STATISTIQUE & SOCIETE

Statistique et société

14 | 1 | 2026
Autour de Michel Armatte

Comment estimer la durée d'un questionnaire en ligne dès sa conception ?

A Simple and Reliable Method for Estimating the Duration of an Online Questionnaire at the Design Stage

Pierre Mercklé et Laurine Thizy

OpenEdition Journals

Édition électronique
URL : https://journals.openedition.org/statsoc/5142
DOI : 10.4000/16e3d
ISSN : 2269-0271

Éditeur
Société Française de Statistique (SFdS)

Référence électronique
Pierre Mercklé et Laurine Thizy, « Comment estimer la durée d'un questionnaire en ligne dès sa conception ? », *Statistique et société* [En ligne], 14 | 1 | 2026, mis en ligne le 04 mai 2026, consulté le 30 juin 2026. URL : http://journals.openedition.org/statsoc/5142 ; DOI : https://doi.org/10.4000/16e3d

Ce document a été généré automatiquement le 30 juin 2026.

Le texte seul est utilisable sous licence CC BY-NC 4.0. Les autres éléments (illustrations, fichiers annexes importés) sont susceptibles d'être soumis à des autorisations d'usage spécifiques.

Comment estimer la durée d'un questionnaire en ligne dès sa conception ?

A Simple and Reliable Method for Estimating the Duration of an Online Questionnaire at the Design Stage

Pierre Mercklé et Laurine Thizy

NOTE DE L'ÉDITEUR

Des corrections orthotypographiques ont été apportées à cet article après publication.

Nous remercions, pour leurs conseils et leurs relectures de cet article, Valérie Blineau, Rémy Caveng, Paul Cochet, Emmanuelle Duwez, Frédéric Gonthier, Dominique Joye, Olivier Lê Van Truoc, et Arnaud Régnier-Loilier ; pour nous avoir permis d'utiliser les données de leurs enquêtes, Magali Ballatore, Pierre Bataille, Jean-Baptiste Bertrand, Rémy Blans, Guillaume Collain, Sandrine Lefranc, Alice Pavie, Benoît Tudoux et Olivier Zerbib ; et pour nous avoir facilité l'accès aux enquêtes qu'ils ont gérées, Damien Cartron, Mathieu Olivier et Olivier Martzloff.

1 Concevoir une enquête par questionnaire, quelles que soient ensuite ses modalités d'administration aux répondant·es (en face-à-face, par téléphone, en ligne…), ne va pas sans anticiper le temps probablement nécessaire pour y répondre. Si la pertinence d'annoncer ce temps aux répondant·es est controversée dans la littérature méthodologique[1], encore faut-il, pour pouvoir mener le débat, disposer d'un outil fiable pour l'évaluer : c'est ce qui nous préoccupera dans cet article. Cet enjeu méthodologique de prédire et maîtriser le temps de passation d'un questionnaire est commun à tous les organismes concepteurs d'enquêtes. Dans notre équipe de recherche[2], le questionnement a émergé lors de la préparation de questionnaires en ligne pour une enquête portant sur les transformations des relations sociales depuis la

pandémie de Covid-19, administrée à un panel de plusieurs milliers de personnes, constitué en 2020 et renouvelé en 2024[3]. L'enjeu était double : ne pas dépasser une longueur susceptible de décourager des répondant·es (ancien·nes ou nouvellement recruté·es) et répartir équitablement, au sein du questionnaire, le temps dédié à chaque thématique (et donc à chaque équipe) de l'enquête. Cela supposait d'être en mesure d'évaluer *a priori* le temps de réponse aux questions, en amont de la programmation des questionnaires dans un logiciel et des tests empiriques. Mais comment faire concrètement ?

2 On se serait attendu à ce qu'il existe des pratiques systématiques et bien documentées, en particulier dans les manuels sur le sujet. Or, de façon surprenante, les rares articles francophones qui s'intéressent à l'administration des enquêtes par questionnaire (Peneff, 1988 ; Bessière et Houseaux, 1997 ; Caveng, 2012) n'évoquent jamais cette question. La littérature internationale de même que les pratiques d'enquêtes concrètes des organismes français concevant des enquêtes se révèlent quant à elles très variables. Pour savoir à quelle méthode d'estimation nous fier préférentiellement, nous avons donc décidé d'évaluer empiriquement la fiabilité des diverses méthodes rencontrées, en les testant sur un corpus d'enquêtes en lignes auto-administrées.

3 L'objectif de cet article est de présenter les résultats de cette investigation méthodologique, dans le but de définir une méthode à la fois *simple* (c'est-à-dire facile à mettre en œuvre) et *efficace* (c'est-à-dire la plus prédictive possible) pour estimer le temps de passation d'un questionnaire en ligne avant de le tester empiriquement. L'aperçu de la littérature disponible sur le sujet et des échanges avec plusieurs organismes concepteurs d'enquêtes nous ont d'abord permis de dresser un état des lieux des enjeux méthodologiques liés à l'estimation du temps de passation d'un questionnaire (1) et des différentes méthodes susceptibles d'être utilisées pour cela (2). À partir d'un corpus que nous avons réuni, composé de 41 enquêtes pour lesquelles nous disposons du temps de passation effectif, nous avons ensuite évalué empiriquement la capacité de chaque méthode d'estimation à prédire le temps effectivement mesuré (3). Cela nous permet de proposer, à partir de nos résultats, une méthode d'estimation optimisée, à la fois facile à mettre en œuvre et fiable (4). Ces résultats doivent toutefois être modulés en fonction des caractéristiques sociales des répondant·es, qui font elles aussi varier le temps de réponses, indépendamment de la méthode d'estimation utilisée (5). La conclusion synthétise les apports en proposant dix recommandations pour mieux estimer le temps de passation d'un questionnaire.

1. Pourquoi estimer la durée d'un questionnaire ?

4 Estimer *a priori* la durée d'un questionnaire présente un certain nombre d'intérêts : il peut s'agir d'équilibrer ou de répartir équitablement ses différentes sous-sections, comme dans notre situation ; de préparer un devis si le questionnaire est administrée au téléphone ou en face-à-face et si l'enquête est une commande, par exemple à un institut de sondage ; d'anticiper la quantité de travail pour les personnes chargées d'administrer le questionnaire si cela se fait en face-à-face ou par téléphone, etc. D'une façon générale, l'estimation est nécessaire au moins pour une raison simple : définir une durée maximale, puisqu'un questionnaire (trop) long est coûteux pour les concepteur·ices de l'enquête comme pour les personnes enquêtées, ce qui peut dégrader la qualité et la significativité des données produites.

1.1 Les coûts de la longueur pour les enquêteur·ices

5 Plus un questionnaire est long, plus il suppose un temps de travail conséquent pour concevoir, programmer et tester les questions, pour faire passer le questionnaire, préparer et analyser les données recueillies. Ce coût est très important pour les enquêtes réalisées par téléphone ou en face-à-face avec un·e enquêteur·ice rémunéré·e selon son temps de travail : outre l'établissement du protocole de recherche et la recherche de répondant·es, le coût financier et humain de l'enquête augmente avec sa longueur[4], ou, à budget donné, la taille de l'échantillon diminue puisque moins de personnes pourront être interrogées.

6 Mais ce coût n'est pas uniquement financier : il est également temporel et l'argument vaut pour des enquêteur·ices bénévoles (comme les étudiant·es) ou qui ne sont pas payé·es à la tâche (chercheur·ses en poste). Faire passer un questionnaire suppose de lui allouer du temps qui n'est pas alloué à autre chose. Être en mesure d'estimer le temps consacré à la passation des questionnaires est alors indispensable à l'organisation et à la planification de l'enquête (Loosveldt et Beullens, 2013).

7 Dans le cas d'une enquête auto-administrée (par courrier postal ou en ligne), une partie de ces contraintes est levée. Toutefois, l'essentiel de ce coût se déporte alors sur les personnes enquêtées : ce sont alors elles qui consacrent du temps et de l'énergie à répondre aux questions. La longueur du questionnaire peut alors influer à la fois sur le taux de réponse au questionnaire et sur la qualité des réponses obtenues.

1.2 L'impact de la longueur sur le taux de réponse

8 Dans quelle mesure les longs questionnaires en ligne dissuadent-ils les répondant·es de participer à l'enquête ou provoquent-ils plus fréquemment des abandons ? L'analyse de plusieurs enquêtes d'opinion suggère que la durée idéale d'un questionnaire en ligne se situerait entre 10 et 15 minutes, quand la durée maximale acceptable oscillerait plutôt entre 20 et 28 minutes (Revilla et Ochoa, 2017 ; Revilla et Höhne, 2020)[5]. Dans la littérature disponible, l'effet de la longueur n'est cependant pas univoque. De nombreux travaux de psychologie, d'épidémiologie, de démographie ou de marketing se sont penchés sur les différents déterminants du taux de réponse à un questionnaire (postal ou en ligne). Certains suggèrent qu'un long questionnaire provoque plus fréquemment des non-réponses totales quand il est administré par voie postale (Yammarino, Skinner et Childers, 1991) et des abandons quand il est administré en ligne (Deutskens et al., 2004), alors qu'une partie de la littérature pointe au contraire une absence de lien entre longueur du questionnaire et taux de réponse, dans les enquêtes en face-à-face (Heberlein et Baumgartner, 1978) comme en ligne (Cook, Heath et Thompson, 2000), ou bien un effet limité aux seules enquêtes postales (Sandelin, 2022). D'une façon générale, il semble que le taux de réponse soit peu sensible à la longueur du questionnaire (Lynn, 2014), mais bien davantage aux propriétés sociales des répondant·es comme l'âge (Sandelin, 2022), le niveau de scolarité ou de revenu (Dillman, Sinclair et Clark, 1993), et au dispositif de passation – comme un courrier d'invitation ou une notification préalable, des relances multiples, l'utilisation ou non de rétributions (petite rémunération, bons cadeau), le design du questionnaire… L'enjeu porte donc principalement sur la qualité de l'échantillon : la longueur excessive d'un

questionnaire est susceptible de créer des biais de non-réponse au détriment des plus jeunes et des moins diplômé·es (Sandelin, 2022).

1.3 L'impact de la longueur sur la qualité des réponses

9 Si l'effet de la durée du questionnaire sur le nombre total de réponses n'est pas clair, la littérature disponible est en revanche unanime sur un point : l'accélération de la vitesse de réponse avec l'avancée du questionnaire (Crawford, Couper et Lamias, 2001 ; Helgeson et Ursic, 1994 ; Yan et Tourangeau, 2008), qui va de pair avec une dégradation de la qualité des réponses (Zhang et Conrad, 2014). Quels que soient l'ordre et la nature des questions, les répondant·es font preuve de moins de nuances et de variations fines dans les dernières réponses que dans les premières (Hasselhorn et al., 2024). Outre un effet d'apprentissage, se met en place à mesure de l'avancée dans le questionnaire une logique de *satisficing*, identifiée de longue date par la littérature anglophone (Herzog et Bachman, 1981 ; Krosnick, 1991). Elle consiste à donner non pas une réponse *optimale* mais une réponse simplement *satisfaisante* à une question, en limitant l'effort cognitif. Concrètement, la dégradation de la qualité des réponses se manifeste par un choix plus fréquent des premières modalités ou des modalités « Je ne sais pas » (Krosnick et al., 2002 ; Deutskens et al., 2004), par l'adoption de « styles de réponses » (*response styles*) automatisant la façon de répondre (par exemple, en cochant plus systématiquement les réponses extrêmes ou au contraire les réponses médianes), notamment dans les tableaux de questions (Herzog et Bachman, 1981 ; Galesic et Bosnjak, 2009), et dans certains cas par une absence de réponse à certaines questions (Galesic et Bosnjak, 2009).

10 Un questionnaire trop long entraîne ainsi une dégradation de la qualité des réponses, qui augmente à mesure de l'avancée du questionnaire. Être en mesure d'estimer la durée d'un questionnaire permet ainsi d'anticiper et de limiter ces coûts. Mais sur quoi s'appuyer pour obtenir une estimation fiable de la durée totale ?

2. Comment estimer la durée d'un questionnaire ?

11 Les méthodes d'estimation rencontrées dans la littérature se divisent grossièrement en trois grandes méthodes : le comptage des mots, le comptage des questions et le comptage des réponses en tenant compte du type des questions. Ces méthodes n'engagent pas la même représentation de ce que signifie « répondre à une question ». Le processus cognitif de réponse peut se décomposer en une série de tâches – la lecture de la question, son interprétation, la recherche d'information, le choix de la réponse, son éventuelle modification et sa transmission écrite ou orale (Sudman, Bradburn et Schwarz, 1996) – et chacune de ces tâches a un coût cognitif qui se traduit par un coût temporel. Le temps de latence (*response latency*) avant la réponse (Bassili, 1996) est ainsi tributaire de la « quantité d'information à traiter pour répondre à une question[6] » (Bassili et Scott, 1996) et de la qualité de l'information donnée[7]. Selon que l'on estime le temps de réponse à partir du nombre de mots, du nombre de questions ou de leur type, l'accent n'est pas mis sur la même dimension du processus.

2.1 Compter les signes ou les mots

12 Estimer la durée d'un questionnaire à partir du nombre de mots ou du nombre de signes revient à résumer le processus de réponse à un temps de *lecture* plutôt qu'un temps de *réflexion*. Dans cette perspective, le temps de passation serait d'abord tributaire de la vitesse de lecture[8] des répondant·es. C'est ce qu'avancent prudemment Loewen et al. (2022) : en prenant appui sur une grande enquête internationale en ligne sur la consommation de tabac dont le questionnaire a une durée présentant de fortes variations (entre 25 à 50 minutes), ils montrent que celle-ci dépend d'abord de la vitesse de lecture (et des embranchements du questionnaire, nous y reviendrons), laquelle peut alors être estimée par des tests en fonction des caractéristiques de la population ciblée : âge, niveau de diplôme, familiarité avec la langue du questionnaire, ancienneté de la présence dans le panel.

2.2 Compter les questions

13 Une autre méthode fréquemment rencontrée consiste à compter les questions, en considérant que le temps de réponse varie non pas en fonction du temps de lecture, mais du temps de prise de décision associé à la réponse. Cette méthode a pour elle l'avantage de la simplicité, mais il faudrait alors pouvoir proposer une estimation du temps moyen de réponse à une question. Or les valeurs rencontrées varient fortement selon les sources : 2,5 questions par minute selon Puleston (2012), qui doute lui-même de ce chiffre ; 4 questions par minute selon Loosweldt et Beullens (2013)[9] ; 5 par minute selon le service des enquêtes d'un grand service de la statistique publique française, voire 8 par minute pour des questions simples ; mais 5 par minute seulement en face-à-face, selon un billet de blog de Versta Research (2011), chiffre repris tel quel dans celui de Shyamson (2023).

14 Cette méthode ne fait cependant pas de différence selon la nature, la longueur et la complexité de la question, alors même que le temps de réponse peut varier très fortement en fonction de ces critères : on répondra ainsi beaucoup plus vite à une question à choix unique avec peu de modalités (par exemple, le sexe) qu'à une question ouverte ou une question à choix multiple demandant davantage de réflexion et d'implication dans la réponse.

2.3 Compter les *ticks*

15 Ce constat de la variabilité du temps de réponse selon la nature de la question amène la plupart des institutions productrices d'enquêtes et une partie de la littérature méthodologique à proposer une méthode qui est celle des « *ticks* », terme qu'on pourrait traduire par « cochages » et qui correspondant à un décompte des actions élémentaires que doivent effectuer les enquêté·es pour répondre à un questionnaire. Cette méthode consiste à attribuer des points à une question selon son type (simple, choix multiple, numérique, ouverte…) et le nombre éventuel de modalités de réponses, et à attribuer une valeur temporelle à ce point. Cette approche s'inscrit dans le prolongement de la précédente et la précise : le temps de réponse est toujours envisagé comme un temps de prise de décision, mais il varie selon la complexité de la question.

16 On retrouve toutefois, comme pour le comptage des questions, un même problème de variabilité dans l'estimation du nombre de points ou *ticks* par minute (ou de sa valeur en secondes), à quoi s'ajoutent de petites variations, selon les sources, dans le nombre de points que « pèse » une question selon son type. Le Tableau 1 compare ainsi les propositions méthodologiques trouvées dans Versta Research (2011) et les recommandations adressées (dans un document interne non diffusé) par le dispositif Elipss[10] aux concepteur·ices d'enquête.

Tableau 1 : Comparaison de mode de comptage des *ticks* entre deux dispositifs

	Versta Research	**Elipss**
Question simple à choix unique	1 *tick*	1 *tick*
Questions multiples	0,5 *tick* par modalité	0,5 tick par modalité
Batteries ou grille de questions	1 *tick* par ligne	1 *tick* par ligne
Questions numériques	2 *ticks* si cela nécessite du calcul	2 *ticks*
Questions ouvertes	3 *ticks*	3 *ticks*
Autre	1 *tick* toutes les trois phrases pour les consignes longues	non précisé
Nombre de *ticks* par minute	8 *ticks*/minute	4 *ticks*/minute

17 L'attribution des points y est relativement similaire. Cependant, certain·es producteur·ices d'enquêtes (publics et privés) contacté·es nous ont signalé d'autres façons de distribuer les *ticks* : en comptant par exemple un point toutes les trois questions dans une batterie ou grille de questions, ou un point toutes les 3 modalités (et pas toutes les deux modalités) pour les questions ouvertes. Le nombre de *ticks* par minute (soit sa valeur en seconde) peut quant à lui varier du simple au quadruple : certaines estimations tablent sur 2 ou 3 *ticks* maximum par minute (20 à 30 secondes par *tick*), quand d'autres en comptent plutôt 4 (15 secondes par *tick*) voire 8 pour Verta Research (7,5 secondes par *tick*).

18 Cette variabilité dans la durée d'un *tick* trouve peut-être son origine dans le flou qui entoure la qualification du temps de réponse : aucune source ne spécifie clairement si l'estimation du nombre de *ticks* (ou de questions) par minute correspond à une valeur *moyenne*, *médiane* ou plutôt à une valeur *maximale*, comme par exemple celle du dernier décile. Or cette question a toute son importance puisque l'écart entre la médiane et le dernier décile, comme on va le voir, peut varier très fortement.

2.4 Combiner les méthodes

19 Une dernière méthode, plus laborieuse, consiste enfin à combiner les logiques précédentes. Puleston (2012) propose ainsi une formule d'estimation de la « longueur » des questionnaires qui combine le comptage des mots, le comptage des questions et la prise en compte des types des questions, de la façon suivante : la longueur en minutes est définie comme la somme pondéré du nombre de mots (en comptant un temps de lecture de 5 mots par seconde), du nombre de questions (en comptant 5 secondes par question), du nombre de décisions à prendre (en comptant deux secondes par décision, et 1 décision par question unique, 1 décision toutes les deux modalités des questions à choix multiple, et 1 décision par ligne de tableau de questions) et du nombre de questions ouvertes (en comptant 15 secondes par question ouverte).

20 Dans le même billet, l'auteur propose également plusieurs simplifications de cette méthode, qu'il décrit comme moins fiables mais possiblement satisfaisantes sous certaines conditions, et dont l'une (mesure L2) revient à ne plus tenir compte que du nombre de mots (en comptant toujours 5 mots par seconde) et du nombre de lignes du questionnaires (en comptant 1,8 seconde par ligne, autrement dit par modalité de réponse ou ligne de tableau de questions). Ces méthodes hybrides ont été empiriquement évaluées par Carter et al. (2021), qui ont comparé leurs estimations avec les résultats empiriques d'une enquête sur les risques professionnels en santé, pour en conclure que seule la première méthode, la plus complexe, serait fiable.

*

21 En examinant la littérature disponible et les pratiques concrètes de différents producteurs d'enquêtes, l'existence d'une lacune méthodologique devient frappante : du côté des chercheur·ses et des organismes qui produisent des enquêtes, les méthodes empiriques d'estimation de la durée des questionnaires s'avèrent bricolées selon des pratiques locales assez variables, peu documentées formellement et rarement éprouvées empiriquement. Du côté de la littérature, les articles rencontrés – qui ne sont parfois que des articles de blog – concernent très souvent d'autres domaines que les sciences sociales et n'indiquent que rarement les observations empiriques sur la base desquelles les méthodes d'estimation de la durée des questionnaires ont été établies. À la manière de Carter el al. (2021), nous avons donc cherché à proposer un dispositif empirique pour tester les différentes hypothèses et pratiques rencontrées, à partir d'un corpus plus étendu.

3. Existe-t-il une méthode d'estimation plus fiable qu'une autre ?

3.1 Les données utilisées

22 Nous avons réuni 41 enquêtes en ligne auto-administrées (voir Encadré 1) pour lesquelles nous disposons, à défaut de la durée de chaque question, de la durée totale de la réponse à l'ensemble du questionnaire (depuis l'accès à la première page jusqu'à la validation et l'envoi) et de la distribution des questions selon leur nature.

Encadré 1 : Le corpus des 41 enquêtes chronométrées

Le corpus est constitué de **41 enquêtes administrées en ligne entre 2020 et 2025** dans le cadre de recherches en sciences sociales, rassemblant un total de **98 453 réponses individuelles complètes chronométrées** à des questionnaires de longueur réelle variable (les durées médianes des 41 questionnaires sont comprises entre 7 et 31 minutes).

Les questionnaires ont été administrés sur deux plateformes :

19 enquêtes ont été administrées avec LimeSurvey, un outil libre et gratuit largement utilisé dans les enquêtes de recherche : il s'agit des 6 vagues de l'enquête longitudinale auprès du panel des ANR Vico (Mariot, Mercklé et Perdoncin, 2021) et Panelvico (2020-2025), constitué de répondant·es âgés de 18 ans et plus résidant habituellement en France métropolitaine ; de 4 enquêtes auprès des étudiant·es de l'Université Grenoble Alpes, réalisées entre 2021 et 2024 ; et de 8 enquêtes de recherche réalisées en France auprès de publics variés et partagées par des collègues après un appel sur les listes professionnelles.

Les 22 autres enquêtes du corpus ont été administrées avec Qualtrics, dans le cadre du dispositif d'Etude longitudinale par internet pour les sciences sociales (Elipss) du Centre de données socio-politiques (CNRS, Sciences Po). Ces 22 enquêtes ont été administrées à un panel représentatif de la population âgée de 18 ans et plus résidant habituellement en France.

Pour chacune des enquêtes, nous disposons d'une part d'un fichier au format texte contenant la structure du questionnaire (types et textes des questions et des modalités de réponses), et d'autre part d'une base de données contenant au minimum, pour chaque réponse complète au questionnaire, la durée écoulée entre le moment de l'accès à la première page du questionnaire et le moment du clic sur le bouton « Envoyer » à la fin du questionnaire.

Les 41 enquêtes du corpus sont présentées de façon plus détaillée dans l'annexe 1 et dans l'annexe 2 à la fin de l'article.

3.2 Les méthodes testées

23 L'objectif est d'évaluer et comparer la capacité de différentes méthodes à prédire la durée réellement observée des questionnaires du corpus, à partir de leurs caractéristiques (nombres de signes, nombres et types de questions, etc.). Nous avons testé dans un premier temps huit méthodes de prédiction différentes, établies à partir de la littérature méthodologique présentée dans la partie précédente : deux méthodes consistant à mesurer la longueur du texte du questionnaire (hypothèse du temps de lecture), trois méthodes reposant sur le nombre de questions et de modalités (hypothèse du temps de prise de décision), deux méthodes hybrides combinant longueur du texte, nombre de questions et de modalités (il s'agit des méthodes L1 et L2 proposées par Puleston, 2012), et la méthode des *ticks* avec les pondérations proposées par Elipss (voir Tableau 2).

Tableau 2 : Les huit méthodes de prédiction de la durée du questionnaire testées

Méthode	Mode de calcul
Caractères	Nombre de caractères (signes de ponctuations et espaces comprises) proposés à la lecture des répondant·es, dans le préambule du questionnaire, les consignes, les textes des questions et des modalités de réponses.
Mots	Nombre de chaînes de caractères alphanumériques (hors signes de ponctuation) séparées par des espaces dans le texte du questionnaire.
Questions	Nombre de questions que comporte le questionnaire, quel qu'en soit le type. En particulier, dans cette méthode, les questions regroupées dans un tableau à deux dimensions comptent pour une seule question.
Cases à cocher	Nombre de cases à cocher dans le questionnaire (modalités des questions à choix unique et à choix multiple).
Lignes	Somme du nombre de questions et du nombre de cases à cocher tels que calculés ci-dessus.
L1 de Puleston	Méthode hybride proposée par Puleston (2012), la longueur L1 est calculée de la façon suivante : mots / 5 + questions x 3 + questions à choix unique x 2 + modalités de questions à choix multiple + lignes de tableaux x 2 + questions ouvertes x 15.
L2 de Puleston	Méthode hybride simplifiée proposée par Puleston (2012), la longueur L2 est calculée en secondes de la façon suivante : nombre de mots / 5 + nombre de lignes x 1,8. Appliquée à notre corpus d'enquête, elle surestime toutefois systématiquement la durée avec un coefficient multiplicateur de 2. Nous testons donc plutôt le calcul corrigé suivant : mots / 10 + nombre de lignes x 0,9.
***Ticks* Elipss**	Méthode proposée par le Centre de données sociopolitiques (CDSP) aux producteur·ices d'enquêtes pour le dispositif Elipss. Elle consiste pondérer le nombre de questions en fonction de leur type, en comptant : 1 *tick* pour chaque question à choix unique ; 0,5 *tick* pour chaque modalité des questions à choix multiple ; 2 *ticks* pour chaque question numérique et pour chaque question ouverte à réponse courte ; et 3 *ticks* pour chaque question ouverte à texte long.

3.3 Mesurer les durées réelles des questionnaires

24 Le temps de réponse de chaque répondant·e a été calculé en faisant la différence entre le moment précis de l'envoi du questionnaire et celui de l'accès à sa première page – nous disposions de ces deux horodatages pour chaque enquête du corpus. Les réponses incomplètes et les « *outliers* » (c'est-à-dire les réponses avec des durées extrêmes qui s'éloignent significativement de la distribution normale des durées, dans un sens ou l'autre) ont été retirés. L'hypothèse sous-jacente est qu'il y a d'une part un temps minimum de lecture en-dessous duquel le questionnaire n'a pas été lu mais seulement parcouru rapidement ; et que d'autre part les durées les plus élevées sont probablement dues à des interruptions plus ou moins longues. Nous avons ainsi retiré des analyses les

durées inférieures à la moitié de la médiane et les durées supérieures à trois fois la médiane (soit 12 % de l'échantillon initial total), suivant en cela la règle appliquée par Loosveldt et Beullens (2013).

3.4 Mesurer les corrélations entre durées prédites et durées réelles

25 Les durées « réelles » observées ont ensuite été comparées aux différentes durées « théoriques » prédites à partir des caractéristiques des questionnaires. Pour déterminer quelles sont les méthodes de prédiction de la durée les plus efficaces, nous avons donc cherché celles qui « ajustent » le mieux les durées observées, cet ajustement étant mesuré par les coefficients de corrélation de Pearson (pondérés par la taille des échantillons) entre d'une part les durées prédites par chaque méthode, et d'autre part le premier décile, la médiane, la moyenne et le dernier décile des durées observées dans chaque enquête (voir Tableau 3).

Tableau 3 : Tableau des coefficients de corrélation entre les durées prédites et les durées observées

Méthode	Minimum	Premier décile	Médiane	Moyenne	Dernier décile	Maximum
L2 de Puleston	0,67	0,65	0,7	0,7	0,69	0,68
Questions	0,61	0,62	0,69	0,66	0,65	0,62
Ticks Elipss	0,61	0,61	0,68	0,66	0,65	0,62
L1 de Puleston	0,57	0,56	0,65	0,63	0,63	0,58
Lignes	0,66	0,64	0,64	0,66	0,64	0,67
Mots	0,53	0,51	0,6	0,59	0,59	0,54
Cases à cocher	0,6	0,59	0,59	0,6	0,59	0,62
Caractères	0,39	0,37	0,47	0,46	0,48	0,4

Champ : 41 enquêtes par questionnaire en ligne, 2020-2025.
Lecture : Le coefficient de corrélation de Pearson pondéré entre la durée estimée avec la méthode L2 de Puleston et la durée minimum du questionnaire est de 0,64.
Note : Les lignes du tableau sont classées par ordre décroissant du coefficient de corrélation entre la méthode testée et la durée médiane observée. Les couleurs des cases apportent une aide visuelle pour identifier les coefficients les plus élevés : vert >= 0,70 ; bleu >= 0,65 ; jaune >=0,60 ; orange >= 0,50 ; rouge < 0,50.

26 Il ressort de l'analyse des corrélations entre durées prédites et durées observées que dans l'ensemble, les méthodes reposant seulement sur le calcul de la longueur du texte du questionnaire (et donc sur l'hypothèse du temps de lecture) sont les moins efficaces : les coefficients de corrélation entre le nombre de mots ou le nombre de caractères et les durées observées sont tous inférieurs à 0,60, voire inférieurs à 0,50 dans certains cas. À l'opposé, la méthode simplifiée du L2 de Puleston, la méthode encore plus simple du nombre de questions et la méthode du décompte des *ticks*, avec les coefficients proposés par le dispositif Elipss, sont celles qui donnent les meilleurs résultats.

27 Les méthodes les plus prédictives de la durée des questionnaires du corpus suivent des logiques très proches, puisqu'elles consistent toutes les trois à estimer cette durée à partir du nombre de questions : dans la méthode la plus simple, on fait l'hypothèse que chaque question nécessite le même temps de réponse (ou que les questions s'équilibrent en moyenne) ; dans la méthode des *ticks*, on fait l'hypothèse que certaines

questions (les questions à choix multiples, les questions ouvertes…) nécessitent un temps plus long que les questions à choix unique et d'autres (batteries de questions) un peu moins, ce qui ne conduit pas en l'état à une amélioration de l'estimation de la durée.

28 Au demeurant, à notre connaissance, aucune étude empirique n'indique comment déterminer le nombre de *ticks* associé à chaque type de questions – il semble que les choix de coefficients relèvent davantage d'une pratique professionnelle au sein des organismes de recherche et des instituts de sondage que d'un résultat de recherche. Il est pourtant possible de tester empiriquement la robustesse des pondérations proposées selon les questions.

4. Une méthode optimisée est-elle possible ?

29 Pour optimiser l'estimation du temps de passation, nous avons cherché à améliorer les méthodes les plus performantes – celles du L2 de Puleston et celle des *ticks* – en ajustant les coefficients proposés dans ces deux méthodes. Dans un cas comme dans l'autre, la démarche adoptée a consisté à utiliser des régressions linéaires multiples pour estimer de nouveaux coefficients de pondération des différentes caractéristiques prises en compte par ces méthodes, qui permettraient de mieux ajuster les durées réellement observées dans le corpus d'enquêtes utilisé.

4.1 Des régressions linéaires multiples pour mieux ajuster les méthodes les plus performantes

30 Pour améliorer la méthode L2 de Puleston, nous avons ajusté les coefficients du nombre de mots et du nombre de questions à partir de deux régressions linéaires multiples à constante nulle sur la durée médiane et la durée du neuvième décile[11] telles qu'observées dans le corpus. D'après ces deux régressions, dont les résultats sont presque parfaitement convergents, une meilleure méthode peut consister à estimer la durée en secondes non pas à partir de *nombre de mots / 10 + nombre de lignes x 0,9*, mais plutôt à partir de *nombre de mots / 13 + nombre de lignes*[12].

31 Pour améliorer la méthode des *ticks*, nous avons procédé exactement de la même façon, avec cette fois en paramètres les nombres de questions à choix unique, de lignes dans les tableaux de questions, de modalités de questions à choix multiple, de questions numériques et de questions ouvertes. Dans le cas des *ticks*, les durées observées sont normalisées pour fixer la valeur du premier paramètre (nombre de *ticks* des questions à choix unique) à 1 par construction et en faire ainsi l'unité élémentaire du modèle. Les deux modèles (médiane et neuvième décile) donnent des estimations presque parfaitement identiques (Tableau 4) : en raison de la très petite taille de l'échantillon (N = 41), seuls les deux premiers paramètres sont significatifs et il en ressort donc que si chaque question à choix unique compte pour 1 *tick*, alors le modèle suggère que chaque ligne de tableau de questions devrait compter pour 0,44 *tick* (pour la durée médiane, et 0,42 pour le dernier décile), et non pour 1 *tick* comme dans la règle du dispositif Elipss et dans Versta Research. Les trois autres paramètres du modèle ne sont pas significatifs, et les deux derniers sont même négatifs, alors qu'il est peu vraisemblable

qu'une augmentation du nombre de questions numériques ou de questions ouvertes entraîne une baisse du temps mis pour répondre aux questionnaires.

32 Dans la quatrième colonne du Tableau 4, nous présentons notre propre règle de comptage des *ticks* optimisée. Elle est révisée par rapport à celle d'Elipss et Versta Research, simplifiée par rapport aux deux modèles, et se veut à la fois réaliste (sans valeur négative) et facile à mettre en œuvre (sans calculs complexes). Nous proposons, pour chaque type de question, de retenir comme nombre de *ticks* la valeur arrondie non-nulle la plus proche (par défaut ou par excès) de la borne supérieure de l'intervalle de confiance des coefficients estimés par le modèle de régression. Selon cette façon de procéder, si la question à choix unique est l'unité de compte élémentaire et vaut donc toujours 1 *tick*, alors nous proposons de compter 0,5 *tick* pour chaque ligne des tableaux de questions, 0,10 *tick* pour chaque modalité de question à choix multiples (soit 1 *tick* toutes les 10 modalités), 0,25 *tick* pour chaque question numérique et 0,10 *tick* pour chaque question ouverte[13].

Tableau 4 : Modèle de régression linéaire multiple de la durée des questionnaires

	Elipss (rappel)	**Modèle de la médiane**	**Modèle du dernier décile**	**Mercklé & Thizy**
Question à choix unique	1	1,00 *** (0,69-1,31)	1,00 *** (0,69-1,31)	**1**
Ligne de tableau de questions	1	0,44 *** (0,29-0,59)	0,42 *** (0,28-0,57)	**0,50** (2 lignes = 1 *tick*)
Modalité de question à choix multiples	0,50	0,00 (-0,09, 0,10)	0,00 (-0,00, 0,09)	**0,10** (10 modalités = 1 *tick*)
Question numérique	2	-1,05 (-2,44-0,35)	-1,10 (-2,46-0,26)	**0,25** (4 questions numériques = 1 *tick*)
Question ouverte	3	-0,70 (-1,53-0,13)	-0,85 (-1,61-0,02)	**0,10** (10 questions ouvertes = 1 *tick*)

Champ : 41 enquêtes par questionnaire en ligne, 2020-2025.
Lecture : Le tableau présente en première colonne le nombre de ticks en fonction du type de questions d'après la règle proposée par Elipss ; en deuxième et troisième colonnes, les estimations des coefficients des modèles de régression linéaire de la durée de la médiane et du dernier décile, ainsi que leur significativité (*** signifie une p-value inférieure à 0,001) et leur intervalle de confiance ; en quatrième colonne, notre proposition de révision de la règle de décompte des ticks.

4.2 Comparaison des méthodes optimisées

33 Les choix de simplification ainsi effectués ne conduisent pas à un résultat aussi performant que si l'on appliquait tels quels les coefficients des modèles de régression, mais ils permettent d'aboutir à une révision de la méthode des *ticks* qui constitue une amélioration significative de la méthode initiale (Tableau 5) : le coefficient de corrélation entre les nombres de *ticks* calculés d'après notre méthode révisée et les durées médianes observées dans les 41 enquêtes du corpus est de 0,76 (au lieu de 0,69 avec la méthode Elipss initiale), et de 0,72 au lieu de 0,65 pour la durée du dernier décile. Et cette méthode produit également de meilleurs résultats que le L2 de Puleston, même révisé avec les nouveaux coefficients indiqués plus haut.

Tableau 5 : Tableau des coefficients de corrélation entre les durées observées et les durées prédites par les différentes méthodes de décompte des questions

Méthode	Minimum	Premier décile	Médiane	Moyenne	Dernier décile	Maximum
Ticks (régression)	0,75	0,74	0,79	0,77	0,74	0,76
Ticks (Mercklé et Thizy)	**0,7**	**0,7**	**0,76**	**0,73**	**0,72**	**0,72**
L2 de Puleston révisé	0,68	0,66	0,7	0,7	0,69	0,69
L2 de Puleston	0,67	0,65	0,7	0,7	0,69	0,68
Questions	0,64	0,64	0,71	0,68	0,67	0,65
Ticks Elipss	0,61	0,62	0,69	0,66	0,65	0,63

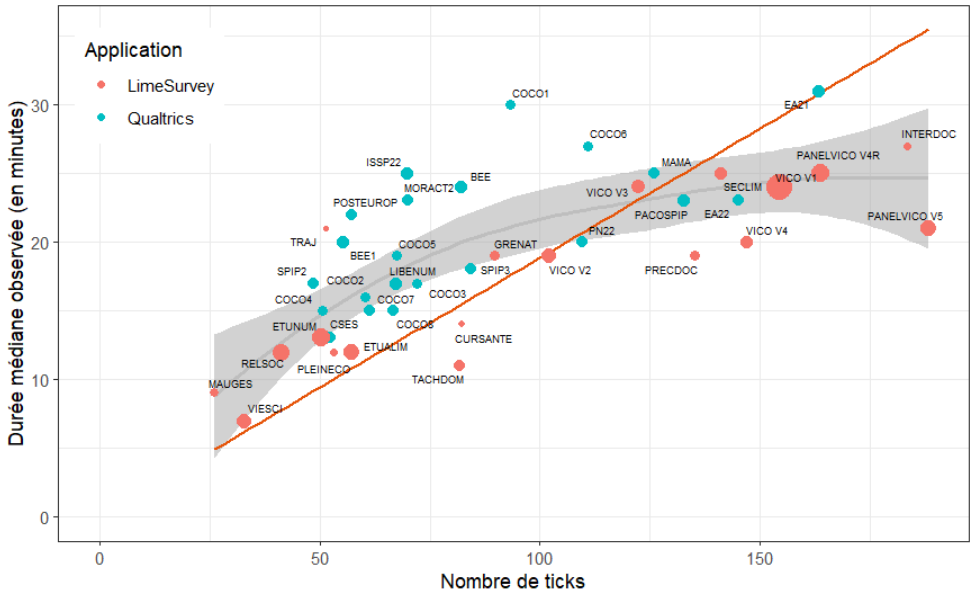
Champ : 41 enquêtes par questionnaire en ligne, 2020-2025.
Lecture : Le coefficient de corrélation de Pearson pondéré entre le nombre de ticks calculé à partir des coefficients du modèle de régression et la durée minimum observée est de 0,78.
Note : Les lignes du tableau sont classées par ordre décroissant du coefficient de corrélation entre la méthode testée et la durée médiane observée. Les couleurs des cases apportent une aide visuelle pour identifier les coefficients les plus élevés : vert >= 0,70 ; bleu >= 0,65 ; jaune >=0,60 ; orange >= 0,50 ; rouge < 0,50.

4.3 Combien de secondes vaut un *tick* ?

34 Une fois ajustée cette pondération du nombre de *ticks* selon le type de questions, une inconnue demeure : comme pour la méthode du comptage des questions, il faut désormais être en mesure d'attribuer une valeur temporelle au *tick*. Pour cela, nous appliquons notre méthode révisée de comptage des *ticks* selon le type de question et nous la rapportons aux durées observées (Figure 1). Le calcul d'une régression linéaire simple à constante nulle permet de déterminer que la durée médiane observée en secondes peut être ajustée par une droite passant par l'origine (figurée en rouge dans le Graphique 1), en multipliant le nombre de *ticks* par 11,3 (avec un R^2 de 0,90), et le dernier décile observé en multipliant le nombre de *ticks* par 20,0 (avec un R^2 de 0,90 également). On propose donc, à nouveau par souci de simplification de la méthode, de retenir la règle suivante : un *tick* (ou une question simple) vaut 10 secondes si on

souhaite estimer la durée médiane du questionnaire (soit 6 *ticks* par minute), et 20 secondes[14] si on souhaite estimer la durée du dernier décile (soit 3 *ticks* par minute).

Figure 1 : Nombre de *ticks* et durée médiane observée (en minutes)



Note : Les enquêtes sous LimeSurvey sont en rouge, et les enquêtes sous Qualtrics en bleu.
Champ : 41 enquêtes par questionnaire en ligne, 2020-2025.
Lecture : Le questionnaire de l'enquête MAMA (Elipss, 2022), réalisée avec Qualtrics, compte 144 ticks, sa durée médiane prédite par la méthode des ticks est donc de 144x10/60 = 24 minutes, et sa durée médiane observée est de 25 minutes.

35 Nous avons choisi cette méthode de régression linéaire par souci d'efficacité, en considérant que la durée d'un *tick* est constante quel que soit la longueur du questionnaire (ou la place de la question correspondante dans le questionnaire). La Figure 1 montre cependant que la liaison entre la durée prédite avec la méthode des *ticks* et la durée observée n'est pas parfaitement linéaire : les durées réelles des enquêtes les plus courtes sont au-dessus de la droite de régression (autrement dit leur durée est sous-estimée par notre modèle), tandis que celles des enquêtes les plus longues sont en-dessous (leur durée est surestimée). De fait, la courbe de régression locale (LOESS) figurée en gris sur le graphique présente une pente qui s'infléchit à mesure que le nombre de *ticks* augmente. Cela s'explique par l'effet d'accélération identifié dans la littérature (Malhotra, 2008 ; Yan et Tourangeau, 2008) : les enquêté·es répondent plus vite aux questions de la fin que du début des questionnaires, ce qui fait baisser en moyenne la durée du *tick* d'autant plus que le questionnaire est long. Cela signifie que dans les faits, la valeur temporelle d'un *tick n'est pas constante* au cours du questionnaire (elle décroît avec sa longueur). Il nous paraît cependant plus pratique, pour garder une règle simple d'estimation, de s'en tenir à une valeur constante.

36 D'autres paramètres sont par ailleurs susceptibles de faire varier la durée d'un *tick* et plus généralement le temps de passation d'un questionnaire. La littérature montre en effet que ce dernier ne dépend pas seulement du nombre et de la nature des questions posées : indépendamment de la structure du questionnaire, le temps de passation varie aussi selon les caractéristiques sociales des répondant·es, selon le matériel utilisé et les filtres du questionnaire.

5. Qu'est-ce qui peut faire varier la durée d'une enquête ?

5.1 Facteurs socio-démographiques de variation de la vitesse des réponses

37 La littérature disponible a nettement mis en évidence que le temps de réponse à un questionnaire varie en fonction des caractéristiques sociales des enquêteur·ices et des enquêté·es. Ainsi, dans le cas d'enquêtes qui ne sont pas auto-administrées, le temps de passation est largement tributaire des pratiques et de l'expérience des enquêteur·ices (Olson et Peytchev, 2007) qui lisent plus ou moins vite, demandent ou non des précisions, reviennent ou non sur certaines questions (Loosveldt et Beullens, 2013).

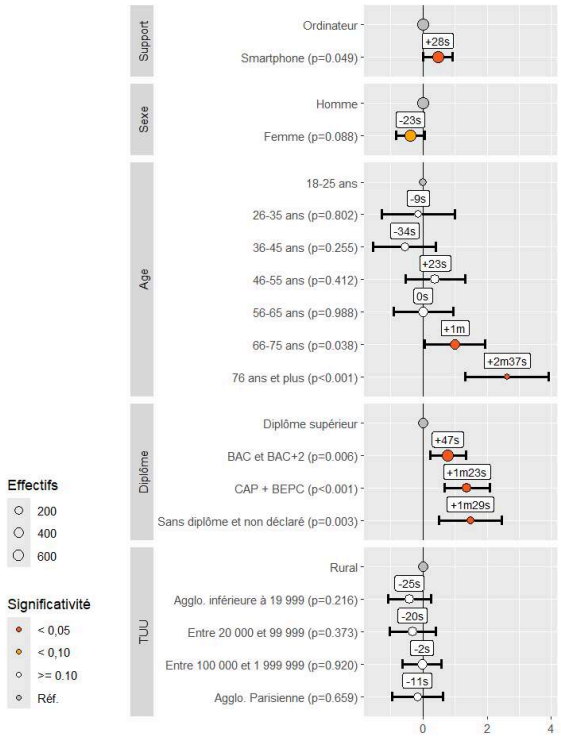
38 Mais la principale variation, à plus forte raison dans les enquêtes auto-administrées, tient aux propriétés sociales des enquêté·es (Yan et Tourangeau, 2008) : les plus diplômé·es, les jeunes, et les personnes ayant une expérience des questionnaires, par exemple des panélistes (Loewen et al., 2022), ont tendance à répondre plus vite que les moins diplômé·es, les plus âgé·es et les novices en matière d'enquête (Fricker et al., 2005). Les questionnaires traduits en plusieurs langues montrent également de fortes différences de vitesse de réponse entre pays, qui s'expliquent par la langue mais aussi par des différences de cultures nationales en matière d'enquête (Loosveldt et Beullens, 2013).

39 Pour tester un certain nombre de ces variations, nous avons utilisé une des enquêtes Elipss à notre disposition, parce qu'elle est la plus récente du corpus qui contienne la totalité des variables que nous voulions introduire dans l'analyse. Les données de la vague 2 de l'enquête « Goûts, morales, groupes sociaux » réalisée auprès du panel Elipss en décembre 2024 et janvier 2025 (MORACT2) permettent en effet de connaître, pour chaque répondant·e (N = 1 841), à la fois le temps total mis pour répondre, le support utilisé, et un certain nombre de caractéristiques socio-démographique (sexe, âge, niveau de diplôme, tranche d'unité urbaine du lieu de résidence)[15]. Après retrait comme précédemment des durées extrêmes les plus aberrantes, puis normalisation de la durée[16], l'analyse du modèle de régression linéaire multiple estimant la durée de réponse montre que de fait, le niveau de diplôme et l'âge sont les principaux déterminants des variations du temps de réponse (Figure 2). Toutes choses égales par ailleurs, si l'enquête durait 10 minutes, les bachelier·es mettraient 47 secondes de plus que les diplômé·es du supérieur pour répondre à l'enquête, et ce temps supplémentaire augmente encore à mesure que baisse le niveau de diplôme, jusqu'à atteindre 1 minute 29 secondes de plus pour les répondant·es sans diplôme. Le capital scolaire, objectivé par le niveau de diplôme, permet ainsi aux enquêté·es d'être plus rapides dans leurs réponses à des questionnaires qui ressemblent souvent à des exercices académiques.

40 L'effet de l'âge sur le temps de réponse, quant à lui, n'est pas linéaire, quoiqu'au-delà de 65 ans les enquêté·es sont de plus en plus lent·es à répondre. La littérature y voit l'effet d'un affaiblissement des capacités cognitives (voir notamment Bassili, 1996 ; Loosveldt et Beullens, 2013), auquel s'ajoute probablement – effet de génération – un affaiblissement de la littératie numérique. Mais ce ne sont pas les seules explications possibles. Dans l'enquête examinée, les adultes entre 36 et 45 ans sont significativement

plus rapides que les plus âgé·es, mais ils sont aussi plus rapides que les plus jeunes, qui maîtrisent pourtant mieux les outils informatiques mais peuvent aussi expérimenter plus de « freins psychosociaux » face à certains usages du numériques (Brice Marandal et al., 2025). Ces variations plaident plutôt pour une interprétation en termes de disponibilité temporelle : il existerait un temps maximum allouable à l'enquête, en concurrence plus ou moins forte avec les autres occupations sociales, qui varie en fonction de la position dans le cycle de vie. À l'âge où le travail salarié, les tâches domestiques et éventuellement parentales laissent moins de temps de loisir que pendant les études ou à la retraite, le temps alloué aux questionnaires est plus contraint. Que les femmes mettent moins de temps à répondre que les hommes plaide également en faveur de cette interprétation en termes de contraintes temporelles liées aux autres occupations.

Figure 2 : Déterminants de la durée de l'enquête MORACT2 (régression linéaire multiple)



Source : Caveng, Rémi ; Darbus, Fanny ; Denord, François ; Serre, Delphine ; Thine, Sylvain, 2025, « Goûts, morales, groupes sociaux - vague 2 (ELIPSS 2024) », https://doi.org/10.21410/7E4/AXZRDI, data.sciencespo, V2, UNF:6:mY/xqHEtp6+p71qdWNdgSg== [fileUNF].
Champ : Personnes résidant habituellement en France âgées de 18 ans et plus (N = 1 489).
Lecture : Toutes choses égales par ailleurs, si l'enquête MORACT2 durait en moyenne 10 minutes, les répondant·es remplissant le questionnaire sur téléphone mettraient 27 secondes de plus que les répondant·es le remplissant sur ordinateur (estimation significative au seuil de 10 %). Note : le modèle contrôle également l'ancienneté du recrutement dans le panel Elipss, dont l'effet propre n'est pas significatif au seuil de 10 %.

5.2 Facteurs matériels de variation de la vitesse des réponses

41 La figure 2 met également en évidence un effet propre du support sur le temps de réponse : toutes choses égales par ailleurs, répondre sur smartphone plutôt que sur ordinateur augmente le temps de réponse. Ce résultat, déjà étayé dans la littérature

(Couper et Peterson, 2017 ; Cunningham et al., 2013), est assez intuitif si l'on prend en considération que la saisie sur un petit écran tactile[17] est moins ergonomique que celle avec un clavier et une souris : il faut davantage scroller, zoomer ou dézoomer, les textes des réponses aux questions ouvertes sont plus fastidieux à écrire, et les tableaux de questions peuvent être éclatés en question indépendantes par le recours au *responsive design* des deux applications.

42 Un autre facteur matériel qui peut impacter la vitesse de réponse (mais que nous n'avons pas pu tester) est la qualité de l'accès à internet. Une mauvaise connexion peut ralentir la passation du questionnaire, notamment s'il comporte des médias (photos, vidéos, sons) de taille importante. Le temps d'affichage du questionnaire peut ainsi s'avérer plus long dans les zones rurales ou dans les catégories de la population moins bien équipées – cet accès à un équipement informatique et une connexion internet étant lui-même conditionné par des variables socio-démographiques.

43 Enfin, certains articles suggèrent que le design de questionnaire lui-même peut avoir un effet sur le temps de passation (Yan et Tourangeau, 2008). Nous ne disposions cependant pas du temps de réponse par question, ni d'un dispositif permettant de comparer toutes choses égales par ailleurs l'effet d'un design particulier sur une question, mais ce point pourrait être exploré davantage.

5.3 Filtres, boucles et embranchements : des variations du temps de réponse difficiles à anticiper

44 Pour finir – *last but not least* – se pose l'épineuse question des filtres et des embranchements qui peuvent faire varier de façon importante le temps de réponse selon les répondant·es. Quand un questionnaire est construit avec des filtres, tous·tes les répondant·es ne sont pas tous·tes concerné·es par toutes les questions, ce qui, à questionnaire identique, fait varier le nombre de questions soumises à chacun·e, et complexifie l'estimation du temps de passation. Cela vaut également pour les « boucles » de questions, comme les répétitions de modules de questions pour chaque membre d'un ménage, ou pour les membres du réseau personnel dans les enquêtes sur la sociabilité, que les répondant·es doivent remplir un nombre variable de fois, difficilement prévisible. Cette difficulté méthodologique est récurrente dans la littérature anglophone sur la durée des questionnaires, et elle est difficile à surmonter. Loewen et al. (2022) considèrent que l'estimation de la durée des filtres en fonction de chaque profil de répondant·es peut se révéler très fastidieuse et complexe, ce qui les conduit à proposer une technique alternative, reposant sur la pondération des questions par la proportion de répondant·es susceptibles d'y répondre. C'est également la pratique retenue par un des instituts de sondage rencontrés, qui tente d'estimer la proportion de répondant·es susceptibles de répondre à une question à partir de données déjà disponibles. Mais de telles données n'existent pas toujours, ce qui rend difficile d'anticiper les effets d'un filtre, de boucles ou d'un embranchement important du questionnaire sur le temps de passation. Nous avons pour notre part choisi de laisser cette question de côté, d'une part parce que nous étions à la recherche d'une règle d'estimation aussi simple que possible, d'autre part parce que ce point peut être estimé non pas en amont de la programmation dans un logiciel, mais lors d'une phase de test du questionnaire sur un échantillon restreint. Ce point mériterait dans tous les cas de

plus amples investigations pour mieux estimer l'effet de la présence de filtres, de boucles ou de bifurcations.

Conclusion : règles pratiques pour estimer la durée d'un questionnaire

45 Les analyses développées montrent qu'il n'est certes pas possible de prédire de façon absolument exacte la durée d'un questionnaire à partir de ses seules caractéristiques (longueur, nombre et types de questions, etc.). De fait, encoder et tester un questionnaire avant sa diffusion à l'ensemble de la population enquêtée reste un préalable méthodologique indispensable. Notre objectif n'était cependant pas de rendre cette étape évitable, mais plutôt de dégager une méthode aussi simple et efficace que possible d'estimation pour accompagner les concepteur·ices de questionnaire et leur fournir un outil d'aide à la décision dans les nombreux dilemmes qui peuvent se poser lors de la conception de l'enquête. En ce sens, il apparaît que les corrélations observées entre les caractéristiques des questionnaires et les durées sont très bonnes (supérieures à 0,75 avec la révision de la méthode des *ticks* que nous proposons). Nous proposons donc, pour conclure, de résumer les principaux apports de cet article en dégageant une liste de recommandations pour estimer le temps de passation d'un questionnaire :

1. Ne vous appuyez pas sur le nombre de signes ou de mots, mais plutôt sur le nombre de questions et sur leur nature.
2. Si vous êtes est pressé·es, contentez-vous de compter les questions : cela donne, à moindre coût, des résultats plutôt satisfaisants.
3. Si vous voulez améliorer la prédiction, attribuez des points aux questions selon leur nature (méthode des *ticks*). Nous proposons pour cela des coefficients révisés, testés empiriquement :
   1 *tick* pour une question à choix unique ;
   1 *tick* toutes les 2 lignes des tableaux de questions ;
   1 *tick* toutes les 10 modalités de question à choix multiples ;
   0,25 *tick* pour chaque question numérique ;
   0,1 *tick* pour chaque question ouverte (car elles sont souvent laissées vides).
4. Selon que vous souhaitez estimer le temps *médian* ou le temps *maximal* de passation, ayez en tête que le temps mis par le dernier décile peut être le double de celui de la médiane.
5. Pour estimer la médiane, comptez 6 *ticks* ou questions par minute (soit un *tick* = 10 secondes). Pour estimer le dernier décile, comptez plutôt 3 *ticks* ou questions par minute, voire 3 (soit un *tick* = 20 secondes).
6. Effet d'accélération 1 : N'attendez pas la fin du questionnaire pour poser les questions les plus importantes, car la qualité des réponses a tendance à se dégrader à mesure de l'avancée dans le questionnaire.
7. Effet d'accélération 2 : plus le questionnaire est long, plus la durée moyenne du *tick* diminue (dit autrement, plus le nombre de *ticks* par minute augmente). Si votre questionnaire est très court ou très long, vous pouvez donc ajuster le nombre de *ticks* par minute respectivement à la baisse ou à la hausse pour mieux estimer la durée.
8. Effet de composition de l'échantillon : vous pouvez également ajuster le nombre de *ticks* par minute en fonction de la population enquêtée. Les jeunes, les plus diplômé·es (et dans une moindre mesure les panélistes aguerri·es[18]) répondent plus vite que les personnes âgées, moins diplômées ou moins expérimentées.

9. Si vous avez des filtres longs, essayez d'estimer la proportion de répondant·es susceptibles d'y répondre pour pondérer ces questions – sinon, faites sans.
10. Ne renoncez pas, malgré tout, à une phase de tests en conditions réelles sur un sous-échantillon de votre population, la méthode que nous proposons n'ayant pas vocation à se substituer à cette phase, mais bien à la précéder et la préciser.

BIBLIOGRAPHIE

Bassili, J.N. (1996), The How and Why of Response Latency Measurement in Telephone Surveys, *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, Hoboken, NJ, US, Jossey-Bass/Wiley, p. 319-346.

Bassili, J.N. & Scott, B.S. (1996), Response Latency as a Signal to Question Problems in Survey Research, *Public Opinion Quarterly*, vol. 60, n° 3, p. 390.

Beck, F., Castell, L. & al. (2022), Le multimode dans les enquêtes auprès des ménages : une collecte modernisée, un processus complexifié, *Courrier des statistiques*, n° 7, p. 7-28. https://www.insee.fr/fr/information/6035934?sommaire=6035950

Bessière, C. & Houseaux, F. (1997), Suivre des enquêteurs, *Genèses*, n° 29, p. 100-114. https://www.persee.fr/doc/genes_1155-3219_1997_num1_1482

Brice Marandal, L., Paquet, N. & al. (2025), La société numérique française : comprendre les freins psychosociaux à l'usage du numérique, ANCT, CREDOC, Université Rennes 2 CREAD-M@rsouin. https://www.credoc.fr/publications/la-societe-numerique-francaise-comprendre-les-freins-psychosociaux-a-lusage-du-numerique

Brysbaert, M. (2019), How many words do we read per minute? A review and meta-analysis of reading rate, *Journal of Memory and Language*, vol. 109, p. 104047. https://linkinghub.elsevier.com/retrieve/pii/S0749596X19300786

Carter, B., Bennett, J. & al. (2021), Evaluation of Estimated Survey Duration Equations Using a Health Risk Assessment, *Survey Research Methods*, p. 187-194. https://ojs.ub.uni-konstanz.de/srm/article/view/7800

Carver, R.P. (1992), Reading Rate: Theory, Research, and Practical Implications, *Journal of Reading*, vol. 36, n° 2, p. 84-95. https://www.jstor.org/stable/40016440

Caveng, R. (2012), La production des enquêtes quantitatives, *Revue d'anthropologie des connaissances*, vol. 6, n° 1, p. 65-88. https://journals.openedition.org/rac/12828

Conrad, F.G., Couper, M.P. & al. (2010), The Impact of Progress Indicators on Task Completion, *Interacting with Computers*, vol. 22, n° 5, p. 417-427. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2910434/

Cook, C., Heath, F. & al. (2000), A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys, *Educational and Psychological Measurement*, vol. 60, n° 6, p. 821-836.

Couper, M.P. & Peterson, G.J. (2017), Why Do Web Surveys Take Longer on Smartphones?, *Social Science Computer Review*, vol. 35, n° 3, p. 357-377. https://doi.org/10.1177/0894439316629932

Crawford, S.D., Couper, M.P. & al. (2001), Web surveys: Perceptions of Burden, *Social Science Computer Review*, vol. 19, n° 2, p. 146-162.

Cunningham, J.A., Neighbors, C. & al. (2013), Use of Mobile Devices to Answer Online Surveys: Implications for Research, *BMC Research Notes*, vol. 6, n° 1, p. 258. https://doi.org/10.1186/1756-0500-6-258

Deutskens, E., De Ruyter, K. & al. (2004), Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study, *Marketing Letters*, vol. 15, n° 1, p. 21-36.

Dillman, D.A., Sinclair, M.D. & al. (1993), Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys, *Public Opinion Quarterly*, vol. 57, n° 3, p. 289-304.

Duwez, E. & Mercklé P. (éd.) (2021), *Un panel français. L'étude longitudinale par Internet pour les sciences sociales (ELIPSS)*, Paris, Ined Éditions, https://books.openedition.org/ined/16275.

Duwez, E. & Mercklé, P. (2024), Construire un dispositif d'enquêtes longitudinales pour la recherche en sciences sociales : le panel ELIPSS, *L'Année sociologique*, vol. 74, n° 1, p. 251-273. https://www.cairn.info/revue-l-annee-sociologique-2024-1-page-251.htm

Fricker, S., Galesic, M. & al. (2005), An Experimental Comparison of Web and Telephone Surveys, *The Public Opinion Quarterly*, vol. 69, n° 3, p. 370-392. https://www.jstor.org/stable/3521546

Galesic, M. & Bosnjak, M. (2009), Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey, *Public Opinion Quarterly*, vol. 73, n° 2, p. 349-360.

Hasselhorn, K., Ottenstein, C. et al. (2024), The Effects of Questionnaire Length on the Relative Impact of Response Styles in Ambulatory Assessment, *Multivariate Behavioral Research*, p. 1-15. https://www.tandfonline.com/doi/full/10.1080/00273171.2024.2354233

Heberlein, T.A. & Baumgartner, R. (1978), Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature, *American Sociological Review*, vol. 43, n° 4, p. 447-462.

Helgeson, J.G. & Ursic, M.L. (1994), The role of affective and cognitive decision-making processes during questionnaire completion, *Psychology & Marketing*, vol. 11, n° 5, p. 493-510. https://onlinelibrary.wiley.com/doi/abs/10.1002/mar.4220110506

Herzog, A.R. & Bachman, J.G. (1981), Effects of Questionnaire Length on Response Quality, *Public Opinion Quarterly*, vol. 45, n° 4, p. 549-559.

Koitsalu, M., Eklund, M. & al. (2018), Predictors of Participation in Risk-Based Prostate Cancer Screening, *PloS One*, vol. 13, n° 7, p. e0200409.

Krosnick, J.A. (1991), Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys, *Applied Cognitive Psychology*, vol. 5, n° 3, p. 213-236. https://onlinelibrary.wiley.com/doi/abs/10.1002/acp.2350050305

Krosnick, J.A., Holbrook, A.L. & al. (2002), The Impact of "No Opinion" Response Options on Data Quality: Non-Attitude Reduction or an Invitation to Satisfice?, *The Public Opinion Quarterly*, vol. 66, n° 3, p. 371-403.

Loewen, E.R., Bauer, E. & al. (2022), Timing Estimates for Complex Programmed Surveys, *Survey Practice*, vol. 15, n° 1. https://www.surveypractice.org/article/38715-timing-estimates-for-complex-programmed-surveys

Loosveldt, G. & Beullens, K. (2013), "How Long Will It Take?" An Analysis of Interview Length in the Fifth Round of the European Social Survey, *Survey Research Methods*, vol. 7, n° 2, p. 69-78, https://ojs.ub.uni-konstanz.de/srm/article/view/5086

Lynn P. (2014), Longer interviews may not affect subsequent survey participation propensity, *Public Opinion Quarterly*, vol. 78, n° 2, p. 500-509.

Lynn, P., Roberts, C. & al. (2010), Data quality in telephone surveys and the effect of questionnaire length: a cross- national experiment, *ISER Working Paper Series*, n° 36. https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2010-36.pdf

Malhotra, N. (2008), Completion Time and Response Order Effects in Web Surveys, *Public Opinion Quarterly*, vol. 72, n° 5, p. 914-934. https://doi.org/10.1093/poq/nfn050

Mariot, N., Mercklé, P. & al. (éd.) (2021), *Personne ne bouge. Une enquête sur le confinement du printemps 2020*, Grenoble, UGA Editions. https://books.openedition.org/ugaeditions/18372

Nation, P. (2009), Reading Faster, *International Journal of English Studies*, vol. 9, n° 2. https://revistas.um.es/ijes/article/view/90791

Olson, K. & Peytchev, A. (2007), Effect of Interviewer Experience on Interview Pace and Interviewer Attitudes, *Public Opinion Quarterly*, vol. 71, n° 2, p. 273-286.

Peneff, J. (1988), The Observers Observed: French Survey Researchers at Work, *Social Problems*, vol. 35, n° 5, p. 520-535.

Puleston, J. (2012), How to Calculate the Length of a Survey, *Question Science*. https://question-science.blogspot.com/2012/07/how-to-calculate-length-of-survey.html.

Revilla, M. & Höhne, J.K. (2020), How Long Do Respondents Think Online Surveys Should Be? New Evidence from Two Online Panels in Germany, *International Journal of Market Research*, vol. 62, n° 5, p. 538-545. https://doi.org/10.1177/1470785320943049

Revilla, M. & Ochoa, C. (2017), Ideal and Maximum Length for a Web Survey, *International Journal of Market Research*, vol. 59, n° 5, p. 557-565.

Sandelin, F. (2022), The Effects of Questionnaire Length on Response Rate, Non-response Bias, and Data Quality, *The SOM Institute's Notes on Survey Methodology*, n° 1, p. 1-19. https://www.gu.se/sites/default/files/2022-11/2022-1%20Effects%20of%20questionnaire%20length%20%28Sandelin%202022%29%20v2_1.pdf

Shyamson, S. (2023), Proven Methods to Measure Survey Completion Time, *BlockSurvey*. https://blocksurvey.io/research/proven-methods-to-measure-survey-completion-time

Sudman, S., Bradburn, N. & al. (1996), *Thinking about Answers: The Application of Cognitive Process to Survey Methodology*, San Francisco, Jossey-Bass Publishers.

Tourangeau, R., Couper, M.P. & al. (2004), « Spacing, Position, and Order: Interpretive Heuristics for Visual Features of Survey Questions », *Public Opinion Quarterly*, vol. 68, n° 3, p. 368-393.

Trauzettel-Klosinski, S. & Dietz, K. (2012), Standardized Assessment of Reading Performance: The New International Reading Speed Texts IReST, *Investigative Opthalmology & Visual Science*, vol. 53, n° 9, p. 5452. http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.11-8284

Versta Research (2011), How to Estimate the Length of a Survey, *Versta Research Newsletter*. https://verstaresearch.com/newsletters/how-to-estimate-the-length-of-a-survey/

Yammarino, F.J., Skinner, S.J. & al. (1991), Understanding Mail Survey Response Behavior: A Meta-Analysis, *The Public Opinion Quarterly*, vol. 55, n° 4, p. 613-639.

Yan, T. & Tourangeau, R. (2008), Fast Times and Easy Questions: the Effects of Age, Experience and Question Complexity on Web Survey Response Times, *Applied Cognitive Psychology*, vol. 22, n° 1, p. 51-68.

Zhang, C. & Conrad, F. (2014), Speeding in Web Surveys: The tendency to answer very fast and its association with straightlining, *Survey Research Methods*, vol. 8, n° 2, p. 127-135. https://ojs.ub.uni-konstanz.de/srm/article/view/5453

ANNEXES

Annexe 1. Les 41 enquêtes du corpus

Le tableau ci-dessous présente, pour chaque des 41 enquêtes du corpus : l'année d'administration de l'enquête, le nombre de réponses complètes chronométrées, l'application utilisée pour diffuser le questionnaire, le nombre de questions, le nombre de signes du questionnaire, le nombre de *ticks* mesurés avec les règles présentées à la fin de l'article, la durée médiane et la durée du dernier décile observées (en minutes).

Enquête	**Année**	**N**	**Application**	**Questions**	**Signes**	**Ticks**	**Médiane**	**D9**
COCO1	2020	1251	Qualtrics	97	10131	120	30	55
COCO2	2020	1161	Qualtrics	51	7504	77	16	27
COCO3	2020	1189	Qualtrics	60	8509	89	17	28
COCO4	2020	1104	Qualtrics	40	6967	57	15	28
COCO5	2020	1127	Qualtrics	53	7106	79	19	33
COCO6	2020	1179	Qualtrics	93	16205	133	27	49
COCO7	2020	1344	Qualtrics	55	7990	69	15	27
VICO V1	2020	16224	LimeSurvey	180	51927	166	24	40
VICO V2	2020	3624	LimeSurvey	117	31297	109	19	36
COCO8	2021	1511	Qualtrics	62	17985	74	15	25
EA21	2021	1895	Qualtrics	179	41529	200	31	56
ETUNUM	2021	6962	LimeSurvey	52	15131	52	13	22
INTERDOC	2021	503	LimeSurvey	208	46630	195	27	49
VICO V3	2021	3007	LimeSurvey	142	40437	133	24	44
VIESCI	2021	3248	LimeSurvey	49	20272	34	7	14

CSES	2022	1575	Qualtrics	75	26364	80	13	22
EA22	2022	1766	Qualtrics	162	61197	180	23	41
ETUALIM	2022	4291	LimeSurvey	69	14511	59	12	21
PN22	2022	1588	Qualtrics	80	29293	124	20	34
RELSOC	2022	5279	LimeSurvey	60	13322	43	12	21
SPIP2	2022	1614	Qualtrics	69	23319	91	17	32
SPIP3	2022	1627	Qualtrics	88	27165	116	18	31
BEE	2023	2010	Qualtrics	80	24871	94	24	44
CURSANTE	2023	297	LimeSurvey	77	26808	86	14	27
ISSP22	2023	2205	Qualtrics	67	35223	86	25	45
MAMA	2023	1557	Qualtrics	65	37252	144	25	43
MAUGES	2023	600	LimeSurvey	27	7485	27	9	14
PLEINECO	2023	562	LimeSurvey	16	12518	54	12	20
PRECDOC	2023	992	LimeSurvey	148	40676	139	19	36
TACHDOM	2023	1563	LimeSurvey	103	23461	87	11	19
VICO V4	2023	2013	LimeSurvey	161	141756	154	20	36
BEE1	2024	1882	Qualtrics	50	19056	64	20	39
GRENAT	2024	1068	LimeSurvey	114	32583	94	19	32
LIBENUM	2024	1920	Qualtrics	74	36482	94	17	29
MORACT2	2024	1841	Qualtrics	73	20566	75	23	41
PACOSPIP	2024	1914	Qualtrics	144	89818	189	23	40
PANELVICO V4R	2024	6132	LimeSurvey	187	150888	173	25	45
POSTEUROP	2024	1811	Qualtrics	46	20531	66	22	40
SECLIM	2024	2188	LimeSurvey	149	23538	153	25	35
PANELVICO V5	2025	4694	LimeSurvey	172	60909	206	21	38
TRAJ	2025	182	LimeSurvey	69	16117	60	21	32

Annexe 2. Présentation détaillée des 41 enquêtes

Les 19 enquêtes réalisées avec LimeSurvey

6 vagues de l'enquête longitudinale en population générale âgée de 18 ans et plus, réalisées en ligne entre 2020 et 2025 à l'aide du panel des ANR Vico (Mariot, Mercklé et Perdoncin, 2021) et Panelvico (2020-2025).

4 enquêtes auprès des étudiant·es de l'Université Grenoble Alpes (UGA), réalisées entre 2021 et 2024 :

ETUNUM : Enquête sur les usages numériques des étudiant·es, réalisée en TD de L2 de sociologie en 2021, N = 6 962 réponses complètes.

ETUALIM : Enquête sur les pratiques alimentaires des étudiant·es, en français, réalisée en TD de L2 de sociologie en 2022, N = 4 291.

RELSOC : Enquête sur les relations sociales des étudiant.es, réalisée en TD de L2 en 2023, N = 5 279.

PRECDOC : Enquête sur la précarité en doctorat, réalisée auprès des doctorant·es de l'UGA en 2023 dans le cadre d'une recherche commandée par le collège doctoral de l'UGA et coordonnée par Pierre Bataille et Pierre Mercklé, N = 992.

9 enquêtes de recherche réalisées en France auprès de publics variés et partagées par des collègues, pour certaines après un appel sur les listes professionnelles.

INTERDOC : Enquête en français et en anglais auprès des doctorants internationaux en France, réalisée en 2021 par Magali Ballatore et Jean-Baptiste Bertrand, N = 503 réponses complètes.

VIESCI : Vie de père/mère, vie de scientifique, enquête réalisée en 2021 par Sandrine Lefranc auprès d'un échantillon de chercheurs et chercheuses, N = 3 248.

CURSANTE : Enquête sur les doubles cursus santé/science réalisée en 2023, conjointement par l'équipe PULWAR (Centre nantais de sociologie, Nantes Université) et par la commission Parité de l'Association nationale des doubles cursus en santé (ANDCS). Les données nous ont été transmises par Elsa Boulet, N = 297.

PLEINECO : Pleine conscience et traits de personnalité des étudiant(es). Enquête réalisée en 2023 sur le serveur LimeSurvey de l'Université Toulouse Jean Jaurès, dans le cadre d'une étude du CERPPS (Université Toulouse Jean Jaurès). L'enquête a été réalisée sous la responsabilité de Patrick Raynal, auprès d'étudiant(e)s âgé(e)s de 18 à 28 ans, N = 562.

SECLIM : Les élèves de seconde face au changement climatique, enquête réalisée par les étudiant·es du Master Quantifier en sciences sociales (ENS, EHESS) en 2023, et mise à disposition par Damien Cartron, N = 2 188.

TACHDOM : Enquête sur le partage des tâches domestiques dans la famille, réalisée en 2023 par les étudiant·es du Master Quantifier en sciences sociales (ENS, EHESS), et mise à disposition par Damien Cartron, N =1 563.

GRENAT : Enquête sur les Grenoblois et la nature, réalisée en TD de L2 de sociologie en 2024 auprès d'un échantillon d'habitant·es de la ville de Grenoble, N = 1 068.

MAUGES : Enquête réalisée en 2025, dans le cadre d'une commande de Mauges Communauté à l'Institut de Géographie et d'Aménagement Régional de l'Université de Nantes (IGARUN), et réalisée par des étudiants de deuxième année de master de géographie sociale et aménagement, N = 600. Les données ont été transmises par et utilisées avec l'autorisation de Rémy Blans.

TRAJ : L'enquête a été réalisée auprès d'élèves de l'enseignement professionnel (N = 182) et porte sur leurs trajets entre le domicile et l'établissement scolaire. Elle a été réalisée en 2025 et transmise par Rémy Blans, doctorant en géographie de l'Université de Nantes.

Les 22 enquêtes Elipss réalisées avec Qualtrics

Les 22 autres enquêtes du corpus ont été administrées avec Qualtrics, dans le cadre du dispositif d'Etude longitudinale par internet pour les sciences sociales (Elipss) du Centre de données socio-politiques (CNRS, Sciences Po) (Duwez et Mercklé, 2021, 2024). Le catalogue complet des enquêtes du dispositif Elipss peut être consulté en ligne ici : https://data.sciencespo.fr/dataverse/elipss.

COCO : 8 vagues de l'enquête « Faire face au Covid-19 », réalisée en 2020 et 2021, N = 1 251 (vague 1) à 1 511 (vague 8) réponses complètes.

EA21 et EA22 : 2 vagues de l'enquête annuelle Elipss, réalisé en 2021 (N = 1 891) et 2022 (N = 1 766).

CSES : vague 4 de l'enquête French Election Study 2022, qui reprend l'intégralité du module 6 de l'enquête internationale CSES (Comparative study of electoral systems). L'enquête CSES a été réalisée en 2022, N = 1 575.

PN22 : Vague 9 de l'enquête Pratiques numériques Elipss, réalisée en 2022, N = 1 588.

SPIP2 et SPIP3 : 2 vagues de l'enquête Sociologie Politique de l'Insécurité durant les élections Présidentielles de 2022, réalisées en avril 2022 (vague 2, N = 1 614) et octobre 2022 (vague 3, N = 1 627).

ISSP22 : Enquête française de la vague 2022 de l'International Social Survey Program (module Famille), N = 2 205.

MAMA : Enquête Monde d'avant, monde d'après, sur les transformations des pratiques culturelles après la crise sanitaire, réalisée en 2023, N = 1 557.

BEE et BEE1 : Questionnaire d'entrée (décembre 2023, N = 2 010) et vague 1 (octobre 2023, N = 1 882) du Baromètre Environnement et Écologie.

LIBENUM : Normes de la parole et liberté d'expression sur les réseaux sociaux, enquête réalisée en avril 2024, N = 1 920.

PACOSPIP : 3 modules d'enquêtes administrés en septembre 2024 : Parentalités et styles pédagogiques, Accès à l'information sur les collèges en France, et vague 4 de l'enquête Sociologie Politique de l'Insécurité durant les élections Présidentielles de 2022, N = 1 914.

POSTEUROP : Enquête post-élections européennes, réalisée en 2024, N = 1 811.

MORACT2 : Vague 2 de l'enquête Goûts, morales, groupes sociaux, réalisée en décembre 2024, N = 1 841.

NOTES

1. Des questionnaires variant de fait du simple au double pour une même durée annoncée obtiennent des taux de réponses quasi équivalents (Koitsalu et al., 2018). Certains auteurs suggèrent ainsi que minorer volontairement le temps de passation annoncé par rapport au temps réel augmente le taux de réponse aux questionnaires (Galesic et Bosnjak, 2009), mais ces résultats sont à nuancer : si plus de personnes commencent un questionnaire à la durée annoncée plus courte, davantage de personnes abandonnent si la durée effective s'avère supérieure à celle annoncée (Crawford, Couper et Lamias, 2001). Les études sur l'affichage d'une barre montrant la complétion progressive du questionnaire suggèrent quant à elles que leur effet dépend des anticipations des répondant·es : elle peut encourager à poursuivre si la vitesse de remplissage correspond à celle attendue, mais décourager si la progression semble trop lente (Conrad et al., 2010).

2. L'équipe de l'ANR Panelvico est divisée en trois axes portés par trois laboratoires : le LISST à Toulouse, Pacte à Grenoble et le LEST à Aix-Marseille. Elle réunit une quarantaine de chercheur·ses de différents statuts – voir la présentation détaillée sur notre carnet de recherche : https://vico.hypotheses.org/lequipe.

3. L'enquête menée par l'ANR VICO puis PANELVICO est une enquête en méthodes mixtes auprès d'un panel initialement composé de plus de 5 000 répondant·es, qui porte sur les conséquences à long terme de la crise sanitaire de 2020 sur les relations sociales, le rapport au travail, au logement, à l'Etat. Entre 2020 et 2025, cinq vagues de questionnaires ont été administrées, complétées par plusieurs séries d'entretiens qualitatifs auprès du panel. Pour une présentation détaillée de l'enquête et de ses résultats, voir notamment Mariot, Mercklé et Perdoncin, 2021, et le carnet Hypothèses du projet : https://vico.hypotheses.org.

4. Dans cette situation, l'effet de la durée du questionnaire porte plutôt sur la qualité des données recueillies : on peut supposer que lorsque le salaire est indépendant de la durée du questionnaire, les enquêteur·ices sont incité·es à passer plus rapidement sur les questions, au détriment de la qualité des réponses. Cette remarque vaut lorsque les enquêteur·ices sont rémunéré·es à l'heure, et non à la pièce comme dans les entreprises de sondage (Caveng, 2012). Si cela a été le cas aussi à l'Insee dans le passé (Bessière et Houseaux, 1997), les enquêteur·ices n'y sont plus rémunéré·es à la pièce depuis 2012, mais dans le cadre de contrats de la fonction publique.

5. Les questionnaires administrés par téléphone ou en face-à-face peuvent être plus longs (Beck et al., 2022) : jusqu'à 30 minutes par téléphone (Lynn et al., 2010), et jusqu'à une heure ou plus en face-à-face.

6. « the amount of information processing necessary to answer a question ».

7. La qualité intrinsèque du questionnaire, c'est-à-dire la façon dont il est structuré, a un effet certain sur la vitesse de réponse : il est également plus facile et donc plus rapide de répondre à un questionnaire bien construit, logiquement articulé et aux questions bien formulées qu'à un questionnaire mal conçu, à la progression peu claire

(Tourangeau, Couper et Conrad, 2004 ; Yan et Tourangeau, 2008). Les enquêtes sur le temps de latence avant une réponse montrent ainsi qu'une hésitation ou une difficulté face à une question, en augmentant le temps de réponse, peut refléter sa mauvaise formulation (Bassili et Scott, 1996). Le temps de réponse à chaque question dépend de la nature de celle-ci, de sa formulation et de la façon dont elle s'articule avec les questions précédentes.

8. Pour des travaux estimant le temps de lecture, voir Brysbaert, 2019 ; Carver, 1992 ; Nation, 2009 ; Trauzettel-Klosinski et Dietz, 2012.

9. Plus précisément, les auteurs indiquent une durée de 45 minutes pour 180 questions, c'est-à-dire environ 15 secondes par question, soit 4 questions par minute.

10. ELIPSS est un panel en ligne destiné à la recherche en sciences sociales qui propose un questionnaire chaque mois à plusieurs milliers de personnes représentatives de la population française (Duwez et Mercklé, 2021, 2024).

11. Nous avons retenu pour les régressions ces deux mesures de la durée observée, parmi toutes celles calculées dans le Tableau 1, parce que ce sont elles qui sont généralement invoquées dans la littérature, et celles qui importent le plus aux producteur·ices d'enquêtes.

12. Nous avons également testé cette même approche sur la méthode L1 de Puleston, supposée plus complète et plus performante que la méthode simplifiée dite L2. Mais d'une part, de façon en partie contre-intuitive, la méthode L1 donne de moins bons résultats sur notre corpus que la méthode simplifiée L2 ne tenant compte que du nombre de mots et du nombre de questions. Et d'autre part, la modélisation indique que pour améliorer la méthode L1, il conviendrait d'affecter un coefficient nul au nombre de mots, ce qui revient donc à ne plus tenir compte du temps de lecture et à transformer cette méthode en celle des *ticks*.

13. Ces choix peuvent sembler contre-intuitifs, en particulier pour les questions ouvertes, qui prennent par construction plus de temps que les questions fermées. Mais dans les enquêtes en ligne, elles sont en réalité très fréquemment, voire parfois systématiquement laissées vides par les répondant·es, et ne leur prennent donc au total que très peu de temps.

14. Dans son document interne plusieurs fois cité ici, Elipss indique de son côté que le temps d'administration d'une question à choix unique est de 15 secondes, sans préciser s'il s'agit d'une durée médiane, moyenne ou maximum. La Figure 1 et l'introduction de l'application utilisée dans le modèle de régression indiquent, cela dit, que les enquêtes réalisées avec Qualtrics (c'est-à-dire les enquêtes Elipss) durent plus longtemps, à nombre de *ticks* donné, que les enquêtes réalisées avec LimeSurvey. Il reste donc possible de maintenir une durée de 15 secondes par *tick* dans l'estimation de la durée des enquêtes Elipss, et de prévoir plutôt une durée de 10 secondes dans les autres enquêtes sous LimeSurvey. Nous faisons l'hypothèse que cet allongement n'est pas dû à des différences de logiciels, que nous ne pouvons pas contrôler en l'absence d'enquêtes de chercheurs réalisées sous Qualtrics dans notre corpus. Il nous semble plutôt dû à des caractéristiques spécifiques des répondant·es du panel Elipss, que nous ne pouvons pas contrôler non plus parce que, pour presque la moitié des enquêtes du corpus, nous ne disposons pas des caractéristiques sociodémographiques des répondant·es.

15. L'examen des résultats obtenus avec plusieurs autres enquêtes Elipss du corpus aboutit à peu près exactement aux mêmes conclusions, ce qui n'est pas surprenant

étant donné que ces enquêtes sont administrées auprès des mêmes répondant·es de ce panel.

16. L'enquête MORACT2 dure en réalité en moyenne 26 minutes, mais nous avons divisé toutes les durées individuelles par 2,6 pour ramener cette durée moyenne à 10 minutes et faciliter ainsi la lecture des résultats et la compréhension de l'ampleur des variations.

17. Le test d'un modèle alternatif contrôlant plus de variables (en particulier la taille de l'écran et l'heure de la journée) montre que la taille de l'écran absorbe la plus grande partie de l'effet du support, et non l'heure de début du questionnaire, même si les réponses en soirée se font de fait plus souvent sur smartphone.

18. À ce titre, notre échantillon de 41 enquêtes étant composé presque pour moitié de panélistes relativement expérimenté·es, il n'est pas impossible que nous ayons sous-estimé la durée du *tick* (c'est-à-dire surestimé le nombre de *ticks* par minutes). Nous pensons toutefois que cet effet « panéliste » est limité par rapport à l'effet d'âge ou de diplôme, comme le montre le caractère non-significatif de l'effet propre de la date de recrutement dans le panel sur le temps de réponse dans l'enquête MORACT 2.

RÉSUMÉS

Estimer *a priori* la durée d'un questionnaire présente un certain nombre d'intérêts, pour équilibrer ses différents modules, anticiper la quantité de travail pour les personnes chargées de l'administrer, prévoir et maîtriser le temps demandé aux enquêté·es pour y répondre, etc. Pourtant, les rares articles francophones qui s'intéressent à l'administration des enquêtes par questionnaire n'évoquent jamais cette question, et la littérature méthodologique internationale ne converge pas sur une règle unique et empiriquement étayée. L'objectif de cet article est donc double. Dans un premier temps, il passe en revue les différentes méthodes d'estimation de la durée des questionnaires recensées dans la littérature, puis il en évalue empiriquement la fiabilité en testant leur capacité à produire correctement la durée réellement observée des questionnaires d'un corpus de 41 enquêtes en ligne. Dans un second temps, l'article propose une méthode alternative simple (c'est-à-dire facile à mettre en œuvre) et fiable (c'est-à-dire la plus prédictive possible) pour estimer le temps de passation des questionnaires en ligne auto-administrés avant de les tester empiriquement, tout en indiquant comment celui-ci est susceptible de varier en fonction des caractéristiques de l'enquête (support, structure du questionnaire) et des caractéristiques sociales des répondant·es. La conclusion synthétise les apports en proposant dix recommandations pour mieux estimer la durée des questionnaires.

Estimating the duration of a questionnaire *a priori* serves several purposes: balancing its different modules, anticipating the workload for those responsible for administering it, planning and controlling the time required from respondents, and so on. Yet the few French-language articles that address the administration of questionnaire surveys never discuss this issue, and the international methodological literature does not converge on a single empirically grounded rule. This article therefore has a twofold objective. First, it reviews the various methods for estimating questionnaire duration identified in the literature, and empirically assesses their reliability by testing their ability to accurately predict the actual observed duration of questionnaires from a

corpus of 41 online surveys. Second, it proposes a simple (*i.e.*, easy to implement) and reliable (*i.e.*, as predictive as possible) alternative method for estimating the completion time of self-administered online questionnaires prior to empirical testing, while also indicating how completion time may vary depending on survey characteristics (mode, questionnaire structure) and respondents' social characteristics. The conclusion synthesizes the main contributions by offering ten recommendations for improving the estimation of questionnaire duration.

INDEX

Mots-clés : Enquête, questionnaire, durée des questionnaires
Keywords : Survey, questionnaire, questionnaire duration

AUTEURS

PIERRE MERCKLÉ

Professeur des universités, Université Grenoble Alpes, Laboratoire Pacte, pierre.merckle@univ-grenoble-alpes.fr

LAURINE THIZY

Chercheuse post-doctorante en sociologie, Université Grenoble Alpes, Laboratoire Pacte, laurine.thizy@univ-grenoble-alpes.fr