# Sondages d'opinion : mesurer l'incertitude

Nicolas Sauger

# Sondages d'opinion : mesurer l'incertitude

*Nicolas Sauger*[*1]

***Résumé***

Parmi les modes de connaissance de la société moderne sur elle-même, les sondages sont apparus tardivement, au tournant du XXe siècle. Conçus pour quantifier et légitimés par leur capacité de prédiction, les sondages entretiennent une relation paradoxale à l'incertitude. La notion d'intervalle de confiance a été clé dans le développement de la technique mais reste discutée. La perspective de « l'erreur totale » montre toute la difficulté d'une quantification effective de cette incertitude.

***Abstract***

Among the tools for the self-understanding of modern societies, opinion polls emerged lately, by the turn of the 20th century. Designed to quantify and legitimized by their predictive capacity, polls have a paradoxical relation to uncertainty. The notion of confidence interval has been key in the development of the technique but is still discussed. The total survey error perspective shows how difficult it is to provide a realistic quantification if this uncertainty.

L'avènement de la société moderne s'est accompagné d'un besoin croissant de connaissance de la société sur elle-même. Au-delà de l'optique de dénombrer une population, il s'est agi de comprendre les dynamiques sociales et politiques, questions pour lesquelles les recensements sont apparus comme une méthode trop lourde et trop onéreuse dans bien des cas. Quetelet (1997), dans son œuvre fondatrice, est ainsi très explicite quant à l'une des motivations de sa proposition d'appliquer des techniques statistiques aux phénomènes sociaux. Il s'agit bien de connaître les faits sociaux pour « éviter les

* Professeur des Universités à Sciences Po, Directeur du Centre de données socio-politiques (UAR828)

[1] Nicolas Sauger est le responsable scientifique de plusieurs séries de sondages en France. Il est notamment actuellement le directeur scientifique du panel web probabiliste Elipss (https://cdsp.sciences-po.fr/fr/projets/panel-elipss/), le coordonnateur national pour la France de la série Enquête sociale européenne (ESS-ERIC, https://www.europeansocialsurvey.org/) et la *French Electoral Study* dans le cadre du projet international CSES (https://cses.org/). Il est également le conseiller scientifique du panel probabiliste *Public Voice* de Kantar – France.

coûteuses révolutions ». Pour autant, pour Quetelet comme pour ses contemporains, l'essentiel de cette utilisation des statistiques s'applique aux données administratives ou aux données de recensement. Et il faut attendre le début du XX^e^ siècle pour voir entrer dans la panoplie des technologies de mesure cette idée du sondage (Desrosières 2005).

Comme le montre Desrosières (1998), la naissance des sondages d'opinion est longtemps repoussée puis apparaît très rapidement dans les premières décennies du XX^e^ siècle. Alors que la plupart des concepts nécessaires, notamment la formalisation du calcul des probabilités, étaient disponibles (et utilisés) dès la fin du XVIII^e^, les enquêtes par sondage sur échantillon représentatif naissent véritablement entre 1906 et 1935. En 1925, l'Institut international de statistiques consacre ainsi la possibilité de construire l'estimation d'une valeur d'intérêt par échantillonnage d'une population (Didier 2012). Les sondages doivent ainsi s'imposer dans un monde dominé par les recensements mais également les enquêtes monographiques. Les sondages apportent dans cet univers la possibilité de caractérisation de populations hétérogènes dans un temps et pour un coût réduits. Ils permettent également d'interroger sur une plus grande variété de questions et d'échapper ainsi aux contraintes inhérentes à un recensement. Les sondages sont ainsi d'abord développés à grande échelle dans une logique de prospection des marchés économiques, par exemple pour cibler les publicités. Mais ils sont véritablement légitimés en 1936 quand l'entreprise créée un an plus tôt par Gallup prédit à raison (et contre les références de l'époque) la victoire de Roosevelt contre Landon lors de l'élection présidentielle américaine (Converse 1987, Blondiaux 1998).

Dès le départ, les sondages nouent ainsi une relation complexe à l'incertitude. Ils ont acquis une reconnaissance publique initialement par leur capacité prédictive, sur le court terme, et leur portée directement pratique pour l'action en raison d'une revendication de représentativité. Mais les sondages sont également dès l'abord fondés sur la capacité à quantifier l'incertitude, les premiers calculs d'intervalle de confiance remontant à Bowley, en 1906 (Desrosières 1988). C'est cette relation que cet article explore plus avant en revenant sur ce principe de quantification de l'incertitude.

## *L'intervalle de confiance*

Les sondages ont pu prendre leur essor, intellectuel et industriel, en affirmant leur caractère de représentativité, sous réserve de les assortir d'une marge d'erreur acceptable. Cette mesure de l'incertitude est largement démocratisée aujourd'hui par la notion d'in-

tervalle de confiance. Cette notion repose largement sur le théorème central limite, montrant que la distribution des erreurs liées à un échantillonnage probabiliste suit une distribution gaussienne (Alwin 2007).

L'intervalle de confiance se calcule aisément (par exemple Ardilly 2006), avec les méthodes statistiques classiques. En pratique, cet intervalle ne dépend que de la distribution de la variable observée, de la taille de l'échantillon et de la marge d'erreur. Aujourd'hui, la marge d'erreur la plus courante est un seuil fixé à 95 pour cent, signifiant que l'on accepte que notre estimation puisse sortir de l'intervalle de confiance calculé dans 5 pour cent des cas. On peut représenter ainsi une table de référence de cet intervalle, pour différentes valeurs observées, comme le propose le tableau I. Dans ce tableau, on lit ainsi qu'au seuil de 95 pour cent, pour obtenir un score de 30 pour cent et un échantillon de 1 000 personnes, il y a 95 pour cent de chances que la valeur réelle dans la population de référence soit comprise entre 27,2 (30-2,8) et 32,8 (30+2,8). On voit bien dans ce tableau que l'intervalle de confiance est d'autant plus petit que l'échantillon est de taille importante. L'étendue de cette marge d'erreur est ainsi quatre fois plus petite quand notre score est de 30 pour cent et que l'on fait varier la taille de l'échantillon de 500 à 4 000. Cette plage est de 2,8 points de pourcentage pour un échantillon de 4 000, 7,6 points pour un échantillon de 500. En d'autres termes, plus l'échantillon est grand, meilleure est la précision. On observe néanmoins que le gain en précision s'amenuise au fur et à mesure que l'échantillon grandit.

Tableau I. – Intervalle de confiance au seuil de 95 pour cent suivant différentes tailles d'échantillon.

| | Score obtenu | | | | | | |
|---|---|---|---|---|---|---|---|
| Taille de l'échantillon | 2 % | 5 % | 10 % | 20 % | 30 % | 40 % | 50 % |
| 500 | 1,2 | 1,9 | 2,6 | 3,5 | 3,8 | 4,3 | 4,4 |
| 1000 | 0,9 | 1,4 | 1,9 | 2,5 | 2,8 | 3,0 | 3,1 |
| 2000 | 0,6 | 1,0 | 1,3 | 1,8 | 2,0 | 2,1 | 2,2 |
| 4000 | 0,4 | 0,7 | 0,9 | 1,2 | 1,4 | 1,5 | 1,5 |

Cette notion d'intervalle de confiance est aujourd'hui suffisamment entrée dans les mœurs pour qu'elle ait même été intégrée dans l'appareil législatif encadrant les sondages dans le cadre des élections. La loi 2016-508 du 16 avril 2016 de modernisation des règles applicables aux élections stipule ainsi dans son article 6 (modifiant l'article 2 de loi n° 77-808 du 19 juillet 1977) que tout sondage

électoral publié doit comporter « une mention précisant que tout sondage est affecté de marges d'erreur » et que « les marges d'erreur des résultats publiés ou diffusés, le cas échéant [le sont] par référence à la méthode aléatoire ». On notera que cette loi formalise plus largement l'ensemble de la publication des sondages électoraux, mais seulement les sondages électoraux, sous l'égide de la Commission des sondages[2]. Les sondages dans ce domaine n'ont ainsi jamais été aussi encadrés et transparents quant à leur construction. Les différentes polémiques sur la qualité de prédiction de ces sondages ont en effet conduit à vouloir renforcer le professionnalisme et l'exigence de qualité qui les entourent.

L'intervalle de confiance ne reste néanmoins qu'une approche limitative de l'incertitude pour les sondages d'opinion. La loi elle-même le reconnaît implicitement puisqu'elle propose son calcul « par référence » à la méthode aléatoire. Mais ce complément laisse lui-même beaucoup à préciser.

### *De l'intervalle de confiance à « l'erreur totale »*

La recherche sur les sondages d'opinion propose un cadre global de compréhension de l'incertitude, notamment au travers de l'approche de « l'erreur totale de sondage » (Biemer 2010, Groves & Lyberg 2010). Comme son nom le laisse deviner, cette approche propose une approche plus complète de l'incertitude dans les sondages. L'intervalle de confiance représente en effet seulement la prise en compte d'un risque d'erreur spécifique, celui lié au tirage d'un échantillon dans une distribution aléatoire. Or, en pratique, les sondages sur les opinions sont sujets à beaucoup plus de risques d'erreur. On sait par exemple que les personnes interviewées peuvent se tromper ou ne pas bien comprendre une question. Ce risque vient donc augmenter le risque lié à l'échantillonnage lui-même. Par ailleurs, le calcul classique de l'intervalle de confiance implique que l'incertitude n'est que liée au hasard. Or les sondages, comme tout autre méthode de mesure, sont également sujets de biais systématiques. Ces biais sont systématiques au sens où ils sont déterminés, au moins pour partie de manière causale. Chacun peut en effet par exemple ne pas comprendre une question de sondage, mais il est probable que le diplôme ou la langue maternelle soient des caractéristi-

[2] La Commission des sondages assure ainsi un dispositif de surveillance mais également de transparence pour l'ensemble des sondages électoraux publiés en France. Cf. https://www.commission-des-sondages.fr/ (consulté le 28 avril 2022).

ques des répondants qui vont influencer systématiquement ce risque de se tromper. Et si une partie de la population a plus de risques de se tromper qu'une autre, l'estimation sur l'ensemble de la population sera forcément biaisée. L'approche par l'erreur totale reprend ces intuitions pour les systématiser et identifier ainsi les risques principaux. Par commodité, il est possible pour cela de distinguer deux niveaux : celui de la mesure d'une part et celui de la représentation d'autre part.

*Concernant la mesure*. L'interrogation centrale est de savoir comment les réponses reçues au questionnaire reflètent effectivement le concept pensé par l'observateur. Pour plus de simplicité, on peut distinguer pour cela d'un côté la correspondance entre les questions posées et l'intention de l'observateur et, de l'autre côté, la réaction de la personne interrogée aux questions posées.

Un questionnaire de sondage doit en effet d'abord être en mesure des fournir les indications recherchées par l'observateur. Mesurer des intentions de vote, par exemple, est un exercice est un peu plus compliqué qu'en apparence. Il faut savoir comment combiner une intention de participation à l'élection et une intention de vote. Il faut déterminer également quels candidats proposer ou comment les présenter (avec une affiliation partisane ou non, dans l'ordre des panneaux électoraux, l'ordre alphabétique, ou aléatoirement). L'ensemble de ces choix – et bien d'autres encore – est susceptible d'exercer une influence sur le résultat produit. Et la question se complique encore si l'on cherche à mesurer une attitude (par exemple la xénophobie), c'est-à-dire la propension psychologique à évaluer des objets de telle ou telle façon. Il s'agit en effet alors de construire une batterie d'indicateurs indirects demandant à être combinés pour mesurer cette dimension latente (Dargent 2011).

À un autre niveau, l'interprétation et l'usage que font les répondants des questions qui leur sont proposées dans un sondage sont eux aussi susceptibles d'introduire de l'incertitude. Une réaction classique des répondants est celle du « satisficing ». Si ce concept est utilisé avec des perspectives différentes (Roberts *et al.* 2019), il renvoie néanmoins généralement à une tendance suivant laquelle les répondants cherchent, dans un sondage, à suivre une norme sociale plutôt que de donner leur opinion. Cela se manifeste par le fait de ne pas révéler les comportements qui sont perçus comme ne se conformant pas aux normes dominantes. On surestime ainsi par exemple généralement la participation aux élections ou on sous-estimait le vote pour le Front National pour cette raison. Le « satisficing » peut aussi consister à vouloir à donner la réponse « juste », c'est-à-dire celle dont le répondant estime qu'elle est attendue.

L'incertitude dans les sondages peut donc en réalité être située non pas seulement au niveau de l'échantillonnage mais au niveau du questionnaire lui-même. Cette incertitude peut être importante en termes de taille. Michaud *et al.* (2022) utilisent par exemple une expérimentation sur la formulation de questions sur la perception de l'équité de la distribution des revenus dans le pays. Cette perception est mesurée sur une échelle en 7 positions, 6 proposant une graduation de l'injustice et 1 la constatation d'une distribution juste. Ils montrent l'existence d'une différence allant jusqu'à 40 points de pourcentage quant au soutien à la distribution des revenus (sélection de la modalité « juste »). Offrir plus de nuances (proposer par exemple des modalités offrant des options comme « plutôt injuste » en plus de la dichotomie juste / injuste) fait ainsi décroître très significativement le constat de l'acceptation, en moyenne, de la distribution des revenus. Ce type d'effet lié à la conception des sondages est bien connu et en partie systématisé. C'est ce qui permet par exemple à une équipe située à Barcelone de proposer une évaluation de la qualité d'un sondage à partir d'une analyse systématique de la dimension formelle du questionnaire[3].

*Concernant la représentation*. La question est de savoir dans quelle mesure l'échantillon utilisé est effectivement représentatif de la population d'intérêt, au niveau de laquelle est faite l'inférence. On note là aussi plusieurs problèmes potentiels. Ils peuvent se situer sur la définition de l'échantillon cible, sur le recrutement de l'échantillon effectif, et sur la non-réponse. Un exemple concret permet d'illustrer ces difficultés. L'Enquête sociale européenne[4] (ESS), qui sert souvent de standard méthodologique dans le monde académique, est une enquête en face-à-face qui se déroule tous les deux ans en France. Cette enquête est construite autour d'un échantillon dit représentatif des résidents en France de plus de quinze ans. Un échantillon d'individus est tiré par l'INSEE dans les bases fiscales (Fidéli) puis mis à disposition d'un institut de sondages qui envoie son réseau d'enquêteurs exploiter ces contacts suivant une méthodologie bien spécifiée (visites en personne uniquement, au moins quatre visites à des horaires différents, dont certains le week-end, avant de déclarer un contact infructueux, etc.). Malgré la qualité du dispositif, de nombreux angles morts persistent. Les bases fiscales tout d'abord couvrent de manière partielle certaines populations, exclues d'un logement fixe par exemple ou en cours d'étude. La base de tirage est utilisée par ailleurs plus d'un an après sa constitution, impliquant

[3] Une analyse peut être ainsi fournie suivant une méthode dite SQP (« Survey predictor quality »), voir http://sqp.upf.edu/ [consulté le 20 mai 2022].

[4] https://www.europeansocialsurvey.org/, consulté le 28 avril 2022.

une déformation de l'image (déménagements, décès, etc.). L'outremer et la Corse sont absents. Les personnes souffrant d'une déficience visuelle ou auditive exclus de l'échantillon, de même que ceux ne maîtrisant pas suffisamment le français, etc. Et, une fois les enquêteurs sur le terrain, d'autres difficultés apparaissent encore, qu'il s'agisse de localiser une adresse ou de convaincre une personne de participer à une étude. Pour une étude de référence pour les sciences sociales comme ESS, on sait que le taux de réponse est dans les cas les meilleurs d'environ 50 pour cent. Ce taux de réponse cumulé aux autres difficultés signifie qu'il faut en moyenne avoir au moins 2,5 individus dans son échantillon cible pour obtenir 1 personne effectivement intégrée dans son échantillon. Et l'on comprend bien qu'une partie des difficultés rencontrées peuvent être dues au hasard (une personne absente systématiquement de son foyer aux passages de l'enquêteur malgré la diversité des horaires). Mais un biais systématique est également présent. Un indépendant a par exemple moins de chances d'être présent qu'une personne au foyer. Les autres modes d'enquêtes présentent par ailleurs des sources de biais généralement considérées comme beaucoup plus importantes. Les taux de réponse, quand ils sont calculés, sont beaucoup plus faibles. Le ratio échantillon cible sur entretiens effectifs peut être très nettement supérieur, jusqu'à 100 numéros pour 1 interview par exemple dans le cadre d'une enquête par téléphone suivant la méthode aléatoire. La méthode par quotas utilisée en France de manière dominante est également discutée parce qu'elle repose sur l'hypothèse que les caractéristiques sociodémographiques de base sont plus prédictives des autres caractéristiques de l'individu que sa propension à accepter de répondre rapidement à une étude. De même, le passage de plus en plus généralisé aux sondages par internet pose par exemple la question non seulement de la familiarité avec les technologies numériques inégalement répartie dans la population, mais également celle de la base de tirage puisque les échantillons sont systématiquement tirés dans des viviers de volontaires pour répondre à ces études[5], souvent motivés par la rémunération, même faible, qu'elles peuvent octroyer[6].

[5] Une majorité de sondages en France est par exemple aujourd'hui réalisée à partir d'un fournisseur d'adresse dominant, Bilendi. Certains instituts ont sinon un panel exclusif à leurs fins.

[6] Un article publié par *Le Monde* en 2021 a par exemple suscité une certaine émotion au travers du partage d'une expérience de participations multiples, et fictives, à différents sondages. Le journaliste dit ainsi par exemple avoir participé à plus de 200 sondages dans l'intervalle de 6 semaines. Cf. https://www.lemonde.fr/politique/article/2021/11/04/dans-la-fabrique-opaque-des-sondages_6100879_823448.html, consulté le 28 avril 2022.

### *Une quantification de l'erreur impossible*

L'approche par l'erreur totale de sondage qui vient d'être utilisée montre ainsi globalement l'existence de biais multiples et parfois importants en plus d'une erreur aléatoire à plusieurs niveaux. Si elle induit des stratégies possibles pour la correction de ces biais, elle ne suggère en revanche pas de solution pour une quantification exacte et complète de l'incertitude. En effet, les techniques de correction du biais sont disponibles, suivant deux stratégies possibles et complémentaires. L'une consiste à corriger pour les effets d'échantillonnage connus. De manière classique, mais seulement pour les échantillons aléatoires, il est possible de prendre en compte les probabilités inégales de sélection des répondants liées par exemple à la taille variable des foyers. De même, quand des informations préalables sont connues sur les adresses de tirage, il est également possible d'opérer des corrections pour la non-réponse, donnant ainsi un poids supplémentaire à ceux les moins susceptibles de donner suite à une invitation à répondre à un questionnaire. L'autre technique de correction du biais est *a posteriori*. Elle consiste en la correction des résultats bruts à l'aune d'autres grandeurs connues de la population (grâce aux données issues de recensement par exemple). Dès lors, il est possible de « caler » l'échantillon sur les marges de références, en supposant que les corrections ainsi obtenues soient corrélées aux paramètres d'intérêt (technique dite de post-stratification).

L'approche par l'erreur totale n'offre pas en revanche de technique effective de calcul des erreurs totales. De nombreuses études quantifient des erreurs sur tel ou tel aspect des processus, qu'il s'agisse de la formulation des questions, des biais cognitifs des répondants, des effets d'échantillonnage, etc. Ces études mettent bien sûr en lumière des régularités dans le type d'effet des choix de spécification d'un sondage. Mais l'ampleur de ces effets reste le plus souvent variable. Cela est d'autant plus vrai que des interactions complexes se nouent entre les différentes spécifications d'un même sondage. Les effets de formulation des questions dépendent par exemple du mode de passation de ce sondage. Ces effets restent par ailleurs contextuels, liés aux caractéristiques d'une population spécifique.

Les études de qualité de suivi des intentions de vote ont été de ce point de vue un terrain particulièrement fertile. Pour la période récente et le terrain français, C. Durand montre ainsi comment, encore aujourd'hui, tant les instituts de sondage que le législateur, essaient d'apprendre des décalages observés entre les sondages et les votes passés pour faire évoluer leurs méthodologie (Durand, 2008). Par exemple, l'élection présidentielle de 2002 avait conduit à une surestimation forte des candidats de gauche, de l'ordre de 5

à 6 points. Les instituts de sondage avaient à la suite de cela révisé leur stratégie de calage et le législateur encadré plus particulièrement ce processus (obligation de transparence et de constance sur le traitement des données brutes par exemple). À l'élection présidentielle suivante, c'est au contraire le score de la candidate Marine Le Pen qui fut surestimé. Mais un simple exemple contemporain peut venir illustrer de manière encore plus éclairante la situation. Considérons les intentions de vote pour le second tour de l'élection présidentielle de 2022[7], en ne nous préoccupant que des derniers sondages publiés (du 21 au 22 avril 2022). Les intentions de vote publiées en faveur d'Emmanuel Macron oscillaient entre 53 et 57 pour cent des voix (moyenne de 56 pour cent), or le résultat pour un résultat effectif de l'élection fut de 58,5 pour cent. Sur les dix sondages d'intentions de vote considérés, six avait proposé un résultat dont la différence excédait la marge d'erreur indiquée quand celle-ci est calculée uniquement par rapport à la taille d'échantillon. Ces différences sont souvent justifiées, et parfois à raison, par la dynamique électorale pouvant exister jusqu'au jour du vote. Mais le point que nous voulons souligner ici est différent. Il est de dire que les sondages bénéficiant des échantillons les plus importants ne sont pas forcément les meilleurs. Il apparaît même que la taille de l'échantillon n'est pas significativement liée à l'écart enregistré. Ce résultat illustre parfaitement l'analyse par l'erreur totale de sondage. La taille d'échantillon n'est pas le seul élément source de réduction de l'incertitude dans les sondages (Lejeune, 2021). Ceci ne doit être lu ni comme une condamnation de la publication d'intervalles de confiance – même s'ils doivent être utilisés avec précaution, comme bien des notes d'institut le mentionnent elles-mêmes – ni comme un doute sur l'utilité d'échantillons plus importants en nombre. Ceux-ci sont assurément indispensables, notamment pour explorer le comportement de populations spécifiques. Mais un intervalle de confiance raisonnable devrait être considéré autour des estimations données et être plutôt proposé avec pour référence un échantillon type de 1 000 personnes. On pourra pour cela se référer de nouveau au Tableau I. L'intervalle de confiance est ainsi plus grand, réduisant donc le risque que les valeurs vraies en population générale soient à l'extérieur de l'intervalle de confiance. L'usage tel qu'il est fait aujourd'hui d'un intervalle de confiance basé uniquement sur l'erreur aléatoire d'échantillonnage nous paraît en effet risquer de créer une illusion sur l'exactitude des estimations proposées.

[7] Voir par exemple le site : https://poliverse.fr/polls/ (consulté le 3 mai 2022).

## *Conclusion*

Les conséquences à tirer de cette analyse ne doivent pas être celles d'une critique systématique des sondages. On notera d'ailleurs que, bien souvent, les critiques les plus virulentes de sondages ne sont pas construites autour d'une discussion des fondations méthodologiques de ces études mais sur les usages qui sont faits de ces études (par exemple Champagne 2007). Le paradoxe pourrait en effet être inversé, en posant la question de savoir comment les sondages font « aussi bien » quand les considérations théoriques pointent vers autant de sources possibles d'erreur (Sauger 2008). Il faut rendre ici autant hommage à l'expertise des gens du métier qu'à la nature humaine, faite, au-delà des revendications de singularité de chacun, de biens des régularités.

## *Bibliographie*

Alwin, D. F. (2007), *Margins of Error: A Study of Reliability in Survey Measurement*, John Wiley & Sons, New York.

Ardilly, P. (2006), *Les Techniques de sondage*, Technip, Paris.

Biemer, P. (2010), « Total survey error. Design, implementation, and evaluation », *Public Opinion Quarterly*, 74(5): 817-848.

Blondiaux, L. (1998), *La Fabrique de l'opinion : une histoire sociale des sondages*, Le Seuil, Paris, 601 p.

Champagne, P. (2007), « Sondages : des interprétations scientifiquement infondées et politiquement nocives », *Notes de l'observatoire des médias – Acrimed*.

Converse, J.M. (1987), *Survey research in the United States: roots and emergence 1890-1960*, University of California Press, Berkeley, 564 p.

Dargent, C. (2011), *Sociologie des opinions*, A. Colin, Paris.

Desrosières, A. (1988), « La partie pour le tout : comment généraliser ? La préhistoire de la contrainte de représentativité », *Statistique et analyse de données*, 13(2): 93-112.

Desrosières, A. (2005), « Décrire l'État ou explorer la société : les deux sources de la statistique publique », *Genèses*, 58: 4-27.

Didier, E. (2012), « Histoire de la représentativité statistique : quand le politique refait toujours surface », *in* M. Selz (dir.), *La Représentativité en statistique*, Ined éditions, Paris, pp. 15-30.

Groves, R., Lyberg, L. (2010), « Total survey error: past, present, future », *Public Opinion Quarterly*, 74(5): 849-879.

Durand, C. (2008), « The polls of the 2007 French presidential campaign: were lessons learned from the 2002 catastrophe? », *International Journal of Public Opinion Research*, 20(3): 275-298.

Lejeune, M. (2021), *La Singulière Fabrique des sondages d'opinion,* L'Harmattan, Paris, 168 p.

Michaud, A., Bosch, O., Sauger, N. (2022), « Can survey scales affect what people report as a fair income? Evidence from the cross-national probability-based online panel CRONOS », manuscrit.

Quetelet, A. (1997) [1869], *Physique sociale ou essai sur le développement des facultés de l'homme*, Académie royale de Belgique, Bruxelles, 700 p.

Roberts, C., Gilbert, E., Allum, N., Eisner, L. (2019), « Satisficing in surveys: A systematic review of the literature », *Public Opinion Quarterly*, *83*(3): 598-626.

Sauger, N. (2008), « Assessing the accuracy of polls for the French presidential election: the 2007 experience », *French Politics*, 57(3): 116-136.