

RAYMOND LÉVY-BRUHL

**Les sondages d'opinion et l'échec des prévisions
électorales de Gallup en 1948**

Journal de la société statistique de Paris, tome 92 (1951), p. 63-70

http://www.numdam.org/item?id=JSFS_1951__92__63_0

© Société de statistique de Paris, 1951, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

VIII

VARIÉTÉS

Les sondages d'opinion et l'échec des prévisions électorales de Gallup en 1948.

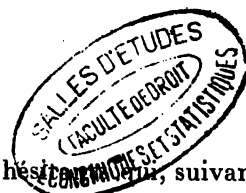
Le 2 novembre 1948, 45 millions d'Américains se sont rendus dans les bureaux de vote pour désigner, entre autres, celui qui, pendant quatre ans, serait le Président des États-Unis. Pour l'immense majorité du peuple américain, ce vote n'était que symbolique puisque le vainqueur était connu d'avance. En effet, les trois grands organismes de mesure de l'opinion publique, Gallup, Crossley et Roper, qui avaient depuis 1936 régulièrement prédit sans erreur le résultat des élections, en utilisant la méthode des sondages, avaient annoncé que le candidat républicain Dewey serait élu. Or, le 3 novembre, le monde entier stupéfié apprit que le candidat démocrate Harry Truman avait été élu Président des États-Unis, ayant recueilli sur son nom 49,5 % des votes alors que Dewey n'en recueillait que 45,1 % et obtenant au collège chargé d'élire le Président 303 voix sur 531 voix. On sait, en effet, que le Président n'est pas élu à la majorité des voix, mais que chacun des 48 États élit un nombre de délégués électoraux égal au nombre de ses députés et sénateurs. Or, Gallup avait prédit 49,5 % des votes pour Dewey et 44,5 % pour Truman; Crossley avait prédit 49,9 % pour Dewey et 44,8 pour Truman, et Roper, qui de 1936 à 1944

ne s'était jamais trompé de plus de 1 % au cours des trois élections présidentielles, avait prédit que Dewey recueillerait 52,2 % des voix et Truman 37,1 % seulement.

Certes, les Instituts de mesure de l'opinion publique ont toujours proclamé que leur but véritable n'était pas d'anticiper de quelques jours sur un vote en prévoyant à l'avance les résultats de celui-ci, mais de faire connaître, dans l'intervalle des élections ou en leur absence, soit l'évolution des attitudes politiques, soit l'état des opinions sur les questions qui ne seront jamais soumises à une consultation électorale. Il n'en demeure pas moins que c'est l'exacte prévision des résultats des principales élections américaines qui avait attiré l'attention du public sur les méthodes de sondage. Même si Gallup rappelait parfois que les résultats de ses enquêtes étaient susceptibles d'être affectés d'un certain pourcentage d'erreur et en dépit des échecs qu'ont connu, durant ces dernières années dans certains pays, de jeunes Instituts utilisant des méthodes identiques, le nom de Gallup était devenu à travers le monde le synonyme de l'infaillibilité en matière de prévision électorale. Cette foi aveugle, habilement entretenue par la publicité, ne devait cependant pas résister au premier échec et la réaction fut d'autant plus forte que la confiance avait été plus totale. Gallup et ses collègues devinrent l'objet de plaisanteries, non seulement de la part de la presse, mais de l'ensemble du public. C'est ainsi, par exemple, que le jour où Truman prenait officiellement ses fonctions de Président à Washington, au Sénat de l'État d'Indiana, tous les sénateurs unanimes, debout, observaient une minute de silence « à la mémoire du D^r Gallup ».

Les milieux scientifiques s'émurent cependant du discrédit qui risquait d'atteindre une technique, qui, lorsqu'elle se borne à mesurer et qu'elle utilise des méthodes scientifiquement correctes, a fait la preuve de son efficacité et de sa précision. Il importait donc de trouver pourquoi Gallup s'était trompé pour pouvoir faire comprendre au public les causes de cet échec en même temps que les limites des diverses techniques de sondage et leurs champs d'application qui restent considérables. Aux États-Unis, le « Social Science Research Council » nomma un comité composé d'éminents statisticiens, psychologues et sociologues pour rechercher les causes de l'échec des prévisions de Gallup, Crossley et Roper. De même, les sociétés de statistique, de sociologie ou d'étude de l'opinion publique mirent cette question à l'ordre du jour de leurs congrès ou la discutèrent dans leurs périodiques. Des expériences furent tentées, des hypothèses furent édifiées. Il peut être intéressant de faire le point de ces recherches et de dégager les conclusions qu'il est possible d'en tirer.

Tout d'abord il semble nécessaire de souligner qu'il a été unanimement reconnu que l'honnêteté des divers Instituts d'Opinion Publique ne pouvait être mise en doute. Il serait d'ailleurs difficile de déterminer lequel des deux grands partis avait intérêt à faire croire à la certitude d'une victoire de Dewey; mais on peut penser, en définitive, que ce fut plutôt le parti démocrate. Il semble, en effet, que le nombre des républicains qui se sont dispensés de voter parce que Dewey était sûr de l'emporter et de démocrates qui se sont finalement décidés à aller apporter leurs voix à Truman, puisque celui-ci avait besoin de rassembler toutes les voix démocrates pour tenter de triompher fut plus important que le nombre des démocrates qui renoncèrent à voter pour un



candidat qui n'avait aucune chance de vaincre ou des hésitations, suivant « l'appel du vainqueur », votèrent pour Dewey. Les expériences tentées aux États-Unis semblent d'ailleurs montrer que la publication des prévisions influe beaucoup moins sur les résultats qu'on ne le pense communément. Il y a d'autre part lieu de rappeler que l'élection présidentielle n'était pas seule en cause en ce premier mardi de novembre 1948, mais qu'en même temps que le Président des États-Unis, on élisait des représentants aux Parlements Fédéral et de l'État, ainsi qu'un grand nombre de fonctionnaires locaux. Le déplacement, souvent long, nécessaire pour se rendre jusqu'au lieu de vote pouvait donc avoir de multiples autres causes que le désir de voter pour Dewey ou Truman, mais une fois arrivé, pour un motif quelconque, dans l'isolement, l'électeur n'avait guère de raison de s'abstenir de voter dans l'élection présidentielle, ou d'y voter, en contradiction avec ses propres sentiments, soit pour le vainqueur, soit pour le vaincu présomptif

De nombreuses causes d'erreur peuvent intervenir dans un sondage d'opinion publique et amener une prédiction erronée des résultats d'une élection. On distinguera quatre causes principales.

1° L'échantillon interrogé peut ne pas être représentatif de la population des électeurs;

2° La technique de l'enquête et du questionnaire peut ne pas être suffisante pour permettre, à partir des résultats de l'enquête, de déterminer quels seraient les résultats d'un vote ayant lieu le même jour;

3° Il peut y avoir des variations dans l'opinion du corps électoral entre le jour de l'enquête et le jour de l'élection;

4° La présentation des prévisions tirées de l'enquête peut ne pas être accompagnée d'explications suffisantes sur les limites de leur précision.

1. Toute méthode scientifique de sondage comporte l'élaboration d'un plan d'échantillonnage.

Ce plan devra indiquer comment seront choisies les unités de l'univers à étudier qui feront partie de l'échantillon, de manière à ce que cet échantillon soit représentatif de l'univers.

Deux grandes méthodes sont utilisées. La première, qui fait application des lois du calcul des probabilités, consiste en un tirage au hasard des unités à prélever. L'utilisation la plus élémentaire de cette méthode consiste à affecter à chacune des N unités de l'univers à sonder un numéro variant de 1 à N et à tirer au hasard, comme dans une loterie n numéros compris entre 1 et N . Les n unités portant ces numéros tirés formeront l'échantillon. Cette méthode, à laquelle le développement des techniques modernes a apporté des améliorations constantes, permet toujours, simplement ou plus ou moins difficilement suivant la complexité du plan d'échantillonnage adopté, de calculer l'erreur qu'entraîne, pour une mesure donnée effectuée sur l'échantillon, le remplacement d'un univers de N unités par un échantillon de n unités. Cependant, toutes les méthodes basées sur un choix strictement au hasard des unités seront en général coûteuses tant par les dépenses d'investissement nécessaires à la constitution d'une base permettant le choix des unités de l'échantillon que par la nécessité d'effectuer une enquête respectant un choix rigide des unités. Pour cette raison, les organismes gouvernementaux ont été jusqu'à présent presque les seuls, dans

les différents pays qui effectuent des enquêtes par sondage, à utiliser des techniques basées sur un choix au hasard des unités de l'échantillon.

Dans la seconde méthode, au contraire, le choix des unités de l'échantillon repose sur le jugement. L'échantillon est alors constitué par un choix raisonné, et non plus au hasard, de chaque unité. C'est une telle méthode, dont le principal avantage est le prix de revient peu élevé, qu'utilisent les Instituts d'Opinion Publique sous la forme de la méthode des quotas. Le choix de chaque unité est alors fait de manière à ce que l'ensemble des unités de l'échantillon représente une image réduite de la population à étudier. On s'assure, au moyen de différents contrôles, du caractère représentatif de cette image. Dans des enquêtes portant sur des populations humaines, par exemple, on pourra choisir pour contrôle : le sexe, l'âge et la classe sociale ou économique à laquelle appartient l'individu. Si on dispose de statistiques précises, pour une population donnée, de la distribution des individus suivant le sexe, les tranches d'âge et la classe économique-sociale, on pourra s'assurer que les distributions de ces caractères pour les individus de l'échantillon sont bien les mêmes que pour l'ensemble de la population. L'enquêteur, suivant une telle méthode, n'aura pas à interroger une personne qui lui aura été désignée à la suite d'un tirage au hasard, mais il pourra choisir lui-même des personnes auprès de qui il effectuera son enquête, sous la seule condition de respecter les quotas qui lui ont été assignés, c'est-à-dire, — si les contrôles retenus sont le sexe, l'âge et la classe économique-sociale, — d'interroger le nombre d'individus de chaque sexe, de chaque tranche d'âge et de chaque classe économique-sociale qui lui aura été indiqué.

De telles méthodes où l'on construit l'échantillon de façon à ce qu'il soit représentatif, — au lieu de s'appuyer sur un tirage au hasard qui, par application de la loi des grands nombres, rend de plus en plus invraisemblable, à mesure que croît la taille de l'échantillon, une divergence importante, pour quelque caractère que ce soit, entre l'échantillon et l'univers d'où il est tiré, — peuvent être critiquées à plusieurs points de vue. Elles ne permettent pas, en effet, de fixer des limites probables à l'erreur entraînée par la substitution d'un échantillon à l'univers à étudier. Elles n'ont, d'autre part, de chance de fournir une image représentative de cet univers, relativement au caractère à mesurer que si, d'abord, on dispose de statistiques précises et récentes sur la distribution dans la population des contrôles choisis et si, ensuite, le caractère à mesurer a une très forte corrélation avec le caractère retenu pour contrôle. Enfin, et c'est là une des causes principales d'erreurs de ces méthodes, on est obligé de laisser à l'enquêteur une liberté totale dans le choix des personnes à interroger pourvu qu'il ait recueilli, à la fin de l'enquête, des réponses provenant d'individus répartis suivant différents contrôles (sexe, âge, etc...) comme il le lui a été prescrit. Tout naturellement l'enquêteur interrogera les personnes qui lui paraîtront les plus faciles à joindre, c'est-à-dire en particulier celles qui appartiennent au même milieu que celui dans lequel il a l'habitude de vivre.

Les enquêteurs font, en général, partie de la « classe moyenne » et, même en imposant un contrôle basé sur la notion subjective de classe économique-sociale, il est certain que les éléments les plus déshérités de la population et les plus difficiles à joindre (étrangers récemment naturalisés et parlant mal anglais, personnes vivant à l'écart de toute agglomération, etc...) seront rarement choisis

par des enquêteurs qui leur substitueront des personnes plus faciles à interroger. Il en résulte que les éléments les plus pauvres, et qui, généralement votent pour le parti démocrate, sont insuffisamment représentés dans l'échantillon.

Si on étudie, par exemple, la répartition suivant le degré d'instruction des échantillons interrogés depuis 1936 par les Instituts d'Opinion Publique, on trouve que le groupe correspondant à l'instruction la plus faible est de 12 à 20 % trop peu représenté. L'erreur apparaît du même ordre quand on prend pour contrôle la possession du téléphone ou d'une voiture automobile. Cette erreur n'a pas échappée aux directeurs des organismes de mesure de l'opinion publique et, en 1940 et 1944, ils tentèrent de corriger l'échantillon en se basant sur la manière dont les personnes interrogées déclaraient avoir voté, lors de la précédente élection présidentielle.

Mais une telle méthode ne pourrait être utilisée qu'avec beaucoup de précautions, car, au bout de quatre ans, la mémoire peut ne plus être fidèle, et on peut, en toute bonne foi, déclarer qu'on avait voté pour le vainqueur alors qu'on s'était abstenu ou qu'on avait voté pour son concurrent.

En fait, aussi bien en 1940 qu'en 1944, les prévisions de Gallup et Crossley, malgré les corrections apportées, ont fait apparaître que l'échantillon comportait une proportion légèrement trop élevée de Républicains. C'est ainsi qu'en 1940, pour 16 États sur 48, Gallup fit une erreur de plus de 3 points dans l'estimation de la proportion des voix de chacun des partis. Pour 15 de ces 16 États, il sous-estima le nombre de votants démocrates, tandis que pour un seul des États, il sous-estima l'importance des votants républicains.

Quel que soit le soin apporté à la formation et au recrutement des enquêteurs, la méthode de choix de l'échantillon adoptée jusqu'à présent, — qui représente un progrès considérable par rapport aux méthodes grossières utilisées avant 1936 et consistant à recueillir le nombre maximum de réponses sans aucun souci de la représentativité de l'échantillon, — est cependant susceptible d'engendrer des erreurs. Il est certain que toute méthode de sondage ne peut fournir que des résultats entachés d'une certaine erreur. Cependant, les techniques basées sur le choix au hasard des unités permettent de connaître l'erreur probable qui est commise et de la réduire à la limite désirée en augmentant l'importance de l'échantillon, tandis qu'il apparaît que la technique utilisée par les Instituts d'Opinion publique peut introduire des erreurs systématiques non mesurables. Il semble donc que la méthode utilisée puisse être tenue responsable, pour une part, de l'erreur totale.

Il est difficile de chiffrer cette part de l'erreur, mais il semble qu'on puisse estimer qu'elle est comprise entre $1/5$ et $3/5$ de l'erreur totale pour Gallup et Crossley.

Le cas de Roper, qui fit la plus mauvaise estimation, est spécial. Il ne fournit qu'une estimation nationale et son échantillon a toujours compris une trop faible proportion de démocrates dans le Nord et une trop forte proportion dans le Sud. Au cours des élections de 1936, 1940 et 1944, ces erreurs se compensèrent parfaitement, mais en 1948 la candidature dans les États du Sud d'un démocrate dissident détruisit cet équilibre et rendit très mauvaise l'estimation de Roper.

Quoique les méthodes de sondage basées sur un choix au hasard eussent pu

introduire des erreurs équivalentes à celles qu'ont entraînées, en fait, les plans d'échantillonnage de Gallup ou Crossley, il est certain que l'utilisation de ces méthodes marquera un progrès scientifique et sera susceptible, à l'avenir, d'améliorer la valeur des travaux des Instituts d'Opinion Publique. Il faut donc souhaiter qu'une diminution de leur coût élevé rende cette utilisation possible.

2. Même en supposant qu'on puisse interroger un échantillon parfaitement représentatif de la population, il est hors de doute que de nombreuses causes d'erreur peuvent intervenir et que les résultats de l'enquête ne correspondraient pas exactement aux résultats d'un vote effectué le même jour.

Il est évident, tout d'abord, que l'enquêteur peut avoir une influence, par sa présentation, la manière de poser les questions ou d'enregistrer les réponses sur les résultats de l'enquête. Des expériences effectuées en 1940 et 1944 ont montré que si, au lieu d'enregistrer lui-même les réponses, l'enquêteur demande aux personnes interrogées de remplir un bulletin et de le déposer dans une boîte scellée que l'enquêteur transporte, on n'améliore pas d'une façon notable la valeur de l'enquête. En 1940, les essais effectués avaient indiqué que la proportion des votes favorable à Roosevelt était de 4 % plus élevée avec la méthode des bulletins secrets. Au contraire, en 1944 l'utilisation du bulletin secret fournit une proportion de vote favorable à Roosevelt de 2 % plus faible qu'avec la méthode d'enregistrement des réponses par l'enquêteur.

Il est certain, cependant, que la proportion des indécis est beaucoup plus faible lorsqu'on utilise la méthode du bulletin secret, ce qui prouve que nombre de ces indécis ont, en fait, pris une décision, mais n'aiment pas l'exprimer devant un étranger. Aucune étude ne permet de mesurer exactement l'importance de l'erreur qui peut être due au milieu social ou aux préférences politiques de l'enquêteur. Il semble qu'à l'avenir on pourra tenter d'estimer ces erreurs en utilisant la méthode des réseaux superposés de sondage, — c'est-à-dire qu'on comparera les résultats d'enquêtes menées par deux groupes d'enquêteurs différent soit par le milieu social, soit par les opinions politiques, soit par toute autre caractéristique dont on a désiré mesurer l'effet, et, ayant à interroger deux échantillons semblables de la même population.

Une autre cause d'erreur peut résider dans la recherche de la partie de l'échantillon qui ira effectivement voter. S'il est simple d'éliminer de l'échantillon toutes les personnes âgées de moins de vingt et un ans qui ne sont pas électeurs, il est extrêmement difficile, en revanche, de savoir si la personne interrogée et qui déclare qu'elle votera pour Truman ou pour Dewey, se dérangera ou non pour aller voter; et il est certain que la réponse à la question « Avez-vous l'intention d'aller voter? » qui est en général posée, ne permettra pas de savoir avec exactitude si cette personne ira voter. C'est par conséquent vers une technique plus poussée du questionnaire que devront s'orienter les Instituts d'Opinion Publique pour tenter de réduire les erreurs qui peuvent ainsi s'introduire.

Enfin, même si on a pu déterminer les personnes qui n'iront pas voter, il reste, parmi celles qui ont l'intention de voter, une certaine proportion de personnes qui ne savent pas, au moment où on les interroge, ou qui ne veulent pas révéler le nom du candidat qui figurera sur le bulletin qu'elles déposeront dans l'urne. La solution qui consiste à répartir ces indécis proportionnellement à la population qui a pris une décision peut être cause d'importantes erreurs. Ici encore

des études approfondies de psychologie et de technique du questionnaire pourront permettre des améliorations de la technique actuellement utilisée. En 1948, la proportion des indécis dans l'échantillon de Gallup fut particulièrement élevée, atteignant 15 % des personnes ayant l'intention de voter. On put déterminer pour la moitié de ces personnes vers quel candidat allaient leurs préférences, mais près de 8 % des bulletins durent être classés comme indécis. Or, il semble que plus de la moitié de ces indécis, d'après les enquêtes faites après les élections, votèrent effectivement et votèrent pour Truman dans la proportion de 65 %. Il ne semble pas cependant qu'on puisse imputer à cette cause une erreur de plus de 1,5 points.

3. Les Instituts d'Opinion Publique ont laissé entendre que la cause principale de leur échec était dans une variation de l'opinion qui serait survenue au dernier moment. Il semble certain qu'il ait existé une tendance nette dans les dernières semaines favorisant Truman, soit en raison de la campagne menée par le Président sortant, soit par la peur qu'ont eue certains de voir se terminer la prospérité économique que les démocrates leur avaient apportée durant seize ans de gouvernement, soit parce que le candidat progressiste Wallace n'a pu finalement se présenter dans certains États, soit pour toute autre cause. Cette variation de l'opinion est très difficile à mesurer. Des enquêtes ont été effectuées après les élections. Gallup a pu montrer, en particulier, que les électeurs de Dewey avaient pris leur décision plus tôt que ceux de Truman, ainsi que le prouvent les résultats reproduits dans le tableau ci-dessous :

ÉPOQUE à laquelle les électeurs ont pris leur décision	ÉLECTEURS de Truman	ÉLECTEURS de Dewey	ENSEMBLE
	%	%	%
Avant l'ouverture de la campagne électorale	46	54	54
Au début de la campagne électorale	11	12	12
Première quinzaine d'octobre	4	2	3
Seconde quinzaine d'octobre	13	5	9
Jour de l'élection	5	3	4
Non précisée	21	14	18
	100	100	100

Il semble cependant que la variation de l'opinion durant les deux dernières semaines, c'est-à-dire entre la dernière enquête de Gallup ou Crossley et le jour de l'élection, n'ait pu engendrer une erreur dans l'estimation de la proportion de voix de Truman supérieur à 2 points. De plus, la tendance qui se manifestait très nettement en faveur de Truman en comparant les enquêtes successives effectuées au cours des derniers mois aurait dû inciter à plus de prudence dans le pronostic ou montrer la nécessité d'effectuer des enquêtes à une période encore plus rapprochée du jour de l'élection.

4. Toutes ces causes d'erreur étaient connues de Gallup et, comme il tient à le dire, il est probable que ses enquêtes de 1948 n'étaient ni meilleures ni plus mauvaises que celles des années précédentes. Il apparaît simplement que pour des élections aussi disputées que l'élection présidentielle de 1948, les méthodes utilisées ne sont pas encore assez précises pour permettre de désigner le vainqueur avec une très forte probabilité de ne pas se tromper.

Gallup l'avait compris et c'est pourquoi en 1944 il avait refusé sur la base de ses enquêtes de désigner le vainqueur. En 1948, les journalistes, omettant bien souvent ses réserves, l'ont entraîné à effectuer un pronostic malheureux. Il est certain qu'il n'a pas suffisamment averti son lecteur des possibilités d'erreur de ses enquêtes et les méthodes journalistiques américaines ont transformé un pronostic basé sur une enquête de sondage en une certitude mathématique.

Si on se souvient que Dewey aurait été élu président s'il avait obtenu la majorité dans les États d'Ohio, d'Illinois et de Californie, États où le nombre des voix de Truman ne fut pas supérieur de 1 % à son nombre de voix, on pensera que Gallup n'eut pas de chance. Mais, en revanche, on se réjouira en espérant que son retentissant échec amènera tous les Instituts d'Opinion à faire, dans le domaine du plan d'échantillonnage, du questionnaire, de l'interview et de la prévision, des progrès tels qu'ils reconquièrent rapidement la confiance un instant perdue du public.

Raymond LÉVY-BRUHL.

*
* *