

## Une petite histoire de sondages

L'idée de s'intéresser à l'opinion publique, de tenter de la comprendre, de la mesurer et de l'influencer n'est pas récente :

« L'avènement [...] du suffrage universel en a fait un acteur de principe. Aujourd'hui, l'opinion publique est une donne inéluctable des gouvernements soucieux de fonder leur domination sur un minimum de légitimité apparente, quels que soient le régime et le degré réel de démocratie. »<sup>1</sup>

Les premiers exemples de sondages de l'opinion publique peuvent être retracés dans les pratiques des journaux locaux américains qui recouraient au « vote de paille » (*straw vote* ou *straw poll* en anglais) pour tenter de prédire les résultats électoraux. Le vote de paille consiste essentiellement à distribuer des *similis* bulletins que les lecteurs du journal peuvent retourner en y indiquant leur intention de vote. Les résultats ainsi compilés servent de prévisions aux vrais résultats électoraux.

Cette façon de sonder l'opinion publique serait apparue dans les années 1820. Le Harrisburg Pennsylvanian, en 1824, aurait ainsi correctement estimé la domination au vote populaire d'Andrew Jackson (41,3%) sur son rival, John Quincy Adams (30,9%). Même si, au final, le système des grands électeurs a fait en sorte que la présidence est allée à Adams, le journal pouvait légitimement revendiquer un certain succès pour sa méthodologie.

La pratique du vote de paille s'est par la suite répandue, devenant de plus en plus populaire. En 1916, la publication *Literary Digest*<sup>2</sup> (1890 – 1938 ; voir Boorstin, 2012<sup>3</sup>) jouissait d'une très grande popularité et d'une large diffusion sur l'ensemble du territoire américain. Cette situation était favorable à un projet de vote de paille à l'échelle nationale

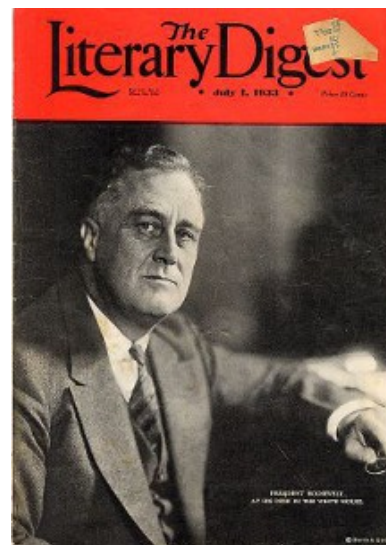
<sup>1</sup> [http://fr.wikipedia.org/wiki/Opinion\\_publicque](http://fr.wikipedia.org/wiki/Opinion_publicque)

<sup>2</sup> Les *digests* sont des publications qui se spécialisent dans les abrégés (résumés de livres), les histoires courtes et les articles qui vulgarisent, dans certains cas, des publications scientifiques. Un des *digests* les plus connus est évidemment le fameux *Readers Digest*.

<sup>3</sup> Boorstin D.J., *Le triomphe de l'image*, Lux Éditeur, 2012. Cet ouvrage est une traduction du titre original *The Image: A guide to Pseudo-Events in America* paru en 1961. Plusieurs pages sont consacrées à une critique sociale de ce type de publications ; malgré le fait que les propos tenus datent de plus de 50 ans, on ne peut s'empêcher de les trouver des plus actuels.

et la publication prédit ainsi correctement l'élection du président Woodrow Wilson (2<sup>e</sup> mandat à titre de 28<sup>e</sup> président des États-Unis). Entre 1916 et 1932, le *Literary Digest* s'est enquis des intentions de vote des Américains en postant et en distribuant des millions de cartes, les prévisions étant simplement basées sur le résultat du décompte des cartes retournées. De 1920 à 1932, les présidents Harding, Coolidge, Hoover et Roosevelt ont tous été prédits gagnants par le *Literary Digest*.

Quand  $n = 2\,266\,566$



L'édition du 31 octobre 1936 du *Literary Digest* prédit une victoire de Alfred M. Landon (candidat Républicain) aux dépens de Franklin D. Roosevelt (candidat Démocrate). Près de 2 300 000 *simili* votes ont été retournés : 1 293 669 (57%) pour Landon contre 972 897 (43%) pour Roosevelt<sup>4</sup>. Fort de cinq prédictions confirmées (1916, 1920, 1924, 1928 et 1932), rien ne pouvait ébranler la confiance du *Literary Digest* dans sa capacité à annoncer le nom du prochain président des États-Unis.

Par ailleurs, soulignons que ces votes de paille étaient une opération payante pour la revue : en plus de créer la nouvelle, chaque carte de vote était accompagnée d'un formulaire d'abonnement à la publication.

<sup>4</sup> <http://historymatters.gmu.edu/d/5168>

En 1936, le *Literary Digest* est une publication très populaire et respectée. Forte de cette renommée et de sa capacité à rejoindre des millions d'Américains, elle distribue 10 000 000 (environ le quart de la population des 48 états américains de l'époque) de cartes afin que l'opinion publique se prononce sur les résultats de la présidentielle à venir. Le taux de retour a été d'environ 25%. En fin de compte, c'est un échantillon faramineux de près de 2 300 000 répondants qui a permis aux éditeurs du *Literary Digest* d'annoncer la victoire à venir de Landon.

Pour avoir une idée de ce que représente un tel nombre de répondants dans le domaine des sondages, mentionnons qu'en 2006 les résultats d'un sondage mondial réalisé en 2005 (par la *Gallup International Association* et Léger Marketing) ont été publiés<sup>5</sup> : 68 pays faisaient partie de l'échantillon et une équipe de plus de 5 000 interviewers a été mise à contribution. Malgré l'ampleur de la tâche et de sa portée géographique, le nombre de répondants n'a été que de 53 749, soit seulement à peine plus de 2% du nombre de répondants du vote de paille du *Literary Digest*.

#### ***Une tendance prend fin ...***

*A trend is a trend. But the question is, will it bend? Will it alter its course through some unforeseen force and come to a premature end?* – Sir Alexander K. Cairncross (1911 – 1998), économiste.

Toute bonne chose a une fin ! La série de prédictions heureuses du *Literary Digest* ne s'étirera pas jusqu'à six. Roosevelt sort grand gagnant des élections et l'erreur de prévision est énorme : une différence de 21 points de pourcentage<sup>6</sup>. La prévision de la revue était de 57% du vote populaire pour Landon alors qu'au final, il n'en aura reçu que 36%<sup>7</sup>. Mais qui plus est, le fait d'annoncer Landon comme gagnant est un écart encore plus gigantesque par rapport à la réalité lorsque l'on tient compte du fait que le système américain des grands électeurs ne répartit pas forcément ceux-ci au *pro rata* du vote populaire global. C'est ainsi qu'au chapitre des grands électeurs, Roosevelt a carrément anéanti son opposant en ralliant 98% des grands électeurs contre un presque insignifiant – au sens de non significatif – 2% pour Landon.

La situation est des plus ironique et catastrophique pour le *Literary Digest*. Son sondage était l'un des plus grands (nombre de répondants) et coûteux à avoir été mené. Les attentes étaient évidemment en rapport avec les efforts ; compte tenu du nombre de répondants, la fiabilité des résultats n'était pas questionnable, *a priori* du moins. *A posteriori*, on ne peut qu'admettre qu'un mauvais échantillonnage est une source potentielle de biais qui ne peut être compensée par une augmentation du nombre des répondants. L'aspect critique était donc l'échantillonnage et non la taille de l'échantillon. Au moins quatre sources d'erreur (de biais) ont pu contribuer à l'échec de cette sixième tentative de prédire les résultats d'une présidentielle américaine : la sélection des

<sup>5</sup> Léger Marketing, *L'Opinion du monde*, Les Éditions Transcontinental, 2006.

<sup>6</sup> Les sources ne donnent pas toutes exactement le même chiffre, mais l'ordre de grandeur de 20% est généralement admis.

<sup>7</sup> [http://fr.wikipedia.org/wiki/Élection\\_présidentielle\\_américaine\\_de\\_1936](http://fr.wikipedia.org/wiki/Élection_présidentielle_américaine_de_1936)

destinataires pour les cartes de *simili* vote, le taux de non-réponse, la non-représentativité sociodémographique et la non-représentativité idéologique.

#### La sélection des destinataires pour les cartes de simili vote

La liste d'envoi (échantillon initial de 10 000 000) a été constituée à partir des registres des abonnés de la revue, des bottins téléphoniques, des registres de propriétaires d'automobiles, des registres d'autres revues et publications, etc... Compte tenu du contexte économique de l'époque, une telle liste favorisait donc forcément les personnes à revenus moyens et élevés. Pensons qu'en 1936, la téléphonie résidentielle était loin d'être un service accessible à tous, tout comme le fait de posséder une automobile.

#### Le taux de non-réponse

Malgré le fait que le nombre de répondants ait été de près de 2 300 000, il n'en demeure pas moins que cela ne représente qu'à peine plus du quart de l'échantillon initial de 10 000 000. Ceci constitue un faible taux de réponse et dans un tel cas l'échantillon souffre d'un biais de non-réponse. Qui plus est, le type de sondage retenu – postal – est probablement celui qui favorise le plus un tel type de biais compte tenu que dans une telle situation rien ne vient encadrer la décision de participer ou non au sondage et rien ne peut être fait pour améliorer le taux de réponse.

#### La non-représentativité sociodémographique

*These figures are exactly as received from more than one in every five voters polled in our country – they are neither weighted, adjusted, nor interpreted.* – *Literary Digest*, 31 octobre 1936.

Que dire de plus ? Les données auraient probablement dû être pondérées (selon le revenu des répondants par exemple) et ajustées pour tenir compte du taux de non-réponse et de ce qui différencie les répondants des non-répondants.

#### La non-représentativité idéologique

Compte tenu de la surreprésentation des classes sociales moyennes-élevées et élevées dans l'échantillon (et donc dans les répondants), les intentions de vote recueillies étaient celles d'électeurs disposant de revenus au-dessus de la moyenne. Hors, de tels électeurs étaient réputés être des supporteurs de l'idéologie républicaine (peut-être est-ce toujours le cas). Alors même qu'une large partie de l'électorat avait peu de chances de s'exprimer dans le vote de paille du *Literary Digest*, leur idéologie politique ne s'est donc pas reflétée dans les prévisions proclamant vainqueur le candidat Landon :

*But 1936 was the height of the depression. At that time, phones, cars, and country clubs were more often available only to rich. The economy was the central campaign issue of the election with Roosevelt advocating the New Deal that was attractive to lower income people, so the opinions of the rich and poor were bound to differ on this issue. As such, their sample turned out to be incredibly biased, and ultimately, useless.*<sup>8</sup>

Le *Literary Digest* disparut en 1938 ; cette éventualité n'était évidemment pas envisagée en juillet 1936 !

<sup>8</sup> <http://www.oxfordmathcenter.com/drupal7/node/251>

*... une autre voit le jour*

*Il n'existe que les atomes et l'espace vide ; tout le reste est doxa*<sup>9</sup>. – Démocrite (460-370 av. J.-C.)

L'histoire du *Literary Digest* est encore plus intéressante si la grille d'analyse inclut un autre jalon important de l'histoire des sondages : la fondation, par le statisticien George H. Gallup (1901 – 1984), de l'American Institute of Public Opinion en 1935. La même élection qui causa la perte du *Literary Digest* consacra la nouvelle entreprise de Gallup. Avec un échantillon beaucoup plus petit – 50 000 – et surtout, avec une méthodologie solidement ancrée sur des principes scientifiques (statistiques et probabilités), Gallup a été capable de prédire la victoire de Roosevelt, contredisant ainsi directement les résultats de la revue populaire et remettant en question, du coup, une approche méthodologique qui avait pourtant connu du succès, mais qui n'était qu'un géant aux pieds d'argile<sup>10</sup>.



Gallup a été une inspiration pour plusieurs autres en Amérique, mais aussi en Europe. Jean Stoetzel (1910 – 1987), alors qu'il était professeur à l'Université Columbia de New York, rencontra Gallup et il porta un fort intérêt à l'approche de ce dernier envers la mesure de l'opinion publique.

De retour en France, Stoetzel fonda, en 1938, l'Institut français d'opinion publique (IFOP)<sup>11</sup> et adopta le modèle de Gallup d'enquêtes indépendantes et de publication des résultats dans les médias.

En tant que statisticien, Gallup a aussi reconnu les applications possibles des principes scientifiques aux études de marché : il a travaillé pour l'agence de publicité Young & Rubicam où il a été directeur de la recherche et ensuite vice-président (1937 à 1947), et il a fondé avec Claude E. Robinson, en 1948, l'agence de publicité *Gallup and Robinson*.

En 1945, c'est en Grande-Bretagne que Gallup implante une filiale qui connut aussi un succès rapide, ayant produit l'un des seuls sondages prédisant la surprenante victoire du parti Travailliste aux élections générales de 1945. Lors de cette élection, le candidat Travailliste Clement Attlee avait en effet obtenu la majorité aux dépens du Conservateur Winston Churchill.

Par la suite, George H. Gallup fonda l'agence *Gallup International* (1947) ; il en fut le président pendant plus de 35 ans. En 2010, la *Gallup International Association* (GIA) fusionna avec la *Worldwide Independent Network of market research* (WIN) ; cette entité regroupe aujourd'hui des membres provenant de plus de 100 pays.

<sup>9</sup> Du grec *doksa* : opinion, rumeur.

<sup>10</sup> Voir [http://en.wikipedia.org/wiki/Historical\\_polling\\_for\\_U.S.\\_Presidential\\_elections](http://en.wikipedia.org/wiki/Historical_polling_for_U.S._Presidential_elections) pour prendre connaissance des sondages pour les élections américaines de 1936 à 2012, incluant les comparaisons avec les résultats réels des élections.

<sup>11</sup> [http://www.ifop.com/?option=com\\_publication&type=publication&id=37](http://www.ifop.com/?option=com_publication&type=publication&id=37)

L'héritage de Gallup est certainement la démocratisation des sondages et le recours à une approche méthodologique et scientifique qui a eu un immense impact sur la fiabilité des résultats et la confiance du public, des entreprises, des organisations et des gouvernements envers ceux-ci.

*René Gélinas – Directeur général, Takto marketing*