



ined
INSTITUT
NATIONAL
D'ÉTUDES
DÉMOGRA
PHIQUES



Répliquer une enquête face-à-face sur un access panel web ? Une comparaison multipanels.

**Stéphane Legleye^{1,2}, Sandra Hoibian³, Jérôme Cubillé⁴, Patricia Croutte³,
Géraldine Charrance¹**

1 Institut national d'études démographiques

2 Univ. Paris-Saclay, Univ. Paris-Sud, UVSQ, CESP, INSERM, *Villejuif*, France

3 Centre de recherche pour l'étude et l'observation des conditions de vie (CREDOC)

4 EDF R&D

- L'enquête Aspirations et conditions de vie des français existe depuis 38 ans
- C'est un face-à-face par quota auprès de 2000 personnes de 18 ans et plus
- La durée du questionnaire est d'environ 1 heure en face-à-face
- Questions :
 - Peut-on le passer sur Internet ?
 - Quels sont les écarts de mesure à prévoir ?
- Contrainte :
 - pas de base de sondage possible donc recours à un panel commercial

Effets de mode et de sélection (1)

- La littérature montre que :
 - Les panels commerciaux sont issus de processus de sélection largement non-aléatoires
 - Équipement Internet
 - Exposition à des invitations (mode d'usage d'Internet)
 - La participation à une enquête dans le panel est également en partie non-aléatoire
 - Filtre du thème
- Les compositions sociodémographiques peuvent varier d'un panel à l'autre
- Il y a donc un effet de sélection majeur

Effets de mode et de sélection (2)

- La présence d'un enquêteur est généralement bénéfique pour (en plus du taux de participation):
 - Les questionnaires longs, répétitifs ou peu intéressants pour l'enquêté (satisficing)
 - Possibilité d'explicitier, de relancer etc.
- L'absence d'un enquêteur est généralement bénéfique pour :
 - Les sujets sensibles (désirabilité sociale)

Les principales techniques de mesure / neutralisation de l'effet mode de collecte

- Repondération
 - Calages d'un échantillon sur un autre (ou des deux sur une base externe)
 - Pondération par l'inverse d'un score de propension modélisant la probabilité d'être sur un mode plutôt qu'un autre
 - Mix des deux...
- Appariement sur score de propension

Objectifs de la présentation

- Présenter
 - les deux volets de l'enquête
 - Les tentatives de mesure/neutralisation
 - Notamment tester l'usage de variables attitudinales censées neutraliser les effets de satisficing et de désirabilité sociale
 - Les extensions de la méthode : comparaison de plusieurs panels

L'enquête face-à-face

- 2000 personnes, 18 ans et plus
- Quotas sur
 - ZEAT (9 grandes régions)
 - Taille de l'agglomération urbaine (9 modalités)
 - Sexe
 - Âge (5 modalités)
 - PCS (8 modalités)
- Collecte de 13 décembre 2014 au 29 janvier 2015 (attentats de janvier)

L'enquête en ligne

- OpinionWay, choisi après des discussions avec plusieurs instituts; plus de 100 000 panélistes
- 2045 personnes, 18 ans et +
- Quotas identiques
 - Collecte du 23 décembre 2014 au 20 janvier 2015, très peu de questionnaires après le 7 janvier (attentats)
- Taux de retour de 13.8% (questionnaires complets sur questionnaires sollicités)
- Abandons : 12.9% (ordinaire chez OpinionWay)

Les écarts et les accords (1)

- Sociodémographique (hors quotas)
 - Diplômés, hauts revenus surreprésentés en ligne
 - Personnes seules et locataires surreprésentées en face-à-face
- Variables d'intérêt
 - Accords sur :
 - Cadre de vie, rapport à l'emploi, état de santé perçu, opinions sur les homosexuels, la mondialisation, la pauvreté et le chômage ainsi que son évolution, la perception des grandes marques commerciales...
 - Désaccords et surreprésentation en ligne des :
 - Equipement informatique, usages d'internet
 - Opinions politiques radicales et d'une empathie réduite à l'égard des défavorisés
 - Expression de moindre bonheur

Les écarts et les accords (2)

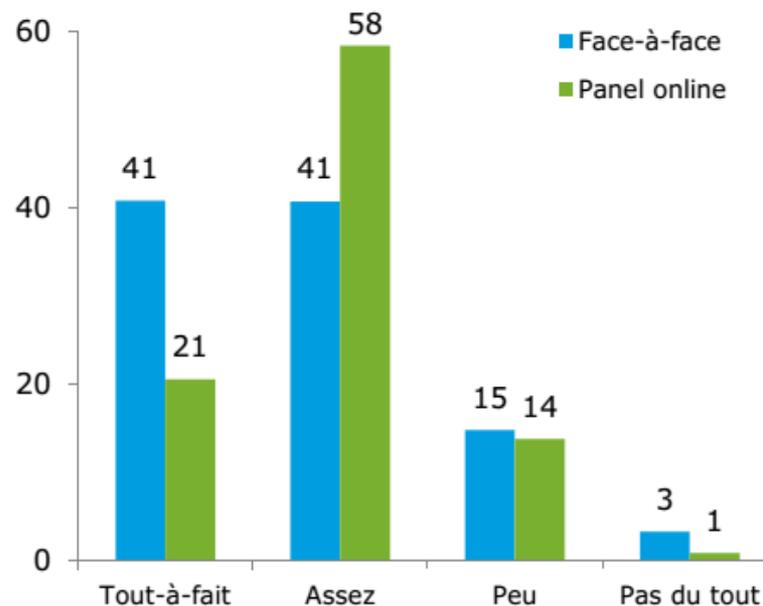
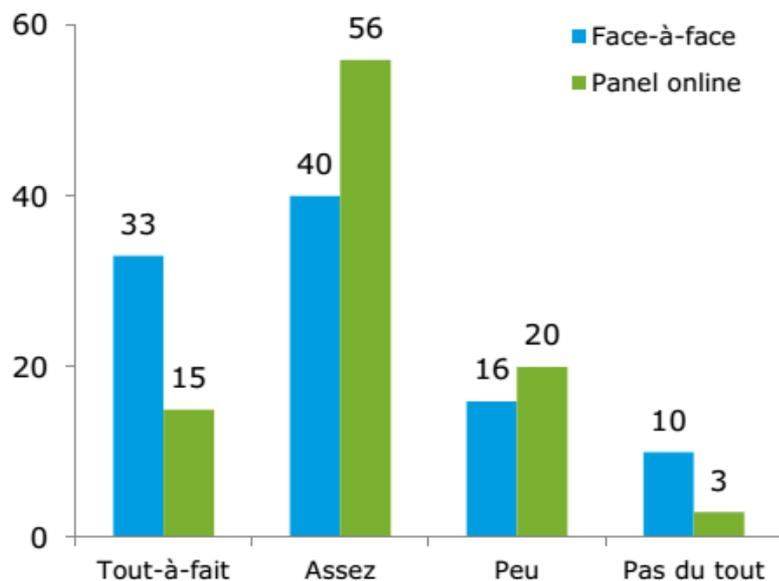
Sommes des écarts absolus moyens par bloc de questions

	Ecart moyen
Situation d'emploi	1,6
Santé	2,8
Moral économique	2,9
Opinions sur la famille	3,4
Sociodémographique	3,4
Logement	3,6
Valeurs	4,2
Opinions sur la société	4,7
Sociabilité	5,0
Opinions sur la solidarité	5,2
TIC	7,7
	3,9

Les écarts et les accords (3)

variables attitudinales (ipsographiques)

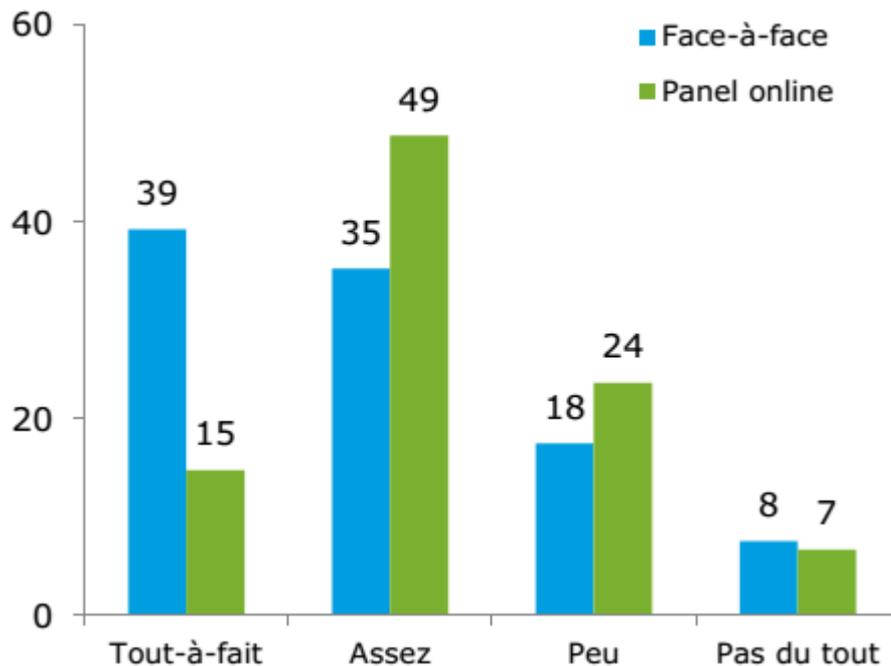
- Dans la vie, vous faites toujours attention à ce que vous dites pour ne pas froisser les gens
- Dans la vie, quoi que vous fassiez, vous êtes toujours très exigeant avec vous-même



Les écarts et les accords (4)

variables attitudinales (ipsographiques)

- Dans la vie, vous vous comportez toujours de la même façon lorsque vous êtes seul ou avec les autres



Les redressements (1)

Les calages

- Calages successifs:
 - H0 : âge*sexe, type de logement, PCS et taille d'agglomération
 - H1 : utiliser Internet pour comparer avant achat
 - H2 : diplôme (4 modalités)
 - **H3 : variables attitudinales**
 - H4 : typologie des niveaux de vie
 - H5 : tout ensemble

Les redressements (2)

Les scores de propension et méthodes mixtes

- Score de propension à partir de 11 variables (sélectionnées par une procédure automatique robuste):
 - Situation d'emploi ; diplôme ; typologie des niveaux de vie ; recours à Internet avant achat ; 3 variables attitudinales (ipsographiques) ; participation à un réseau social, achat en ligne, internet sur smartphone ; opinion sur le versement de prestations sociales aux plus démunis
- Méthodes subséquentes :
 - H6 : inverse du score de propension
 - H7 : calage H0 sur pondération tirée du score de propension
 - H8 : calage H0+diplôme sur pondération tirée du score de propension

Les redressements (3)

Les résultats

	poids initial	max/min	médiane	Ecart moyen résiduel
H0	1	19.5	0.9	3.9
H1	1	47.3	0.6	3.8
H2	1	94.5	0.7	3.8
H3	1	1468.0	0.6	3.5
H4	1	42.2	0.8	3.9
H5	1	15980.0	0.4	3.5
H6	1/p	87.4	27.2	3.7
H7	1/p	160.0	0.7	3.6
H8	1/p	303.0	0.6	3.7

Les appariements (1)

Les variables

- Etude systématique des écarts entre échantillons : 4 familles de variables
 - Sociodémographiques : sexe, diplôme, type de logement, agglomération, revenu, situation professionnelle, enfants
 - Webographiques (internet et matériel connecté) : micro-ordinateur, tablette, internet au domicile, smartphone connecté, durée hebdomadaire d'usage d'internet
 - Enquêtographiques : participé à des enquêtes sur Internet, au téléphone, en face-à-face (2 composantes d'ACP retenues)
 - Ipsographiques (présentation de soi) : superstitieux, prudent, 3 variables attitudinales

Les appariements (2)

Les variables

- 6 variables d'intérêt
 - Tabagisme quotidien
 - consommation régulière d'alcool
 - expérimentation de cannabis
 - jugement négatif sur son cadre de vie
 - sentiment d'être en mauvaise santé
 - conduite automobile
- 6 scores de propension :
 - PSD ($_i$): Sociodémo. +/- ipsographiques
 - PSDW ($_i$): PSD + webographiques +/- ipsographiques
 - PSDWE ($_i$): PSDW + enquêtographiques +/- ipsographiques
- Hypothèse : les variables ipsographiques diminuent l'écart

Les appariements (4)

Les résultats

- Les variables ipsographiques ne réduisent pas les écarts...

	Calage	PSD	PSD_i	PSDWE	PSDWE_i
	N=4045	N=2714	N=2136	N=980	N=936
cannabis	1.41	1.56	1.69	2.15	1.95
tabac	0.43	0.45	0.51	0.53	0.51
alcool	1.06	0.96	0.89	0.99	0.99
santé	1.06	1.44	1.42	1.29	1.49
cadre de vie	0.99	1.22	1.38	1.55	1.39
auto	2.12	1.54	1.33	1.09	1.12

- Pourquoi les variables ipsographiques (attitudinales) ne marchent pas ?
 - Pas les bonnes variables ?
 - Choix très contraint ici...
 - Effet de mesure même sur ces variables ?
 - Très possible... et en fait vrai quel que soit l'appariement (PSD, PSDW, PSDWE)=> le corriger avant de l'intégrer dans la modélisation ?
 - Petit échantillon ?
 - Discutable...

Peut-on extrapoler à d'autres panels (1)



L'expérimentation

- Trois panels questionnés en 2016
 - Toluna (n=1000)
 - Panel on the web (n=1000)
 - OpinionWay (n=3050) : panel de l'enquête...
- Quotas :
 - Sexe, âge (6 modalités)
 - PCS (8 modalités)
 - Taille d'unité urbaine (5 modalités)
 - Régions
 - Type de logement
- Redressement :
 - Sexe, taille d'unité urbaine, région, type de logement, âge (en 5) x diplôme (en 3)

Peut-on extrapoler à d'autres panels (2)



Les résultats

- D'importants écarts entre les panels...
- Qui ne disparaissent pas avec le redressement
- Tous ne se valent pas... il faut expérimenter !

Références

Erens B, Burkill S, Couper MP, Conrad F, Clifton S, Tanton C, Phelps A, Datta J, Mercer CH, Sonnenberg S, Prah P, Mitchell KR, Wellings K, Johnson AM, Copas AJ. Nonprobability web surveys to measure sexual behaviours and attitudes in the general population: a comparison with a probability sample interview survey. *J Med Internet Res.* 2014;16(12):e276. doi: /jimr.3382 <http://www.jmir.org/2014/12/e276/>

[Www.credoc.fr](http://www.credoc.fr)

Régis Bigot, Patricia Crouette, Fanette Recours, Enquêtes en ligne : peut-on extrapoler les comportements et les opinions des internautes à la population générale ?, CREDOC, Cahier de recherche n°273, décembre 2010, <http://www.credoc.fr/pdf/Rech/C273.pdf>