

A S S E M B L É E N A T I O N A L E

1 7 ^e L É G I S L A T U R E

Compte rendu

Commission d'enquête concernant l'organisation des élections en France

- Table-ronde, ouverte à la presse, réunissant des
représentants des entreprises Bilendi, Dynata et Toluna. 2
- Présences en réunion..... 14

Mardi
29 avril 2025
Séance de 17 heures 20

Compte rendu n° 35

SESSION ORDINAIRE DE 2024-2025

**Présidence de
M. Pierre-Yves Cadalen,
Vice-président de la
commission**



La séance est ouverte à dix-sept heures vingt.

M. Pierre-Yves Cadalen, président. Nous poursuivons les travaux de la commission d'enquête avec une table ronde consacrée à la question de la fabrique des sondages en nous intéressant plus précisément à la constitution des panels de personnes sondées. Nous recevons des représentants de trois entreprises exerçant cette activité : M. Marc Bidou, président-directeur général de Bilendi, accompagné de Mme Lucie Deliry, directrice France ; M. Renaud Farrugia, directeur commercial Europe de Dynata, accompagné de Mme Carla Dib-Roblès ; MM. Jean-Daniel Lévy, Patrick Van Bloeme et Paul Hambly, représentants de Toluna Harris Interactive. Nous vous remercions de vous être rendus disponibles.

Je vous rappelle que l'article 6 de l'ordonnance du 17 novembre 1958 relative au fonctionnement des assemblées parlementaires impose aux personnes auditionnées par une commission d'enquête de prêter le serment de dire la vérité, toute la vérité, rien que la vérité.

(M. Marc Bidou, Mme Lucie Deliry, M. Renaud Farrugia, Mme Carla Dib-Roblès, MM. Jean-Daniel Lévy, Patrick Van Bloeme et Paul Hambly prêtent serment.)

M. Marc Bidou, président-directeur général de Bilendi. Je vous remercie de nous avoir conviés aujourd'hui à cette table ronde et de nous offrir l'opportunité de parler de notre métier, qui est relativement peu connu. J'ai créé la société Bilendi en 2000, société française indépendante dont le siège social est situé à Paris. Bilendi s'est développée par croissance organique et par un certain nombre d'acquisitions. Nous opérons aujourd'hui des panels dans quarante pays du monde, dont la France, et nous disposons de bureaux dans dix-neuf pays. La société rassemble 650 collaborateurs, dont 105 en France, et travaille pour environ 2 500 clients dans le monde. Ces clients sont composés d'un très grand nombre d'instituts d'études, d'universités, de centres de recherche, de cabinets de conseil, d'entreprises. L'essentiel de notre activité porte sur des études de marché, des études marketing, des études d'opinion ; les enquêtes électorales représentent une très faible proportion de notre chiffre d'affaires.

Notre métier porte sur la collecte de données et les technologies dans le secteur des études de marché. Nous sommes donc une société de technologie, nous ne sommes pas une société de conseil. Nous n'employons pas de sociologues, de politologues, de consultants, mais plutôt des développeurs informatiques, des personnes chargées de traiter la donnée. En ce qui concerne les sondages électoraux, le rôle de l'institut consiste à rédiger le questionnaire, de déterminer l'échantillon, puis nous réalisons les travaux qu'il nous demande d'effectuer. Notre rôle consiste ainsi à collecter les données et les réponses qui correspondent parfaitement aux questions et aux attentes des clients.

Nous livrons les réponses à l'institut, qui peut procéder ensuite à d'éventuels redressements. Nous travaillons à partir de panels, plus précisément d'*access panels*, de grandes bases de données où sont inscrites des personnes ayant accepté d'être sollicitées pour des études. Nous leur posons 300 questions sur leur profil, qui nous fournissent environ 2 000 réponses et nous permettent de répondre aux attentes de nos clients.

Nous gérons à peu près 4 millions de panélistes dans nos panels dans le monde, dont 811 000 en France, que nous suivons tout au long de l'année. Ces panels subissent de nombreux contrôles de qualité. Ils sont hébergés en France et nous menons uniquement des enquêtes d'opinion à partir de nos propres panels. Les dossiers d'*access panel* nous donnent la possibilité

de livrer des échantillons parfaitement représentatifs à nos clients et notre activité est extrêmement régulée par des lois françaises, par la réglementation générale sur les données personnelles – qui impacte fortement notre activité – et par de nombreuses règles déontologiques du métier et des associations professionnelles, locales et internationales. Enfin, nous sommes certifiés ISO.

M. Renaud Farrugia, directeur commercial Europe de Dynata. La société Dynata est une société française née en 2006, membre d'un groupe américain. J'ai été à l'origine de la création de la société en France, dont l'objet porte également sur la collecte de données. Notre métier consiste ainsi à fournir des répondants à nos clients, qui sont des instituts d'études, des marques, des cabinets de conseil ou des instituts de sondage en particulier, dans plus de quarante pays dans le monde. Nous donnons accès à plus de 70 millions de personnes à travers nos *access panels*. En France, la taille de notre panel est légèrement supérieure à un million de personnes et nous avons vingt-cinq salariés, soit une petite entité ; l'ensemble de nos activités de gestion de panels et de technologies sont situées dans d'autres pays que la France.

M. Patrick Van Bloeme, co-président de Toluna Harris Interactive. Nous représentons aujourd'hui Toluna, maison-mère de plusieurs sociétés, dont Harris Interactive. Toluna a été créé en 2000 et j'y travaille depuis 2014, lorsque Harris Interactive, dont j'étais l'un des deux cofondateurs, a rejoint le groupe. À ce titre, je co-dirige aujourd'hui le groupe Toluna Harris Interactive en France.

Notre activité de panel est également très importante, car il s'agit de l'activité fondatrice du groupe Toluna. Aujourd'hui, nous gérons plus de 70 millions de panélistes dans le monde, dont plus d'un million en France. Le groupe conduit également des activités d'instituts d'études, la partie opinion étant une partie visible, mais relativement marginale en termes de chiffre d'affaires du groupe, au niveau mondial ou en France.

M. Antoine Léaument, rapporteur. Je vous remercie d'avoir accepté de participer à cette audition, qui nous semble très importante. En effet, l'un des objets de cette commission d'enquête vise à analyser les facteurs d'erreur potentiels sur les sondages et les instituts de sondage, en particulier sur les sondages politiques. Pour ce qui me concerne, je ne connaissais pas l'activité de panéliste.

Je souhaite d'abord vous interroger sur la manière dont vous réalisez les contrôles qualité sur vos panels. Comment vous assurez-vous que les gens qui participent à vos panels vous disent bien la vérité lors des descriptions sociodémographiques ?

Ensuite, de quelle manière se déroulent vos relations avec les instituts de sondage ? À partir des demandes que ces derniers formulent, quel est le processus, sur le plan technique ? Produisez-vous l'outil sur lequel les réponses seront récoltées ? L'institut de sondage vous fournit-il un logiciel de récolte des réponses ?

Enfin, je veux vous interroger concernant les rémunérations. Le législateur a précisé que lorsque les panélistes étaient rémunérés, cette information devait être mentionnée. Rémunérez-vous ces personnes par des points qui leur fournissent des avantages ? Les instituts de sondage contribuent-ils à cette rémunération ? Enfin, quelle part de cette rémunération concerne les questions d'ordre politique du sondage ? Les données publiées évoquent des points obtenus en avantages, dont les équivalences en euro sont assez faibles, de l'ordre de moins d'un euro.

M. Jean-Daniel Lévy, directeur délégué de Toluna Harris Interactive. Je vais commencer par répondre à votre première question, en insistant sur un point qui me paraît essentiel et que j'avais déjà évoqué lors de la table-ronde du 26 mars dernier.

En tant que panéliste, l'immense majorité de notre activité professionnelle, de nos revenus, du chiffre d'affaires et, encore plus, de la marge échappe aux études d'opinion d'un point de vue général, aux études politiques de manière plus particulière et aux études publiées de manière encore plus précise. Les études politiques publiées représentent ainsi une partie extrêmement faible de notre activité.

La démarche initiale dans le recrutement et l'entretien des panélistes ne constitue pas une vision dont l'objet cherche à « satisfaire » le marché de l'opinion, le marché politique ou le marché des études publiées, mais le marché de nos clients, c'est-à-dire des instituts ou des clients finaux. Ces derniers doivent ainsi prendre des décisions essentielles pour leurs investissements, leurs campagnes de communication, leurs organisations marketing ou pour la plupart de leurs choix stratégiques, lesquels peuvent engager des conséquences majeures pour le développement de leur entreprise.

Le premier « contrôle » qualité de la donnée que nous leur fournissons provient donc de nos clients finaux, du marché. Cette donnée s'appuie sur un échantillon de panélistes identique, quel que soit le type de question qui est posée : une personne inscrite dans le panel Toluna peut très bien répondre à une enquête d'opinion, à une question politique ou à une question marketing.

Ensuite, nous sommes pertinemment conscients que nous ne pouvons pas savoir si la personne qui nous déclare son genre, son âge, sa catégorie sociale, sa localisation géographique nous dit la vérité. À l'heure actuelle, aucun moyen légal ne permet de s'en assurer. Mais nous nous permettons de considérer qu'il ne s'agit pas d'un problème majeur, même s'il peut toujours exister des plaisantins qui s'inscrivent dans les différents panels. Cela deviendrait un problème si le phénomène était massif, mais rien ne le laisse penser.

Comme nous l'avions indiqué lors de la précédente audition, lorsque nous réalisons des baromètres, nous posons les mêmes questions à des personnes absolument différentes, d'une vague d'interrogation à une autre vague d'interrogation. Or nous observons que les évolutions sont faibles entre chaque vague d'enquête. Mais surtout, lorsque nous regardons dans le détail les réponses qui nous sont données selon le genre, la catégorie sociale, l'âge, la localisation géographique, voire la proximité politique, les variations sont en général assez faibles. En conséquence, les phénomènes de mensonge volontaire quant aux caractéristiques sociodémographiques sont malgré tout assez marginaux dans l'ensemble de nos échantillons.

Des contrôles qualité existent et je rappelle l'existence du code ICC/ESOMAR, qui constitue la référence mondiale en termes de bonnes pratiques. Les panélistes ne sont pas « lâchés dans la nature » dans leurs opérations, dans leurs méthodes de collecte et de conservation des données qui peuvent être délivrées par les personnes interrogées. Ces bonnes pratiques sont rassemblées dans des codes nationaux et codes internationaux.

M. Renaud Farrugia. Je partage cette réponse. Dans l'exercice de notre métier, nous sommes encadrés par un organisme comme ESOMAR ou d'autres associations dont nous sommes membres.

La principale motivation des personnes qui s'inscrivent dans ces panels correspond à leur volonté de donner leur avis. Dans notre industrie, la qualité des données est un sujet très important. Tant que nous pouvons nous assurer que cette intention prédomine dans leur participation à des *access panels*, nous nous prémunissons contre toute tentative de masse de tromperie. Pour autant, ces comportements peuvent exister, pour un certain nombre de raisons, notamment lorsque la gratification est plus importante que celle que vous avez évoquée, ce qui n'est pas le cas pour les sondages politiques, ou dans certains cas très précis où l'on cherche des cibles rares. Or dans les études d'opinion ou les études politiques, l'*incentive* est en général très faible, car les populations que nous cherchons n'ont pas de caractéristiques de rareté particulière.

Ensuite, bien évidemment, la qualité des données que nous fournissons est gage de notre succès économique : les clients nous achètent la garantie de constituer des informations ou de disposer d'informations fondées sur des données de qualité. *A priori*, ici aussi, le principe de réalité économique laisse penser que la qualité qui est délivrée par Dynata à ses clients est suffisamment bonne, éprouvée et reconnue puisque nos clients continuent de travailler avec nous et que d'autres nous rejoignent.

Enfin, nous conduisons de nombreux contrôles de qualité. Il est impossible de garantir la véracité des propos qui sont tenus par les répondants à l'intérieur d'une enquête, mais plusieurs centaines de contrôles sont effectués, au moment du recrutement des répondants, à chacune de leur interaction avec nous dans le cas d'une prise d'enquête, au moment où ils renseignent le questionnaire, à la fin de l'enquête et également au moment où ils récupèrent leur gratification. À chacun de ces moments, nous collectons de l'information qui nous permet d'éduquer un modèle et de détecter des comportements suspects. Si vous êtes intéressés, je serais très heureux de vous transmettre un papier vous détaillant un peu plus ces contrôles, lesquels nous permettent de garantir un minimum de rejet de la part de nos clients sur le type de répondants.

M. Antoine Léaument, rapporteur. Nous l'acceptons volontiers. D'une enquête à l'autre, vérifiez-vous bien que la même personne vous transmet les mêmes informations concernant ses caractéristiques ?

M. Renaud Farrugia. Je vous apporterai une réponse plus précise par écrit à ce propos. Nous procédons à des vérifications d'ordre technique comme l'adresse IP ou la localisation géographique, qui nous conduisent à dresser un profil de validité de la personne. Ces données techniques sont souvent plus fiables que des données humaines, car des répondants peuvent être inattentifs durant les enquêtes.

M. Marc Bidou. J'adhère aux explications qui ont été formulées. Au-delà, quelle serait la motivation de nos répondants à nous mentir ? Un homme de 30 ans ne gagnerait pas plus de points s'il prétendait être une femme de 50 ans ; au contraire, il recevrait des dizaines d'invitations pour des enquêtes qui ne le concernent pas. Si une prétendue jeune maman n'est pas en mesure de citer deux à trois marques de lait premier âge, des suspicions peuvent apparaître et être confirmées par d'autres questions. En réalité, il n'existe pas d'intérêt à vouloir tromper les enquêteurs.

M. Antoine Léaument, rapporteur. Sur le plan politique, il pourrait exister un intérêt à créer des profils fictifs pour essayer de biaiser les sondages à la marge, en déclarant de fausses intentions de vote. C'est la raison pour laquelle il m'apparaissait souhaitable de vous interroger sur votre capacité à vérifier ces informations.

M. Marc Bidou. L'électeur d'un parti ou d'un candidat n'a pas vraiment d'intérêt à annoncer une intention de vote erronée ou à travestir son opinion. Un parti ne pourrait pas influencer avec succès un panel, pour de simples raisons statistiques. Nous avons 800 000 membres, nous en tirons 1 000 dans une réponse. Pour n'avoir ne serait-ce que 5 % de « faux » invités, il faudrait créer 40 000 faux comptes. Si tel devait être le cas, nous le verrions puisque nous vérifions toutes les adresses IP ou l'usage de VPN par une série de contrôles. Si vingt comptes étaient créés, nous ne le verrions pas, mais cela n'aurait pas d'influence statistique sur un échantillon de cette taille.

M. Antoine Léaument, rapporteur. Votre réponse me conforte, il est important de dissiper tout fantasme concernant les panélistes.

Vous n'avez pas répondu à la question concernant l'aspect technique. Le sondeur vous envoie-t-il une série de questions, un lien à partir duquel il faut répondre ? Constituez-vous l'outil à partir des questions qui sont posées ?

Mme Lucie Deliry, directrice France de Bilendi. Deux cas de figure existent. Dans le premier cas, l'institut programme en interne son questionnaire et nous livre un lien que nous insérons dans nos outils. Dans le deuxième cas, il rédige un document Word avec les questions, que nous devons ensuite programmer pour pouvoir l'échantillonner et l'administrer aux panélistes. Tous nos outils sont développés en interne par nos spécialistes IT, sauf l'outil de *scripting* qui est un outil externe.

Un chef de projet est ensuite assigné sur l'enquête. À partir de toutes les informations dont il dispose, il programme le questionnaire. Il teste le questionnaire, l'envoie au client qui le teste à son tour et il procède ensuite à l'échantillonnage, grâce à un algorithme développé en interne qui poursuit deux objectifs : sélectionner les bonnes personnes pour qu'elles répondent au bon moment et s'assurer qu'elles ne soient contactées ni trop souvent ni trop rarement. En fonction de son taux de participation, du nombre de sollicitations qu'il a reçues dans les derniers mois et du nombre d'enquêtes auxquelles il a répondu, un poids est attribué à chaque panéliste. Ce poids est ensuite agrégé au niveau de sa catégorie, de son sexe, de son âge, de sa CSP.

Plusieurs vagues d'envoi du questionnaire sont réalisées jusqu'à atteindre les 1 000 répondants souhaités. Ensuite, en cours de terrain, des contrôles qualité sont effectués de manière automatique, notamment via une intelligence artificielle (IA) développée en interne et de manière humaine, par le chef de projet. Une fois que le terrain est terminé, nous procédons à des vérifications sur la base de données, des contrôles de cohérence.

M. Antoine Léaument, rapporteur. Vous avez évoqué deux cas de figure. Je pense qu'une question peut se poser sur la fiabilité des données. Si le sondeur fournit le lien, comment peut-il vérifier qu'à un individu correspond une réponse et que votre contrôle qualité s'applique sur ce sondage ?

Mme Lucie Deliry. Nous pouvons fournir différents types de lien (lien de *speeder*, de *straightliner*, de *quality*) qui peuvent être insérés, que la programmation soit effectuée par l'institut ou par nous-mêmes. Quand un questionnaire de dix minutes est répondu en trois ou quatre minutes, il y a peu de chances que le répondant ait lu toutes les questions et surtout tous les items de réponse. Dans ce cas, le panéliste est exclu, de même que celui qui cliquerait systématiquement sur les premières réponses ou les dernières réponses, par exemple. Par ailleurs, l'équipe de traitement de données de l'institut d'études réalise des contrôles de

cohérence à la fin du terrain, à partir de la base de données, pour déterminer si des « filous » ont répondu de manière non consciencieuse.

M. Patrick Van Bloeme. Les contrôles de cohérence sont systématisés sur l'ensemble de la filière, qu'il s'agisse des instituts ou des panélistes. Il ne sert à rien qu'une personne fasse n'importe quoi dans une enquête, le jeu n'en vaut pas la chandelle, en matière d'*incentives*, c'est-à-dire de dédommagements plutôt que de rémunération : en moyenne, sur une année, un panéliste ne gagne que quelques euros. La motivation d'un *access panel* est liée en réalité à son souhait de s'exprimer, de s'assurer que sa voix puisse compter, essentiellement pour les sujets marketing et, à la marge, pour des sujets d'opinion. Je précise à ce titre que Paul Hambly, qui nous accompagne aujourd'hui, a inventé l'activité panel chez Toluna.

Il existe également des contrôles à l'entrée sur les cohortes, sur des phénomènes que nous pouvons identifier. Nous allons chercher les panélistes, à travers des campagnes de recrutement, ce qui permet déjà de limiter le risque ; et nous pouvons surveiller effectivement l'activité. Depuis de longues années, nous avons expérimenté des formules avec ou sans dédommagements et sommes parvenus à la conclusion qu'ils permettent de toucher une population plus large et donc de produire des études plus représentatives.

M. Renaud Farrugia. Selon que la programmation soit effectuée par l'institut de sondage ou par nos soins, il existe une différence technique, mais le niveau de contrôle qualité minimal est déjà très élevé, y compris dans le cas où l'institut contrôle la programmation. À la fin de l'enquête, si les clients ne sont pas satisfaits de la qualité de certaines réponses, nous récupérons les identifiants de ces répondants et notre système de qualité en est informé, pour ensuite déterminer si ces répondants doivent être placés en « quarantaine », sans qu'ils en soient informés. Des vérifications supplémentaires sont effectuées et certains d'entre eux réintègrent finalement les panels, quand le problème portait plus sur de l'inattention que sur une fraude ou un mensonge caractérisé.

M. Antoine Léaument, rapporteur. Ce sujet est passionnant, dans la mesure où ces questions sont rarement abordées. Comment vous assurez-vous d'avoir des panels réellement représentatifs de la population, pour les sondages politiques ? En effet, certains échantillons politiques sont chroniquement sous-représentés et d'autres chroniquement surreprésentés, en lien avec les déterminants sociodémographiques. Par exemple, les CSP+ ou les retraités, électeurs fréquents d'Emmanuel Macron, y sont plus représentés que des ouvriers, qui votent plus pour Jean-Luc Mélenchon. Une difficulté supplémentaire est liée à la notion centrale de souvenir de vote, lequel n'est pas forcément fiable.

Il existe également sans doute une évolution des caractéristiques sociodémographiques : l'Insee fournit des caractéristiques sociodémographiques peut-être un peu plus éloignées aujourd'hui de la réalité matérielle des conditions d'existence. Je pense notamment aux personnes « ubérisées », c'est-à-dire officiellement des chefs d'entreprise, mais qui vivent en réalité une relation de salariat déguisé. Comment parvenez-vous à percevoir ces éléments ? La modification des conditions sociales d'existence est-elle prise en compte dans la constitution de vos panels ?

M. Jean-Daniel Lévy. Nous cherchons toujours à interroger un échantillon national représentatif, quel que soit le type de client et quel que soit l'objet de l'étude. En conséquence, nous ne parlons pas d'échantillon politique, mais d'échantillon sociologique, à partir de données sociodémographiques. Nous nous fondons effectivement sur les données de l'Insee, notamment en matière de genre, d'âge, de catégorie sociale, de localisation géographique, celles-ci pouvant

être déclinées en deux parties, avec d'un côté la taille d'agglomération et de l'autre la région ou éventuellement le regroupement de régions administratives. Ces aspects, même s'ils peuvent apparaître comme incomplètement efficaces au regard par exemple de l'évolution du salariat que vous avez évoquée, constituent malgré tout une donnée fiable et, à certains égards, exemplaire par rapport à ce qui peut exister dans d'autres pays, y compris occidentaux.

S'agissant du souvenir de vote, que l'on s'en réjouisse ou qu'on le déplore, les Français ne se rappellent pas toujours leurs comportements électoraux passés. On se base alors sur les données qui nous sont fournies.

M. Patrick Van Bloeme. Dans la relation entre l'institut et le panéliste, le premier est le donneur d'ordre : l'institut fixe les règles de représentativité, les « quotas » qu'il attend. L'institut a tendance à confier à son panel ou à son prestataire des directives assez claires sur la partie qui est maîtrisée, c'est-à-dire la partie sociodémographique. Il revient ensuite à l'institut, à partir des données recueillies, de procéder éventuellement à des redressements, lesquels ne concernent pas trop les panélistes.

Comme M. Levy l'a indiqué, nous devons être cohérents : quel que soit le sujet, qu'il s'agisse d'une enquête d'intention de vote ou une enquête d'opinion, nous devons nous assurer que nous maîtrisons l'échantillon sur lequel nous travaillons, que nous pouvons le comparer à un autre échantillon sur un sujet connexe. Nous avons la chance, en France, de pouvoir disposer de nombreuses statistiques et de données. Ensuite, elles peuvent évoluer avec le temps, comme M. le rapporteur l'a indiqué ; mais de notre côté, nous devons posséder une matière disponible, pour pouvoir l'adapter à ces évolutions.

M. Renaud Farrugia. En tant qu'entreprise, notre responsabilité en matière de représentativité consiste à recruter des panélistes et des membres les plus divers possible, élément sur lequel nous concentrons nos efforts. Il revient à l'institut de nous fournir le profil des personnes qu'il souhaite interroger ; il serait audacieux de notre part de proposer quelque chose de différent des données de l'Insee, qui font référence.

Mme Carla Dib-Roblès, responsable des opérations de Dynata. Même si nous cherchons à être le plus représentatifs possible, nous ne sommes pas certains que les panélistes soient disponibles pour répondre à l'enquête. En conséquence, les quotas constituent vraiment la garantie de disposer d'une certaine représentativité.

M. Antoine Léaument, rapporteur. Je vous remercie. Madame Deliry, vous avez indiqué plus tôt que vous procédiez par vagues. Comment procédez-vous dans le détail, de manière concrète ? Retenez-vous un même groupe de panélistes pour leur envoyer plusieurs fois le questionnaire jusqu'à ce qu'ils répondent ? Si l'on vous demande d'interroger 1 000 personnes, en contactez-vous juste 1 000, ou un plus grand nombre ? Ces éléments techniques apparaissent importants, car ils me semblent être susceptibles de comporter des effets de bord sur le résultat même des enquêtes.

Mme Lucie Deliry. Nous procédons par étapes. Pour atteindre 1 000 répondants, nous cherchons au début à obtenir une centaine de répondants, ne serait-ce que pour valider que tout se passe bien dans l'administration du questionnaire. Nous continuons ensuite, jusqu'à atteindre le chiffre de 1 000 ; raison pour laquelle je parlais de plusieurs vagues d'invitations. L'objectif consiste notamment à couvrir toutes les phases d'une journée : nous savons par exemple qu'un ouvrier ne pourra pas répondre à 11 heures du matin. Nous étalons ainsi le nombre d'envois, afin d'optimiser et d'en maximiser la qualité.

Ensuite, une première extraction représentative est réalisée en sexe, en âge, en CSP, en catégorie d'agglomérations et de régions, puis nous surpondérons malgré tout quelques populations. Nous savons par exemple que les femmes répondent plus rapidement que les hommes ou que les jeunes répondent moins rapidement que les 50 ans et plus, ou les CSP+.

En résumé, nous agissons de manière progressive, parfois en surpondérant, afin d'arriver au chiffre de 1 000, de manière homogène. Telle est la manière dont nous procédons à l'échantillonnage classique, chez Bilendi.

M. Patrick Van Bloeme. Je précise que la surpondération concerne la sélection et l'envoi, et non les réponses : les réponses sont pondérées pour refléter exactement les quotas. Pour éviter que les 300 derniers répondants appartiennent tous à la même cible, nous faisons en sorte de bien répartir l'ensemble de l'échantillon sur la durée des terrains, qui peuvent être plus ou moins longs selon leur difficulté. À cet effet, nous disposons d'algorithmes pour procéder de manière simple et fonctionnelle et éviter tout biais. Je précise qu'il n'existe pas de biais à avoir répondu en début ou en fin d'enquête : qu'il réponde le lundi après-midi ou le mardi matin, le répondant est exposé à la même expérience. En revanche, nous nous assurons que chaque profil puisse trouver « son » moment pour répondre et ne pas introduire de biais dans ce domaine.

Mme Carla Dib-Roblès. Chez Dynata, nous procédons de manière légèrement différente. Nous ne fonctionnons pas vraiment par un système d'invitations par email. Les invitations par email ont juste pour objet de rappeler aux personnes qu'elles font partie d'un panel et leur demander de se connecter pour répondre aux enquêtes. Mais la majorité des répondants se connectent par eux-mêmes sur notre plateforme, où un système de routeur les dirige en fonction des besoins du terrain vers les enquêtes disponibles, qu'il s'agisse d'enquêtes commerciales ou de sondages. Nous tenons également compte des rythmes horaires des personnes : par exemple, les retraités sont plus disponibles le matin que dans l'après-midi ; les jeunes le sont plus en fin de journée.

M. Renaud Farrugia. Il existe donc des mécanismes techniques différents pour conduire les répondants dans l'enquête, mais les processus sont identiques, en fonction des disponibilités connues de chaque catégorie.

M. Antoine Léaument, rapporteur. Je vous remercie de ces précisions, je n'avais pas forcément en tête ces modalités concernant les disponibilités de chaque catégorie en fonction des moments de la journée. D'une certaine manière, vous allez même plus loin que les requêtes des sondeurs en matière de données brutes de caractéristiques sociodémographiques.

Comment procédez-vous pour recruter des panélistes ? Envoyez-vous des emails ? Avez-vous accès à des bases de données particulières ?

Par ailleurs, pour les questions d'ordre politique, il peut sembler que des effets de halo peuvent parfois intervenir. Par exemple, si un répondant est interrogé sur des questions portant sur le partage de la richesse, sur la santé, sur les services publics et qu'il est ensuite questionné sur une intention de vote politique, peut-être se dirigera-t-il plus naturellement vers un candidat de gauche. La mécanique peut être inverse si les premières questions ont concerné la sécurité, l'immigration. Peut-il exister de tels effets de halo associés à des questions totalement différentes ? Au cours d'un même questionnaire, mélangez-vous des questions sur la consommation et des questions d'ordre politique ?

M. Paul Hambly, Toluna Harris Interactive. Nous disposons d'une grande expérience dans ce domaine, depuis vingt-cinq ans, avec des équipes dédiées au marketing digital, aux moteurs de recherche, aux réseaux sociaux, afin de toucher tous types de catégories sociodémographiques. Dans plus de quatre-vingt-dix pays, nous effectuons des campagnes marketing pour promouvoir notre panel et inciter les gens à s'inscrire. Il ne s'agit pas uniquement de campagnes d'emails, mais de bannières qui peuvent s'afficher sur un site comme lemonde.fr, une publicité sur Facebook ou X. Nous les adaptons en fonction des audiences dont nous avons besoin pour nourrir les échantillons.

M. Renaud Farrugia. Chez Dynata, nous utilisons à peu près les mêmes méthodologies. Il peut nous arriver également de nouer des partenariats avec des influenceurs pour aller recruter des cibles très précises de jeunes. Nous codéveloppons également des panels avec des marques, uniquement sur invitation. Ces marques disposent de programmes de fidélité et les répondants sont rémunérés dans la « devise » du programme de fidélité. Il s'agit de constituer un seul même panel virtuel à partir de toutes ces informations et de tous ces membres que nous recrutons par le biais des divers moyens que je viens d'évoquer.

M. Patrick Van Bloeme. Le questionnaire politique est unique, dédié, il n'est pas intégré dans un autre questionnaire qui évoquerait par exemple des habitudes de consommation. Nous ne mélangeons pas les genres. Ensuite, nous sommes très vigilants concernant les effets de halo ; il s'agit d'une question importante que nous nous posons systématiquement sur l'ensemble de nos projets. À titre d'illustration, en moyenne, un paneliste répond à huit enquêtes sur une année, soit une quantité modérée, vous en conviendrez. Ces enquêtes s'influencent-elles les unes les autres ? Nous maîtrisons ces aspects : si nous estimons que le risque existe, nous intégrons des critères d'exclusion ; par exemple, tel panel de 1 000 personnes n'est pas éligible à telle enquête ci.

Cela est encore plus marqué en marketing. Il n'y a aucun sens à demander à un panel de se prononcer sur la notoriété spontanée d'une marque s'il a été interrogé trois jours auparavant sur une liste précise de marques.

Dans le domaine politique, l'impact sur le comportement électoral d'un panéliste des quelques questionnaires auxquels il a pu répondre par le passé peut être considéré comme marginal dès lors que l'on ne parle pas d'une même enquête électorale qui serait intervenue auparavant. Il est marginal par rapport à l'ensemble du temps qu'il passe sur Internet, sur les réseaux sociaux, sur les chaînes d'information. Nous partons donc du principe que nous neutralisons ce risque, si tant est qu'il existe.

Il existe des protocoles où nous avons besoin de neutraliser telle ou telle enquête, dans la mesure où nous maîtrisons toutes les enquêtes que nous effectuons auprès de nos panélistes, en termes de sujet et de risque d'influence. Encore une fois, les sondages d'opinion représentent une extrêmement faible proportion des enquêtes que nous menons. La collusion est donc quasi nulle par rapport aux études de notoriété de marques, pour lesquelles nous procédons à un véritable travail de contrôle.

M. Antoine Léaument, rapporteur. Vous me confirmez donc bien qu'une enquête d'opinion ou une enquête politique n'est jamais intégrée dans une liste de questions portant sur d'autres sujets, n'est-ce pas ? Je vous pose la question à dessein, dans la mesure où il y a quelques années, j'avais participé à des questionnaires qui commençaient par des questions sur telle ou telle marque de savon, puis enchaînaient sur des questions d'ordre politique ou sur la

sécurité, avant de revenir sur des lessives. Une évolution est-elle intervenue ? Certains panélistes ne respectent-ils pas les règles que vous avez énoncées ?

M. Marc Bidou. Un certain nombre d'instituts mènent des questionnaires de type « omnibus ». Ils vendent des questions à des clients ; certains clients achèteront deux questions, d'autres clients en achèteront trois ou quatre. Au sein de ces questions peut figurer une question d'opinion ; il revient aux instituts de décider du contenu de leurs questionnaires. Ensuite, il existe des enquêtes électorales dédiées.

M. Antoine Léaument, rapporteur. Des questions politiques peuvent-elles être intégrées au sein d'un sondage omnibus ?

M. Marc Bidou. Oui.

Mme Lucie Deliry. Lorsque cela est le cas, il s'agit du choix de l'institut et non du nôtre.

M. Antoine Léaument, rapporteur. J'en conviens naturellement. Au sein d'un même questionnaire, différents sujets peuvent être abordés, dont celui d'une intention de vote pour une élection, n'est-ce pas ?

M. Jean-Daniel Lévy. Chez Toluna, nous n'organisons quasiment plus d'omnibus. Mais quand nous procédions de la sorte, les questions politiques étaient toujours placées en tête de questionnaire, afin d'éviter précisément l'effet de halo que vous mentionniez. J'ignore la manière dont les autres entreprises procédaient.

M. Renaud Farrugia. De notre côté, nous ne participons pas à des études omnibus. En dehors du domaine politique, il peut arriver que pour un motif de qualification, nous posions des questions préalables à l'enquête nous permettant de déterminer le profil de la cible et si celui-ci correspond à l'enquête. Mais il ne s'agit pas d'une étude, d'une enquête à proprement parler, ni d'une question relative à une intention de vote.

M. Pierre-Yves Cadalen, président. Je souhaite vous poser une dernière question avant de conclure cet après-midi qui nous a vus recevoir deux instituts de sondage avant votre audition. Il s'agit d'un sujet qui nous semble central dans le cadre des travaux de cette commission au sujet des mesures de l'opinion, des intentions de vote. Nous avons eu l'occasion d'évoquer avec vos prédécesseurs les méthodes de redressement. Tout le monde s'accorde sur un point : les échantillons de base ne sont pas représentatifs. En effet, si tel était le cas, il ne serait pas nécessaire de les redresser. De fait, ces échantillons sont toujours assez inexacts par rapport une représentativité qui serait parfaite sur le plan sociodémographique. Il nous a par exemple été indiqué que dans les échantillons de vote, les ouvriers sont sous-représentés quand les retraités sont surreprésentés.

M. Jean-Daniel Lévy. Il peut exister des différences par rapport aux objectifs initiaux, mais en général, la différence entre la population recherchée idéalement et la population véritablement interrogée n'est pas extraordinairement élevée. Si un redressement intervient, il ne s'agit que d'être le plus fin possible, de l'ordre de quelques points ; il ne s'agit pas de procéder à une modification profonde de la nature de l'échantillon.

Ensuite, face à des critères multiples en termes de genre, d'âge, de catégorie sociale, de localisation géographique, nous savons bien qu'à la fin du terrain, nous cherchons des « moutons à cinq pattes ». Nous acceptons donc des interviews, quitte à ce que les répondants

soient « hors quotas », c'est-à-dire surreprésentés par rapport à la moyenne. En conséquence, une pondération est effectuée pour redresser non pas des individus, mais l'ensemble de l'échantillon.

Enfin, de manière encore plus importante, il s'agit pour nous de nous assurer que l'effet du redressement ne crée pas de biais statistique ; c'est-à-dire que le fait d'accorder 1,1 point ou 1,2 point à un individu et éventuellement 0,8 point ou 0,9 point à un autre individu ne permette pas de remettre en cause la base sur laquelle nous appuyons collectivement.

M. Pierre-Yves Cadalen, président. Pour être tout à fait clair, si je comprends bien, il n'existe pas de marge d'amélioration de constitution de l'échantillon. Lors d'une audition précédente, nous avons constaté que sur un sondage d'intention de vote, un échantillon de base présentait une surpondération de six points de la catégorie retraités par rapport à leur représentation dans la population générale. Si je comprends bien vos propos, ces quelques points n'entraînent donc pas selon vous des conséquences considérables.

M. Antoine Léaument, rapporteur. En l'espèce, il s'agissait d'un cas particulier, c'est-à-dire un sondage politique, à partir d'un échantillon très important, de l'ordre de 10 000 personnes. La Commission des sondages avait ainsi pointé un problème dans ce sondage, qui comportait 5,5 % de retraités en plus ; 4 % de cadres en plus et 3,5 % d'ouvriers en moins. En matière électorale, ces différences ne sont pas forcément négligeables.

M. Patrick Van Bloeme. Il nous est difficile de nous exprimer sur un sondage produit par nos confrères. Quoi qu'il en soit, notre travail consiste à respecter les quotas qui sont donnés par le donneur d'ordre, y compris au sein d'un même groupe. Il peut être nécessaire de procéder à des redressements lorsque sur les 1 000 personnes, les vingt ou trente dernières ne correspondent pas exactement à l'échantillon parfait sur certains critères, comme le sexe, l'âge, la région.

M. Marc Bidou. Lorsqu'il nous est demandé de produire un panel de 1 000 personnes représentatif sur les quotas de l'Insee, nos résultats sont presque parfaits : sur les 52 % de femmes dans la population, nous parvenons à produire un échantillon de 52 % ou 52,1 %. Il en va de même pour les CSP ou les tranches d'âge. Même si notre *access panel* n'est pas totalement représentatif, notre livrable est représentatif à 0,1 % ou 0,2 % près. Dès lors, on ne peut pas affirmer que nos échantillons ne sont pas représentatifs. Ensuite, il est toujours possible de s'interroger pour savoir si les critères de l'Insee sont parfaitement nécessaires pour une enquête, mais il s'agit là d'un autre débat.

M. Antoine Léaument, rapporteur. Vos propos soulèvent une autre question. Les sondeurs eux-mêmes indiquent qu'ils redressent sur la base de plusieurs éléments, notamment les données sociodémographiques et les données politiques. Je déduis de votre intervention que vous livrez aux instituts des données qui s'approchent de la perfection quant à la représentation sociodémographique. Mais vos données sont redressées sur ces bases sociodémographiques. Fournissez-vous les coefficients de pondération aux instituts de sondage ou les instituts font-ils intervenir par la suite de nouveaux critères de redressement ? Si vos données sont les plus proches possible de la réalité, pourquoi effectueraient-ils à leur tour des correctifs ? Le seul correctif qui me semble possible est le correctif politique, qui apparaît encore plus important que ce que j'envisageais de prime abord.

M. Renaud Farrugia. Il est plus facile pour nous, sur le plan statistique, de fournir un échantillon de 1 000 personnes, puisque nous disposons de plus d'un million de personnes en base. Notre rôle se limite à respecter les échantillons et les quotas que le client, en l'occurrence l'institut, nous demande de fournir. Il est très rare de ne pas pouvoir y parvenir. Nous livrons des échantillons représentatifs de la population. En revanche, nous ne touchons pas du tout à l'aspect politique.

M. Marc Bidou. Il est sans doute nécessaire de produire des redressements. Nous livrons des échantillons quasi parfaits, mais ensuite, la science des instituts consiste à savoir s'il faut redresser à partir de l'élection n-1, n-2 ou n-3, s'il s'agit de l'élection présidentielle, des élections européennes ou des élections législatives par exemple. Les redressements n'interviennent pas sur l'échantillon en tant que tel, mais font appel à d'autres éléments, comme le biais de mémorisation des derniers votes, par exemple. Il s'agit d'un autre débat, qui ne nous incombe pas.

M. Pierre-Yves Cadalen, président. Je vous remercie.

La séance s'achève à dix-huit heures trente cinq.

Membres présents ou excusés

Présents. – M. Pierre-Yves Cadalen, M. Antoine Léaument