

Comment les Français s'informent-ils sur Internet ?

ANALYSE DES COMPORTEMENTS D'INFORMATION
ET DE DÉSINFORMATION EN LIGNE

Une étude de la Fondation Descartes — Mars 2021

Pour citer cette étude :

Laurent Cordonier & Aurélien Brest (2021). *Comment les Français s'informent-ils sur Internet ? Analyse des comportements d'information et de désinformation en ligne.*
Étude de la Fondation Descartes (mars 2021), www.fondationdescartes.org/nos-rapports/

Sommaire

1 – L'étude en bref	02
1.1 Présentation de l'étude	02
1.2 Les enseignements clés de l'étude	03
2 – Méthode	08
2.1 Participants à l'étude	09
2.2 Sélection des sources d'information suivies	09
2.3 Mesure de la consultation des sources d'information suivies	10
2.4 Catégorisation des sources d'information comme fiables ou non	10
2.5 Questionnaire de fin d'étude	12
3 – Résultats	13
3.1 Connexions à Internet	14
3.2 Consommation d'informations sur Internet	15
3.2.1 Panorama de la consommation effective	15
3.2.2 Utilisation des applications pour s'informer	17
3.2.3 Utilisation de YouTube pour s'informer	18
3.2.4 Types de sources d'information consultées	19
3.2.5 Les leaders du marché de l'information sur Internet	20
3.3 Caractéristiques des consommateurs d'informations sur Internet	21
3.4 Comparaison entre consommation d'informations déclarée et effective	22
3.5 Rapport entre la confiance en des sources d'information et leur usage	24
3.6 Consommation de désinformation	26
3.6.1 Les facteurs de risque de l'exposition à de la désinformation généraliste sur Internet	26
3.6.2 Les facteurs de risque de l'exposition à de la désinformation en santé sur Internet	28
3.6.3 Les facteurs de risque de l'exposition aux « pièges à clics » sur Internet	30
4 – Discussion des résultats	32
4.1 Consommation d'informations sur Internet	33
4.1.1 S'informer en papillonnant sur le Net	34
4.1.2 La place des médias traditionnels sur Internet	34
4.2 Facteurs de risque de l'exposition à de la désinformation et aux « pièges à clics »	35
Bibliographie	37
Auteurs de l'étude	39
La Fondation Descartes	40

L'étude en bref

1.1 — Présentation de l'étude

Cette étude de la Fondation Descartes a pour objectif de décrire la manière dont les Français consomment de l'information de type médiatique sur Internet¹⁻². Pour y parvenir, nous avons demandé à la société ResponDi, spécialiste des données digitales, d'enregistrer durant 30 jours consécutifs – du 20 septembre au 19 octobre 2020³ – l'intégralité de l'activité Internet d'un panel de personnes représentatif de la population nationale⁴. Les données de connexion des 2 372 participants qui composent ce panel ont été enregistrées, urls visitées par urls visitées, sur leurs différents objets connectés personnels (ordinateurs, téléphones portables, tablettes). Nous avons alors pu analyser la consultation par ces personnes de 2 946 sources d'information en ligne préalablement identifiées à l'aide de l'algorithme de classification des sites Internet de la société Storyzy.

Les sources d'information identifiées et suivies dans cette étude recouvrent et excèdent de beaucoup celles qui sont habituellement retenues dans les études sur la consommation d'informations sur Internet en France. Cela va des pages web de la presse papier nationale et des chaînes d'information du service public à celles de sites plus confidentiels de « réinformation », régulièrement accusés de publier de fausses informations, en passant par les pages Internet de la presse féminine, masculine, people, culturelle, sportive ou de divertissement, ainsi que par celles des agrégateurs

d'informations en ligne et des médias présents uniquement sur Internet (du type L'Internaute, Doctissimo, 01.net, etc.).

Au terme des 30 jours de l'étude, nous avons adressé aux participants un questionnaire permettant d'établir leur profil sociodémographique. Ce questionnaire les interrogeait également sur leur rapport à l'information médiatique, sur les sources d'information en ligne et hors ligne qu'ils consultent, ainsi que sur la confiance qu'ils leur accordent. La méthode innovante suivie dans cette étude nous a permis de dresser un panorama riche et inédit de la manière dont les Français s'informent sur Internet.

1. Nous remercions L'Observatoire Société et Consommation (<https://lobsoco.com>) qui nous a fait bénéficier de son expérience dans la conception initiale de la présente étude.

2. Par « information médiatique », nous entendons l'ensemble des informations et actualités politiques, économiques, sociales, culturelles, sportives, internationales ou portant sur des thèmes plus spécifiques – actualités people, scientifiques, technologiques, etc. – ainsi que les billets ou blogs d'opinion à portée sociale ou politique, essentiellement.

3. Cette période a été choisie car elle se situe entre le premier et le second confinement sanitaire de 2020, le comportement d'information de Français sur Internet y étant dès lors plus représentatif qu'il ne l'aurait été durant un confinement.

4. L'activité Internet enregistrée correspond à tout transfert de données entre les objets connectés des participants et le reste du réseau mondial (hors transferts Bluetooth). Les activités Internet considérées comme relevant de la consultation d'informations correspondent à la connexion aux urls rattachées à une source d'information suivie dans cette étude ainsi qu'à la consultation des applications mobiles associées à ces sources d'information. À noter qu'un système dit de « time-out » interrompt la comptabilisation du temps passé sur une source d'information après trois minutes d'inactivité de la part des participants (cela ne concernant pas les contenus audio ou vidéos en cours de lecture).

Les sources d'information retenues

Les sources intégrées à cette étude regroupent 2 295 sites Internet et 651 chaînes YouTube francophones identifiés par l'algorithme de Storyzy comme relevant de l'information et ayant enregistré au minimum 15 000 visiteurs dans le monde durant le mois précédant notre étude ou, pour les chaînes YouTube, possédant au minimum 15 000 abonnés au moment du lancement de l'étude.

Certains des 2 295 sites Internet suivis disposent d'applications pour téléphone connecté et/ou sont associés à des pages Facebook et Twitter ainsi qu'à des chaînes YouTube. Nous avons enregistré la fréquentation de ces sources d'information par les participants à notre étude au travers de tous ces canaux d'accès. L'ensemble des sites Internet, des applications mobiles, des pages Facebook et Twitter et des chaînes YouTube suivis dans cette étude couvre un très large spectre des sources d'information francophones disponibles en ligne.

Soulignons que le temps potentiellement consacré par les participants à s'informer sur les sources ou *via* les supports suivants n'a pas été comptabilisé dans cette étude : les sources d'information en langues étrangères, celles ayant été consultées par les participants sur d'autres objets connectés que leurs objets personnels, les podcasts d'information consultés hors des sources suivies, les contenus d'information téléchargés en ligne mais consultés hors ligne, les informations partagées *via* des newsletters, des notifications push, des messageries privées (type WhatsApp) ou des groupes privés sur les réseaux sociaux qui ne nécessitent pas de se rendre sur une page Internet externe pour les lire, ainsi que les informations partagées directement sur le « mur » Facebook ou le « fil » Twitter des participants, si la lecture de ces informations peut se faire sans avoir à se rendre sur le site des médias les ayant publiées (voir Encadré 3).

— ENCADRÉ 1

Une approche hybride novatrice

Deux sondages annuels semblent incontournables pour comprendre les habitudes de consommation médiatique des Français sur Internet : le Digital News Report du Reuters Institute⁵ et le Baromètre du numérique réalisé par le CREDOC et l'ARCEP⁶. Les résultats de ces sondages sont régulièrement interprétés comme des indicateurs de la consommation effective d'informations en ligne par les Français, alors qu'ils nous renseignent en réalité sur la représentation que se font les répondants de leurs pratiques informationnelles sur Internet. Or, il n'est pas exclu qu'il existe un décalage plus ou moins marqué entre les pratiques informationnelles effectives des Français sur Internet et la représentation qu'ils en ont.

Un moyen de pallier cette importante limite des études déclaratives est de mesurer le trafic des internautes sur les sites d'information, comme le fait en France la société Médiamétrie. Les données comportementales – et non plus déclaratives – collectées par Médiamétrie ont le mérite de fournir une représentation globale de la

consultation par les Français de diverses sources d'information en ligne. Cependant, cette méthode ne permet pas de savoir qui sont les personnes qui composent le trafic mesuré sur ces sites. Cela interdit notamment de rendre compte de la diversité des pratiques d'information individuelles et de déterminer si des facteurs sociodémographiques (sexe, âge, catégories socioprofessionnelles, etc.) ou cognitifs (intérêt pour l'actualité ou niveau de confiance dans les médias, par exemple) modulent ces pratiques.

La présente étude de la Fondation Descartes est, à notre connaissance, la première à croiser mesures comportementales et déclaratives pour étudier les pratiques informationnelles des Français sur Internet. Grâce à cette méthodologie hybride, il nous a été possible de dévoiler ces pratiques au plus proche de ce qu'elles sont réellement, en tenant compte de leur diversité et en les rapportant au profil sociocognitif des individus ainsi qu'à la représentation que ces derniers ont de la manière dont ils s'informent sur Internet.

— ENCADRÉ 2

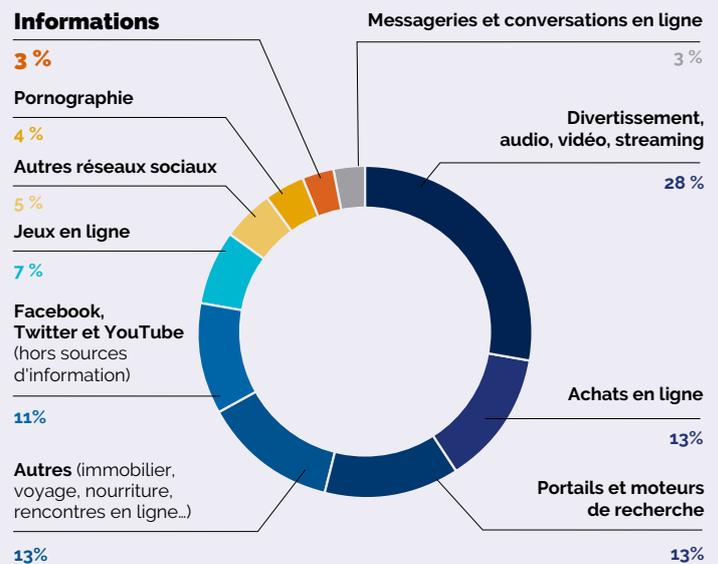
1.2 Les enseignements clés de l'étude

Notre étude met en évidence que la part du temps de connexion des Français consacrée à la consultation d'informations médiatiques est, dans l'ensemble, faible. En moyenne, 3 % seulement du temps total passé en ligne par les participants à notre étude l'a été sur les sources d'information suivies – ce qui, par participant, correspond à un peu moins de 5 minutes par jour de connexion à Internet. Précisons que ces valeurs reflètent le temps passé par les participants à s'informer au moyen de leurs objets connectés personnels sur les sources d'information que nous avons identifiées et retenues dans cette étude. Elles n'intègrent donc pas le temps qu'ils sont susceptibles d'avoir consacré à s'informer sur d'autres objets connectés que les leurs (ordinateurs professionnels, par exemple) ou sur des sources d'information que nous avons écartées de nos analyses, dont les sites Internet en langues étrangères (voir Encadré 1).

3 % du temps total passé en ligne par les participants à notre étude l'a été sur des sources d'information.

Il importe en outre de souligner que si, *en moyenne*, les Français s'informent peu sur Internet, les comportements d'information en ligne sont en réalité extrêmement variables. Par exemple, si 17 % des participants à notre étude n'ont

Répartition du temps d'activité sur Internet (durant les 30 jours de l'étude)



LECTURE : Durant les 30 jours de l'étude, 3 % du temps total passé sur Internet par l'ensemble des participants l'a été sur des sources d'information.

NOTE : Un « portail » est une porte d'entrée commune à un éventail de ressources et de services Internet : pages web, forums de discussion, boîtes de courrier électronique, espaces de publication, moteurs de recherche, etc.⁷ Certains portails, tels que Yahoo! ou Orange, comprennent des espaces informations ou actualités. Dans notre étude, le temps passé par les participants sur de tels espaces au sein des portails a été comptabilisé dans la catégorie « Informations ».

5. www.digitalnewsreport.org/

6. www.arcep.fr/uploads/tx_gspublication/rapport-barometre-num-2019.pdf

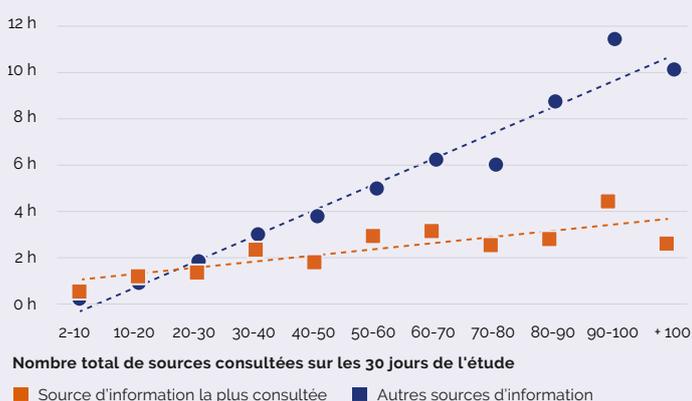
7. https://fr.wikipedia.org/wiki/Portail_web

consulté aucune source d'information sur Internet en 30 jours, ils sont 5 % à en avoir consulté durant plus de dix heures au total. Nous sommes parvenus à identifier un certain nombre de caractéristiques individuelles et sociales associées au fait de s'informer plus ou moins intensément sur Internet. La plus importante d'entre elles semble être l'âge des participants, les 50 ans et plus s'informent davantage que les autres sur Internet. Cependant, les liens observés entre les caractéristiques sociodémographiques des participants et l'usage qu'ils font d'Internet pour s'informer sont tous relativement faibles – ces caractéristiques ne peuvent dès lors pas expliquer, à elles seules, la forte variance interindividuelle des comportements d'information en ligne que nous avons constatée. D'autres facteurs qu'il resterait à identifier entrent ainsi nécessairement en jeu, tels que la curiosité des individus ou le temps dont ils disposent pour s'informer, par exemple.

17 % des participants n'ont consulté aucune source d'information sur Internet en 30 jours, tandis qu'ils sont **5 %** à en avoir consulté durant plus de dix heures au total.

On observe par contre que les participants qui consacrent beaucoup de temps à s'informer sur Internet le font moins en augmentant le temps total passé sur une même source d'information qu'en diversifiant les sources qu'ils consultent. En effet, plus les participants ont passé de temps à s'informer en ligne durant les 30 jours de l'étude, plus le nombre de sources d'information différentes qu'ils ont consultées est important.

Évolution du temps total passé sur la source la plus consultée en fonction du nombre de sources visitées



LECTURE : La ligne pointillée orange (droite de régression) représente l'augmentation du temps total passé sur la source d'information la plus consultée par chaque participant à mesure que s'accroît le nombre de sources visitées. La ligne pointillée bleue représente l'augmentation correspondante pour ce qui est du temps total passé sur les autres sources d'information consultées.

Les 50 ans et plus s'informent davantage que les autres sur Internet.

Cela traduit un comportement de « papillonnage » chez les consommateurs d'informations sur Internet, qui semblent passer d'une source d'information à l'autre sans souvent la consulter longtemps. En témoigne le fait que le temps passé consécutivement sur une source d'information à chaque consultation de cette source est en moyenne inférieur à 2 minutes. Même si ce temps moyen varie sensiblement entre les participants, il n'est de 4 minutes ou plus que chez 6 % des participants.

On observe un comportement de « papillonnage » chez les consommateurs d'informations sur Internet. Ainsi, les participants qui s'informent le plus en ligne le font en multipliant les sources consultées.

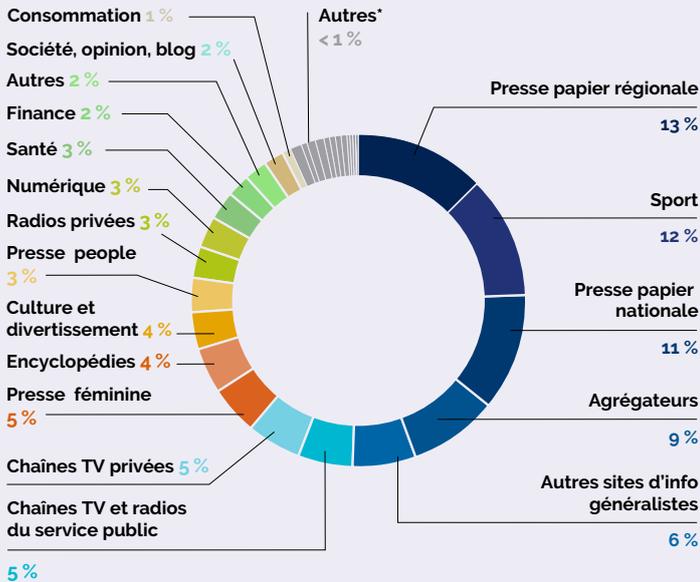
Un autre élément va dans le sens d'un tel comportement de « papillonnage » : au total, nos participants ont visité 1 290 sources d'information différentes au cours des 30 jours de l'étude – un chiffre qui excède de beaucoup le nombre des grands médias présents sur Internet et qui traduit donc le fait que des sources d'information très secondaires en termes de visibilité sont également consultées par certains participants. Le fait que les internautes présentent ce comportement de « papillonnage » dans leur consommation d'informations semble aller contre l'idée répandue, mais discutée⁸, selon laquelle Internet enfermerait ses utilisateurs dans des « chambres d'écho » à l'intérieur desquelles ils ne seraient exposés qu'à des informations allant dans le sens de leurs attentes et de leurs croyances.

Les sources d'information les plus consultées par les participants au cours des 30 jours de l'étude sont des sources de la presse papier régionale. Les sources d'information dédiées à l'actualité sportive arrivent en seconde position du temps consacré par les participants à s'informer en ligne.

Les sources d'information en ligne les plus consultées sont celles de la presse papier régionale, suivies par les sources dédiées à l'actualité sportive et celles de la presse papier nationale.

8. www.fondationdescartes.org/2020/07/bulles-de-filtre-et-chambres-decho/

Types de sources d'information consultées



* Automobile, Lifestyle, Économie, Programme TV, Médias francophones étrangers, Science, Informations professionnelles, Presse masculine, Cuisine, Actualités locales, Environnement, Infos animalières, Parentalité, Spiritualité, Aéronautique, Fact-checking

LECTURE : Les participants ont dans l'ensemble passé 13% de leur temps d'information en ligne sur des sources en ligne de la presse papier régionale.

NOTE : Graphique ne prenant pas en compte les sources en ligne de prévision météo.

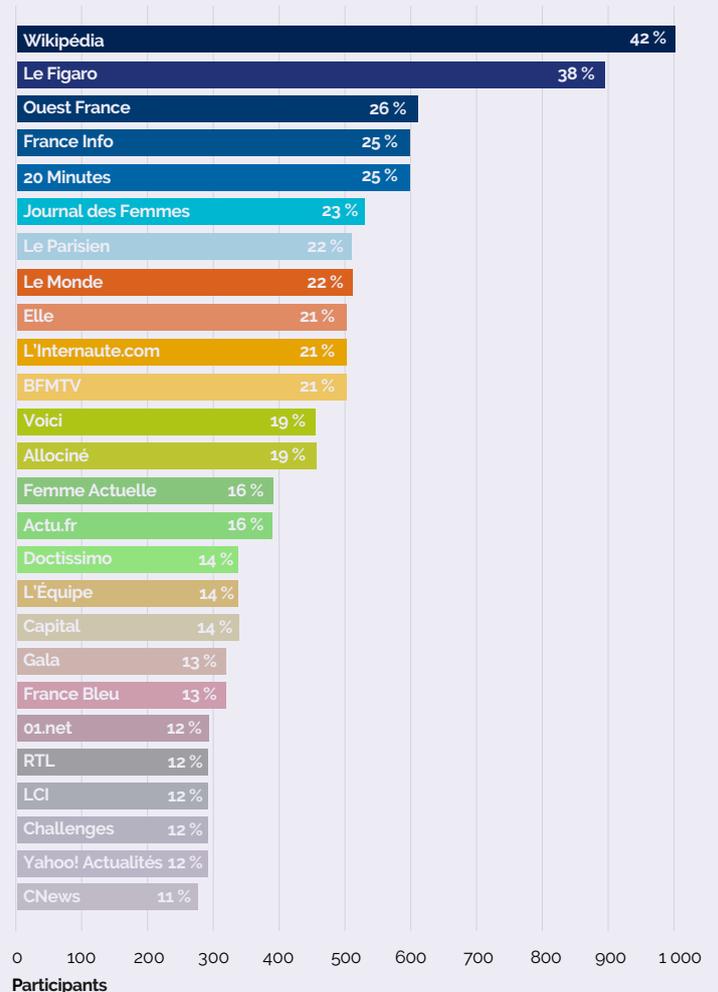
Les sources en ligne de la presse papier nationale arrivent en troisième position de ce classement. Quant aux agrégateurs de médias, ils arrivent en quatrième position, suivis par les autres sources d'information généralistes.

En nous intéressant aux sources d'information en ligne les plus consultées par les participants à notre étude, il est apparu que les 26 premières d'entre elles totalisent à elles seules 40 % du temps qu'ils ont consacré à s'informer en ligne. Ce « top 26 » est largement composé de sources Internet de médias traditionnels – une exception notable étant Wikipédia, qui se positionne comme la source d'information consultée par le plus grand nombre de participants au cours des 30 jours de l'étude.

Les grands médias traditionnels ont donc réussi à exploiter la visibilité dont ils jouissent hors ligne pour s'imposer comme des acteurs majeurs de l'information sur Internet. Notons que les médias purement numériques présents dans ce « top 26 » (L'Internaute, Doctissimo, etc.), outre Wikipédia, appartiennent tous à de grands groupes médiatiques.

Les **26** sources d'information en ligne les plus consultées captent **40 %** du temps passé sur Internet pour s'informer. Il s'agit principalement de sources web de médias traditionnels.

Classement des principales sources d'information par nombre de participants les ayant consultées



LECTURE : 42 % de l'ensemble des participants (soit 990 d'entre eux) ont consulté au moins une fois la source d'information Wikipédia au cours des 30 jours de l'étude.

NOTE : Graphique ne prenant en compte ni les sources en ligne de prévision météo, ni celles de programmes TV.

Un résultat de notre étude qu'il nous faut encore souligner est que l'usage que les participants déclarent faire d'Internet pour s'informer – que ce soit quant à la nature des sources qu'ils disent consulter ou quant à la fréquence avec laquelle ils rapportent le faire – ne correspond que faiblement à leur activité effective d'information en ligne, telle que nous l'avons mesurée. En d'autres termes, les Français semblent avoir une perception relativement floue de la manière dont ils s'informent sur Internet.

Les Français semblent avoir une perception relativement floue de la manière dont ils s'informent sur Internet.

De même, nous n'avons observé qu'un lien extrêmement ténu entre la consommation de sources d'information données et le niveau de confiance que les participants disent avoir en ces sources. Il est possible qu'il s'agisse là d'une conséquence du comportement de « papillonnage » des Français sur Internet. Il se peut en effet que ces derniers passent d'une source d'information à une autre, ou qu'ils cliquent sur un titre qui s'affiche sur leur « fil » Twitter ou leur « mur » Facebook, moins pour rechercher une information précise que par divertissement ou curiosité intellectuelle.

Nous n'avons observé qu'un très faible lien entre la confiance déclarée envers des sources d'information et la consultation de ces sources.

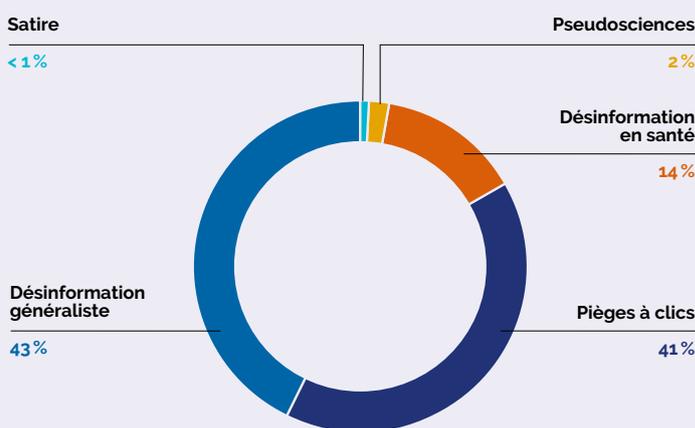
Finalement, nous avons utilisé dans cette étude un algorithme de la société Storyzy afin de classer les sources d'information suivies selon une catégorisation binaire « fiable »/« non fiable ». Nous avons comparé la catégorisation ainsi obtenue à celle proposée par Les Décodeurs du journal Le Monde. Il ressort de cette comparaison que les deux méthodes d'évaluation de la fiabilité des sources d'information aboutissent à des résultats largement cohérents entre eux. Cette catégorisation des sources d'information nous a permis d'évaluer l'exposition des Français à de la désinformation en ligne, qu'elle soit de nature généraliste (des sites jugés non fiables proposant des contenus d'information et d'actualité principalement sociaux et politiques), axée sur des contenus de santé, de pseudosciences, satiriques ou de type « pièges à clics » (des sites dont les contenus n'ont pas pour but d'informer le lecteur, mais uniquement d'attirer son attention en vue de générer du trafic sur leurs pages).

39 % des participants ont consulté des sources d'information jugées non fiables au cours des 30 jours de l'étude.

Nous avons ainsi pu constater que 39 % des participants ont consulté des sources d'information jugées non fiables au cours des 30 jours de l'étude. Ces participants y ont passé en moyenne 11 % de leur temps quotidien d'information sur Internet, soit 0,4 % de leur temps total de connexion. Sur l'ensemble des participants, le temps passé sur des sources d'information jugées non fiables représente 5 % du temps total d'information en ligne, soit 0,16 % du temps total de connexion à Internet. Les sources de désinformation les plus visitées sont les sources généralistes non fiables, suivies des sources « pièges à clics » et des sources de désinformation en santé.

Sur l'ensemble des participants, le temps passé sur des sources d'information jugées non fiables représente **5 %** du temps total d'information en ligne, soit **0,16 %** du temps total de connexion à Internet.

Répartition du temps passé sur des sources d'information non fiables



LECTURE : 43 % du temps passé sur des sources d'information non fiables l'a été sur des sources de désinformation généralistes.

NOTES : N = 921 participants. Il s'agit des participants ayant consulté des sources d'information non fiables au cours des 30 jours de l'étude.

Il ressort des analyses statistiques que nous avons conduites que les hommes sont surreprésentés dans le public des sources de désinformation généralistes. On y retrouve aussi, en proportion, davantage de personnes qui vivent seules, qui se disent favorables au mouvement des Gilets Jaunes ou membres de ce mouvement. Ces personnes présentent en outre un niveau de confiance moyen particulièrement faible à l'égard des institutions, du gouvernement et des médias, tandis qu'elles se disent en moyenne plus intéressées que les autres participants par les questions d'actualité sociale et politique.

Ce sont au contraire les femmes qui sont surreprésentées dans le public des sources de désinformation en santé, et on y retrouve en proportion davantage de personnes inactives, dont des retraités. Les participants à notre étude qui ont visité des sources de désinformation en santé présentent par ailleurs un score moyen plus élevé que les autres sur une échelle standardisée d'évaluation du niveau de complotisme.

Le public des sources « pièges à clics » se caractérise avant tout par un âge moyen plus élevé que celui de l'ensemble des participants. Les 65 ans et plus y sont en particulier surreprésentés, ainsi que les inactifs et les retraités. On observe également que le public des sources « pièges à clics » présente, en moyenne, un niveau de diplôme et de revenu du foyer moins élevé que celui du reste des participants.

Dans l'ensemble, la part du temps de connexion passé par les participants sur des sources d'information jugées non fiables peut sembler faible. Ce résultat, en apparence encourageant, pourrait en réalité indiquer que la désinformation en ligne diffuse aujourd'hui davantage par le biais des réseaux sociaux que par celui de sources web ayant « pignon sur rue », tels que les sites Internet. On ne compte en effet plus le nombre d'« infox » autour du Covid-19, par exemple, qui ont été massivement diffusées sur Facebook ou

Twitter. Des messages trompeurs de ce type y circulent de proche en proche, en étant partagés par les utilisateurs des réseaux sociaux sur leur « mur » ou leur « fil » personnel. Bien souvent, ces fausses informations se réduisent à un court texte ou à une photographie commentée, sans renvoyer à une source Internet externe aux réseaux sociaux. Cela fait que nous n'avons pas été en mesure dans notre étude de tenir compte de l'exposition des participants à ce type de désinformations.

Les réseaux sociaux dans cette étude

Les réseaux sociaux constituent à de nombreux égards des « boîtes noires » pour les chercheurs. Il est en effet difficile d'étudier la nature des contenus apparaissant dans les groupes Facebook privés ainsi que sur le « mur » Facebook ou le « fil » Twitter personnel des utilisateurs des réseaux sociaux, que ce soit pour des raisons techniques ou de confidentialité des données.

N'échappant pas à cette difficulté, nous n'avons eu accès ni au « fil » ou au « mur » des réseaux sociaux des participants, ni aux groupes privés qu'ils pourraient fréquenter sur Facebook. Mais cela ne veut aucunement dire que les résultats de notre étude n'intègrent pas la consultation d'informations par les participants au travers de leurs réseaux sociaux. En effet, premièrement, nous avons comptabilisé le temps que les participants ont passé sur les pages Facebook, Twitter et sur les chaînes YouTube associées aux 2 295 sources d'information que nous avons identifiées. Nous en avons fait de même pour le temps qu'ils ont passé sur les 651 chaînes YouTube « indépendantes » suivies dans cette étude (rappelons que YouTube est un réseau social).

Deuxièmement, le fait pour l'utilisateur d'un réseau social de cliquer sur un article ou une vidéo apparaissant sur son « fil » ou son « mur » personnel ou au sein d'un groupe privé le conduit généralement à être redirigé sur la page Internet hébergeant ce contenu. Ainsi, ceux des participants à notre étude qui ont cliqué sur un article de presse présent sur leur « mur » Facebook, par exemple, ont été redirigés sur le site du média ayant publié cet article. Le temps passé par ces participants à lire cet article sur le site du média en question a dès lors été comptabilisé dans notre étude.

En revanche, ce que la méthodologie de notre étude ne nous a pas permis de prendre en compte, ce sont les informations que les participants sont susceptibles d'avoir vues directement sur leur « fil » Twitter, leur « mur » Facebook ou dans les groupes privés qu'ils fréquentent, sans donc être sortis du réseau social pour les lire. Ces informations sont essentiellement de deux types. Premièrement, les titres d'articles de presse que les participants ont pu voir passer, mais sur lesquels ils n'ont pas cliqué – de ces articles, ils n'ont donc lu au mieux que le titre, voire le chapeau dans les cas où il apparaissait également sur le réseau social. Deuxièmement, les courtes vidéos spécifiquement prévues pour être visionnées directement sur les réseaux sociaux que produisent certains médias en ligne, tels que Brut par exemple.

Il faut cependant indiquer que cette limitation de la présente étude n'en affecte probablement les résultats qu'à la marge. En effet, d'après les rares données disponibles sur la question, il semblerait que les informations de type médiatique ne représentent qu'une faible part des contenus en circulation sur les réseaux sociaux. Ainsi, sur Facebook – qui est, avec YouTube, le réseau social le plus utilisé en France⁹ – seuls 5 % en moyenne des contenus présents sur le « mur » des utilisateurs relèveraient de l'information en un sens large¹⁰. Si ce chiffre masque des disparités liées à la personnalisation des contenus effectuée par l'algorithme de Facebook¹¹, il indique bien que l'information de type médiatique est dans l'ensemble relativement marginale sur le principal réseau social en France. Cette situation découle pour partie d'une politique de Facebook visant à limiter activement la part d'informations visibles sur ses pages¹².

— ENCADRÉ 3

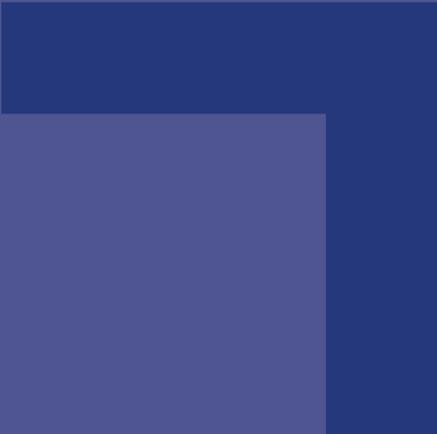
9. Près de 2/3 des internautes français utilisent Facebook (<http://fr.wikipedia.org/wiki/Facebook>). Sur les 30 jours de notre étude, 67 % des participants ont utilisé Facebook, contre 24 % pour Twitter.

10. <https://www.buzzfeednews.com/article/alexkanrowitz/facebook-wont-release-its-5-of-news-feed-content-is-news>. Cette estimation englobe les informations liées à l'actualité politique, sociale, économique, internationale mais aussi au divertissement ou au « lifestyle ».

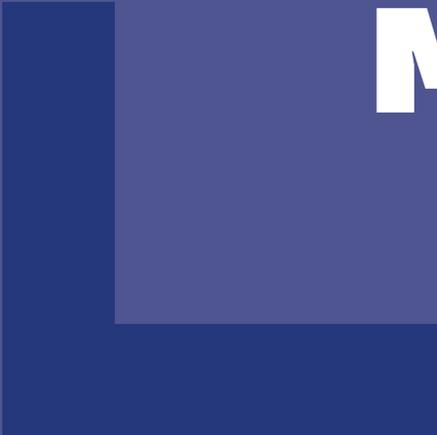
11. Selon la seule étude qui, à notre connaissance, adresse directement la question de la part de l'information sur Facebook, 50 % des « murs » personnels de l'échantillon d'utilisateurs analysé ne présentaient aucun contenu d'information parmi les dix premiers contenus affichés, 23 % un seul contenu d'information, 16 % deux contenus d'information et les 11% restant plus de deux.

<https://www.niemanlab.org/2017/12/how-much-news-makes-it-into-peoples-facebook-feeds-our-experiment-suggests-not-much/>

12. <https://www.buzzfeednews.com/article/alexkanrowitz/facebook-wont-release-its-5-of-news-feed-content-is-news>.



2



Méthode

2.1 — Participants à l'étude

Les participants à notre étude sont les 2 372 personnes qui composent l'*Access panel* français de la société Respondi¹³. Respondi enregistre en continu, avec leur consentement, l'intégralité des données de connexion Internet par ordinateurs, tablettes et téléphones mobiles de ces personnes. Ces dernières ont été recrutées par Respondi de manière à constituer un panel représentatif de la population française âgée de 18 ans et plus. Le panel est régulièrement testé et renouvelé pour assurer sa représentativité.

Pour cette étude, nous avons analysé les données de connexion des 2 372 membres de l'*Access panel* français de Respondi sur une période de 30 jours consécutifs, du 20 septembre au 19 octobre 2020. À noter que, sur la période de l'étude, seuls 48 % des membres du panel ont été suivis sur leur téléphone mobile en plus de leurs autres objets connectés, 52 % ne désirant pas l'être ou ne possédant pas de téléphone connecté¹⁴. Ils sont au contraire 24 % à n'avoir été suivis que sur leur téléphone connecté, ce qui correspond très exactement à la part des Français ne disposant pas d'un ordinateur personnel¹⁵.

2.2 — Sélection des sources d'information suivies

Afin de décrire les comportements d'information en ligne des Français, nous avons cherché à mesurer l'accès des membres de notre panel au plus large éventail possible de sources d'information en langue française. Pour ce faire, nous avons demandé à la société Storyzy¹⁶, spécialisée dans l'analyse des sites Internet, de nous fournir la liste de toutes les sources francophones détectées par son algorithme comme traitant d'information. Storyzy nous a communiqué une base de données regroupant 5 715 noms de domaines, ainsi que leurs chaînes YouTube, leurs pages Facebook et Twitter, et leurs applications mobiles associées, lorsqu'elles existent.

Pour trouver un équilibre entre le volume des données à traiter et la représentation la plus fidèle possible du comportement en ligne des Français, nous avons choisi de ne retenir que les domaines ayant enregistré au minimum 15 000 visiteurs dans le monde durant le mois précédant notre étude. Pour établir le nombre de visiteurs de chaque domaine, nous nous sommes basés sur les estimations fournies par Alexa, un service d'Amazon notamment dédié à cette fonction¹⁷. En suivant ce critère d'inclusion, la liste des sources d'information suivies dans cette étude est passée de 5 715 à 2 295 (rappelons que cette liste ne contient que

des sources d'information en langue française). À noter l'absence dans cette liste des sites de prévision météo ainsi que d'un certain nombre de sites de programmes TV, les sources concernées n'atteignant pas, selon les données Alexa, le seuil des 15 000 visiteurs mensuel.

Le critère d'inclusion de 15 000 visiteurs mensuel minimum que nous avons fixé, bien qu'arbitraire, permet d'alléger la base de données de cette étude et ainsi de faciliter son traitement tout en garantissant d'inclure une très large part des sources d'information potentiellement consultées par les participants. Les 2 295 sources d'information que nous avons retenues constituent une base de données plus large que celles utilisées dans les principales études sur le sujet, que ce soit en langue anglaise (Allen *et al.*, 2020 ; Yang *et al.*, 2020 ; Tewksbury, 2005) ou française (Institut Montaigne, 2019). Le nombre important de sources d'information que nous avons retenues offre notamment l'avantage d'augmenter la probabilité de détecter du trafic sur des médias de niche, plus susceptibles de relayer des informations de faible qualité que les principaux médias nationaux et régionaux (Allcott & Gentzkow, 2017 ; Martens *et al.*, 2018).

Afin d'accroître encore notre couverture des sources d'information en ligne, nous avons ajouté à notre base de données des chaînes YouTube d'information qui n'étaient pas déjà comprises parmi celles associées aux 2 295 sites Internet retenus pour cette étude (près de 30 % de ces derniers possèdent une chaîne YouTube associée). En effet, cette plateforme numérique est devenue incontournable en raison de l'ampleur de sa fréquentation et du nombre important de chaînes qu'elle héberge. Au mois d'août 2020, par exemple, YouTube a vu passer 43 087 000 visiteurs provenant de France et de nombreux médias y sont désormais présents.

Pour sélectionner les chaînes YouTube ajoutées à notre base de données, nous avons demandé à la société Storyzy de nous fournir la liste de toutes les chaînes YouTube francophones détectées par son algorithme comme traitant

13. www.respondi.com/FR/access-panel

14. En 2019, 77 % des Français possédaient un téléphone connecté. www.vie-publique.fr/en-bref/272039-barometre-du-numerique-95-des-francais-disposent-dun-telephone-mobile

15. En 2019, 76 % des Français possédaient un ordinateur personnel. www.vie-publique.fr/en-bref/272039-barometre-du-numerique-95-des-francais-disposent-dun-telephone-mobile

16. <https://storyzy.com/> Storyzy a entraîné un algorithme à reconnaître et classer automatiquement, en fonction de la nature de leur contenu, plusieurs dizaines de milliers de sites web, de blogs et de chaînes vidéo en plusieurs langues. Cette classification porte également sur la fiabilité de ces sources en ligne (cf. *infra*). La base de données de Storyzy regroupant les sources Internet jugées non fiables contient à ce jour plus de 30 000 sites, blogs et chaînes vidéos.

17. www.alexa.com/siteinfo

d'information et d'actualité. Storyzy nous a communiqué les urls de 3142 chaînes YouTube, en plus de celles associées à l'un des 2295 sites Internet initialement retenus. Là encore, afin de ne pas alourdir inutilement notre base de données avec de chaînes qui n'auraient probablement enregistré aucune visite de la part des participants à notre étude, nous n'avons conservé que celles d'entre elles possédant au minimum 15 000 abonnés le mois précédant la récolte des données de notre étude. Au final, nous avons ainsi ajouté 651 chaînes YouTube à notre base de données.

Dans la suite de ce rapport, et sauf mention contraire, quand nous parlons de « **sources d'information** », nous désignons : 1) les 2295 sites Internet d'information qui composent notre base de données avec, le cas échéant, leurs chaînes YouTube associées, leurs pages Facebook et Twitter associées et leurs applications mobiles associées et 2) les 651 chaînes d'information YouTube « indépendantes », c'est-à-dire, non associées à ces sites.

L'ensemble des sites Internet, des applications mobiles, des pages Facebook et Twitter et des chaînes YouTube suivis dans cette étude couvre un très large spectre des sources d'information francophones disponibles en ligne. Cela va des pages Internet de la presse papier nationale et des chaînes d'information du service public à celles de sites plus confidentiels de « réinformation », régulièrement accusés de publier de fausses informations, en passant par les pages Internet de la presse féminine, masculine, people, culturelle, sportive ou de divertissement, ainsi que par celles des agrégateurs d'informations en ligne et des médias présents uniquement sur Internet. Les 2 946 sources d'information identifiées et suivies dans cette étude recouvrent et excèdent de beaucoup celles qui sont habituellement retenues dans les études sur la consommation d'informations sur Internet en France, dont l'étude ONENEXT 2021 de l'Alliance pour les Chiffres de la Presse et des Médias (ACPM). Les seules sources d'information en ligne qui se retrouvent dans l'étude de l'ACPM et qui ne figurent pas dans la nôtre sont celles ayant reçu la visite de moins de 15 000 personnes dans le monde le mois précédant la collecte des données de connexion de nos participants.

Il importe de souligner que, dans la présente étude, le temps potentiellement consacré par les participants à s'informer sur les sources ou *via* les supports suivants n'a pas été comptabilisé : les sources d'information consultées par les participants sur d'autres objets connectés que leurs objets personnels, les sources d'information en langues étrangères, les podcasts d'information consultés hors des sources suivies, les contenus d'information téléchargés en ligne mais consultés hors ligne, les informations partagées *via* des newsletters, des notifications push, des messageries privées (type WhatsApp) ou des groupes privés sur les réseaux sociaux qui ne nécessitent pas de se rendre sur une page Internet extérieure pour les lire, ainsi que les informations partagées directement sur le « mur » Facebook ou le « fil » Twitter des participants, si la lecture de ces informations peut se faire sans avoir à se rendre sur le site des médias les ayant publiées (voir l'Encadré 3 de la première section de ce rapport).

2.3 Mesure de la consultation des sources d'information suivies

Nous avons retenu deux variables pour mesurer l'utilisation par les participants à notre étude des sources d'information qui composent notre base de données. Premièrement, nous avons enregistré l'accès à ces sources. Ainsi, quand un participant s'est rendu sur un site d'information donné (ou sur sa chaîne YouTube, sa page Facebook ou Twitter ou son application mobile associée, lorsqu'elle existe), ce participant a été comptabilisé comme ayant fréquenté cette source d'information. Le nombre de fois que ce participant s'est rendu sur cette même source d'information durant la période de l'étude est également comptabilisé. Nous avons procédé de même pour les chaînes d'information YouTube « indépendantes » : chaque fois qu'un participant a visionné une vidéo appartenant à l'une de ces chaînes, il a été comptabilisé comme ayant fréquenté la chaîne YouTube en question.

Deuxièmement, nous avons enregistré le temps passé par chaque participant sur chacune des sources d'information suivies. Le temps passé par un participant sur chaque source d'information à chacune de ses visites sur cette source a été additionné pour refléter le temps total que ce participant est resté sur cette source durant la période de l'étude.

À noter que pour éviter de faire gonfler artificiellement le temps passé sur une source d'information par un participant qui aurait ouvert l'une de ses pages sur son navigateur, mais qui ne serait pas resté devant son écran pour la consulter, l'enregistrement du temps passé s'interrompt automatiquement après 3 minutes d'inactivité de sa part sur la page en question (cela ne concernant pas les contenus audio ou vidéos en cours de lecture).

Soulignons encore que pour certaines sources d'information suivies, nous avons distingué leurs sous-domaines qui relèvent effectivement de l'information de ceux qui n'en relèvent pas. Par exemple, les sites Internet des grandes chaînes de télévision (sites de TF1, M6, etc.) permettent de visionner en direct ou en différé tant de l'information et de l'actualité que des séries ou des films de fiction, ainsi que des émissions de divertissement non liées à des sujets d'actualité. Dans cette étude, chaque fois que cela était possible – et ça l'a été pour toutes les sources d'information les plus utilisées par notre panel – nous n'avons comptabilisé que les visites et le temps passé par les participants sur les sous-domaines de ces sources spécifiquement consacrés à de l'information.

2.4 Catégorisation des sources d'information comme fiables ou non

Afin de pouvoir évaluer l'exposition des participants à de l'information en ligne jugée non fiable, nous avons catégorisé chacune des sources d'information retenues pour cette étude comme fiable ou non. Pour le faire, nous avons demandé à la

société Storyzy de nous communiquer sa classification en termes de fiabilité de l'intégralité des sources d'information que nous avons retenues (y compris, donc, des chaînes YouTube « indépendantes »).

Storyzy a développé un algorithme permettant de classer les domaines d'information sur Internet en fonction de leur fiabilité. L'algorithme procède à cette classification principalement sur la base des relations existant entre les domaines (auteurs de contenus, liens, citations, reprises de contenus, etc.). Ainsi, un domaine étroitement lié à d'autres domaines initialement jugés non fiables par une évaluation manuelle sera lui-même considéré comme non fiable. L'algorithme de Storyzy propose en outre une classification de la nature des contenus qui posent problème sur les domaines considérés comme non fiables : complotisme, fausses informations, « pièges à clics » et satire.

Nous avons repris la catégorisation binaire fiable / non fiable proposée par Storyzy pour l'intégralité des sources suivies dans cette étude. Après avoir visité une à une les sources jugées non fiables par Storyzy, nous avons précisé la thématique principale de chacune d'entre elles. Nous sommes ainsi arrivés à la classification suivante pour les sources d'information jugées non fiables¹⁸ :

- **Désinformation généraliste.** Il s'agit des sources considérées comme non fiables par Storyzy qui proposent des contenus d'information et d'actualité sur des sujets essentiellement sociaux et politiques. Ces domaines sont considérés comme non fiables en raison du fait qu'une partie au moins des contenus publiés sont jugés complotistes ou consistent en de fausses informations.
- **Désinformation en santé.** Il s'agit des sources considérées comme non fiables par Storyzy qui proposent des contenus axés sur la santé humaine. Ces sources sont jugées non fiables en raison du fait qu'une partie au moins des contenus publiés sont de nature complotiste ou consistent en de fausses informations.
- **Pièges à clics.** Il s'agit des sources considérées comme non fiables et classées comme relevant de la catégorie « pièges à clics » par Storyzy. Les sources « pièges à clics » sont des sources dont les contenus n'ont pas pour but d'informer le lecteur mais uniquement d'attirer son attention en vue de générer du trafic sur leurs pages. Les « pièges à clics » recourent à différents procédés trompeurs : le titre de l'article est volontairement choquant et sensationnaliste, le contenu de l'article ne correspond pas au titre, l'image qui accompagne le titre est hors sujet ou truquée, etc. Ces sites attirent les internautes en achetant des espaces publicitaires pour leurs articles, notamment sur les pages Internet de médias jugés fiables, telles que celles de certains grands journaux nationaux.
- **Pseudosciences.** Il s'agit des domaines considérés comme non fiables par Storyzy qui proposent essentiellement des contenus à caractère prétendument scientifique.
- **Satire.** Les sources satiriques sont considérées comme non fiables par l'algorithme de Storyzy. La pertinence de cette

catégorisation est discutable, dans la mesure où les sources satiriques se distinguent généralement des autres sources jugées non fiables par l'avertissement explicite qu'ils destinent aux lecteurs sur la présence de fausses informations à caractère humoristique sur leurs pages. Cela n'a cependant pas d'effet sur les résultats de notre étude, dans la mesure où, comme nous l'exposerons dans la section Résultats de ce rapport, la catégorie « satire » ne réunit que moins d'un pourcent du trafic des participants sur l'ensemble des sources jugées non fiables.

Ne voulant pas nous fier exclusivement à la catégorisation algorithmique des sources d'information proposée par Storyzy, nous l'avons comparée à celle effectuée par Les Décodeurs du journal *Le Monde*. Les Décodeurs sont un groupe de journalistes professionnels spécialisés dans le *fact checking* qui mettent à disposition du public un outil en ligne permettant d'évaluer le degré de fiabilité de sources Internet : le Décodex¹⁹. La classification des sources Internet proposée sur le Décodex est le fruit de l'évaluation de leurs contenus par les journalistes des Décodeurs. Cette classification n'est pas binaire (fiable/non fiable), mais distingue quatre catégories selon un code couleur²⁰ :

- rouge : « pour les sites diffusant régulièrement de fausses informations » ;
- orange : « pour les sites dont la fiabilité ou la démarche est douteuse (sources peu mentionnées, démarche militante cachée, etc.) » ;
- bleu : « les sites parodiques » ;
- sans couleur : sites fiables.

Pour mener à bien notre comparaison entre la classification proposée par Storyzy, d'un côté, et celle du Décodex, de l'autre, nous avons procédé de la manière suivante. Premièrement, nous avons sélectionné dans la liste des sources retenues pour notre étude les 366 d'entre elles qui ont fait l'objet d'une évaluation par Les Décodeurs (ces derniers nous ont fourni en décembre 2020 la base de données à jour de leur classification des sources Internet)²¹.

18. Chaque source ne se retrouve que dans une seule de ces catégories. À noter que certaines sources « pièges à clics » comprennent de nombreux contenus relevant de questions de santé. Elles ont cependant été considérées comme appartenant à la catégorie des « pièges à clics », dans la mesure où elles peuvent également traiter d'autres sujets et qu'elles présentent les caractéristiques typiques de ce genre de sites : titres sensationnalistes, contenus ne correspondant pas au titre, illustrations trompeuses, etc.

19. www.lemonde.fr/verification/

20. www.lemonde.fr/les-decodeurs/article/2017/01/23/l-annuaire-des-sources-du-decodex-mode-d-emploi_5067719_4355770.html

21. La base de données du Décodex contient au total un peu plus de mille sources Internet évaluées. Le fait que nous n'ayons retrouvé que 366 d'entre elles dans notre propre base de données s'explique par trois raisons : 1) de nombreuses sources Internet évaluées par le Décodex n'atteignaient pas le moins précédent notre étude le seuil des 15 000 visiteurs uniques que nous avons fixé comme critère d'inclusion à notre base de données ; 2) la base de données du Décodex contient aussi des sources en langues étrangères, sources que nous n'avons donc pas prises en compte dans notre étude ; 3) la base de données du Décodex contient finalement, en plus des sites Internet, des pages Facebook ou des chaînes YouTube n'entrant pas dans le cadre de notre étude.

Deuxièmement, nous avons recodé le code couleur utilisé par le Décodex en un code binaire: fiable/autre. Nous avons classé les sites sans couleur dans la catégorie « fiable » et les sites « rouges », « orange » et « bleus » dans la catégorie « autre ».

Nous avons alors été en mesure de comparer statistiquement la catégorisation de ces 366 sites par le Décodex et par Storyzy (catégorisation Décodex : autre = 0, fiable = 1 ; catégorisation Storyzy : non fiable = 0, fiable = 1).

Il ressort de cette analyse statistique que la cohérence entre la catégorisation du Décodex et celle de Storyzy sur cet échantillon commun de 366 sites est excellente (Cohen's kappa = 0,896). Il nous a donc paru légitime d'utiliser la catégorisation de Storyzy pour l'ensemble des sources d'information Internet retenues pour notre étude.

2.5 Questionnaire de fin d'étude

Au terme des 30 jours de collecte des données Internet des participants à notre étude, nous les avons invités à répondre à un questionnaire auto-administré en ligne permettant d'établir leur profil sociodémographique et politique. Ce questionnaire comportait également des questions sur leur rapport à l'information, sur les sources d'information sur et hors Internet qu'ils utilisent ainsi que sur la confiance qu'ils leur accordent. Sur les 2 372 participants à notre étude, 1 614 ont complété ce questionnaire de fin d'étude.

Comme nous le verrons dans la section Résultats de ce rapport, nous avons exploité les réponses à ce questionnaire afin de déterminer : 1) si les participants ont une représentation correcte de la manière dont ils s'informent sur Internet ; 2) si la consultation de sources d'information par les participants est fonction du niveau de confiance qu'ils disent leur accorder ; 3) si les participants qui s'informent plus que les autres sur Internet, ou ceux qui consultent des sources d'information non fiables, présentent des caractéristiques particulières.

Tests statistiques utilisés dans cette étude

Nous avons eu recours à deux types de tests statistiques simples pour décrire et analyser certaines données de la présente étude :

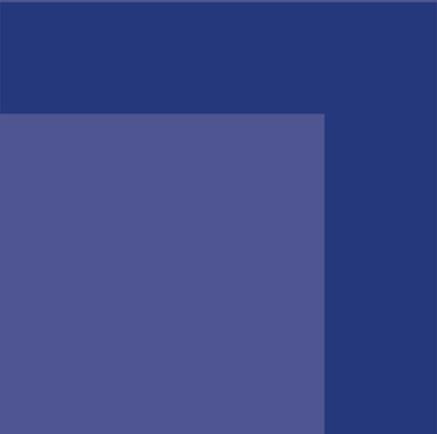
1 — Analyses de corrélations linéaires.

Ces analyses permettent de déterminer si deux variables entretiennent une relation linéaire entre elles et de quantifier leur degré d'association. Le coefficient de corrélation (exprimé R), qui se situe entre 0 et 1, indique ce degré d'association : plus il est proche de 0, plus le lien entre les deux variables est faible, plus il est proche de 1, plus ce lien est fort. Lorsque les deux variables corrélées évoluent dans une direction opposée, la valeur R est exprimée négativement. Indépendamment de son coefficient de corrélation, le lien entre deux variables peut être considéré comme statistiquement significatif ou non. Il l'est lorsque la probabilité que les variables n'entretiennent pas de relation linéaire se situe en dessous d'un seuil conventionnel, exprimé par la valeur p. On considère généralement qu'une corrélation entre deux variables est significative en ce sens lorsque $p < 0,05$ (le seuil de $p < 0,10$ est aussi parfois utilisé).

2 — T-tests.

Ces tests permettent de déterminer si les moyennes de deux groupes de données diffèrent significativement entre elles, à savoir, si $p < 0,05$ (ou $p < 0,10$). Lorsqu'on observe une telle différence statistique, on peut en mesurer l'ampleur, dite « taille d'effet ». La taille d'effet peut notamment se quantifier au moyen du g de Hedges (lorsque les deux groupes de données sont de tailles différentes, ce qui sera le cas dans nos analyses). Par convention, on considère généralement qu'un g de Hedges autour de 0,2 correspond à un effet faible, autour de 0,5 à un effet moyen et autour de 0,8 à un effet fort.

— ENCADRÉ 4



3



Résultats

3.1 – Connexions à Internet

Le temps total de connexion à Internet de chaque participant durant les 30 jours de l'étude est, en moyenne, de 65,3 heures (médiane = 41 heures). Ce temps de connexion, qui correspond à celui que les participants à notre étude ont passé connectés et actifs sur Internet, quel que soit le support utilisé (ordinateur, tablette ou téléphone mobile) et quelle que soit l'activité à laquelle ils s'y adonnaient, varie grandement entre les participants, comme on peut le constater sur la figure suivante.

Temps total de connexion à Internet

(durant les 30 jours de l'étude)

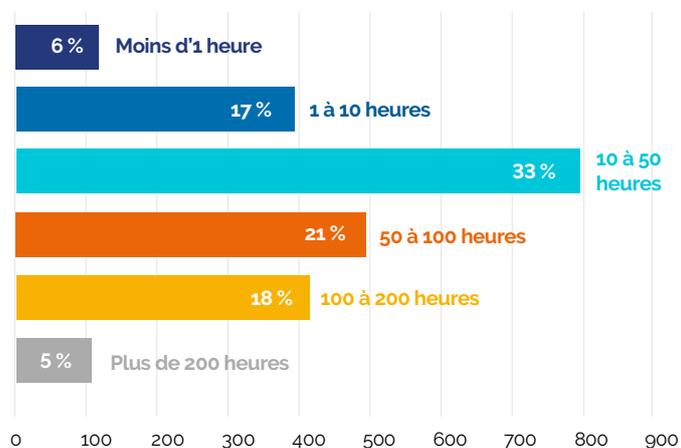


FIG. 1 – LECTURE : 33 % des participants ont passé de 10 à 50 heures sur Internet durant les 30 jours de l'étude.

Tous les participants ne se sont pas connectés quotidiennement à Internet. En moyenne, ils se sont connectés au cours de 21 jours sur les 30 de l'étude (médiane = 25 jours).

La durée quotidienne moyenne de connexion des participants pris individuellement, en ne considérant que les jours au cours desquels ils ont utilisé Internet, est de 2,6 heures (médiane = 1,9 heure). Là encore, les variations interindividuelles sont importantes, comme le montre la figure suivante.

Il faut souligner que les membres du panel de Respondi ont la possibilité de désactiver l'enregistrement des données lorsqu'ils le souhaitent. Nous ne pouvons donc pas exclure la possibilité que nos participants se soient connectés à Internet davantage de temps que ce que nous avons enregistré au cours des 30 jours de l'étude. Cependant, si l'on se réfère aux chiffres avancés par la société de mesure d'audience

Durée quotidienne de connexion

(en ne considérant que les jours d'utilisation d'Internet)

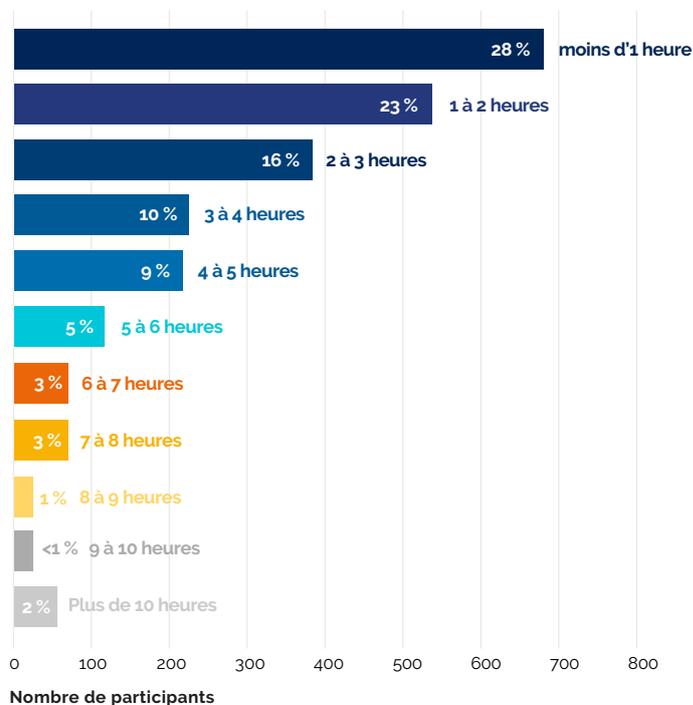


FIG. 2 – LECTURE : 23 % des participants ont passé en moyenne entre 1 et 2 heures sur Internet les jours au cours desquels ils ont utilisé Internet.

Médiamétrie, les valeurs d'utilisation d'Internet enregistrées chez nos participants sont très similaires à celles de l'ensemble des Français. Selon Médiamétrie, en effet, les Français se seraient connectés à Internet en moyenne 22 jours par mois en 2019, pour une durée quotidienne moyenne de 2,12 heures²². Les données de connexion que nous avons enregistrées chez les participants à notre étude sont ainsi cohérentes avec celles de la population française dans son ensemble.

À quelles activités nos participants se sont-ils consacrés sur Internet durant les 30 jours de l'étude ? La figure suivante représente la manière dont le temps total de connexion de l'ensemble des participants est réparti selon la nature des sources consultées.

22. www.mediametrie.fr/fr/lannee-internet-2019

Répartition du temps d'activité sur Internet (durant les 30 jours de l'étude)

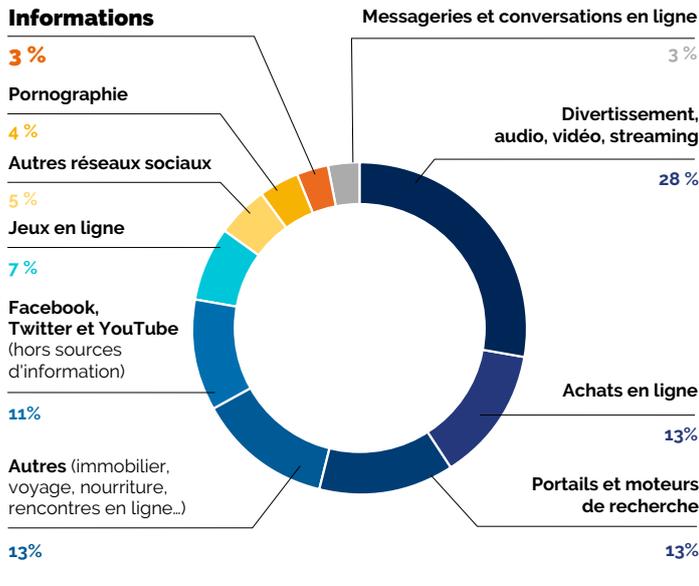


FIG. 3 — LECTURE : Durant les 30 jours de l'étude, 3 % du temps total passé sur Internet par l'ensemble des participants l'a été sur des sources d'information.

NOTE : Un « portail » est une porte d'entrée commune à un éventail de ressources et de services Internet : pages web, forums de discussion, boîtes de courrier électronique, espaces de publication, moteurs de recherche, etc.²³ Certains portails, tels que Yahoo! ou Orange, comprennent des espaces informations ou actualités. Dans notre étude, le temps passé par les participants sur de tels espaces au sein des portails a été comptabilisé dans la catégorie « Informations ».

3.2 — Consommation d'informations sur Internet

3.2.1 – Panorama de la consommation effective

Comme le montre la figure précédente, 3 % du temps total passé sur Internet par l'ensemble des participants l'a été sur les sources d'information retenues dans cette étude (cette valeur est de 3,5 % chez les participants n'ayant pas été suivis sur leur téléphone connecté²⁴). Cela correspond à 2,1 heures en moyenne par participant sur les 30 jours de l'étude, le temps médian étant de 43,3 minutes. Si l'on ne considère que les jours au cours desquels les participants se sont connectés à Internet, il ressort qu'ils ont passé chacun quotidiennement 4,9 minutes en moyenne sur des sources d'information en ligne (médiane = 1,3).

23. https://fr.wikipedia.org/wiki/Portail_web

24. Les participants n'ayant pas été suivis sur leur téléphone connecté en plus de leur ordinateur passent donc une plus grande proportion de leur temps de connexion à Internet sur des sources d'information en ligne que les participants qui ont été suivis sur leur ordinateur et leur téléphone connecté. Cependant, en valeur absolue, les seconds ont passé 1,8 fois plus de temps que les premiers sur des sources d'information, leur temps total de connexion à Internet, tous objets connectés confondus, étant en moyenne 2,3 fois supérieur.

3 % du temps total passé sur Internet par l'ensemble des participants l'a été sur des sources d'information.

Ces valeurs moyennes masquent cependant une très grande variabilité des comportements d'information en ligne de la part de nos participants. Ainsi, sur les 30 jours de l'étude, 17 % d'entre eux (N = 409) n'ont consulté aucune source d'information sur Internet, 46 % (N = 1 099) ont passé en tout moins d'une heure à en consulter, tandis qu'ils sont 5 % (N = 118) à en avoir consulté durant plus de dix heures au total.

Il existe une grande variabilité des comportements d'information en ligne.

Nombre d'heures passées au total sur des sources d'information en ligne (durant les 30 jours de l'étude)

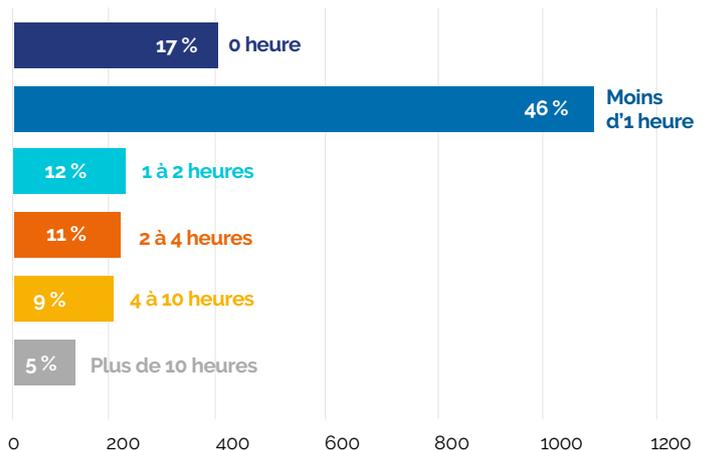


FIG. 4 — LECTURE : 46 % des participants ont passé moins d'une heure au total sur des sources d'information durant les 30 jours de l'étude.

Si l'on se penche sur la fréquence à laquelle les participants ont consulté des sources d'information, on constate qu'en moyenne ils en ont visité au moins une 45 % des jours au cours desquels ils se sont connectés à Internet (médiane = 40 %). La figure suivante détaille cette valeur pour chaque participant.

Au total, 43,8 % des 2 946 sources d'information qui composent notre base de données ont été visitées au moins une fois sur les 30 jours de l'étude par au moins l'un des 2 372 participants – ce qui correspond, dans le détail, à 51 % des 2 295 sources d'information centrées autour d'un site Internet et 18,3 % des 651 chaînes YouTube « indépendantes » ajoutées à notre base de données (voir section 2.2).

Pourcentage de jours de connexion de chaque participant au cours desquels il a consulté de l'information en ligne

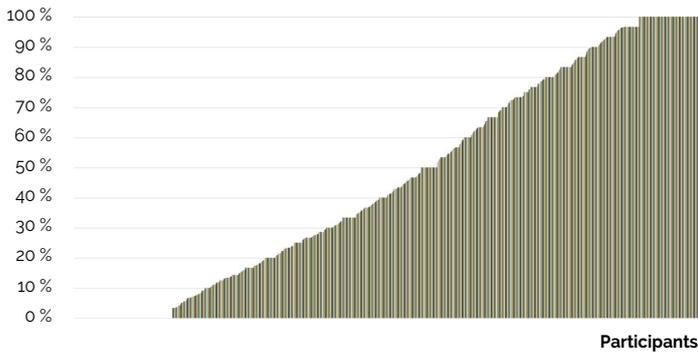


FIG. 5 — LECTURE : Chaque ligne verticale de couleur représente un participant. La hauteur d'une ligne indique le pourcentage de jours de connexion à Internet au cours desquels le participant correspondant a consulté au moins une source d'information. Les participants atteignant 50 %, par exemple, sont ceux qui ont consulté une ou plusieurs source(s) d'information un jour sur deux de connexion à Internet.

NOTE : La plage blanche à gauche du graphique correspond aux 409 participants qui, durant toute l'étude, n'ont consulté aucune source d'information sur Internet.

43,8 % des 2 946 sources d'information suivies dans cette étude ont été visitées au moins une fois sur 30 jours par au moins l'un des 2 372 participants.

Les participants dans leur ensemble ont consulté en moyenne 15 sources d'information différentes sur la durée de l'étude (médiane = 8). La figure suivante illustre l'hétérogénéité du comportement des participants quant au nombre de sources d'information qu'ils ont consultées.

Nombre de sources d'information consultées (au cours des 30 jours de l'étude)

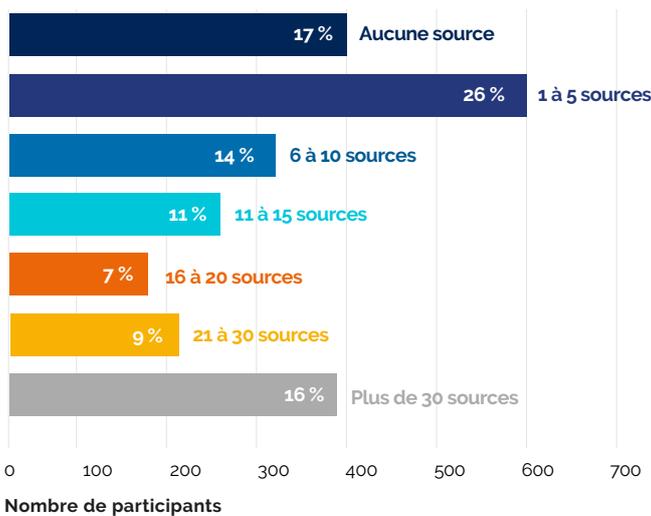


FIG. 6 — LECTURE : 26 % des participants ont consulté de 1 à 5 sources d'information différentes au cours des 30 jours de l'étude.

En moyenne, la durée de consultation d'une source d'information est de 1,9 minute par visite (médiane = 1,4 minute). Comme le montre la figure suivante, cette durée moyenne varie sensiblement entre les participants.

Durée moyenne de consultation d'une source d'information

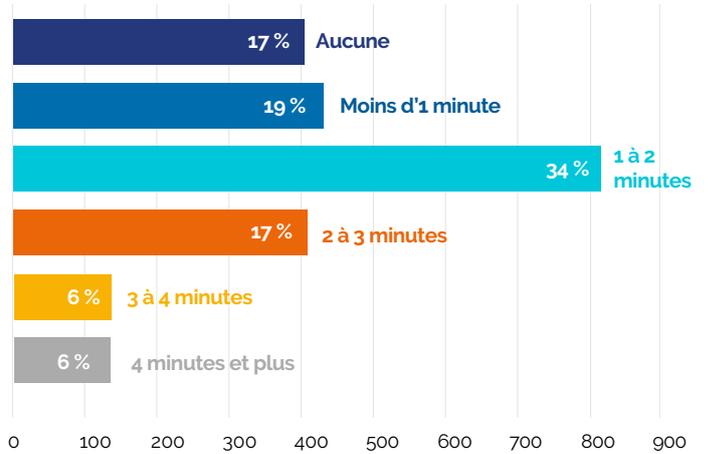


FIG. 7 — LECTURE : 34 % des participants ont passé en moyenne entre une et deux minutes consécutives par visite sur une source d'information.

Sans surprise, on observe que plus les participants ont passé de temps à s'informer en ligne durant les 30 jours de l'étude, plus le nombre de sources d'information différentes qu'ils ont consultées est important. Cela suggère que les personnes qui consacrent beaucoup de temps à s'informer sur Internet le font moins en augmentant le temps total passé sur une même source d'information qu'en diversifiant les sources qu'ils consultent.

Rapport entre temps d'information et nombre de sources d'information consultées

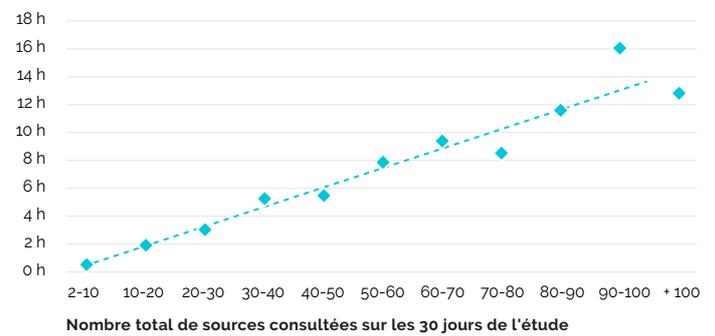


FIG. 8 — LECTURE : Les participants ayant consulté de 2 à 10 sources d'information différentes sur les 30 jours de l'étude ont passé en moyenne 0,74 heure à s'informer sur Internet. Ceux qui ont consulté de 90 à 100 sources d'information différentes ont passé en moyenne 16,01 heures à s'informer. Il existe une relation quasi-linéaire entre le temps moyen passé à s'informer en ligne et le nombre de sources d'information différentes consultées.

Cette interprétation est corroborée par le fait que, comme le montre la figure suivante, la proportion du temps total passé sur la source d'information la plus consultée par chaque participant tend à décroître à mesure que le nombre de sources d'information visitées augmente.

Proportion de temps passé au total sur la source d'information la plus consultée en fonction du nombre de sources visitées

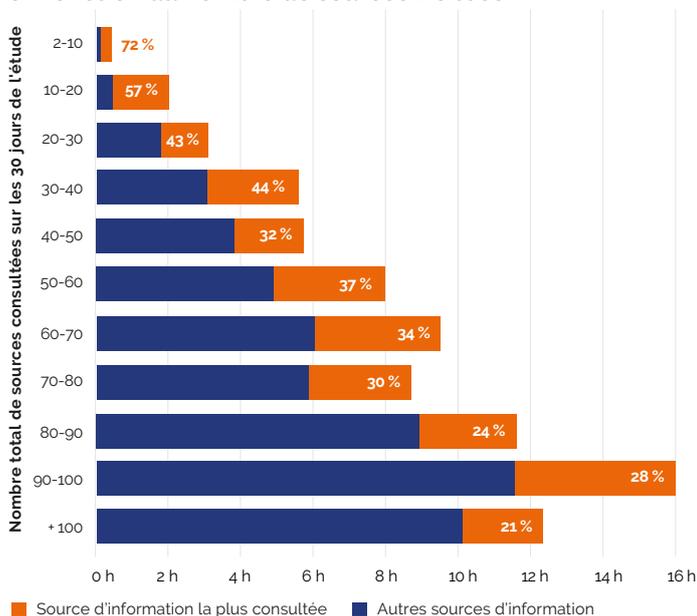


FIG. 9 — LECTURE : Les participants ayant visité sur les 30 jours de l'étude entre 50 et 60 sources d'information différentes ont passé 37 % de leur temps sur la source qu'ils ont la plus consultée.

À noter que les participants qui ont consacré beaucoup de temps à s'informer en ligne ont tout de même passé en moyenne davantage de temps au total que les autres sur leur source d'information principale. C'est ce que l'on peut observer sur la figure suivante.

Évolution du temps total passé sur la source la plus consultée en fonction du nombre de sources visitées

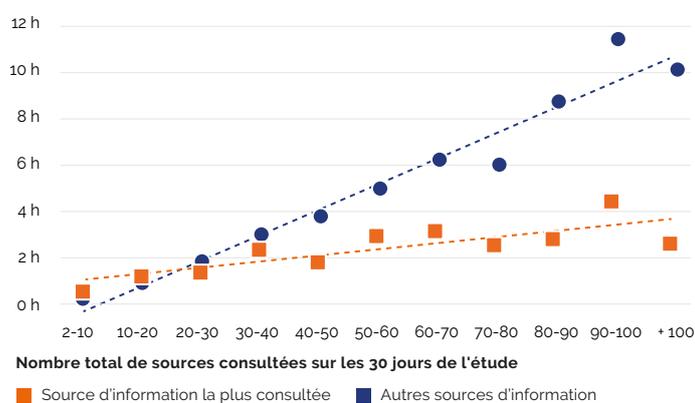


FIG. 10 — LECTURE : La ligne pointillée orange (droite de régression) représente l'augmentation du temps total passé sur la source d'information la plus consultée par chaque participant à mesure que s'accroît le nombre de sources visitées. La ligne pointillée bleue représente l'augmentation correspondante pour ce qui est du temps total passé sur les autres sources d'information consultées.

3.2.2 — Utilisation des applications pour s'informer

Comme nous l'avons exposé dans la section Méthode de cette étude, les 2 295 sources d'information retenues (hors des 651 chaînes YouTube « indépendantes » que nous avons ajoutées à notre base de données) correspondent à un site Internet ainsi que, le cas échéant, à la chaîne YouTube, à la page Facebook et Twitter et à l'application mobile qui lui sont associées. La figure suivante montre que la grande majorité du temps passé par les participants sur ces sources d'information l'a été sur leur site Internet.

La grande majorité du temps passé par les participants sur les sources d'information l'a été sur leur site Internet.

Mode de consultation des sources d'information

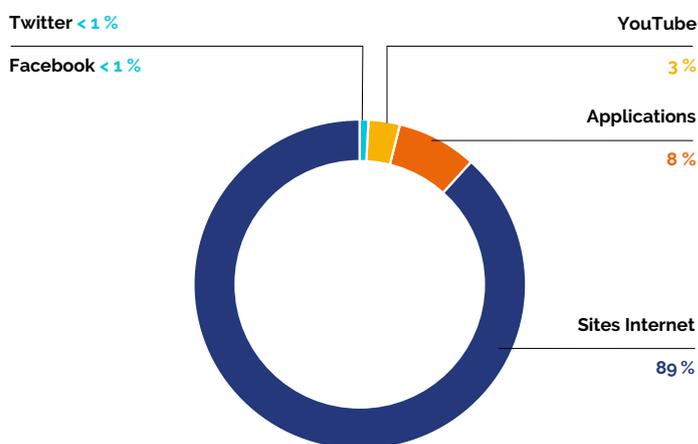


FIG. 11 — LECTURE : 89 % du temps passé par les participants sur des sources d'information l'a été sur le site Internet de ces sources, 8 % sur leurs applications mobiles.

NOTE : Graphique excluant les applications des grandes chaînes de télévision généralistes.

Il importe de souligner que nos analyses sous-estiment partiellement l'utilisation d'applications d'information par les participants, cela pour deux raisons. Premièrement, seuls 48 % des participants ont été suivis sur leur téléphone mobile en plus de leurs autres objets connectés. Autrement dit, nous n'avons pu mesurer l'utilisation des applications d'information que chez un peu moins de la moitié des participants.

Deuxièmement, nous avons écarté de nos analyses les applications mobiles des grandes chaînes de télévision généralistes (M6 et TF1, en particulier). En effet, l'outil de mesure de connexion aux applications mobiles ne permet pas de déterminer ce que les participants y ont consulté et, donc, de savoir si ces contenus relèvent ou non de l'information. Or, si l'on se réfère à nos données provenant des sites Internet

des chaînes de télévision généralistes, on observe que la très grande majorité du trafic s'y concentre sur les programmes de divertissement et sur les séries télévisées, non pas sur l'information. Contrairement à ce que nous avons pu faire avec les données en provenance des sites Internet de ces chaînes, il ne nous a pas été possible de « nettoyer » les données des applications mobiles afin de ne conserver que celles correspondant à de la consultation d'informations. Dans ce contexte, il nous a paru préférable d'écarter de nos analyses les applications mobiles de ces chaînes afin d'éviter de gonfler artificiellement le rôle joué par les applications dans la consommation d'informations en ligne.

Sur les 48 % des participants ayant été suivis sur leur téléphone mobile, seuls 215 d'entre eux (18,7 %, soit 9,1 % du panel total) ont utilisé au moins une application d'information sur les 30 jours de l'étude. Si on limite nos analyses à cette catégorie de participants, on constate que, pour eux, les applications mobiles constituent un moyen d'accès privilégié à l'information. En effet, ces participants ont passé en moyenne 48,1 % de leur temps d'information en ligne sur des applications dédiées (médiane = 45 %).

Mode de consultation des sources d'information chez les utilisateurs d'applications

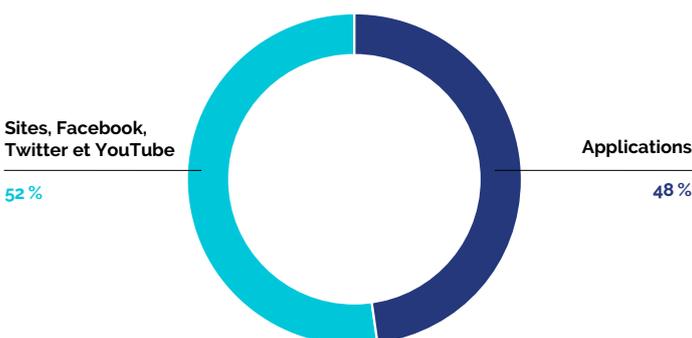


FIG. 12 — LECTURE : 48 % du temps passé sur des sources d'information par les utilisateurs d'applications (N = 215) l'a été *via* l'application de ces sources.

NOTE : Graphique excluant les applications des grandes chaînes de télévision généralistes.

Dans l'ensemble, seul **4,4 %** du temps de connexion consacré par nos participants à la consultation d'informations en ligne l'a été sur YouTube.

La figure suivante présente le classement des dix applications d'information les plus utilisées par ces 215 participants.

Classement des applications d'information par nombre d'utilisateurs

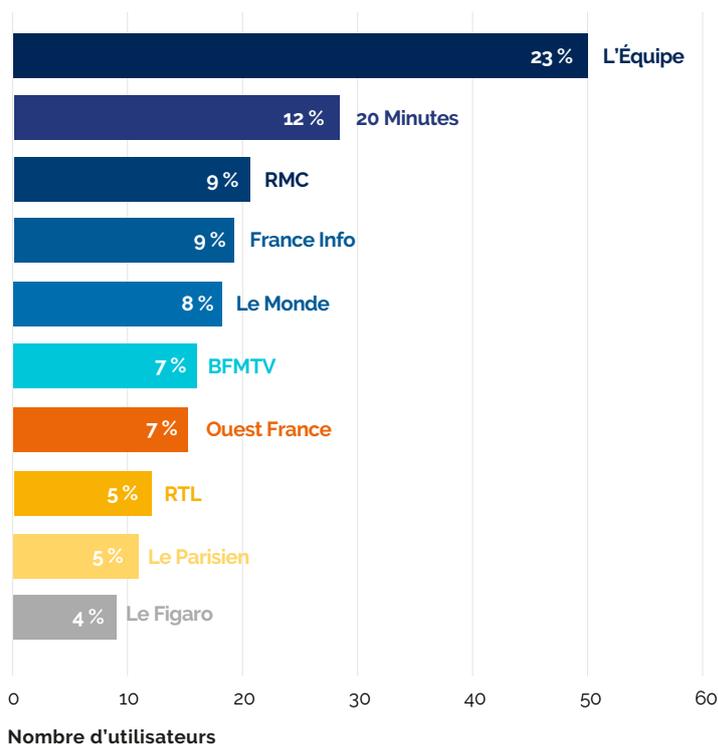


FIG. 13 — LECTURE : Sur les 215 participants qui ont utilisé une application d'information pour smartphone, 23 % d'entre eux ont utilisé l'application du journal L'Équipe, 9 % celle de France Info.

NOTE : Graphique ne prenant pas en compte les applications des grandes chaînes de télévision généralistes ni celles de prévision météo.

3.2.3 — Utilisation de YouTube pour s'informer

Dans l'ensemble, seul 4,4 % du temps de connexion consacré par nos participants à la consultation d'informations en ligne l'a été sur YouTube. Cette proportion de temps passé sur YouTube inclut tant la consultation de chaînes appartenant aux sources d'information prises en compte dans cette étude que celle de chaînes « indépendantes » ajoutées à notre base de données (voir section 2.2). Dans le détail :

- 325 participants (13,7 % de notre panel) se sont connectés au moins une fois à une chaîne YouTube appartenant à l'une des sources d'information sélectionnées. En moyenne, ces participants y ont passé 64 secondes par jour de connexion à Internet sur la durée de l'étude.
- 143 participants (6 % de notre panel) se sont connectés au moins une fois à une des 651 chaînes YouTube « indépendantes ». En moyenne, ces participants y ont passé 89 secondes par jour de connexion à Internet.

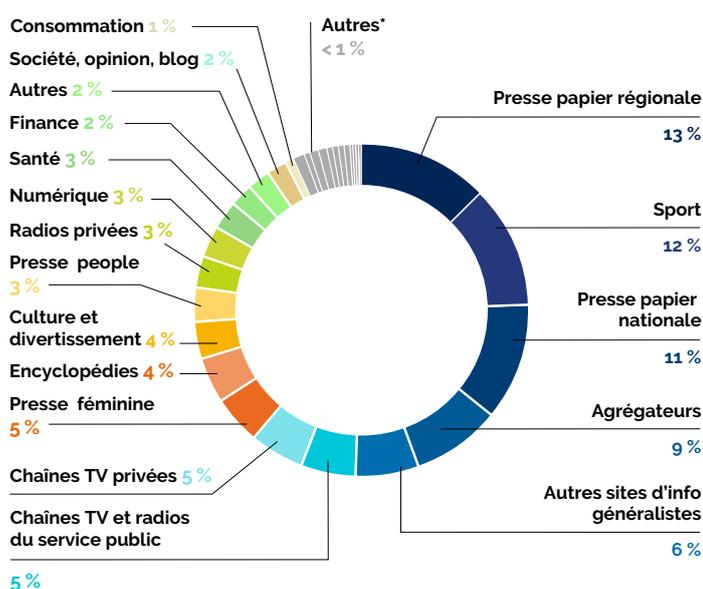
Contrairement à ce que l'on aurait pu anticiper, la consultation d'informations sur YouTube ne semble pas être liée à l'âge des participants. C'est ce que fait apparaître une analyse de corrélation que nous avons conduite entre l'âge des

participants et le temps passé sur l'intégralité des chaînes YouTube incluses dans la base de données de notre étude : $R = -0,04$, $p = 0,06$, $N = 2\,372$.

3.2.4 — Types de sources d'information consultées

Afin de pouvoir évaluer le type d'informations consultées par les participants, nous avons catégorisé manuellement chacune des 2 946 sources d'information qui composent notre base de données en fonction de la nature de leurs contenus. Les catégories identifiées ainsi que la proportion de temps passé sur chacune d'entre elles par les participants sont présentées dans la figure suivante.

Types de sources d'information consultées



* Automobile, Lifestyle, Économie, Programme TV, Médias francophones étrangers, Science, Informations professionnelles, Presse masculine, Cuisine, Actualités locales, Environnement, Infos animalières, Parentalité, Spiritualité, Aéronautique, Fact-checking

FIG. 14 — LECTURE : Les participants ont dans l'ensemble passé 13 % de leur temps d'information en ligne sur des sources en ligne de la presse papier régionale.

Comme on peut le constater, **les sources d'information les plus visitées par les participants au cours des 30 jours de l'étude sont des sources de la presse papier régionale.** Les sources en ligne d'Ouest France, de la Dépêche et de l'Est Républicain arrivent en tête de cette catégorie, suivies par celles du Parisien, du Télégramme et de Sud-Ouest. Les sources Internet de ces 6 quotidiens réunissent à elles seules plus de 50 % du trafic de leur catégorie. **Les sources d'information dédiées à l'actualité sportive arrivent en seconde position du temps consacré par les participants à s'informer en ligne.** Le succès de cette catégorie est pour bonne part dû à la forte fréquentation du site et, dans une moindre mesure, de l'application du journal l'Équipe. **Les sources en ligne de la presse papier nationale arrivent en troisième position de ce classement.**

Les sources du 20 Minutes, du Figaro et du Monde se démarquent fortement des autres sources d'information de cette catégorie en termes de fréquentation par les participants. Quant aux agrégateurs de médias, ils arrivent en quatrième position, suivis par les autres sites d'information généralistes.

Le tableau suivant détaille la fréquentation des huit types de sources d'information en ligne les plus consultées par les participants à notre étude. On y observe en particulier que si les sources d'information sportive arrivent en seconde position du classement par le temps de consultation total, elles ont été visitées par moins de la moitié du nombre de participants ayant consulté des sources de la presse papier régionale ou de la presse papier nationale, pourtant troisième de ce classement. Cela indique donc que le temps passé en moyenne par les participants sur une source d'information dépend dans une certaine mesure de sa nature. La même remarque vaut pour les agrégateurs d'informations, dont le temps de consultation total semble particulièrement élevé par rapport au nombre de participants s'y étant rendus durant les 30 jours de l'étude.

Cette situation s'explique certainement, en partie au moins, par le fait que les articles rassemblés par les agrégateurs sont en libre accès, contrairement à ceux de la presse papier, dont la lecture en ligne nécessite souvent un abonnement. Les internautes qui arrivent sur un article payant sans être abonnés au média correspondant n'y resteront donc que le temps de réaliser qu'ils n'ont pas accès à son contenu.

	Nombre de visiteurs	Nombre de visites	Temps de consult. moyen (en minutes)	Temps de consult. total (en heures)	Sources les plus consultées (selon le temps de consultation total)
Presse papier régionale	1 211	16 114	2,4	633,7	• Ouest France • La Dépêche • L'Est Républicain
Sport	560	13 404	2,6	589,5	• L'Équipe • sports.fr • footmercato.net
Presse papier nationale	1 300	14 495	2,4	567,9	• 20 minutes • Le Figaro • Le Monde
Agrégateurs	511	5 377	4,8	431,1	• MSN Actualités • Orange Actualités • Yahoo! Actualités
Autres sites d'information généralistes	1 076	8 458	2,2	313,0	• Planet.fr • Actu.fr • L'Internaute.fr
Chaînes TV et radio du service public	862	7 685	2,1	264,4	• France Info • France Inter • France Bleu
Chaînes TV privées	780	5 185	3	258,9	• BFMTV • CNEWS • M6
Presse féminine	1 024	7 043	2,1	242,6	• Le Journal des Femmes • Elle • Femme Actuelle

TAB. 1

3.2.5 — Les leaders du marché de l'information sur Internet

La figure suivante montre le « top 26 » des sources d'information qui ont été consultées par le plus grand nombre de participants au cours des 30 jours de l'étude.

Classement des principales sources d'information par nombre de participants les ayant consultées

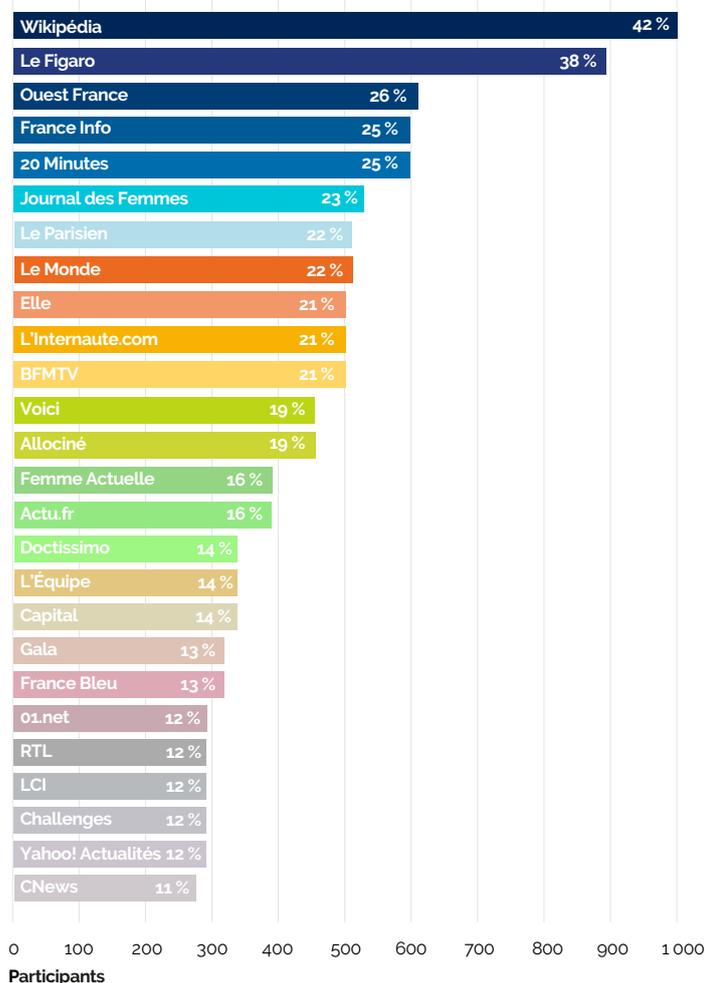


FIG. 15 — LECTURE : 42 % de l'ensemble des participants (soit 990 d'entre eux) ont consulté au moins une fois la source d'information Wikipédia au cours des 30 jours de l'étude.

NOTE : Graphique ne prenant en compte ni les sources en ligne de prévision météo, ni celles de programme TV.

On remarque que Wikipédia arrive en tête de ce classement. Si Wikipédia est une encyclopédie, et non pas une source d'information médiatique à proprement parler, nous avons décidé de le prendre en compte dans notre étude car, fort d'une communauté de contributeurs très active, le site est actualisé en temps réel, ou presque, sur de très nombreux sujets relevant de l'information de type médiatique (affaires judiciaires, actualité politique et sociale, biographies de personnalités, etc.). Au-delà de sa vocation initiale d'encyclopédie universelle et participative en ligne, Wikipédia peut ainsi être utilisé comme une source d'information médiatique majeure par de nombreux internautes.

Une part de son succès auprès des participants à notre étude s'explique probablement par le fait que Wikipédia arrive en tête des pages listées par les moteurs de recherche en réponse à de nombreuses requêtes liées ou non à l'actualité. **Dans l'ensemble, ce « top 26 » des sources d'information consultées par le plus grand nombre de participants est largement composé de sources Internet de médias traditionnels, et il totalise à lui seul 40 % du temps consacré par les participants à s'informer en ligne.**

Si l'on procède au classement des principales sources d'information non plus en termes de nombre de participants les ayant consultées mais en termes de temps de consultation total sur les 30 jours de l'étude, la composition du « top 26 » change peu, tandis que l'ordre des sources est passablement réarrangé, comme on peut le constater sur la figure suivante. Cette fois, les 26 sources d'information présentes dans ce classement totalisent 50,4 % du temps consacré par les participants à s'informer sur Internet.

Classement des principales sources d'information par temps de consultation total

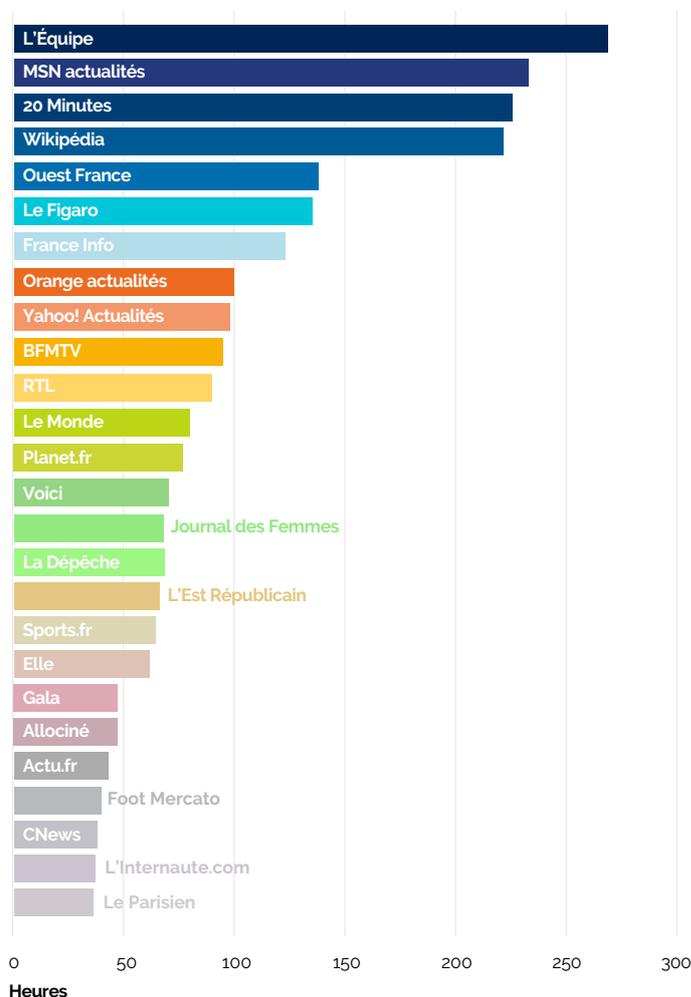


FIG. 16 — LECTURE : Les participants dans leur ensemble ont passé 266 heures au total sur la source d'information sportive L'Équipe au cours des 30 jours de l'étude.

NOTE : Graphique ne prenant en compte ni les sources en ligne de prévision météo, ni celles de programme TV.



3.3 — Caractéristiques des consommateurs d'informations sur Internet

Comme nous l'avons exposé dans la section 3.2.1 de ce rapport, les comportements de consommation d'informations sur Internet varient grandement entre les participants à notre étude. Existe-t-il des caractéristiques individuelles ou sociales associées au fait de recourir plus ou moins intensément à Internet pour s'informer ?

Pour chercher à le déterminer, nous avons procédé à des analyses de corrélation entre la proportion du temps de connexion à Internet consacrée à la consultation de sources d'information en ligne et :

- ▮ le sexe des participants,
- ▮ leur âge,
- ▮ leur catégorie socioprofessionnelle (CSP),
- ▮ leur niveau de diplôme,
- ▮ leur niveau de revenu,
- ▮ le niveau de revenu de leur foyer,
- ▮ la composition de leur foyer (le fait de vivre seul ou non et le nombre d'enfants dans le foyer),
- ▮ le fait de se déclarer plus ou moins intéressé par l'information en général et par l'information politique et sociale en particulier,
- ▮ le fait de dire s'informer plus ou moins régulièrement *via* la télévision, la radio, la presse papier ou Internet,
- ▮ le fait de déclarer une confiance plus ou moins forte à l'égard des institutions (armée, police, justice, école), à l'égard des médias, à l'égard du Gouvernement et à l'égard de la communauté scientifique,
- ▮ leur positionnement politique déclaré (sur un continuum allant de l'extrême gauche à l'extrême droite),
- ▮ leur attitude à l'égard du mouvement des Gilets Jaunes (sur un continuum allant de l'hostilité déclarée à l'égard de ce mouvement au fait de s'en dire soi-même membre),
- ▮ leur score sur une échelle standardisée de mesure du niveau d'adhésion à des propositions complotistes (Bruder *et al.*, 2013),
- ▮ leur temps quotidien de connexion à Internet,
- ▮ le nombre de sources d'information qu'ils ont consultées par jour.

Les résultats de ces analyses de corrélation sont présentés dans les tableaux suivants.

Corrélations entre la proportion de temps de connexion à Internet consacrée à s'informer et les caractéristiques des participants suivantes :

Caractéristiques	R	Caractéristiques	R
Femmes	- 0,10	Temps quotidien de connexion à Internet	-0,01
Âge	0,20	Nombre de sources d'information consultées par jour	0,35
18-24 ans	- 0,07	Intérêt déclaré pour l'actualité générale	0,13
25-34 ans	- 0,10	Intérêt déclaré pour l'actualité sociale et politique	0,13
35-49 ans	- 0,05	Fréquence d'information déclarée <i>via</i> la télévision	0,02
50-64 ans	0,10	Fréquence d'information déclarée <i>via</i> la radio	0,05
65 ans et plus	0,11	Fréquence d'information déclarée <i>via</i> la presse papier	0,08
CSP supérieure	- 0,02	Fréquence d'information déclarée <i>via</i> Internet	0,15
CSP intermédiaire	- 0,02	Confiance déclarée dans les institutions	0,04
CSP inférieure	- 0,06	Confiance déclarée dans les médias	0,03
Inactifs (tous)	0,09	Confiance déclarée dans le Gouvernement	0,04
Élèves / Étudiants	- 0,07	Confiance déclarée dans la communauté scientifique	0,04
Chômeurs	0,00	Positionnement politique	- 0,03
Retraités	0,15	Extrémisme politique (extrême gauche et extrême droite)	- 0,02
Niveau de diplôme	0,02	Attitude déclarée à l'égard du mouvement des Gilets Jaunes	- 0,05
Niveau de revenu	0,09	Se dit membre du mouvement des Gilets Jaunes	- 0,03
Niveau de revenu du foyer	0,07	Niveau d'adhésion à des propositions complotistes	- 0,01
Vit seul	0,04		
Nombre d'enfants foyer	- 0,06		

TAB. 2 — LECTURE : L'âge est positivement corrélé avec la proportion de temps de connexion à Internet consacrée à s'informer ($R = 0,20$).

NOTE : Les coefficients de corrélation (R) en gris sont non significatifs ($p \geq 0,05$). En rouge : $R \geq 0,10$. Analyses conduites sur les 1 614 participants ayant répondu au questionnaire de fin d'étude.

Il ressort de ces analyses qu'un certain nombre de caractéristiques individuelles et sociales sont associées au fait de s'informer plus ou moins intensément sur Internet. La plus importante d'entre elles semble être l'âge des participants, les 50 ans et plus, dont les retraités, s'informent davantage que les autres sur Internet. Cependant, les liens observés entre les caractéristiques sociodémographiques des participants et l'usage qu'ils font d'Internet pour s'informer sont tous relativement faibles – ces caractéristiques ne peuvent dès lors pas expliquer, à elles seules, la forte variance interindividuelle des comportements d'information en ligne que nous avons constatée.

3.4 — Comparaison entre consommation d'informations déclarée et effective

Au terme des 30 jours de notre étude, nous avons adressé un questionnaire aux participants les interrogeant notamment sur leur rapport à l'information, sur les canaux qu'ils utilisent pour s'informer ainsi que sur leur comportement d'information en ligne (fréquence et type de sources consultées, entre autres). Sur les 2 372 personnes suivies dans cette étude, 1 614 d'entre elles ont complété et nous ont retourné ce questionnaire. Les analyses suivantes portent donc sur ces 1 614 participants.

Par quel canal de diffusion ces participants disent-ils s'informer principalement ? La figure suivante montre que le canal d'information préféré des participants demeure la télévision, 72 % d'entre eux affirmant la regarder « plus ou moins tous les jours » pour s'informer. Internet arrive en seconde position, devant la radio et la presse papier, avec **56 % des participants qui déclarent se connecter « plus ou moins tous les jours » pour consulter des informations**. 23 % disent s'informer sur Internet « au moins une fois par semaine », 11 % « au moins une fois par mois » et 10 % « jamais ».

Fréquence de la consommation d'informations déclarée par les participants en fonction du canal de diffusion

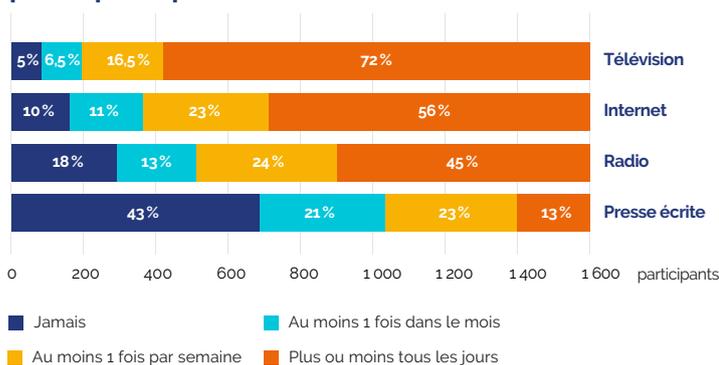


FIG. 17 — LECTURE : 56 % des participants déclarent utiliser Internet « plus ou moins tous les jours » pour s'informer, 23 % « au moins une fois par semaine », 11 % « au moins une fois par mois » et 10 % « jamais ».

NOTE : N = 1 614 participants.

La représentation que les participants se font de la fréquence à laquelle ils recourent à Internet pour s'informer correspond-elle à leur comportement effectif, tel que nous l'avons mesuré dans cette étude ? Afin de le déterminer, nous avons calculé pour chacun des 1 614 participants le ratio du nombre de jours au cours desquels il a consulté au moins une source d'information sur son nombre de jours de connexion à Internet durant la période de l'étude. Nous avons alors pu comparer ce ratio individuel à la fréquence à laquelle chaque participant déclare s'informer sur Internet. Pour le faire, nous avons conduit une analyse de corrélation sur l'ensemble des

1 614 participants entre ratios de consultation effective d'informations sur Internet, d'un côté, et fréquences déclarées de consommation d'informations en ligne, de l'autre.

Comme le montre la figure suivante, il existe une corrélation positive et statistiquement significative entre ces deux variables ($R = 0,17$, $p < 0,001$). Il faut pourtant souligner que cette corrélation est faible. Cela indique que de nombreux participants surestiment la fréquence à laquelle ils s'informent sur Internet, tandis que d'autres la sous-estiment. En d'autres termes, dans l'ensemble, **les participants semblent avoir une perception relativement floue de la fréquence réelle à laquelle ils recourent à Internet pour consulter de l'information**. À noter, cependant, que comme nous l'avons indiqué dans l'Encadré 3 de ce rapport, nous n'avons pas pu tenir compte dans la présente étude de la consultation d'informations (notamment sous la forme de courtes vidéos ou de titres d'articles) directement sur le « mur » ou le « fil » des réseaux sociaux de chaque participant. Il est donc possible que la corrélation entre fréquence déclarée et fréquence effective d'information sur Internet soit en réalité un peu moins faible que celle que nous observons ici.

Corrélation entre fréquence déclarée et fréquence effective d'information sur Internet

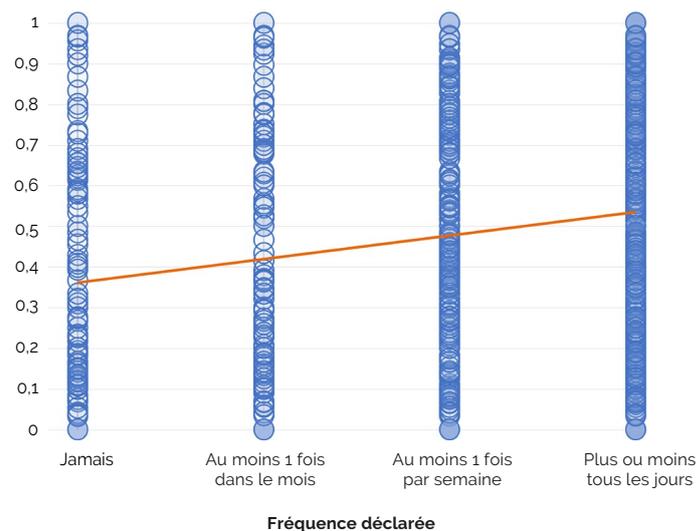


FIG. 18 — LECTURE : Un petit disque bleu positionné sur « jamais » sur l'axe horizontal et sur 0 sur l'axe vertical correspond à un participant ayant déclaré ne jamais s'être informé sur Internet au cours des 30 jours précédant la passation du questionnaire et qui n'a effectivement consulté aucune source d'information en ligne durant cette période. A contrario, un disque positionné sur « plus ou moins tous les jours » sur l'axe horizontal et sur 1 sur l'axe vertical correspond à un participant ayant déclaré s'être informé plus ou moins tous les jours sur Internet au cours des 30 jours précédant la passation du questionnaire et qui a effectivement visité au moins une source d'information chaque jour où il s'est connecté à Internet durant cette période. Lorsque plusieurs disques sont superposés, leur surface apparaît plus foncée. La droite rouge indique le rapport global entre fréquences déclarées et effectives d'information sur Internet pour l'ensemble des participants ($R = 0,17$, $p < 0,001$, $N = 1 614$).

Fréquence déclarée : réponse à la question « Au cours des 30 derniers jours, à quelle fréquence vous êtes-vous informé sur Internet, y compris sur les sites web des TV, radios et de la presse papier ? »

Fréquence de consultation effective : de 0 = aucune source d'information consultée durant les jours de connexion à 1 = au moins une source d'information consultée chaque jour de connexion.



Afin d'aller plus avant dans ces analyses, nous avons demandé aux participants de nous indiquer lesquelles parmi 15 sources d'information que nous leur avons présentées ils consultaient sur Internet et à quelle fréquence ils le faisaient. Les 15 sources présentées aux participants correspondent, pour 12 d'entre elles, aux sources d'information apparaissant dans l'étude Reuters de 2020 comme étant celles que les Français disent le plus consulter sur Internet et, pour 3 d'entre elles, aux sources d'information qui n'étaient pas déjà présentes parmi les précédentes et qui étaient le plus consultées par les membres du panel suivi dans notre étude au cours du mois précédant la conduite de l'étude.

La figure suivante permet de comparer visuellement le classement de consultation de ces 15 sources d'information selon les déclarations des participants, d'un côté, et selon leur fréquentation effective au cours de l'étude, de l'autre. Chacune de ces sources d'information a été considérée comme ayant été effectivement consultée par un participant si ce dernier s'y est connecté au moins une fois au cours des 30 jours de l'étude²⁵.

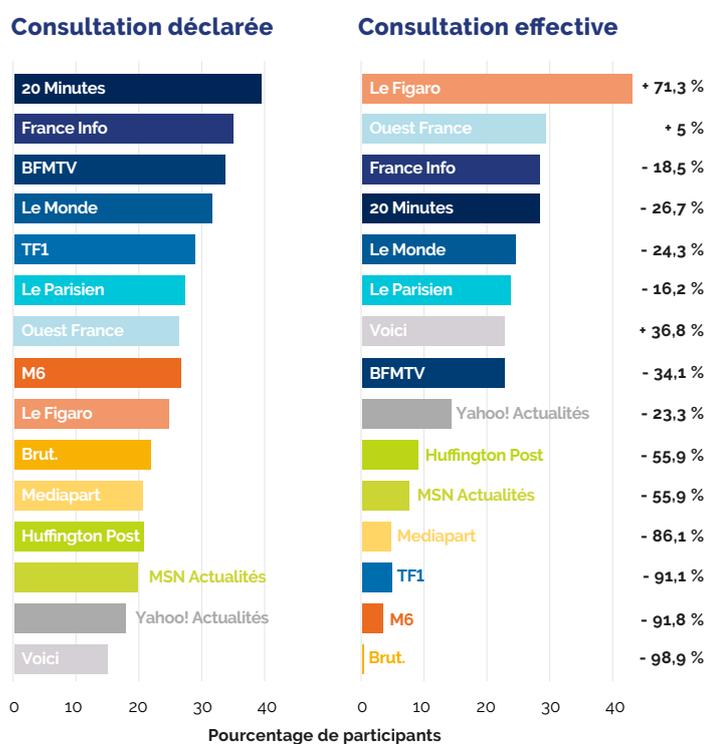


FIG. 19 – LECTURE : Figure de gauche : 39 % des participants ont déclaré avoir consulté 20 Minutes sur Internet pour s'informer au cours des 30 jours précédant la passation du questionnaire. Figure de droite : 28 % des participants ont effectivement consulté au moins une fois 20 Minutes sur Internet durant cette période, soit une diminution de 26,7 %.

NOTE : N = 1 614 participants. La consultation effective de Brut est sous-estimée dans notre étude. En effet, Brut propose de courtes vidéos prévues pour être visionnées directement sur le « mur » ou le « fil » personnel des utilisateurs des réseaux sociaux (contrairement aux articles « partagés » sur les réseaux sociaux dont la lecture se fait, lorsque l'utilisateur clique dessus, sur le site Internet du média qui les a produits). Or, dans notre étude, nous n'avons pas pu tenir compte de la consultation d'informations directement sur le « mur » ou le « fil » des réseaux sociaux de chaque participant (voir Encadré 3 de ce rapport).

25. À noter que pour TF1 et M6, le direct et le replay ont été pris en compte. Les sites web de TF1 et de M6 proposant une classification de leurs programmes, nous avons retenu dans cette étude ceux catégorisés « Info, mag et sport » sur MYTF1.fr et « Info et Société » et « Sport » sur 6play.fr. Les programmes diffusant du sport et non des informations sportives ont été écartés des données.

On observe d'importantes variations pour certaines sources d'information entre consultation déclarée et consultation effective. Si le cas de Brut s'explique en partie au moins par le fait que la méthodologie suivie dans cette étude en sous-estime la consultation effective (voir la note de la figure précédente), il n'en va pas de même pour les sources en ligne de TF1 et M6. La consultation de ces sources en ligne est très fortement surdéclarée par les participants. Cela provient probablement du fait qu'un certain nombre d'entre eux s'informent sur ces chaînes *via* la télévision et non sur Internet. Il est dès lors possible que ces participants aient déclaré s'informer sur les sources en ligne de ces chaînes, quand bien même ils le font en réalité *via* leur poste de télévision. Notons encore que les trois seules sources d'information dont la consultation a été sous-déclarée sont *Ouest France*, *Voici* et *Le Figaro*.

Nous avons ensuite procédé à une analyse de corrélation sur le même modèle que la précédente : nous avons calculé pour chacun des 1 614 participants le ratio du nombre de jours au cours desquels il a consulté chacune des 15 sources d'information sur son nombre de jours de connexion à Internet durant la période de l'étude. Nous avons alors pu comparer ce ratio individuel à la fréquence à laquelle chaque participant déclare consulter chacune de ces 15 sources.

Corrélation entre fréquence déclarée et fréquence effective de consultation des 15 sources d'information

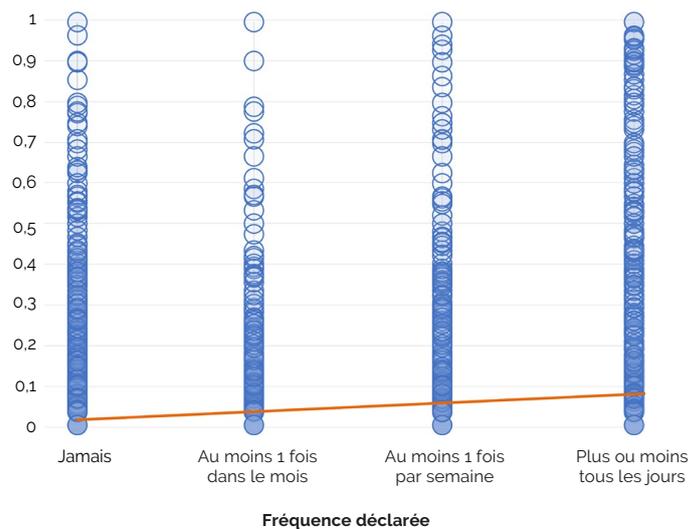


FIG. 20 – LECTURE : Un petit disque bleu positionné sur « jamais » sur l'axe horizontal et sur 0 sur l'axe vertical correspond à un participant ayant déclaré ne jamais s'être informé sur l'une des 15 sources d'information en ligne au cours des 30 jours précédant la passation du questionnaire et qui n'a effectivement jamais consulté la source en question au cours de cette période. Lorsque plusieurs disques sont superposés, leur surface apparaît plus foncée. Les participants ayant évalué leur fréquentation en ligne de 15 sources d'information différentes, la figure comporte 15 disques pour chaque participant. La droite rouge indique le rapport global entre fréquences déclarées et effectives d'information sur les 15 sources pour l'ensemble des participants ($R = 0,22$, $p < 0,001$, $N = 24\ 210$ réponses, 1 614 participants). Fréquence déclarée pour chacune des 15 sources d'information : réponses à la question « Au cours des 30 derniers jours, à quelle fréquence vous êtes-vous informé sur les sources d'information suivantes sur Internet : [liste des 15 sources] ? »

Fréquence de consultation effective : de 0 = source jamais consultée durant les jours de connexion à 1 = source consultée chaque jour de connexion.

Comme le montre la figure précédente, il existe une corrélation positive et statistiquement significative entre les deux variables ($R = 0,22$, $p < 0,001$). Là encore, il faut relever que cette corrélation est plutôt faible, ce qui confirme que les participants n'ont pas une perception très claire de leur comportement de consommation d'informations sur Internet.

3.5 — Rapport entre la confiance en des sources d'information et leur usage

La fréquence à laquelle les participants consultent une source d'information donnée est-elle liée au niveau de confiance qu'ils disent avoir en cette source ? Pour le savoir, nous avons mesuré au moyen du questionnaire de fin d'étude le niveau de confiance des participants à l'égard des 15 sources d'information pour lesquelles nous leur avons demandé de nous indiquer la fréquence à laquelle ils les consultaient (voir section précédente pour la modalité de sélection de ces sources d'information).

Afin d'établir cette mesure, nous avons invité les participants à indiquer leur niveau de confiance à l'égard de chacune de ces sources sur une échelle allant de 1 = « pas du tout confiance » à 4 = « tout à fait confiance ». Les participants avaient également la possibilité de répondre pour chacune des sources qu'ils étaient « sans avis ». La figure suivante expose les résultats moyens de cette mesure.

Niveau de confiance déclarée dans 15 sources d'information

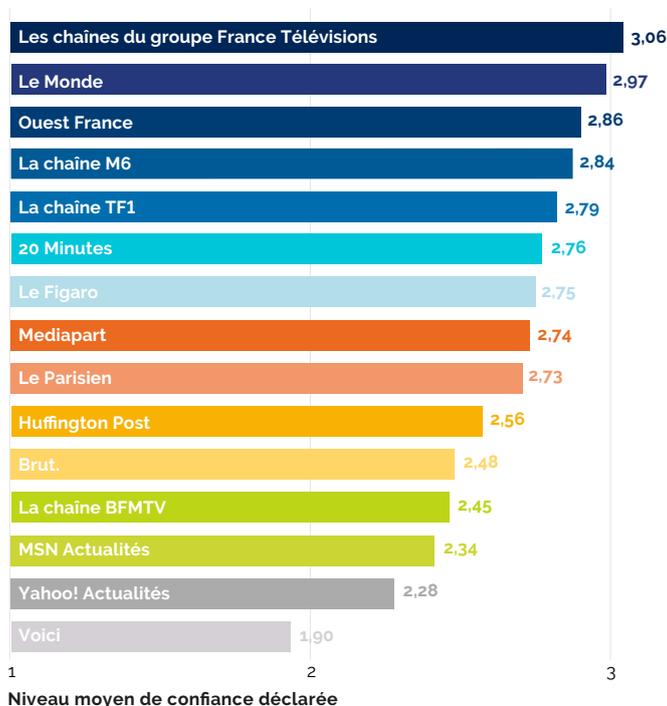


FIG. 21 — LECTURE : Les chaînes du groupe France Télévisions ont reçu la meilleure évaluation de confiance de la part des participants, avec une note moyenne de 3,06 (3 = « plutôt confiance »). Échelle de confiance : de 1 = « pas du tout confiance » à 4 = « tout à fait confiance ».

NOTE : N = 17 931 réponses, 1 614 participants.

Dans son étude de 2020, Reuters a également sondé le niveau de confiance de Français envers un certain nombre de sources d'information. On retrouve 10 sources d'information communes parmi celles évaluées par l'étude de Reuters et la nôtre. Il est intéressant de constater que, comme le montre la figure suivante, il existe une corrélation très étroite entre les notes moyennes de confiance attribuées à chacune de ces 10 sources par les participants à l'étude Reuters et celles attribuées par les participants à notre étude.

Confiance déclarée envers 10 sources d'information : comparaison des résultats de Reuters 2020 et de la Fondation Descartes

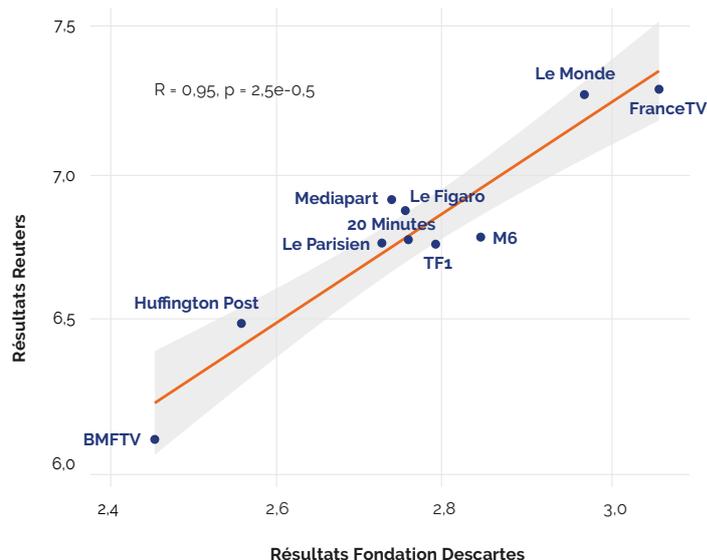


FIG. 22 — LECTURE : Corrélation entre les notes moyennes de confiance attribuées à chacune des 10 sources communes par les participants à l'étude Reuters 2020 et celles attribuées par les participants à notre étude. Échelle de confiance utilisée dans l'étude Reuters : 0 = « pas du tout digne de confiance » à 10 = « totalement digne de confiance ». Échelle de confiance utilisée dans notre étude : de 1 = « pas du tout confiance » à 4 = « tout à fait confiance ». La droite rouge représente la corrélation entre les deux variables ($R = 0,95$, $p < 0,001$, $N = 10$).

Quelle relation existe-t-il entre le niveau de confiance déclarée par les participants envers chacune des 15 sources d'information évaluées dans la présente étude et la fréquence à laquelle ils déclarent les consulter en ligne. Pour le déterminer, nous avons conduit une analyse de corrélation entre ces deux variables déclaratives. Comme le montre la figure suivante, il existe une corrélation positive et statistiquement significative, mais modérée ($R = 0,35$), entre ces deux variables.



Corrélation entre confiance et fréquence de consultation déclarée

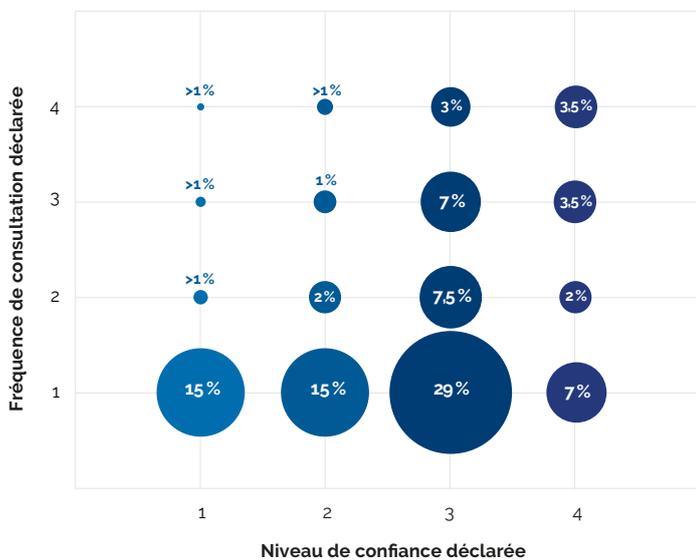


FIG. 23 — LECTURE : Dans 29 % de leurs réponses au questionnaire de fin d'étude, les participants ont déclaré avoir « plutôt confiance » (3 sur l'axe horizontal) à l'égard d'une source d'information qu'ils avaient par ailleurs indiqué ne jamais avoir consultée au cours des 30 jours précédant la passation du questionnaire (1 sur l'axe vertical). Niveau de confiance déclarée : de 1 = « pas du tout confiance » à 4 = « tout à fait confiance ». Fréquence déclarée : de 1 = « jamais » à 4 = « plus ou moins tous les jours ». Corrélation globale entre le niveau de confiance déclarée par les participants envers chaque source d'information et la fréquence à laquelle ils déclarent consulter ces sources : $R = 0,35$, $p < 0,001$, $N = 17\,931$ réponses (1 614 participants interrogés).

Cette corrélation positive entre confiance et fréquence de consultation en ligne d'une source d'information s'amenuise drastiquement lorsque l'on s'intéresse non plus à la fréquence de consultation déclarée, mais à la fréquence de consultation effective. Comme le montre la figure suivante, le coefficient de corrélation passe en effet de $R = 0,35$ à $R = 0,07$ ($p < 0,001$).

Corrélation entre confiance et fréquence de consultation effective

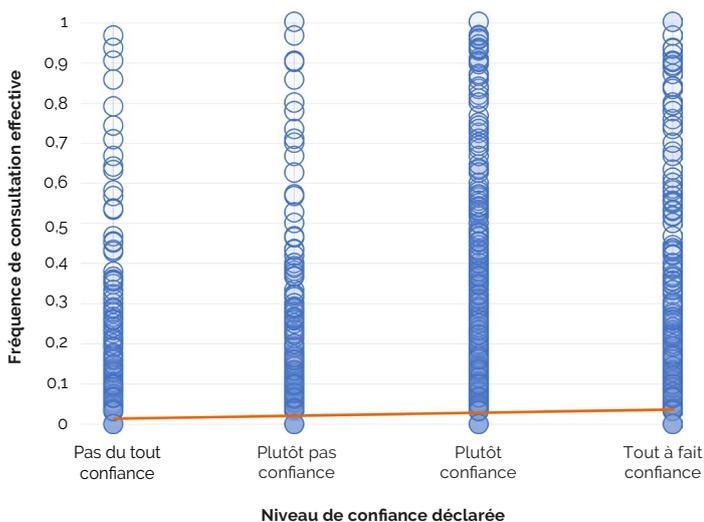


FIG. 24 — LECTURE : Un petit disque bleu positionné sur « pas du tout confiance » sur l'axe horizontal et sur 0 sur l'axe vertical correspond à un participant ayant déclaré ne pas faire du tout confiance à l'une des sources

d'information en ligne testées et qui n'a jamais consulté la source en question au cours des 30 jours précédant la passation du questionnaire. Lorsque plusieurs disques sont superposés, leur surface apparaît plus foncée. La droite rouge indique le rapport global entre confiance et fréquence de consultation effective ($R = 0,07$, $p < 0,001$, $N = 17\,931$ réponses, 1 614 participants).

Niveau de confiance déclarée : réponses à la question « À quel point avez-vous confiance dans chacune des sources d'information suivantes : [liste des 15 sources] ? »

Fréquence de consultation effective : de 0 = source jamais consultée durant les jours de connexion à 1 = source consultée chaque jour de connexion.

Comme le montre la figure suivante, si l'on compare la confiance attribuée à chacune de 15 sources d'information non plus à la fréquence de consultation effective de ces sources sur Internet, mais à la proportion du temps total d'information en ligne passé sur chacune d'entre elles, le coefficient de corrélation diminue encore pour tomber à $R = 0,05$ ($p < 0,001$).

Corrélation entre confiance en une source d'information et part du temps passé sur cette source



FIG. 25 — LECTURE : Un petit disque bleu positionné sur « pas du tout confiance » sur l'axe horizontal et sur 0 sur l'axe vertical correspond à un participant ayant déclaré ne pas faire du tout confiance à l'une des sources d'information en ligne testées et qui n'a jamais consulté la source en question au cours des 30 jours précédant la passation du questionnaire. Lorsque plusieurs disques sont superposés, leur surface apparaît plus foncée. La droite rouge indique le rapport global entre confiance et part du temps total de consultation ($R = 0,05$, $p < 0,001$, $N = 17\,931$ réponses, 1 614 participants). Niveau de confiance déclarée : réponses à la question « À quel point avez-vous confiance dans chacune des sources d'information suivantes : [liste des sources que les participants disent avoir fréquentées] ? » Part du temps total d'information passé sur la source : de 0 = source jamais consultée durant les jours de connexion à 1 = intégralité du temps d'information passé sur cette source.

On le voit, le lien entre la confiance ressentie à l'égard d'une source d'information et le fait de s'informer en ligne plus ou moins régulièrement ou longtemps sur cette source semble au mieux ténu.

3.6 — Consommation de désinformation

Comme indiqué dans la section Méthode de ce rapport, nous avons catégorisé comme non fiables certaines sources d'information comprises dans la base de données de notre étude. Cette catégorisation reprend celle proposée par l'algorithme de Storyzy, dont nous avons vu qu'il donne des résultats cohérents avec la classification manuelle effectuée par les journalistes du Décodex.

Après avoir visité une à une les sources d'information catégorisées comme non fiables, nous avons précisé la thématique principale de chacune d'entre elles. Nous sommes ainsi arrivés à la classification suivante pour les sources d'information jugées non fiables : désinformation généraliste, désinformation en santé, pièges à clics, pseudosciences et satire (voir section Méthode).

Sur l'ensemble des participants (N = 2 372), le temps passé sur des sources d'information jugées non fiables représente 5 % du temps total d'information en ligne, soit 0,16 % du temps total de connexion à Internet. Tous les participants n'ont cependant pas consulté de sources d'information non fiables au cours des 30 jours de l'étude : ils sont 39 % à l'avoir fait. Ces participants ont visité de telles sources durant 40 secondes par jour de connexion en moyenne (médiane = 7 secondes), ce qui représente, toujours en moyenne, 10,9 % de leur temps quotidien d'information sur Internet (médiane = 3,7 %), soit 0,4 % de leur temps total de connexion.

La figure suivante illustre la répartition du temps passé par ces participants sur les différents types de sources d'information non fiables.

Répartition du temps passé sur des sources d'information non fiables

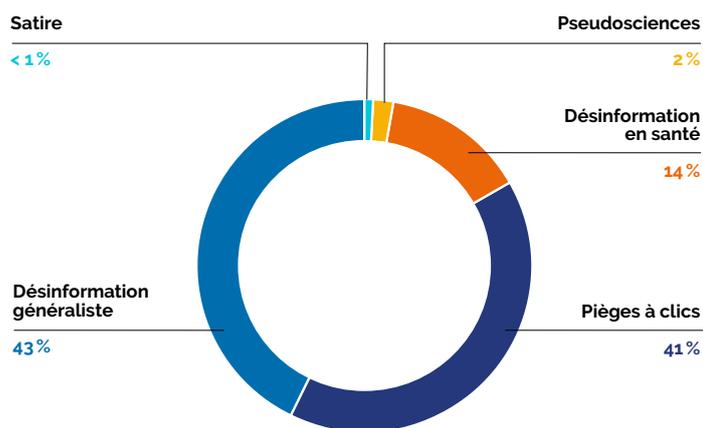


FIG. 26 — LECTURE : 43 % du temps passé sur des sources d'information non fiables l'a été sur des sources de désinformation généralistes.

NOTE : N = 921 participants. Il s'agit des participants ayant consulté des sources d'information non fiables au cours des 30 jours de l'étude.

39 % des participants ont consulté des sources d'information jugées non fiables au cours des 30 jours de l'étude.

3.6.1 — Les facteurs de risque de l'exposition à de la désinformation généraliste sur Internet

Nous avons cherché à déterminer si les participants qui se sont rendus *au moins une fois* durant les 30 jours de l'étude sur une source de désinformation généraliste présentent un profil particulier. Pour ce faire, nous avons procédé à des analyses statistiques portant sur les 1 614 participants qui ont répondu au questionnaire de fin d'étude. Sur ces 1 614 participants, 163 d'entre eux (10,1 %) se sont rendus au moins une fois sur une source de désinformation généraliste au cours des 30 jours de l'étude. Le temps moyen qu'ils y ont passé est de 1 616 secondes (26,93 minutes), le temps médian étant, lui, de 120 secondes (2 minutes). L'écart important entre les valeurs moyenne et médiane indique qu'une partie des participants concernés a passé beaucoup plus de temps que les autres sur ces sources. De fait, si 49 participants y ont passé moins d'une minute en tout au cours des 30 jours de l'étude, les 20 participants qui les ont fréquentées le plus longtemps l'ont fait durant 4 heures en moyenne.

Nos analyses font premièrement apparaître une corrélation positive entre le fait de s'être rendu *au moins une fois* sur une source de désinformation généraliste et :

- ▮ le temps quotidien moyen passé sur Internet ($R = 0,25$, $p < 0,001$);
- ▮ le nombre de sources d'information (toutes catégories confondues) visitées en moyenne par jour de connexion à Internet ($R = 0,38$, $p < 0,001$);
- ▮ le fait de s'être rendu *au moins une fois* sur un site d'information généraliste *fiable* ou sur un site ou blog d'opinion et société *fiable* ($R = 0,10$, $p = 0,001$);
- ▮ le fait de s'être rendu *au moins une fois* sur une source « piège à clics » ($R = 0,16$, $p < 0,001$, voir Tableau 8), sur une source de désinformation en santé ($R = 0,16$, $p < 0,001$, voir Tableau 8) et sur un réseau social²⁶ (Facebook : $R = 0,17$, $p < 0,001$; Twitter : $R = 0,22$, $p < 0,001$; YouTube : $R = 0,20$, $p < 0,001$).

Les personnes passant le plus de temps sur Internet, celles qui « papillonnent » sur un grand nombre de sources d'information et les utilisateurs des réseaux sociaux semblent ainsi être particulièrement à risque de fréquenter des sources de désinformation généralistes.

²⁶. Sont considérés comme s'étant rendus au moins une fois sur un réseau social donné les participants qui y ont consulté au moins une page hors de celles associées aux sites d'information suivis dans cette étude.

Deuxièmement, nous avons comparé le profil des 163 participants s'étant rendus au moins une fois sur une source de désinformation généraliste avec celui des 1451 participants qui n'en ont pas visitée durant les 30 jours de l'étude. Les critères de comparaison portent sur :

- ▮ le sexe des participants,
- ▮ leur âge,
- ▮ leur niveau de diplôme,
- ▮ leur catégorie socioprofessionnelle,
- ▮ leur niveau de revenu,
- ▮ le niveau de revenu de leur foyer,
- ▮ la composition de leur foyer (le fait de vivre seul ou non et le nombre d'enfants dans le foyer),
- ▮ leur positionnement politique (sur un continuum allant de l'extrême gauche à l'extrême droite),
- ▮ leur attitude à l'égard du mouvement des Gilets Jaunes (sur un continuum allant de l'hostilité déclarée à l'égard de ce mouvement au fait de s'en dire soi-même membre),
- ▮ le fait de déclarer une confiance plus ou moins forte à l'égard des institutions (armée, police, justice, école), à l'égard des médias, à l'égard du Gouvernement et à l'égard de la communauté scientifique,
- ▮ le fait de se déclarer plus ou moins intéressé par l'information en général et par l'information politique et sociale en particulier,
- ▮ le fait de dire s'informer plus ou moins régulièrement *via* la télévision, la radio, la presse papier ou Internet,
- ▮ leur score sur une échelle standardisée de mesure du niveau d'adhésion à des propositions complotistes (Bruder *et al.*, 2013).

Le tableau suivant expose les différences statistiquement significatives entre ces deux groupes de participants qui ressortent de nos analyses.

On le voit, les hommes sont surreprésentés dans le public des sources de désinformation généralistes, et on y retrouve en proportion davantage de personnes qui vivent seules, qui se disent favorables au mouvement des Gilets Jaunes ou membres de ce mouvement. Ces personnes présentent en outre un niveau de confiance moyen particulièrement faible à l'égard des institutions, du gouvernement et des médias, tandis qu'elles se disent en moyenne plus intéressées que les autres participants par les questions d'actualité sociale et politique.

Les personnes qui consultent des sources de désinformation généralistes présentent un niveau de confiance particulièrement faible à l'égard des institutions, du Gouvernement et des médias.

Afin de contrôler si ces caractéristiques sont propres aux participants qui ont consulté des sources de désinformation généralistes et qu'elles ne sont pas, plus généralement, celles des participants qui consultent de l'information généraliste sur des sites ou des blogs, nous avons sélectionné parmi les 1614 participants ayant répondu au questionnaire de fin d'étude tous ceux d'entre eux qui se sont rendus au moins une fois sur une source d'information généraliste, indépendamment du fait qu'elle soit considérée comme fiable ou non, et/ou sur un site ou un blog d'opinion et société, également fiable ou non. En tout, 824 participants remplissent ce critère de sélection.

Nous avons alors séparé ces 824 participants en deux groupes : d'un côté, ceux qui ne sont jamais allés sur une source de désinformation généraliste (N = 661) et, de l'autre, les 163 participants qui ont fréquenté une source de désinformation généraliste au moins une fois sur la durée de

Par rapport aux autres participants (N = 1451), ceux qui se sont rendus au moins une fois sur une source de désinformation généraliste (N = 163)...		Significativité (p-value)	Taille d'effet (Hedges' g)
Sexe	... sont proportionnellement plus nombreux à être des hommes (63,2%, contre 49,1% dans l'autre groupe)	$p < 0,001$	$g = 0,28$
Composition du foyer	... sont proportionnellement plus nombreux à vivre seul (28,2%, contre 18,5% dans l'autre groupe)	$p = 0,003$	$g = 0,25$
	... ont en moyenne moins d'enfants	$p = 0,014$	$g = 0,20$
Attitude envers les Gilets Jaunes	... déclarent être en moyenne davantage favorables au mouvement des Gilets Jaunes	$p < 0,001$	$g = 0,28$
	... se disent en moyenne davantage eux-mêmes membres des Gilets Jaunes (12,9%, contre 5,2% dans l'autre groupe)	$p < 0,001$	$g = 0,32$
Confiance déclarée	... déclarent en moyenne une confiance moindre :		
	• dans les institutions	$p = 0,007$	$g = 0,22$
	• dans le Gouvernement	$p < 0,001$	$g = 0,33$
	• dans les médias	$p = 0,004$	$g = 0,24$
Intérêt actualité déclaré	... se disent être en moyenne davantage intéressés :		
	• par l'actualité en général	$p = 0,008$	$g = 0,22$
	• par l'actualité sociale et politique	$p < 0,001$	$g = 0,28$
Canaux information déclarés	... affirment en moyenne s'informer davantage sur Internet	$p < 0,001$	$g = 0,34$

TAB. 3 — NOTE : T-test bilatéraux. Aucune des autres caractéristiques mesurées ne diffère significativement (à $p < 0,10$) entre ces deux groupes de participants.

l'étude (101 d'entre ces derniers, 62 %, se sont aussi rendus au moins une fois sur un site d'information généraliste fiable ou sur un site ou un blog d'opinion et société fiable). Observe-t-on des différences entre les membres de ces deux groupes ?

Lorsqu'on focalise nos analyses sur ces deux groupes de participants (N = 824 au total), on observe premièrement une corrélation positive entre le fait de s'être rendu au moins une fois sur une source de désinformation généraliste et :

- 1 le temps quotidien moyen passé sur Internet ($R = 0,24, p < 0,001$);
- 1 le nombre de sources d'information (toutes catégories confondues) visitées en moyenne par jour de connexion à Internet ($R = 0,36, p < 0,001$);
- 1 le fait de s'être rendu au moins une fois sur une source « piège à clics » ($R = 0,12, p < 0,001$), sur une source de désinformation en santé ($R = 0,16, p < 0,001$) et sur

un réseau social (Facebook : $R = 0,17, p < 0,001$;
Twitter : $R = 0,21, p < 0,001$; YouTube : $R = 0,22, p < 0,001$).

Deuxièmement, nous avons comparé les caractéristiques des membres de ces deux groupes de participants sur la base des mêmes critères que ceux utilisés précédemment. Le tableau suivant présente les différences statistiquement significatives entre ces deux groupes de participants qui ressortent de nos analyses.

Comme on peut le constater, les résultats de ces analyses sont très similaires à ceux reportés précédemment. Cela indique que les participants qui se sont rendus au moins une fois au cours de l'étude sur une source de désinformation généraliste présentent en moyenne des caractéristiques qui les distinguent non seulement de l'ensemble des participants à l'étude, mais également de ceux qui ont consulté sur Internet des sources d'information de format comparable, mais jugées fiables.

Par rapport aux participants ne s'étant rendus que sur des sites ou des blogs d'information et d'opinion jugés fiables (N = 661), ceux qui ont (aussi) fréquenté au moins une fois une source de désinformation généraliste (N = 163)...

		Significativité (p-value)	Taille d'effet (Hedges' g)
Sexe	... sont proportionnellement plus nombreux à être des hommes	$p = 0,012$	$g = 0,22$
Composition du foyer	... sont proportionnellement plus nombreux à vivre seul	$p = 0,039$	$g = 0,18$
	... ont en moyenne moins d'enfants	$p = 0,014$	$g = 0,22$
Attitude envers les Gilets Jaunes	... déclarent être en moyenne davantage favorables au mouvement des Gilets Jaunes	$p < 0,001$	$g = 0,31$
	... se disent en moyenne davantage eux-mêmes membres des Gilets Jaunes	$p < 0,001$	$g = 0,30$
Confiance déclarée	... déclarent en moyenne une confiance moindre :		
	• dans les institutions	$p = 0,005$	$g = 0,25$
	• dans le Gouvernement	$p < 0,001$	$g = 0,33$
	• dans les médias	$p = 0,001$	$g = 0,29$
Intérêt actualité déclaré	... se disent être en moyenne davantage intéressés :		
	• par l'actualité en général	$p = 0,076$	$g = 0,16$
	• par l'actualité sociale et politique	$p = 0,021$	$g = 0,20$
Canaux information déclarés	... affirment en moyenne s'informer davantage sur Internet	$p = 0,002$	$g = 0,27$
	... affirment en moyenne moins s'informer via la télévision	$p = 0,052$	$g = 0,17$

TAB. 4 — NOTE : T-test bilatéraux. Aucune des autres caractéristiques mesurées ne diffère significativement (à $p < 0,10$) entre ces deux groupes de participants.

3.6.2 — Les facteurs de risque de l'exposition à de la désinformation en santé sur Internet

Nous avons procédé aux mêmes analyses pour chercher à déterminer si les personnes qui se sont rendues au moins une fois durant les 30 jours de l'étude sur une source de désinformation en santé, cette fois, présentent un profil particulier. Sur les 1 614 participants qui ont répondu au

questionnaire de fin d'étude, 298 d'entre eux (18,5 %) se sont rendus au moins une fois sur une source de désinformation en santé. Le temps moyen qu'ils y ont passé est de 339 secondes (5,65 minutes), le temps médian étant, lui, de 102 secondes (1,70 minute).

Nos analyses font premièrement apparaître une corrélation positive entre le fait de s'être rendu *au moins une fois* sur une source de désinformation en santé et :

- ▮ le temps quotidien moyen passé sur Internet ($R = 0,25$, $p < 0,001$);
- ▮ le nombre de sources d'information (toutes catégories confondues) visitées en moyenne par jour de connexion à Internet ($R = 0,31$, $p < 0,001$);
- ▮ le fait de s'être rendu *au moins une fois* sur un site d'information en santé fiable ($R = 0,13$, $p < 0,001$);
- ▮ le fait de s'être rendu *au moins une fois* sur une source « piège à clics » ($R = 0,22$, $p < 0,001$, voir Tableau 8), sur une source de désinformation généraliste ($R = 0,16$, $p < 0,001$, voir Tableau 8) et sur un réseau social (Facebook : $R = 0,21$, $p < 0,001$; Twitter : $R = 0,15$, $p < 0,001$; YouTube : $R = 0,22$, $p < 0,001$).

Là encore, les personnes passant le plus de temps sur Internet, celles qui « papillonnent » sur un grand nombre de sources d'information et les utilisateurs des réseaux sociaux semblent être particulièrement à risque de fréquenter des sources de désinformation en santé.

Deuxièmement, nous avons comparé le profil des 298 participants s'étant rendus au moins une fois sur une source de désinformation en santé avec celui des 1 316 participants qui n'y sont jamais allés, en nous penchant sur les mêmes caractéristiques que celles listées dans la section précédente. Le tableau suivant expose les différences statistiquement significatives entre ces deux groupes de participants qui ressortent de nos analyses.

Par rapport aux autres participants (N = 1 316), ceux qui ont fréquenté au moins une fois une source de désinformation en santé (N = 298)...		Significativité (p-value)	Taille d'effet (Hedges' g)
Sexe	... sont proportionnellement plus nombreux à être des femmes (56 %, contre 48 % dans l'autre groupe)	$p = 0,012$	$g = 0,16$
Âge	... sont en moyenne plus âgés	$p = 0,024$	$g = 0,15$
	... sont proportionnellement plus nombreux à se trouver dans la tranche des 50-64 ans	$p = 0,042$	$g = 0,13$
	... sont proportionnellement moins nombreux à se trouver dans la tranche des 25-34 ans	$p < 0,001$	$g = 0,22$
CSP	... sont proportionnellement plus nombreux à être des inactifs, dont des retraités	$p = 0,051$	$g = 0,13$
Complotisme	... présentent un score moyen plus élevé sur l'échelle de complotisme	$p = 0,021$	$g = 0,15$
Canaux information déclarés	... affirment en moyenne s'informer davantage sur Internet	$p = 0,079$	$g = 0,11$

TAB. 5 — NOTE : T-test bilatéraux. Aucune des autres caractéristiques mesurées ne diffère significativement (à $p < 0,10$) entre ces deux groupes de participants.

Comme on peut le voir, **les femmes sont surreprésentées dans le public des sources de désinformation en santé** et on y retrouve en proportion davantage de personnes âgées et d'inactifs, dont des retraités. Finalement, les participants à notre étude qui ont visité des sources de désinformation en santé présentent un score moyen plus élevé que les autres sur une échelle standardisée d'évaluation du niveau de complotisme.

Afin de contrôler si ces caractéristiques sont propres aux participants qui ont consulté des sources de désinformation en santé et qu'elles ne sont pas, plus généralement, celles des participants qui s'intéressent aux questions de santé, nous avons sélectionné parmi les 1 614 participants ayant répondu au questionnaire de fin d'étude tous ceux d'entre eux qui se sont rendus au moins une fois sur une source d'information en santé, indépendamment du fait qu'elle soit considérée comme fiable ou non. En tout, 671 participants remplissent ce critère de sélection.

Nous avons alors séparé ces 671 participants en deux groupes : d'un côté, ceux qui ne sont jamais allés sur une source de désinformation en santé (N = 373) et, de l'autre, les 298 participants qui ont fréquenté une source de désinformation en santé au moins une fois sur la durée de l'étude (129 d'entre ces derniers, 43,3 %, se sont aussi rendus

au moins une fois sur une source d'information en santé fiable). Observe-t-on des différences entre les membres de ces deux groupes ?

Lorsqu'on focalise nos analyses sur ces deux groupes de participants (N = 671 au total), on observe une corrélation positive entre le fait de s'être rendu au moins une fois sur une source de désinformation en santé et :

- ▮ le temps quotidien moyen passé sur Internet ($R = 0,20$, $p < 0,001$);
- ▮ le nombre de sources d'information (toutes catégories confondues) visitées en moyenne par jour de connexion à Internet ($R = 0,26$, $p < 0,001$);
- ▮ le fait de s'être rendu *au moins une fois* sur une source « piège à clics » ($R = 0,19$, $p < 0,001$), sur une source de désinformation généraliste ($R = 0,17$, $p < 0,001$) et sur un réseau social (Facebook : $R = 0,20$, $p < 0,001$; Twitter : $R = 0,11$, $p = 0,004$; YouTube : $R = 0,22$, $p < 0,001$).

Nous avons ensuite comparé les caractéristiques des membres de ces deux groupes de participants sur la base des mêmes critères que ceux utilisés précédemment. Le tableau suivant montre les différences statistiquement significatives entre ces deux groupes de participants qui ressortent de nos analyses.

Par rapport aux participants ne s'étant rendus que sur des sources d'information en santé jugées fiables (N = 373), ceux qui ont (aussi) fréquenté au moins une fois une source de désinformation en santé (N = 298)...

		Significativité (p-value)	Taille d'effet (Hedges' g)
Sexe	... sont proportionnellement plus nombreux à être des femmes	p = 0,011	g = 0,20
Âge	... sont proportionnellement moins nombreux à se trouver dans la tranche des 25-34 ans	p = 0,010	g = 0,20
CSP	... sont proportionnellement plus nombreux à être des inactifs, dont des retraités	p = 0,069	g = 0,14
Diplôme	... ont en moyenne un niveau de diplôme moins élevé	p = 0,080	g = 0,14
Complotisme	... présentent un score moyen plus élevé sur l'échelle de complotisme	p = 0,076	g = 0,14

TAB. 6 — NOTE : T-test bilatéraux. Aucune des autres caractéristiques mesurées ne diffère significativement (à p < 0.10) entre ces deux groupes de participants.

Les facteurs qui ressortent de ces analyses permettent de préciser les résultats obtenus précédemment. En particulier, il semblerait que, chez les personnes qui consultent des sources en ligne sur les questions de santé, le fait d'être jeune (dans la tranche des 25-34 ans) protège contre le risque de tomber dans le piège des sources de désinformation en santé, tandis que le fait d'être peu diplômé surexpose à ce risque.

3.6.3 — Les facteurs de risque de l'exposition aux « pièges à clics » sur Internet

Pour finir, nous avons procédé à des analyses similaires pour chercher à déterminer si les personnes qui se sont rendues au moins une fois durant les 30 jours de l'étude sur une source « piège à clics » présentent un profil particulier. Sur les 1 614 participants qui ont répondu au questionnaire de fin d'étude, 491 d'entre eux (30,4 %) se sont rendus au moins une fois sur une source « piège à clics ». Le temps moyen qu'ils y ont passé est de 560 secondes (9,33 minutes), le temps médian étant, lui, de 168 secondes (2,80 minutes).

Nos analyses font premièrement apparaître une corrélation positive entre le fait de s'être rendu au moins une fois sur une source « piège à clics » et :

- ▮ le temps quotidien moyen passé sur Internet ($R = 0,31$, $p < 0,001$) ;
- ▮ le nombre de sources d'information (toutes catégories confondues) visitées en moyenne par jour de connexion à Internet ($R = 0,47$, $p < 0,001$) ;
- ▮ le fait de s'être rendu *au moins une fois* sur un site de désinformation généraliste ($R = 0,16$, $p < 0,001$, voir Tableau 8), sur un site de désinformation en santé ($R = 0,22$, $p < 0,001$, voir Tableau 8) et sur un réseau social (Facebook : $R = 0,26$, $p < 0,001$; Twitter : $R = 0,24$, $p < 0,001$; YouTube : $R = 0,22$, $p < 0,001$).

On observe ici aussi que les personnes passant le plus de temps sur Internet, celles qui « papillonnent » sur un grand nombre de sources d'information et les utilisateurs des réseaux sociaux semblent être particulièrement à risque de fréquenter des sources « pièges à clics ».

Deuxièmement, nous avons comparé le profil des 491 participants s'étant rendus au moins une fois sur une source « piège à clics » avec celui des 1 123 participants qui n'y sont jamais allés, en nous penchant sur les mêmes caractéristiques que celles listées dans la section 3.6.1. Le tableau suivant expose les différences statistiquement significatives entre ces deux groupes de participants qui ressortent de nos analyses.

Par rapport aux autres participants (N = 1 123), ceux qui ont fréquenté au moins une fois une source « piège à clics » (N = 491)...

		Significativité (p-value)	Taille d'effet (Hedges' g)
Âge	... sont en moyenne plus âgés	p < 0,001	g = 0,20
	... sont proportionnellement plus nombreux à se trouver dans la tranche des 65 ans et plus	p = 0,036	g = 0,11
	... sont proportionnellement moins nombreux à se trouver dans la tranche des 25-34 ans	p = 0,021	g = 0,12
CSP	... sont proportionnellement plus nombreux à être des inactifs, dont des retraités	p = 0,002	g = 0,17
	... sont proportionnellement plus nombreux à être des retraités	p = 0,008	g = 0,14
	... sont proportionnellement moins nombreux à appartenir à une CSP supérieure	p = 0,026	g = 0,12
	... sont proportionnellement moins nombreux à appartenir à une CSP intermédiaire	p = 0,043	g = 0,11
	... sont proportionnellement moins nombreux à être des étudiants	p = 0,003	g = 0,16
Diplôme	... ont en moyenne un niveau de diplôme moins élevé	p = 0,046	g = 0,11
Revenu ²⁷	... ont en moyenne un niveau de revenu du foyer moins élevé	p = 0,050	g = 0,11
Canaux information déclarés	... affirment en moyenne s'informer davantage sur Internet	p = 0,002	g = 0,17
	... affirment en moyenne plus s'informer <i>via</i> la télévision	p = 0,038	g = 0,11

TAB. 7 — NOTE : T-test bilatéraux. Aucune des autres caractéristiques mesurées ne diffère significativement (à p < 0.10) entre ces deux groupes de participants.

Ces analyses font ressortir le fait que le public des sources « pièges à clics » se caractérise avant tout par un âge moyen plus élevé que celui de l'ensemble des participants. Les 65 ans et plus y sont en particulier surreprésentés, ainsi que les inactifs et les retraités. On observe également que le public des sources « pièges à clics » présente, en moyenne, un niveau de diplôme et de revenu du foyer moins élevé que celui du reste des participants.

Le public des sources « pièges à clics » se caractérise avant tout par un âge moyen plus élevé que celui de l'ensemble des participants.

27. N = 1 495, car les participants ayant répondu au questionnaire (N = 1 614) avaient la possibilité de ne pas indiquer la tranche de revenu du foyer dans laquelle ils se situaient. En tout, 1 495 participants ont donné cette indication, dont 1 047 appartiennent au groupe des participants ne s'étant jamais rendus sur un site « piège à clics » et 448 appartiennent au groupe des participants étant allés au moins une fois sur un tel site durant la période de l'étude.

Le tableau suivant met en évidence le fait qu'une part plus ou moins importante des participants ayant fréquenté au cours de l'étude l'un des trois types de sources d'information jugées non fiables analysées ici ont également fréquenté des sources d'information non fiables appartenant aux deux autres types.

ont consulté au moins...			
	Une source de désinformation généraliste	Une source de désinformation en santé	Une source « piège à clics »
Une source de désinformation généraliste	-	37,4 %	52,1 %
Une source de désinformation en santé	20,5 %	-	52 %
Une source « piège à clics »	17,3 %	31,6 %	-

TAB. 8 — LECTURE : 52,1 % des participants qui ont consulté au moins une fois au cours des 30 jours de l'étude une source de désinformation généraliste ont aussi consulté au moins une fois une source « piège à clics ».

4

Discussion des résultats

Discussion des résultats

4.1 – Consommation d'informations sur Internet

Notre étude met en évidence que la part du temps de connexion des Français consacrée à la consultation d'informations médiatiques est, dans l'ensemble, faible. En moyenne, 3 % seulement du temps total passé en ligne par les participants à notre étude l'a été sur des sources d'information – ce qui, par participant, correspond à un peu moins de 5 minutes par jour de connexion. À titre de comparaison, le journal de 20 heures de TF1, qui réunit chaque soir plusieurs millions de Français devant leur poste de télévision²⁸, dure à lui seul 30 minutes environ et, chez les plus de 15 ans, la durée d'écoute individuelle quotidienne des principaux journaux TV est de 20 minutes²⁹.

Cette faible consommation d'informations en ligne peut paraître surprenante, surtout quand on sait que 56 % des participants à notre étude déclarent consulter « plus ou moins tous les jours » de l'information sur Internet. Cependant, les résultats de notre étude indiquent qu'il faut se montrer prudent à l'égard de la perception que les individus peuvent avoir de leur propre comportement d'information en ligne. En effet, nous avons vu que l'usage que nos participants déclarent faire d'Internet pour s'informer – que ce soit quant à la nature des sources qu'ils disent consulter ou quant à la fréquence avec laquelle ils rapportent le faire – ne correspond que faiblement à leur activité effective d'information en ligne, telle que nous l'avons mesurée.

Nous ne pouvons bien sûr pas totalement exclure la possibilité qu'une partie des participants à notre étude ont consulté de l'information ailleurs que sur les sources que nous avons retenues dans notre base de données. Que cela ait pu nous conduire à sous-évaluer massivement la consommation d'informations de l'ensemble de nos participants est par contre peu probable. En effet, nous avons couvert un très large spectre de sources d'information, dont plus de la moitié n'a d'ailleurs reçu aucune visite des participants au cours des 30 jours de l'étude. De plus, toutes les sources d'information qui apparaissent dans la liste des sites les plus fréquentés publiée par de l'Alliance pour les chiffres de la presse et des médias³⁰ figurent dans notre base de données.

Il faut cependant souligner que nous n'avons pas pu prendre en compte dans notre étude les informations apparaissant directement sur le « fil » ou le « mur » des réseaux sociaux de nos participants. Or, certains médias en ligne, comme Brut, par exemple, se sont spécialisés dans la production

de courtes vidéos pouvant être directement lues sur les réseaux sociaux. Le temps potentiellement consacré par nos participants à visionner de telles vidéos sur leur « mur » Facebook, par exemple, n'a pas pu être comptabilisé dans notre étude.

Cela ne signifie pas pour autant que les résultats de notre étude n'intègrent aucunement les informations partagées sur les réseaux sociaux. En effet, en plus des 651 chaînes YouTube « indépendantes » suivies dans cette étude, nous avons également comptabilisé le temps passé par les participants sur les pages Facebook et Twitter ainsi que sur les chaînes YouTube associées aux sites d'information que nous avons retenus dans notre base de données. Mais surtout, le fait de cliquer sur un article partagé sur les réseaux sociaux conduit l'internaute à être redirigé sur la page du média qui l'a publié. Le temps passé par un participant à le lire sur le site du média en question aura dès lors été pris en compte dans notre étude, pour peu que ce média fasse partie des sources d'information que nous avons suivies.

En réalité, il est probable que les participants à notre étude qui utilisent des réseaux sociaux y ont vu passer les titres de nombreux articles de presse sans avoir pour autant cliqué dessus pour les lire – une étude a d'ailleurs montré que la majorité des articles *partagés* par des internautes sur Twitter le sont sans avoir été préalablement lus (Gabiolkov *et al.*, 2016). Dans ce contexte, la fréquentation des réseaux sociaux peut peut-être donner l'impression à leurs utilisateurs de consulter de l'information, bien que cette dernière se limite alors souvent à des titres ou, au mieux, à des chapeaux d'articles. Cela pourrait contribuer à expliquer pourquoi les participants ont une perception relativement floue des médias en ligne qu'ils consultent effectivement ainsi que de la fréquence réelle à laquelle ils recourent à Internet pour consulter de l'information.

28. www.ozap.com/tag/audiences_t14/1

29. « *Les jeunes et l'information* », synthèse de juillet 2018, Médiamétrie et le Ministère de la Culture. www.culture.gouv.fr/Presse/Communique-de-presse/Les-jeunes-et-l-information-une-etude-du-ministere-de-la-culture-vient-eclairer-les-comportements-des-jeunes-en-matiere-d-acces-a-l-information

30. www.acpm.fr/Les-chiffres/Frequentation-internet/Classement-des-Sites?periode=202010

4.1.1 – S’informer en papillonnant sur le Net

Un autre résultat important de notre étude est que si, en moyenne, les participants s’informent peu sur Internet, les comportements d’information en ligne sont en réalité extrêmement variables entre participants. Par exemple, si 17 % d’entre eux n’ont consulté aucune source d’information sur Internet au cours des 30 jours de l’étude, ils sont 5 % à en avoir consulté durant plus de dix heures au total. Nous sommes parvenus à identifier un certain nombre de caractéristiques individuelles et sociales associées au fait de s’informer plus ou moins intensément sur Internet. La plus importante d’entre elles semble être l’âge des participants, les 50 ans et plus s’informent davantage que les autres sur Internet. Cependant, les caractéristiques des participants corrélées avec le temps passé sur des sources d’information ne le sont toutes que faiblement. Elles ne peuvent donc pas expliquer, à elles seules, la forte variance interindividuelle des comportements d’information en ligne que nous avons constatée.

On observe par contre que les participants qui consacrent beaucoup de temps à s’informer sur Internet le font moins en augmentant le temps total passé sur une même source d’information qu’en diversifiant les sources qu’ils consultent. En effet, plus les participants ont passé de temps à s’informer en ligne durant les 30 jours de l’étude, plus le nombre de sources d’information différentes qu’ils ont consultées est important. Cela traduit un comportement de « papillonnage » chez les consommateurs d’informations sur Internet, qui semblent passer d’une source d’information à l’autre sans souvent la consulter longtemps.

En témoigne le fait que le temps passé consécutivement sur une source d’information à chaque consultation de cette source est en moyenne inférieur à 2 minutes. Même si ce temps moyen varie sensiblement entre les participants, il n’est de 4 minutes ou plus que chez 6 % des participants. Un autre élément va dans le sens d’un tel comportement de « papillonnage » : au total, nos participants ont visité 1290 sources d’information différentes au cours des 30 jours de l’étude – un chiffre qui excède de beaucoup le nombre des grands médias présents sur Internet et qui traduit donc le fait que des sources d’information très secondaires en termes de visibilité sont également consultées par certains participants.

Le fait que les internautes présentent ce comportement de « papillonnage » dans leur consommation d’informations semble aller contre l’idée répandue, mais discutée³¹, selon laquelle Internet enfermerait ses utilisateurs dans des « chambres d’écho » à l’intérieur desquels ils ne seraient exposés qu’à des informations allant dans le sens de leurs attentes et de leurs croyances. De fait, si l’on s’intéresse aux participants à notre étude qui ont consulté les sources Internet de titres nationaux généralement vus comme étant « de centre gauche », comme *Le Monde* ou, plus à gauche, *Mediapart*, on observe bien qu’une part significative d’entre eux a également consulté les sources Internet de titres nationaux vus comme étant « de droite », comme *Le Figaro* ou, plus à droite, *Valeurs actuelles*, et réciproquement.

	N participants (N total : 1614)	% du panel	Temps moyen (sec.)	Temps médian (sec.)	
ont consulté...	<i>Le Monde</i>	389	24 %	468,15	119
	<i>Mediapart</i>	48	3 %	310,02	40
	<i>Le Figaro</i>	694	43 %	516,79	130
	<i>Valeurs actuelles</i>	79	5 %	703,10	127

TAB. 9 – LECTURE : Description du trafic des participants à notre étude sur les sources Internet du *Monde*, de *Mediapart*, du *Figaro* et de *Valeurs actuelles*, au cours des 30 jours de l’étude.

Ce résultat, présenté dans le tableau suivant, est difficilement compatible avec l’idée qu’Internet se réduirait à une juxtaposition de chambres d’écho au sein desquelles les individus ménageraient leur confort cognitif en évitant de se confronter à des visions politiques du monde qui n’iraient pas dans un seul sens, le leur. La question de l’existence de chambres d’écho informationnelles dans lesquelles certains des participants à notre étude pourraient être enfermés mérite cependant d’être creusée plus avant et fera donc l’objet d’analyses complémentaires.

	ont aussi consulté...				
	<i>Le Monde</i>	<i>Mediapart</i>	<i>Le Figaro</i>	<i>Valeurs actuelles</i>	
ont consulté...	<i>Le Monde</i>	-	9,3 %	72 %	13,9 %
	<i>Mediapart</i>	75 %	-	83,3 %	22,9 %
	<i>Le Figaro</i>	40,4 %	5,8 %	-	9,7 %
	<i>Valeurs actuelles</i>	68,4 %	13,9 %	84,8 %	-

TAB. 10 – LECTURE : 72 % des participants qui ont consulté au moins une fois au cours des 30 jours de l’étude la source Internet du *Monde* ont aussi consulté au moins une fois la source Internet du *Figaro*.

4.1.2 – La place des médias traditionnels sur Internet

En nous intéressant aux sources d’information en ligne les plus consultées par les participants à notre étude, il est apparu que les 26 premières d’entre elles totalisent à elles seules 40 % du temps qu’ils ont consacré à s’informer en ligne. Ce « top 26 » est largement composé de sources Internet de médias traditionnels – une exception notable étant Wikipédia, qui se positionne comme la source d’information consultée par le plus grand nombre de participants au cours des 30 jours de l’étude.

Les grands médias traditionnels ont donc réussi à exploiter la visibilité dont ils jouissent hors ligne pour s’imposer comme des acteurs majeurs de l’information sur Internet. Il faut cependant préciser que seuls 10 % des Français déclarent

31. www.fondationdescartes.org/2020/07/bulles-de-filtre-et-chambres-decho/

payer pour accéder à une source d'information en ligne (Reuters, 2020). L'essor d'Internet a donc fragilisé la presse papier, dont les chiffres d'affaires sont en baisse constante depuis 2008³². La radio et la télévision ont été moins affectées par la transition vers le numérique, mais leurs chiffres d'affaires ont eux aussi globalement baissé durant cette période. La transition numérique des médias traditionnels a donc quelque chose de paradoxal : les recettes ont diminué pour la plupart des groupes médiatiques, alors qu'Internet a parfois permis d'augmenter leurs audiences.

Notons que les médias purement numériques présents dans ce « top 26 », outre Wikipédia, appartiennent tous à de grands groupes médiatiques : le *Journal des femmes* et *L'internaute* au groupe Figaro, *01.net* au groupe NextRadioTV (BFMTV, RMC), *Doctissimo* au groupe TF1, *Allociné* à Webedia, *YahooNews* au groupe américain Verizon Media³³.

Un résultat de notre étude qu'il nous faut encore souligner est le fait que nous n'avons observé l'existence que d'un lien extrêmement ténu entre la consommation de sources d'information et le niveau de confiance que les participants disent avoir en ces sources. Si, dans un article régulièrement commenté, Tsfaty et Cappella (2005) ont montré qu'il n'est pas rare que des individus consultent des sources d'information dont ils se méfient, la raison n'en est pas toujours claire. Dans le cas qui nous concerne, il est possible qu'il s'agisse là encore d'une conséquence du comportement de « papillonnage » des participants. Il se peut en effet que ces derniers passent d'une source d'information à une autre, ou qu'ils cliquent sur un titre qui s'affiche sur leur « fil » Twitter ou leur « mur » Facebook, moins pour rechercher une information précise que par divertissement ou curiosité intellectuelle.

4.2 — Facteurs de risque de l'exposition à de la désinformation et aux « pièges à clics »

Il ressort premièrement de nos analyses sur la consommation de désinformation que, sans grande surprise, le risque d'exposition à des sources d'information non fiables est partiellement fonction du temps passé sur Internet. On observe en effet chez les participants à notre étude une corrélation positive entre le temps moyen de connexion quotidien et le fait de se rendre sur des sources de désinformation ou « pièges à clics ». À noter que ceux des participants qui fréquentent de telles sources affirment recourir en moyenne plus que les autres à Internet pour s'informer, ce que leur comportement confirme. Cela pourrait en partie expliquer pourquoi ils passent plus

de temps sur Internet et indique également que choisir Internet comme canal d'information majeur est susceptible d'augmenter le risque de s'exposer à de l'information non fiable.

Nos analyses montrent, deuxièmement, que l'utilisation des réseaux sociaux est elle aussi corrélée à un risque accru d'exposition à de l'information non fiable sur Internet hors desdits réseaux. Les réseaux sociaux ne seraient ainsi pas uniquement des lieux où circule de l'information non fiable, mais également des portes d'entrée vers d'autres sources Internet d'information malveillantes (sources « piège à clics ») ou trompeuses (sources de désinformation généralistes ou en santé).

Troisièmement, nos analyses permettent de penser que les publics surexposés à de l'information non fiable sur Internet ne sont pas les mêmes selon qu'il s'agit de sources de désinformation généraliste, de désinformation en santé ou de « pièges à clics ». Ainsi, les hommes sont surreprésentés dans le public des sources de désinformation généralistes, et on y retrouve en proportion davantage de personnes qui vivent seules, qui se disent favorables au mouvement des Gilets Jaunes ou membres de ce mouvement. Ces personnes présentent en outre un niveau de confiance moyen particulièrement faible à l'égard des institutions, du gouvernement et des médias, tandis qu'elles se disent en moyenne plus intéressées que les autres participants par les questions d'actualité sociale et politique.

Ce sont au contraire les femmes qui sont surreprésentées dans le public des sources de désinformation en santé, et on y retrouve en proportion davantage de personnes inactives, dont des retraités. Les 25-34 ans y sont, eux, sous-représentés. Finalement, les participants à notre étude qui ont visité des sources de désinformation en santé présentent un score moyen plus élevé que les autres sur une échelle standardisée d'évaluation du niveau de complotisme (Bruder *et al.*, 2013). Il faut souligner que ces caractéristiques distinguent les participants ayant consulté des sources de désinformation en santé de l'ensemble des participants à notre étude, mais aussi de ceux qui ont fréquenté des sources d'information en santé considérées comme fiables. Autrement dit, ces caractéristiques ne sont pas celles des personnes qui s'intéressent aux questions de santé dans leur ensemble, mais bien celles des participants qui se sont spécifiquement rendus sur des sources de désinformation en santé.

Le public des sources « pièges à clics », quant à lui, se caractérise avant tout par un âge moyen plus élevé que celui de l'ensemble des participants à notre étude. Les 65 ans et plus y sont en particulier surreprésentés, ainsi que les inactifs et les retraités. On observe également que le public des sources « pièges à clics » présente, en moyenne, un niveau de diplôme et de revenu du foyer moins élevé que celui du reste des participants.

Il importe de souligner que les propriétés des participants qui sont significativement associées d'un point de vue statistique à la fréquentation de sources d'information non

32. Ministère de la Culture, Chiffres clés 2020.

33. Les agrégateurs de contenu se contentent généralement de proposer des articles provenant d'autres médias, sans eux-mêmes produire de contenus.

fiables d'un certain type ne le sont que faiblement. Cela signifie, par exemple, que si le public des sources « piège à clics » est en moyenne plus âgé que l'ensemble des participants, il n'est que légèrement plus âgé (les Hedges'g rapportés dans les sections précédentes, une mesure statistique de la taille d'effet d'une variable, sont tous faibles). Il ne faudrait donc pas interpréter les résultats que nous venons de présenter comme dressant le « portrait-robot » des personnes consultant des sources d'information non fiables d'un type donné sur Internet. Il s'agit plutôt de facteurs de risque associés à l'exposition à des contenus malveillants ou trompeurs.

Ainsi, avoir 65 ou plus est associé à un risque accru de se retrouver pris dans les filets de sources « pièges à clics ». Cela s'explique probablement par le fait que les personnes de cet âge, n'ayant connu Internet que tardivement dans leur vie, ne sont pas aussi au fait de son fonctionnement, de ses codes et de ses pièges que les 25-34 ans qui, eux, ont grandi avec Internet.

De même, les 25-34 ans sont moins à risque de tomber dans le piège de sources de désinformation en santé que les inactifs, dont les retraités. Les femmes sont au contraire plus à risque de le faire, un résultat à mettre en relation avec le fait que ces dernières sont, en France comme ailleurs, davantage consommatrices de médecines alternatives et complémentaires que les hommes (voir par ex., Grimaldi-Bensouda *et al.*, 2012 ; Mayer-Lévy, 2010 ; Thariat, 2012). Les médecines alternatives et complémentaires reposent sur des théories au mieux non prouvées, mais plus souvent trompeuses ou erronées, voire surnaturelles ou paranormales, et elles peuvent conduire à refuser ou à retarder une prise en charge médicale classique (voir par ex., Cordonier, Cafiero & Bronner, in sub.). La composante surnaturelle de certaines médecines alternatives et complémentaires se retrouve sur nombre de sources de désinformation en santé, ce qui pourrait expliquer pourquoi ces dernières attirent un public présentant un niveau de complotisme plus élevé que la moyenne. Plusieurs études ont en effet établi l'existence d'un lien entre sensibilité aux croyances surnaturelles ou paranormales et adhésion à une grande variété de théories du complot (par ex., Bruder *et al.*, 2013 ; Lobato *et al.*, 2014).

Si l'on se penche sur les facteurs de risque associés à la consommation de désinformation généraliste sur Internet, on observe qu'ils sont de nature plus politique, avec un public au sein duquel les soutiens et les membres des Gilets Jaunes sont surreprésentés. Ce public présente également un niveau de confiance dans les institutions, le gouvernement et les médias significativement plus faible que la moyenne des participants, un point commun avec les membres du mouvement de Gilets Jaunes (Algan *et al.*, 2019 ; Wagner-Egger *et al.*, in sub.).

Une défiance marquée à l'égard des institutions et des médias constitue évidemment une motivation importante pour aller rechercher de l'information sur des supports Internet alternatifs aux sites des médias traditionnels publics

et privés, y compris sur des sites ou des blogs jugés non fiables. Ces supports alternatifs de « réinformation », aux contenus plus que douteux et souvent radicaux, relayent des informations susceptibles de nourrir en retour la défiance de leur public envers les institutions et les médias.

Pour finir, il est intéressant de constater que les participants à notre étude qui se sont rendus sur des sources de désinformation généralistes ne représentent que 9,4 % de l'ensemble des participants. De plus, la moitié d'entre eux n'y a passé que très peu de temps au total (la médiane du temps passé sur ces sources au cours des 30 jours de l'étude est de 1,92 minute). Par ailleurs, les participants qui ont fréquenté des sources de désinformation généralistes ne semblent pas majoritairement enfermés dans une chambre d'écho qui ne les exposerait qu'à des sources d'information non fiables. En effet, 68 % d'entre eux, par exemple, ont également consulté au moins une fois au cours de l'étude la source en ligne d'un titre de la presse nationale traditionnelle.

De manière générale, le temps passé par les Français sur des sources Internet d'information jugées non fiables semble relativement faible. En effet, si 39 % des participants à cette étude ont consulté au moins une source d'information jugée non fiable sur 30 jours, ces derniers n'y ont passé en moyenne que 11 % de leur temps quotidien d'information sur Internet, soit 0,4 % de leur temps total de connexion. Il faut cependant souligner que ces valeurs varient fortement entre les individus – une dizaine de participants à notre étude ayant par exemple passé en tout plus de deux heures et demie sur des sources de désinformation généralistes.

Par ailleurs, il est plus que vraisemblable que la désinformation sur Internet diffuse aujourd'hui davantage par le biais des réseaux sociaux que par celui de sources web ayant « pignon sur rue », tels que les sites Internet. On ne compte en effet plus le nombre d'« infox » autour du Covid-19, par exemple, qui ont été massivement diffusées sur Facebook ou Twitter³⁴. Des messages trompeurs de ce type y circulent de proche en proche, en étant partagés par les utilisateurs du réseau social sur leur « mur » ou leur « fil » personnel. Bien souvent, ces fausses informations se réduisent à un court texte ou à une photographie commentée, sans renvoyer à une source Internet externe au réseau social. Cela fait que nous n'avons pas été en mesure dans notre étude de tenir compte de l'exposition des participants à ce type de désinformations.

34. Par exemple : https://www.francetvinfo.fr/sante/maladie/coronavirus/coronavirus-un-faux-message-d-infirmiere-circule-sur-les-reseaux-sociaux_3887911.html



Bibliographie

Bibliographie

Algan, Y., Beasley, E., Cohen, D., Foucault, M., & Péron, M. (2019). *Qui sont les Gilets Jaunes et leurs soutiens ?* Observatoire du Bien-être du CEPREMAP et CEVIPOF, n° 2019-03.

Allcott, H., & Gentzkow, M. (2017). *Social media and fake news in the 2016 election.* *Journal of economic perspectives*, 31(2), 211-36.

Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). *Evaluating the fake news problem at the scale of the information ecosystem.* *Science Advances*, 6(14), eaay3539.

AVAAZ (2020). *Why is YouTube broadcasting climate misinformation to millions?* Rapport accessible via : https://secure.avaaz.org/campaign/en/youtube_climate_misinformation/

Baye, M. R., De los Santos, B., & Wildenbeest, M. R. (2016). *Search engine optimization: what drives organic traffic to retail sites?* *Journal of Economics & Management Strategy*, 25(1), 6-31.

Bruder, M., Haffke, P., Neave, N., Nouripanah, N., & Imhoff, R. (2013). *Measuring individual differences in generic beliefs in conspiracy theories across cultures: Conspiracy Mentality Questionnaire.* *Frontiers in psychology*, 4, 225.

Cordonier, L., Cafiero, F., & Bronner, G. (in sub.). *The side-effects of Complementary and Alternative Medicines. Monitoring the link between CAM and vaccine coverage using Internet queries data.*

Gabelkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016). *Social clicks: What and who gets read on Twitter?* Proceedings of the 2016 ACM SIGMETRICS international conference on measurement and modeling of computer science, 179-192.

Grimaldi-Bensouda, L., Engel, P., Massol, J., Guillemot, D., Avouac, B., Duru, G., ... & Abenhaim, L. (2012). *Who seeks primary care for sleep, anxiety and depressive disorders from physicians prescribing homeopathic and other complementary medicine? Results from the EPI3 population survey.* *BMJ open*, 2(6).

Institut Montaigne (2019). *Media polarization "à la française"? Comparing the French and American ecosystems.* Rapport accessible via : www.institutmontaigne.org/publications/media-polarization-la-francaise

Li, H. O. Y., Bailey, A., Huynh, D., & Chan, J. (2020). *YouTube as a source of information on Covid-19: a pandemic of misinformation?* *BMJ Global Health*, 5(5), e002604.

Lobato, E., Mendoza, J., Sims, V., & Chin, M. (2014). *Examining the relationship between conspiracy theories, paranormal beliefs, and pseudoscience acceptance among a university population.* *Applied Cognitive Psychology*, 28(5), 617-625.

Martens, Be., Aguiar, L., Gomez-Herrera, E. & Mueller-Langer, F. (2018). *The digital transformation of news media and the rise of disinformation and fake news.* Digital Economy Working Paper 2018-02, Joint Research Centre Technical Reports. <https://ssrn.com/abstract=3164170>.

Mayer-Lévy, C. (2010). *Recours aux Médecines Complémentaires et Alternatives parmi les patients de Médecine Générale à Paris.* Th. Méd. : Université Paris 5.

Newman, N., Fletcher, R., Schulz, A., Andi, S., & Nielsen, R. (2020). *Digital news report 2020.* Reuters Institute for the Study of Journalism.

Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A., & Meira Jr, W. (2020, January). *Auditing radicalization pathways on YouTube.* In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 131-141)

Tewksbury, D. (2005). *The seeds of audience fragmentation: Specialization in the use of online news sites.* *Journal of broadcasting & electronic media*, 49(3), 332-348.

Thiriart, F. (2012). *Le recours aux médecines complémentaires et alternatives en milieu rural.* Th. Méd. : Université de Lorraine.

Tsfati, Y., & Cappella, J. N. (2005). *Why do people watch news they do not trust? The need for cognition as a moderator in the association between news media skepticism and exposure.* *Media psychology*, 7(3), 251-271.

Wagner-Egger, P., Adam-Troian, J., Cordonier, L., Cafiero, F., & Bronner, G. (in sub.). *The Yellow Vests in France. Psychosocial determinants and consequences of the adherence to a social movement in a representative sample of the population.*

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making.* Council of Europe report, 27.

Yang, T., Majo-Vazquez, S., Nielsen, R. K., & González-Bailón, S. (2020). *Exposure to News Grows Less Fragmented with Increase in Mobile Access.* Available at SSRN 3564826.

Auteurs de l'étude



Laurent Cordonier, docteur en sciences sociales, est chercheur à la Fondation Descartes, coordinateur de son Conseil scientifique et chercheur associé à l'Université de Paris.

Contact : lc@fondationdescartes.org



Aurélien Brest, titulaire d'un Master en sciences cognitives, est chargé d'études à la Fondation Descartes.

Contact : ab@fondationdescartes.org

La Fondation Descartes

Initiative apartisane, indépendante, citoyenne et européenne, la Fondation Descartes est un institut de recherche pluridisciplinaire basé à Paris et fondé en 2019, dédié aux questions relatives à l'information et au débat public à l'heure d'Internet et des réseaux sociaux. Sa vocation est de contribuer à la recherche sur les questions relatives à l'information et au débat public et de promouvoir l'exigence d'une information sincère pour une démocratie basée sur la confiance.

La gouvernance de la Fondation Descartes est assurée par un Conseil d'administration composé de huit membres et présidé par Jean-Philippe Hecketsweiler.

Le Conseil scientifique de la Fondation Descartes est présidé par Gérald Bronner.

L'équipe de recherche de la Fondation Descartes est dirigée par Laurent Cordonier.

La Fondation Descartes est constituée sous la forme d'un fonds de dotation de droit français.

Elle est financée par des contributions privées.

L'ensemble des publications de la Fondation Descartes est disponible sur le site

www.fondationdescartes.org

Vous pouvez vous abonner à la newsletter mensuelle de la Fondation Descartes *via* ce lien :

www.fondationdescartes.org/inscription-newsletter/

Vous pouvez aussi soutenir notre action *via* ce lien :

www.fondationdescartes.org/soutenir-la-fondation/



FONDATION
DESCARTES 
Information Confiance Démocratie