Bulletin de méthodologie sociologique
Bulletin of sociological methodology

89 | 2006
January

Absence de rigueur dans les sondages d'opinion

Les leçons d'un retour d'expérience

Absence of rigor in opinion polls, Lessons learned from experience

Patrick Che et Jean-Pierre Pagès

Édition électronique
URL : http://journals.openedition.org/bms/706
ISSN : 2070-2779

Éditeur
Association internationale de méthodologie sociologique

Édition imprimée
Date de publication : 1 janvier 2006
Pagination : 5-30
ISSN : 0759-1063

Référence électronique
Patrick Che et Jean-Pierre Pagès, « Absence de rigueur dans les sondages d'opinion », *Bulletin de méthodologie sociologique* [En ligne], 89 | 2006, mis en ligne le 20 mars 2008, consulté le 19 avril 2019.
URL : http://journals.openedition.org/bms/706

Ce document a été généré automatiquement le 19 avril 2019.

© BMS

Absence de rigueur dans les sondages d'opinion

Les leçons d'un retour d'expérience

Absence of rigor in opinion polls, Lessons learned from experience

Patrick Che et Jean-Pierre Pagès

Introduction

1 Comment procéder pour améliorer la qualité des données recueillies lors des enquêtes d'opinion ?

2 Depuis une vingtaine d'années, les contrôles effectués sur les données dont dispose Agoramétrie ont tous abouti au même constat : le « bidonnage » est loin d'être un phénomène exceptionnel. Le briefing auquel on dit se livrer avant d'aller sur le terrain , les contrôles qu'on dit effectuer en cours d'enquête et les tests de cohérence qui, en principe, sont réalisés lors ou après la saisie des données s'avèrent largement insuffisants pour garantir la qualité du recueil.

3 La constance du diagnostic montre qu'il est réducteur d'incriminer les seuls enquêteurs et enquêtés quand des faiblesses qui ne devraient pas apparaître sont détectées dans les données d'enquête. La critique doit porter sur l'ensemble du processus qui conduit des objectifs et du choix du protocole d'enquête (questionnaire et plan de sondage), jusqu'au traitement final des données et à leur interprétation. Les acteurs de l'amont devraient en particulier se sentir beaucoup plus concernés par cette mission confiée aux sondeurs d'assurer la phase intermédiaire d'organisation, de réalisation et de contrôle *a priori* et *a posteriori* du recueil : dans quelle mesure, compte tenu des objectifs, du protocole et du budget imposé, est-il techniquement possible d'obtenir aujourd'hui des informations non biaisées sur les opinions de nos concitoyens ?

4 Voulant jeter les bases d'un véritable retour d'expérience, on cherche ici a faire mieux qu'une précédente étude non publiée (Che et Pagès, 1998) qui avait porté sur les réponses à plus de deux mille interviews réalisées en face à face. Il s'agit non seulement de revoir la

méthode utilisée pour repérer les anomalies et de la rendre opérationnelle pour des contrôles plus systématiques, mais aussi de l'enrichir de façon qu'un jugement circonstancié puisse être porté sur le processus d'enquête par tous ceux qui, y participant en tant qu'acteurs, devraient avoir la qualité comme préoccupation première. Cette connaissance partagée des insuffisances et de leurs causes est essentielle pour aller de l'avant, l'intention étant bien au final de modifier les conditions dans lesquelles s'exercent les activités « sondage ».

Nombres et durées des entretiens

5 Réalisée principalement en décembre 2003 en face à face avec l'aide de l'ordinateur (système CAPI) auprès de 2520 personnes de plus de 18 ans, la nouvelle enquête a nécessité la collaboration de 149 enquêteurs. En 1997, la précédente enquête, d'importance comparable mais effectuée en face à face sans CAPI, avait mobilisé plus de 230 enquêteurs. Le recours à CAPI semble donc modifier certaines caractéristiques du réseau sur lequel on s'appuie ; le nombre d'enquêteurs est beaucoup moins grand : quelle en est la raison ?

6 Alors qu'en 1997 le nombre d'interviews réalisées par un même enquêteur variait entre 1 et 17, 40 % des enquêteurs en réalisant 10, ce nombre varie en 2003 entre 1 et 50, 32 % des enquêteurs en réalisant 15 (voir figure 1). On peut donc considérer qu'avec CAPI la charge de travail « enquêteur » a augmenté en moyenne d'environ 50 %. Ce changement dans l'organisation du travail pouvant être à l'origine de certaines des anomalies constatées, le souci de qualité exige d'en dire plus : quid, avec CAPI, des relations actuelles entre instituts de sondage et enquêteurs ou réseaux d'enquêteurs ?

Figure 1 : Nombre d'interviews

7 La grande variabilité du nombre d'interviews par enquêteur interpelle par ailleurs sur le lien qu'entretient ce nombre avec la manière de décrocher voire même de pratiquer l'interview : en quoi certains enquêteurs sont-ils plus « expérimentés » que d'autres ? Comment ont procédé les stakhanovistes du sondage (plus de 30 interviews) par comparaison avec ceux qui ont réalisé moins de 5 interviews (17 enquêteurs) ? Quid de ceux qui ont été capables d'aligner dans une même journée plus de 8 interviews (c'est un exploit !) ? A-t-on véritablement imposé des normes aux enquêteurs dans la recherche des répondants ? Quels contrôles l'institut de sondage a-t-il véritablement effectués ? Sont-ils efficaces compte tenu que, connus des enquêteurs, ils peuvent être facilement déjoués ?

8 L'entretien était prévu pour durer en moyenne 45 minutes ... *a posteriori,* on constate qu'il a duré en moyenne 43,30 minutes (voir figure 2). La distribution des durées, de forme approximativement log-normale, réserve malgré tout quelques surprises. À l'évidence, certains entretiens ont été menés tambour battant, le plus rapide ayant été enlevé en 20 minutes ; près d'une dizaine ont pris moins de 25 minutes ; plus de 30 % se sont faits en moins de 35 minutes :comment expliquer de telles performances ? Bien entendu, on trouve aussi des entretiens fort longs, plus de 55 ayant duré plus de deux heures. Tous ces faits méritent éclaircissement...

Figure 2 : Durée des interviews

9 La variabilité forte des durées d'interviews est loin d'être uniquement le fait des interviewés, car le temps moyen mis par l'enquêteur pour réaliser un entretien varie lui aussi fortement, entre 31 et 77 minutes (voir figure 3). On retrouve encore ici une distribution des durées de forme approximativement log-normale, même si la population des enquêteurs semble nettement se diviser en deux catégories, les plus rapides, d'une part, qui tournent en 40 minutes environ, et les plus lents, d'autre part, qui mettent au moins 12 minutes de plus en moyenne. Parmi les enquêteurs, 5 ont mis 31 minutes en moyenne pour réaliser un entretien ; 22 ont mis moins de 35 minutes ; 51 des 149 enquêteurs ont passé en moyenne plus de 45minutes. Si l'on écarte les 12 plus lents qui ont consacré en moyenne plus d'une heure à chaque entretien, la rapidité des enquêteurs à récolter des opinions semble avoir été sous-évaluée. On aimerait donc encore ici en savoir plus : l'institut de sondage a-t-il cherché a comprendre le pourquoi de cette grande dispersion des performances et à savoir si vitesse et qualité étaient antinomiques ?

Figure 3 : Durée moyenne des interviews

10 En résumé, les durées, mais aussi les dates d'interviews qui permettent de décompter le nombre d'interviews réalisées dans une même journée, étant relevées sans intervention de l'enquêteur grâce à CAPI, apparaissent comme fondamentales pour le contrôle de la qualité des interviews ou des séries d'interviews. Mais sans des informations complémentaires que le commanditaire est en droit d'exiger, nul n'est à même de décider s'il y a véritablement « anguille sous roche ».

Usage des modalités de réponse

11 Le questionnaire utilisé en décembre 2003 comprend, entres autres, quatre batteries de propositions ou d'items Q1, Q6, Q15, Q16, auxquelles en répond suivant, soit une échelle d'accord en 5 paliers (Q1, Q6, Q16), soit une échelle de préoccupation en 4 paliers (Q15), une modalité étant réservée pour le « sans opinion » ou la non réponse. La question Q1 qui comprend 87 propositions est située en tête du questionnaire ; vient ensuite Q6 qui comprend 31 propositions, puis Q15 avec 16 items et enfin Q16, composée de 45 propositions, qui apparaît en fin de questionnaire.

12 Les taux moyens d'utilisation des modalités des échelles, calculés sur les 2215 interviews, sont sans surprise (figures 4 à 7), le faible taux de « sans opinion » et de non réponses incitant à croire que le questionnaire n'a pas provoqué chez le plus grand nombre une réaction de rejet.

Figure 4 : Taux d'utilisation des modalités – Q1

Figure 5 : Taux d'utilisation des modalités – Q6

Figure 6 : Taux d'utilisation des modalités – Q15

Figure 7 : Taux d'utilisation des modalités – Q16

13 Pour Q1, les taux d'utilisation des 5 modalités de l'échelle d'accord sont tous voisins de 20 %. Cette batterie « passe bien » au dire des enquêteurs interrogés dans le passé. Comme elle traite de thèmes qui font controverse et sont présentés aux interviewés sous forme de propositions qui expriment indifféremment une adhésion ou un rejet, il est naturel que les « pour » et les « contre » se retrouvent en mêmes proportions . Cette uniformité de la distribution des taux moyens pour Q1 donne en quelque sorte le sentiment que les thèmes ont été bien choisis et que les propositions ont dans leur ensemble été bien construites, le constat pour Q6, Q15 et Q16 étant bien différent.

14 Les manières d'utiliser les modalités de l'échelle, si on les résume aux distributions fournies, sont voisines pour Q6 et Q16, qui traitent respectivement de l'ouverture du marché de l'électricité et du nucléaire. Le faible taux de désaccords (modalités 1 et 2) et le fort taux d'accords « mous » (plus de 30 % pour la modalité 4) confortent l'hypothèse qu'un thème de controverse devient moins impliquant quand on tourne autour avec insistance. Si la portée symbolique du thème (pensons au nucléaire, aux 35 heures ou aux immigrés) n'est plus concentrée sur une seule proposition « bien carrée » (cas de Q1), elle perd de sa force. Diluée sur pléthore, surtout sous forme d'arguments jamais définitifs (cas de Q6 et Q16) , le symbole en tant qu'étendard perd de sa puissance pour orienter la réponse. Surtout en fin de questionnaire où la lassitude est plus grande, la tactique du moindre effort incite alors à déclarer valides, et cela sans trop se mouiller puisqu'elles sont en définitive ni vraies, ni fausses, les assertions du questionneur, ce dernier représentant l'autorité.

15 Pour Q15, le phénomène est encore plus criant si on admet que c'est l'autorité des médias qui manifeste sa présence dans la batterie des 16 risques : comment ne pas répondre « très préoccupé » ou « moyennement préoccupé » (taux d'utilisation supérieur à 70 %) quand on est abreuvé d'informations dramatisantes ? Au-delà d'une angoisse véritable que l'on trouve chez certains, les taux obtenus montrent surtout que le conformisme l'emporte sur le regard critique ou sur le détachement vis-à-vis des malheurs du monde tels qu'ils sont mis en scène. Bref, le propos sur les risques que l'on récupère par le sondage est avant tout stéréotypé.

16 Les remarques précédentes sur les taux d'utilisation des modalités de réponse montrent qu'ils sont susceptibles d'alimenter la réflexion sur le comportement face aux sondages des citoyens et sur la manière d'envisager un questionnement. Mais ces taux sont aussi des indicateurs permettant de repérer certains comportements aberrants aussi bien des enquêtés que des enquêtants.

17 Les taux calculés sur l'ensemble des interviewés et leurs écarts-types permettent de construire des intervalles de confiance pour les taux d'utilisation des modalités de réponse par enquêteur. Pour un enquêteur faisant 16 interviews, les taux devraient varier pour Q1 par exemple dans les intervalles (on a vu « large ») :

$1 \rightarrow (12,27)$; $2 \rightarrow (11,25)$; $3 \rightarrow (11,28)$; $4 \rightarrow (13,28)$; $5 \rightarrow (10,28)$; $6 \rightarrow (0,9)$.

18 Il est donc facile de repérer les enquêteurs « hors-normes » en examinant les tableaux décrivant les taux d'utilisation des modalités intra-enquêteurs. En ce qui concerne l'usage de la modalité « 1 » dans Q1, par exemple, sont hors normes les enquêteurs, d'une part : 10,14,16,118 et 125 (taux trop élevés) ; d'autre part : 33,126 et 136 (taux trop bas). Faut-il déclarer suspects ces 8 enquêteurs qui ont réalisé plus de 7 % des interviews ? Non, car les

limites autorisées ne sont pas trop dépassées et que, pour l'instant, on ne dispose pas de critère pour décider s'il y a véritablement tromperie sur la marchandise. Seul, isolément, l'usage immodéré de la « non-réponse » (modalité 6) peut être considéré comme inadmissible ; on est déjà donc dès à présent en droit de douter du sérieux des interviews réalisées par 7 (21), 19 (25), 69 (15), 91 (17), 92 (5) et 127 (15).

19 L'examen des interviews qui présentent des taux trop forts de recours pour Q1 à l'une des six modalités de réponse (voir figure 8) révèle certains faits surprenants. Il y a quelques perles : les 77 réponses « 1 » sur 87 enregistré par 118 (3) ; les 87 réponses « 3 » ou « 4 » de 102 (5) ; les 87 accords (« 3 », « 4 » ou « 5 ») de 73 (7) ; les 87 « 6 » de 92 (4)... Ces entretiens, qui ne sont pas recevables, montrent que le contrôle (s'il existe !) manque de rigueur. Mais, pour approfondir l'analyse, l'effort doit porter avant tout sur le classement des enquêteurs : quels sont a priori les plus suspects ? Le relevé systématique pour Q1 des bizarreries permet d'établir un premier palmarès : l'enquêteur numéroté 91 remporte la palme avec 9 interviews curieuses sur 11 ; viennent ensuite 92 (4 sur 5) ; puis 73 (3 sur 7), 66 (3 sur 15), 105 (3 sur 15), 126 (3 sur 15), 34 (3 sur 30) et 14 (3 sur 50)... etc...

Figure 8 : Liste des interviews aberrantes - Q1

Figure 8 : Liste des interviews aberrantes - Q1 (suite)

20 Voilà où nous conduit déjà le simple examen des taux d'utilisation des modalités de réponse, les enquêteurs distingués sur Q1 pouvant l'être aussi sur Q6, Q15, ou Q16, comme le sont assurément 7, 19, 91 ... et 118.

Proximités entre interviews

21 Les analyses nombreuses auxquelles on s'est déjà livré ont montré que le bidonnage où le comportement de faiblesse face au répondant ou à la sélection des répondants se traduisait très souvent par des ressemblances fortes entre les différentes données recueillies par l'enquêteur, qui peuvent être dues à un recours trop fréquent à la modalité centrale ou à la non réponse. D'où l'intérêt de mesurer des proximités ou des distances entre interviews, susceptibles d'être respectivement plus grandes ou plus petites en cas d'indélicatesse ou d'incompétence. La distance choisie en 1997 pour tenter de repérer les indélicatesses conduisait à accorder une importance démesurée aux contradictions maximales entre deux interviews : on avait travaillé sur les carrés des écarts entre les numéros des modalités de l'échelle, la non réponse étant assimilée à la modalité centrale (Che et Pagès, 1998). On a préféré cette fois travailler à partir des coïncidences et des discordances.

22 Si deux individus i et i' répondent au hasard aux n items d'une batterie avec une probabilité constante Pj de choisir la modalité de réponse « j », les nombres de réponses identiques (coïncidence) ou différentes (discordance) peuvent être considérés comme des variables binomiales de paramètres : n et respectivement $a = \Sigma_j P_j^2$ ou $d = 1 - a$. Les taux de coïncidences et de discordances sont alors des variables aléatoires d'espérances respectives a ou d et de variance $s^2 = ad/n$.

23 Se référant aux taux moyens d'utilisation des modalités précédemment calculées et après les avoir multiplié par 100, on obtient les résultats arrondis suivants :

- pour Q1 (n = 87) : a = 19, d = 81, s = 4,2

- pour Q6 (n = 31) : a = 20, d = 80 , s = 7,2
- pour Q15 (n = 16) : a = 30, d = 70 , s = 11,5
- pour Q16 (n = 45) : a = 22, d = 78 , s = 6,2

24 Avec a et d, on peut déjà se faire une idée des valeurs que pourraient prendre les « proximités » $a_{ii'}$ et les « distances » $d_{ii'} = 1 - a_{ii'}$ entre interviews, quand elles sont définies à partir des taux de coïncidences et de discordances. Mais pour contrôler le travail des enquêteurs, il faut aller plus loin : disposer de distributions de référence pour les distances $d_{ii'}$. Pour en obtenir, on a prélevé au hasard un entretien sur 10, soit 220 entretiens, et calculé pour les quatre batteries les 24 090 distances $d_{ii'}$.

25 Les moyennes des distances calculées qui sont affichées avec les histogrammes (figures 9 à 12) ne sont pas éloignées des valeurs précédemment obtenues pour le modèle de réponse au hasard le plus simple. Pour une batterie de propositions ou d'items, la distance entre interviews peut être alors considérée comme une variable aléatoire dont les moyennes et écarts-types d et s sont ceux obtenus pour les 220 interviews prélevés au hasard. Si tous les enquêteurs font 16 interviews, la distribution des distances moyenne entre les 16 interviews devrait être alors approximativement normale, de moyenne d et d'écart type s/√120. On constate immédiatement que les 4 distributions obtenues pour les 132 enquêteurs ayant réalisé entre 6 et 50 interviews ne respectent pas ce canon : les distances intra-enquêteurs sont plus faibles, voire beaucoup plus faibles que les distances de référence. Ces distributions, qui sont dissymétriques (celles correspondant à Q1 particulièrement) permettent de classer les enquêteurs en 5 catégories selon les distances moyennes qui leur sont associées :

1 : extrêmement faible ; 2 : très faible ; 3 : faible ; 4 : normale ; 5 : forte.

Figure 9 : Distribution des distances moyennes entre interviewers pour Q1, pour les 132 enquêteurs ayant realisé au moins six entretiens

Figure 10 : Distribution des distances moyennes entre interviewers pour Q6, pour les 132 enquêteurs ayant realisé au moins six entretiens

Figure 11 : Distribution des distances moyennes entre interviewers pour Q15, pour les 132 enquêteurs ayant realisé au moins six entretiens

Figure 12 : Distribution des distances moyennes entre interviewers pour Q16, pour les 132 enquêteurs ayant realisé au moins six entretiens

26 Le croisement des classifications obtenues pour les batteries Q1 et Q16, qui sont situées respectivement en début et en fin de questionnaire, met clairement en évidence les enquêteurs sortant le plus de l'épure et donc que le critère des distances entre interviews suggère d'aller examiner en premier (voir figure 13). Le critère « Q1 » étant plus fiable que le critère « Q16 » pour repérer les déviances, on peut considérer *a priori* que tous les enquêteurs qui se retrouvent dans les deux premières lignes du tableau sont suspects et cela d'autant plus qu'ils se situent plus à gauche. Parmi ceux-ci, on retrouve les enquêteurs : 73, 91, 102 et 126, qui déjà précédemment avaient été distingués.

Figure 13 : Distances moyennes intra-enquêteurs (132 enquêteurs) pour Q1 et Q16

		1	2	3	4	5
	1	6 102 141	80 73	22	10 91	
	2	136	76 126 131	107 135 4 24 64	31 129 55 125 140	
Q1	3		84		7 96 21 118 27 120 48 121 63 139 67	
	4		11 114 51 110 25 30 57 106	33 83 38 94 41 130 52 53 61	76 enquêteurs	62 98
	5				90 132 (idem sur Q15)	69 (idem sur Q6)

Légendé : 1 – extrêmement faible ; 2 – très faible ; 3 – faible ; 4 – normale ; 5 - forte

Visite de quelques enquêteurs

27 Les examens précédents permettent d'orienter les investigations auprès d'un petit nombre d'enquêteurs considérés comme les plus bizarres. Dans les fiches par enquêteur qui ont été établies apparaissent les tableaux simplifiés des distances entre interviews pour les batteries Q1 (partie Agoramétrie) et Q16 (partie nucléaire) et les enregistrements des réponses à partir desquelles elles ont été calculées. Quand la distance $d_{ii'}$ est faible, voire très faible, un signe apparaît dans les cases (i, i') et (i', i), les codes suivants étant utilisés :

$d_{ii'} \leq 0,2 \rightarrow$ ■ ; $0,2 < d_{ii'} \leq 0,3 \rightarrow$ □ ; $0,3 < d_{ii'} \leq 0,4 \rightarrow$ ✖ ; $0,4 < d_{ii'} \leq 0,5 \rightarrow$:

28 Dans les fiches apparaissent aussi les taux d'utilisation des modalités, les durées des interviews et les dates auxquelles elles ont été effectuées. Elles sont rangées en annexe dans l'ordre des numéros affectés aux enquêteurs visités.

Un as de l'interview : l'enquêteur 14

29 L'agent enquêteur numéroté 14 se distingue par le grand nombre d'entretiens réalisés : 50 entre les débuts de décembre 2003 et de janvier 2004, une semaine de repos séparant les 30 premiers des 20 derniers entretiens. Ceux-ci ont été conduits avec célérité, 7 en moins de 30 minutes et 6 seulement sur 50 en plus de 40 minutes, avec des séries de 5,6 et même de 8 entretiens dans une même journée.

30 Par comparaison, on note un usage immodéré de la modalité « 1 » au détriment de la « 2 » dans les réponses à la batterie Q1 (partie Agoramétrie) : sur les 13 entretiens sur 2 215 considérés comme hors normes pour usage excessif de la réponse « 1 », 3 sont issus de l'enquêteur 14.

31 Si on se réfère aux deux tableaux 50x50 fournis, les distances entre enregistrements apparaissent normales à quelques broutilles près. Et pourtant, assimilé au centre de gravité de ses 50 entretiens, l'enquêteur 14 se retrouve positionné sur le plan factoriel 1-2 qui sert de référence dans les enquêtes Agoramétrie excentré dans le quadrant de la transgression, celui où sont en surnombre enseignants et étudiants, cadres supérieurs, hauts diplômés et Parisiens. On se déclare ici, plus qu'ailleurs, favorable aux immigrés, à l'Europe, à la vente libre du haschisch, à l'utilisation des robots dans l'industrie, etc ...

32 Cette radiographie du travail effectué par l'enquêteur 14 ne permet guère à elle seule de porter un jugement ; pour que l'on puisse se prononcer sur la qualité des données recueillies, il faudrait poursuivre l'analyse et répondre à un certain nombre de questions : comment l'enquêteur a-t-il procédé pour réaliser autant d'entretiens dans un temps record ? Quelle explication donne-t-il à cet usage excessif de réponses « 1 » ? Pourquoi se retrouve-t-il autant excentré dans le quadrant de la transgression ? Compte tenu du nombre d'entretiens réalisés, ces questions méritent réponses de la part de l'enquêteur que nul n'a cherché à interroger après coup.

Examen de quelques enquêteurs sortant du lot

33 On commencera la visite par l'enquêteur 69, car les distances mesurées entre ses 15 interviews sont parmi les plus grandes, aussi bien pour Q1 que pour Q16. Il s'est montré particulièrement efficace puisqu'il a bouclé ses entretiens très rapidement (durée moyenne 37 minutes) en 3 jours. Véritable exploit : 6 interviews le dernier jour dont 4 en moins de 29 minutes chacun ! Et pourtant, en dehors d'un recours trop prononcé à la non réponse « 6 » pour Q1 et Q16 (il en est de même pour la batterie Q6), rien de très anormal ne transparaît dans les enregistrements. Seule question posée à l'institut de sondage : fallait-il valider des entretiens qui comportent des taux aussi élevés de non réponse (cas de 1 et 3) ?

34 La visite se poursuit parmi les enquêteurs qui, ayant déjà été repérés grâce aux critères des taux de réponses par modalité, se retrouvent aussi avec des distances très faibles entre leurs interviews :

- 102 : l'usage immodéré des « 3 » et des « 6 » explique à lui seul les fortes proximités constatées pour Q1 et Q16 (il en est de même pour Q6), 5 des 7 interviews ayant été réalisées le premier jour.
- 73 : les trois interviews détectées aberrantes pour Q1 le sont aussi pour Q16 ; il faut ici admirer la performance : plusieurs individus ont affirmé ensemble leur accord, voire leur

total accord, avec toutes ou presque toutes les propositions non seulement de Q1 mais aussi de Q16 !

- 91 : ce champion de l'aberration dont 9 interviews sur 11 sont hors normes ne se distingue que modérément avec le critère des distances ; c'est bien entendu l'usage trop fréquent du « 6 » ou du « 3 » (pour Q6) qui créé la ressemblance entre les interviews.

35 Allons rendre visite maintenant à quelques enquêteurs normaux quant au critère des taux d'usage des modalités et remarquables de par leurs distances inter-interviews :

- 6 : c'est un véritable cas d'école : les 30 interviews sont si proches les unes des autres que l'on pourrait les réduire à une seule.
- 141 : un orfèvre comme 6, mais n'ayant heureusement effectué que 7 interviews.
- 80 : 4 des 11 interviews ont été réalisées « autour du sapin le jour de noël » : est-ce suffisant pour expliquer la forte proximité inter-interviews ?
- 22 : 7 sur les 15 interviews ont été ici bouclées en moins de 28 minutes avec un record en 24 minutes. On note : 87 % d'accord total au rétablissement de la peine de mort (colonne 31), contre 18 % dans la population ; 80 % d'accord total pour être plus sévère avec les jeunes délinquants (colonne 62), contre 45 % dans la population ; 67 % d'accord total pour freiner Internet (colonne 59), contre 8 % dans la population ; etc..., et un seul des 15 interviewés à exprimé son soutien aux intermittents du spectacle (colonne 80), contre 41 % dans la population.

36 Arrêtons nous là, la liste très longue des aberrations peuvent lasser le lecteur ; mais qu'il sache que nous sommes encore loin du compte. Bref, n'est-ce pas un euphémisme que de déclarer après ces quelques visites qu'il a bien fallu écourter que des doutes sont émis sur la qualité du sondage ?

En guise de conclusion

37 Symptomatiques souvent plus d'une véritable indélicatesse que d'un manque de rigueur, les anomalies nombreuses constatées montrent avant tout que le travail des enquêteurs n'a pas été encadré de manière appropriée par le sondeur.

38 Lors de l'étude réalisée en 1998, nous avions plaidé naïvement pour l'autodiscipline : on conseillait à l'institut de sondage de mettre en place un retour d'expérience systématique pour les grosses enquêtes, basé sur un contrôle a posteriori des données recueillies. A notre connaissance, ce travail de fond devant conduire à une amélioration des procédures n'a pas été entrepris.

39 Aussi faut-il aujourd'hui changer de tactique : il est indispensable que les commanditaires et les utilisateurs se mobilisent au nom de la qualité. Fort du constat effectué, on doit exiger d'abord d'en savoir plus sur les tenants et aboutissants du recueil ; pour ensuite, en connaissance de cause, imposer de nouvelles « bonnes pratiques » aux sondeurs et se ménager la possibilité de réaliser un contrôle systématique *a posteriori* des données recueillies pour permettre un retour critique plus efficace sur la réalité du terrain.

BIBLIOGRAPHIE

Patrick Che, Jean-Pierre Pagès (1998). Les nouvelles chasses du Comte Zarov, ou comment débusquer les enquêteurs indélicats. Note confidentielle, B3-Hedera, fin 1998.

ANNEXES

Pour ne pas ennuyer le lecteur, on a évité de lui fournier toutes les fiches des enquêteurs visités pour anomalies détectées. On se contente ici de fournir seulement les fiches spectaculaire de l'enquêteur numéro 6 (Figures 14 et 15). Apparaisant successivement le tableau Q1 de la partie Agoramétrie (Figure 14) et le tableau Q16 de la partie nucléaire (Figure 15). Sont fournis : Les tableaux des proximités entre entretiens pour les parties Q1 et Q16 du questionnaire ; Les enregistrements pour Q1 et Q16 ; Les durées et les dates d'interview ; Les taux d'utilisation des modalités de réponse pour Q1 et Q16

Figure 14 : **Tableau Q1 de la partie Agoramétrie**

Figure 15 : **Tableau Q16 de la partie nucléaire**

RÉSUMÉS

Les auteurs analyse statistiquement les données collectées en 2003 lors d'une enquête CAPI (« computer-assisted personal interview ») de 2.520 personnes. La collecte de données par les interviewers est examinée pour le nombre d'entretiens, durées des entretiens, modalités de réponse utilisées et temps entre les entretiens. Cette analyse identifie des interviewers dont la collecte de données pose des questions. Loin d'incriminer seulement les interviewers, les auteurs critique certaines pratiques dans le milieu des sondages qui encouragent la production des données de mauvaise qualité.

The authors statictically analyze data collected during a 2003 CAPI (computer-assisted personal interview) public opinion poll of 2,520 persons. Interviewer data collection is examined for number of interviews, lenght of interviews, response modality use and time between interviews. This analysis identifies a number of interviewers whose data production raises certain questions. Far from placing responsibility only on interviewers, the authors critique certain practices in polling that encourage the production of low-quality data.

INDEX

Keywords : Data Quality, Interviewer Behavior, Opinion Polls
Mots-clés : Comportement des interviewers, Enquêtes d'opinion, Qualité des données

AUTEURS

PATRICK CHE

B3-Traitment statistique de l'information, Paris

JEAN-PIERRE PAGÈS

Agoramétrie, Paris