche (enquête réalisée, refus, ...).

52 % des individus de l'échantillon ont eu un contact direct ou via un tiers avec un enquêteur. Le premier contact a été effectué dès le premier appel pour 35 % d'entre eux. Pour un individu sur quatre le premier contact a eu lieu lors du 2nd ou 3^{ème} appel. Pour 20 %, il s'est fait entre le 4^{ème} et le 6^{ème} appel. Pour les autres, il aura nécessité jusqu'à 90 appels pour obtenir un premier contact.

Dans 41 % des cas, le premier appel par numéro de téléphone a permis un contact (tiers inclus). Pour 19 % des numéros de téléphone appelés, il a fallu 2 appels, et pour 12 %, 3 appels. Pour les 28 % restants, le nombre d'appels a varié entre 4 et 19.

Concernant le nombre d'appels, tous numéros confondus, effectué avant de réaliser une enquête, le graphique ci-après montre que pour 38 % des répondants, 1 à 3 appels ont suffi. Pour 24 %, 4 à 6 appels auront été nécessaires.

Figure 9 – Nombre d'appels total pour avoir une enquête



Source : Enquête téléphonique menée en 2016 auprès de la Génération 2013, Céreq

Si l'on se restreint au champ des répondants, en moyenne 1,7 numéro a été appelé. Un répondant sur deux a été contacté sur un seul numéro, 37,3 % sur deux et 10 % sur trois. Le nombre de répondants diminue fortement après quatre numéros appelés.

21 % des répondants ont répondu à l'enquête sur un numéro enrichi en cours d'enquête, 63 % l'ont passé sur le premier numéro renseigné dans le fichier d'import, 14 % sur le second et 2 % sur les numéros classés entre la troisième et la sixième position. Les numéros en 7^{ème} et 8^{ème} position, issus du fichier d'import, n'ont pas permis de faire d'enquête.

Concernant le nombre d'appels effectué sur le numéro ayant permis de réaliser l'enquête : un individu sur deux a répondu positivement aux sollicitations de l'enquêteur avec moins de trois appels. 27 % des répondants ont nécessité entre quatre et six appels, et 13 % entre sept et neuf appels. Près de 10 % des répondants ont dû être appelés plus de dix fois sur le même numéro, il est donc important de persévérer afin de relancer les individus.

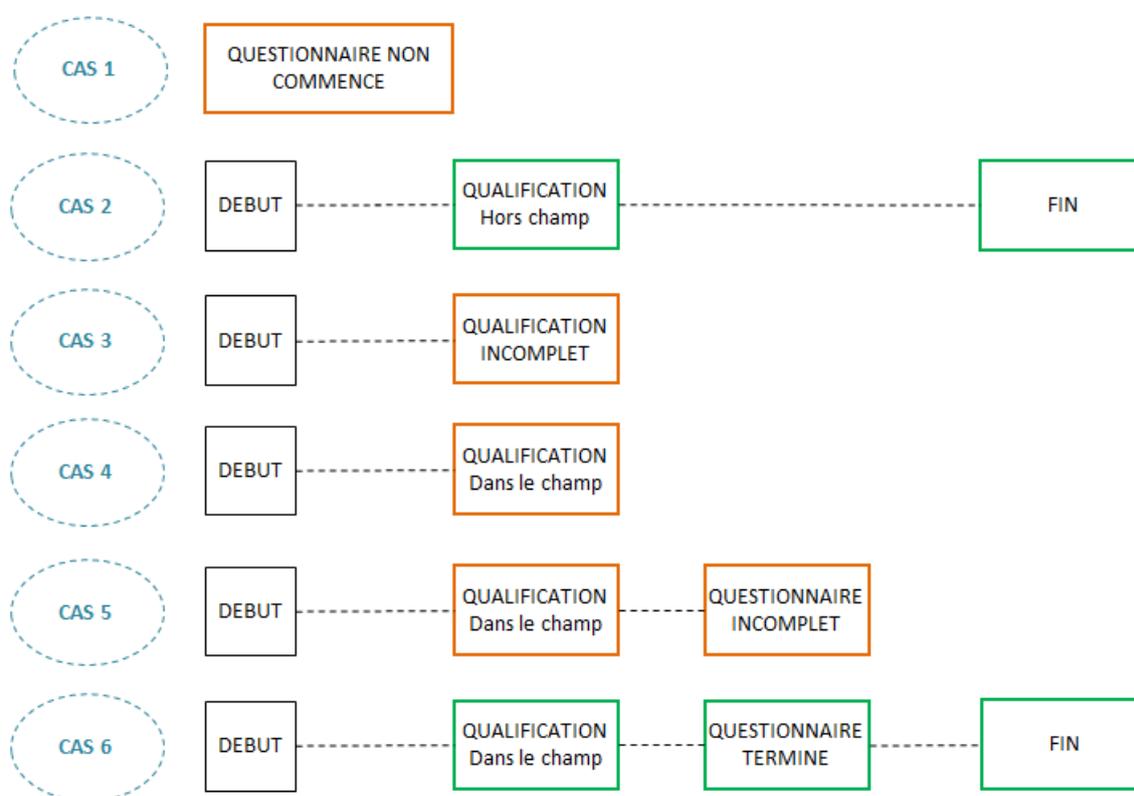
2. Plans de dialogue téléphoniques

2.1. Adaptation de la phase de contact

La collecte multimode impose de repenser la phase de contact par rapport à une collecte monomode téléphone. La phase de contact permet d'identifier le bon individu. Dans le cas de la collecte monomode, la phase de contact dépend essentiellement du fait que le questionnaire soit commencé ou non⁵. La collecte multimode intègre en plus une adaptation selon le mode de connexion.

La seconde expérimentation multimode avait la particularité d'intégrer dans le questionnaire une phase dite de « qualification » visant à déterminer si l'individu est éligible. Plusieurs cas sont possibles. Ils sont indépendants de la phase et du mode de collecte. Ils sont construits selon le dernier évènement de l'individu sur son questionnaire.

Figure 10 – Statut de questionnaire de la seconde expérimentation multimode (2016)



Cas 1 : La personne ne commence pas l'enquête

L'individu ne s'est ni connecté sur internet, ni par téléphone. Son questionnaire n'est pas commencé. Des tentatives de contact par mails, téléphones, sms seront faites pour qu'il participe à l'enquête. Le contact se fera donc de manière classique avec une présentation rapide de l'enquête et une invitation à participer à l'enquête.

Cas 2 : La personne a rempli le questionnaire de qualification et n'est pas éligible à l'enquête

L'individu a pu remplir le questionnaire de qualification sur internet ou téléphone. Il a terminé son questionnaire et ne sera donc plus jamais rappelé.

⁵ Il peut également y avoir une adaptation selon que l'individu répondant à l'appel soit la personne recherchée ou un tiers.

Cas 3 : La personne débute le questionnaire de qualification mais ne le termine pas

L'individu a pu commencer à remplir le questionnaire de qualification sur internet ou téléphone, mais l'a interrompu. Des tentatives de contact par mails, téléphones, sms seront faites pour l'inciter à terminer l'enquête.

Cas 4 : La personne a rempli le questionnaire de qualification et est éligible à l'enquête

L'individu a commencé à remplir le questionnaire et il est dans le champ de l'enquête. S'il a répondu sur internet, il s'agit d'une interruption. S'il a répondu par téléphone, il a soit souhaité interrompre l'enquête, soit il lui a été proposé de poursuivre l'enquête sur internet (et accepté) mais n'a pas terminé l'enquête, auquel cas l'interruption de l'enquête est volontaire. Des tentatives de contact par mails, téléphones, sms seront faites pour l'inciter à terminer l'enquête.

Cas 5 : La personne a rempli le questionnaire de qualification et est éligible à l'enquête mais ne termine pas son questionnaire

Il s'agit d'un cas d'interruption qui sera suivi de relances pour l'inciter à répondre à l'enquête.

Cas 6 : La personne a rempli le questionnaire de qualification et est éligible à l'enquête, puis termine son questionnaire

L'individu a terminé son questionnaire et ne sera donc pas rappelé.

Les cas 2 et 6 ne font pas l'objet d'un plan de dialogue spécifique puisque le questionnaire est terminé.

Dans les cas d'enquêtes commencées, cas 3, 4 et 5, la phase de contact intègre des variables de l'historique d'appels relatives à la dernière connexion du questionnaire en indiquant sur quel mode et quel jour l'individu s'est connecté la dernière fois. L'approche est personnalisée afin de convaincre l'individu.

L'interruption du questionnaire, telle que décrite en cas 4, présente la particularité, dans la plupart des cas, d'être volontaire, l'enquêteur incitant l'individu à poursuivre son questionnaire sur internet après avoir vérifié qu'il est éligible. La partie ci-dessous présente des exemples de plan de dialogue mis en place.

2.2. Plan de dialogue en phase 2 : concurrence des deux modes de collecte

Les expérimentations multimodes visaient à atteindre un taux de réponse au moins égal à celui de l'enquête monomode téléphone, tout en tentant de maximiser le taux de réponse par internet. Afin de répondre à cet objectif, la proposition de répondre sur internet doit être formulée de sorte à convaincre l'individu, sans trop lui donner l'impression que le mode internet lui soit imposé. La proposition du choix du mode de réponse a évolué entre les deux expérimentations multimodes.

Lors de la première expérimentation, la phase 2 intégrait deux plans de dialogue. Le premier (cf. Encadré 1 - Plan de dialogue de la première expérimentation multimode (2015), à l'issue de la phase de contact) incite fortement à répondre sur internet. Il a parfois été interprété par les individus comme une forme d'imposition ; pour d'autres, il s'agissait d'une excuse pour ne pas participer à l'enquête. Pour toutes ces raisons, la mise en production de ce plan de dialogue s'est révélé infructueux et a donc été remplacé plus tôt dans la collecte que prévu. Ce nouveau plan de dialogue (cf. Encadré 2) indique plus clairement que deux options sont possibles – internet ou téléphone – puis propose à l'individu le choix de répondre par internet.

*Encadré 1 - Plan de dialogue de la première expérimentation multimode (2015),
à l'issu de la phase de contact*

Version 1 de la phase 2

Vous avez la possibilité de répondre à ce questionnaire sur internet. Après confirmation de votre adresse email, je vous enverrai un courrier électronique avec le lien pour répondre par internet.

[CHAMP : SI MAIL PRESENT]

Votre email est bien <adresse mail> ?

- Oui.....1
Non.....2
(Ne veut pas répondre par internet).....3

[CHAMP : SI MAIL ABSENT OU MAIL ERRONE]

Quelle est votre adresse email ?

- (Saisie de l'adresse mail)1
(N'a pas d'adresse email)2

[CHAMP : SI MAIL CONFIRME OU SAISI]

Je vous ai envoyé un email contenant le lien pour compléter le questionnaire par internet. En cas de difficulté pour répondre au questionnaire par internet vous avez la possibilité de nous joindre au numéro vert. Je vous remercie pour votre participation et je vous souhaite une bonne fin de journée.

[CHAMP : SI REFUS DE REpondre PAR INTERNET OU N'A PAS D'ADRESSE MAIL]

Dans ce cas, nous allons poursuivre le questionnaire par téléphone.

*Encadré 2 - Plan de dialogue de la première expérimentation multimode (2015),
à l'issu de la phase de contact*

Version 2 de la phase 2

Vous avez la possibilité de répondre par téléphone ou internet à cette enquête. Voulez-vous poursuivre le questionnaire sur internet ?

- Oui.....1
Non.....2

[CHAMP : SI ACCEPTATION ET MAIL PRESENT]

Votre email est bien <adresse mail> ?

- Oui.....1
Non.....2
(Ne veut pas répondre par internet).....3

[CHAMP : SI ACCEPTATION ET MAIL ABSENT OU ACCEPTATION ET MAIL ERRONE]

Quelle est votre adresse email ?

- (Saisie de l'adresse mail)1
(N'a pas d'adresse email)2

[CHAMP : SI MAIL CONFIRME OU SAISI]

Je vous ai envoyé un email contenant le lien pour compléter le questionnaire par internet. En cas de difficulté pour répondre au questionnaire par internet vous avez la possibilité de nous joindre au numéro vert. Je vous remercie pour votre participation et je vous souhaite une bonne fin de journée.

[CHAMP : SI REFUS DE REpondre PAR INTERNET OU N'A PAS D'ADRESSE MAIL]

Dans ce cas, nous allons poursuivre le questionnaire par téléphone.

Au cours de la première expérimentation, des difficultés ont été rencontrées et une adaptation du plan de dialogue a été réalisée en cours d'enquête. Pour la nouvelle expérimentation, le protocole prévoit de proposer les deux modes de collecte avec une variante dans la formulation - toujours avec une incitation à participer par internet - (cf. Encadré 3). Il a été quelque peu modifié par rapport à la version de l'Encadré 2 afin de s'adapter à une première interrogation.

Encadré 3 - Plan de dialogue de la seconde expérimentation multimode (2016) après que l'individu soit identifié dans le champ

À l'issue de ces premières questions, vous faites bien partie des jeunes que nous souhaitons interroger. Vous avez maintenant la possibilité de poursuivre ce questionnaire par internet ou par téléphone. Accepteriez-vous de répondre par internet ?

Oui.....1
Non.....2

[CHAMP : SI ACCEPTATION ET MAIL PRESENT]

Nous allons donc vous envoyer un mail avec le lien qui vous permettra de vous connecter. Nous disposons d'une adresse mail fournie par votre dernier établissement de formation. Votre email est bien <adresse mail> ?

Oui.....1
Non.....2
(Ne veut pas répondre par internet).....3

[CHAMP : SI ACCEPTATION ET MAIL ABSENT OU ACCEPTATION ET MAIL ERRONE]

Nous allons donc vous envoyer un mail avec le lien qui vous permettra de vous connecter. Quelle est votre adresse email ?

(Saisie de l'adresse mail)1
(N'a pas d'adresse email)2

[CHAMP : SI MAIL CONFIRME OU SAISI]

Je vous envoie un email avec le lien pour compléter le questionnaire par internet. Nous comptons sur votre participation. En cas de difficulté pour répondre au questionnaire, vous avez la possibilité de nous joindre au numéro vert. Je vous souhaite une bonne fin de journée.

[CHAMP : SI REFUS DE REpondre PAR INTERNET OU N'A PAS D'ADRESSE MAIL]

Dans ce cas, nous allons poursuivre le questionnaire par téléphone.

Une nuance a été apportée dans la formulation de la question pour tester une autre manière de proposer de répondre sur internet. Par manque de données, il n'a pas été possible de tester statistiquement l'impact de ce changement de question entre 2015 et 2016. Par ailleurs, les retours des enquêteurs sont plutôt positifs, ils ne ressentent pas de réelles objections à cette question.